COMPARISON STRATEGIES IN DIFFERENT TYPES OF GRAPHS

A THESIS SUBMITTED TO

THE GRADUATE SCHOOL OF INFORMATICS

OF

THE MIDDLE EAST TECHNICAL UNIVERSITY

BY

ÖZGE ALAÇAM

IN PARTIAL FULFILLMENT OF THE REQUIREMENTS FOR THE DEGREE OF

MASTER OF SCIENCE

IN

THE DEPARMENT OF COGNITIVE SCIENCE

FEBRUARY 2010

Approval of the Graduate School of Informatics

____

Prof. Dr. Nazife Baykal
Director

I certify that this thesis satisfies all the requirements as a thesis for the degree of Master of Science.

____

Prof. Dr. Deniz Zeyrek
Head of Department

This is to certify that we have read this thesis and that in our opinion it is fully adequate, in scope and quality, as a thesis for the degree of Master of Science.

____________________________	____________________________
Assoc. Prof. Dr. Kürşat Çağıltay	Assist. Prof. Dr. Annette Hohenberger
Co-Supervisor	Supervisor

Examining Committee Members

Assist. Prof. Dr. Bilge Say	(METU, COGS)	____________________
Assist. Prof. Dr. Annette Hohenberger	(METU, COGS)	____________________
Assoc. Prof. Dr. Kürşat Çağıltay	(METU, CEIT)	____________________
Dr. Ceyhan Temürcü	(METU, COGS)	____________________
Assist. Prof. Dr. Mine Mısırlısoy	(METU, PSY)	____________________

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 work.

Name, Last name: Özge Alaçam

Signature : ___________

ABSTRACT

COMPARISON STRATEGIES IN DIFFERENT TYPES OF GRAPHS

Alaçam, Özge

M.S., Department of Cognitive Science

Supervisor: .Assist. Prof. Dr. Annette HOHENBERGER

Co-Supervisor: Assoc. Prof. Dr. Kürşat ÇAGILTAY

February 2010, 126 pages

This study aims to investigate the effects of event type (concepts represented by the graph) in graph comprehension with three graph types (line, bar, area) and two graph designs (linear, round) by means of two different task types (trend assessment, discrete comparison). A novel round graph type was designed for that purpose. Five hypotheses were investigated: H1: Graph type affects comparison strategies; H2:

Event type affects comparison strategies; H3: Graph design affects comparison strategies; H4: Graph design and event type interact; H5: Task type affects comparison strategies.

As a method to collect data on subjects' graph perception and comprehension, behavioral (recollected values, word preferences in the description task) and eye-tracking data (scan paths, gaze length, number of fixation, fixation duration and number of transitions) were collected.

As an outcome of this thesis, while the event type and the task type seemed to affect the graph comprehension, the effect of graph type, the graph design and interaction between graph design and event type were partially observed. These results point out that although round and linear graph designs are informationally equivalent, the round graphs are computationally better suited than linear graphs for the interpretation of cyclic concepts. However, grasping trend information for the linear events and making discrete comparisons were achieved with the same effort in both graph designs. This result is not trivial at all, given the fact that participants were not familiar with the round graph design and were confronted with them in this experiment for the first time.

Keywords: Eye Movements, Graph Comprehension, Cyclic Event comprehension

ÖZ

FARKLI GRAFİK ÇEŞİTLERİNDE KARŞILAŞTIRMA STRATEJİLERİ

Alaçam , Özge

Yüksek Lisans, Bilişsel Bilimler Bölümü

Tez Yöneticisi: Yard. Doç. Dr. Annette Hohenberger

Ortak Tez Yöneticisi: Doç. Dr. Kürşat Çağıltay

Şubat 2010, 126 sayfa

Bu çalışma grafikler tarafından ifade edilen olayların/kavramların grafik algılamadaki rolünü incelemeyi amaçlamaktadır. Bu incelemede, 3 çeşit grafik türü (alan, çizgi ve sütün), ve 2 çeşit grafik tasarımı (doğrusal ve dairesel) ile sunulan kavramlar iki farklı görev stili ile değerlendirilmiştir. Bu amaçla, yeni bir grafik tipi oluşturulmuştur. Bu tezde, beş hipotez incelenmiştir. H1: Grafik türü karşılaştırma

stratejilerini etkilemektedir; H2: Kavram, karşılaştırma stratejilerini etkilemektedir; H3: Grafik tasarımı karşılaştırma stratejilerini etkilemektedir; H4: Grafik tasarımı ve kavram türü etkileşim içindedir. H5: Görev biçimi karşılaştırma stratejilerini etkilemektedir.

Veri toplama yöntemi olarak, hatırlanan veri miktarı ve göz hareketi analizleri (odaklanma süresi, odaklanma sayısı, ortalama odaklanma süresi, grafik birimleri arasındaki geçiş sayısı ve bakış sırası) ölçülmüştür.

Bu çalışmanın sonucu olarak, olayı ifade eden kavramın ve görevin çeşidinin grafik algısını etkilediği gözlemlenirken, grafik türü, grafik tasarımı ve grafik tasarımı ile kavram arasındaki etkileşiminin kısmi etkisi gözlemlenmiştir. Sonuçlar, dairesel ve doğrusal grafik tasarımları bilgi bakımından eşit olsa da dairesel tasarımın döngüsel olayların kavranmasında doğrusal grafiklere göre daha uygun olduğunu göstermektedir. Buna karşılık, doğrusal olaylar, iki grafik tasarımında da aynı başarıyla gerçekleştirilmiştir. Bu sonuçlar, dairesel grafik tipinin kullanıcılara yeni olduğu ve deney sırasında ilk defa karşılaştıkları düşünüldüğünde daha da dikkat çekici olarak değerlendirilebilir.

Anahtar Kelimeler: Döngüsel Kavram Algısı, Grafik Algılama, Göz Hareketleri

This work is dedicated to;

My Grandmother

&

My Parents and Sister

who offered me unconditional love and support throughout the course of this thesis and my whole life.

ACKNOWLEDGMENTS

From the initial of this thesis, to the final level, I owe an immense debt of gratitude to my supervisor, Assist. Prof. Dr. Annette Hohenberger. Additionally, I am also heartily thankful to my co-supervisor, Assoc. Prof. Dr. Kürşat Çağıltay. Their encouragement, careful guidance and support were invaluable for me.

I would also like to thank Cengiz Acartürk, who provided me the opportunity to be part of his study on the graph comprehension. Thus, I have entered to this field, which I really enjoy doing my studies.

I would like to show my gratitude to my colleagues from METU Computer Center, Feride Erdal, Nilgün Öner Tangör, Hasan Nadir Derin, Cihan Yıldırım Yücel, Mustafa Dalcı, Ömer Gözü, Nihan Ocak, Ezgi Akkaya, Gülser Cengiz, Yasemin Oran Saatçioğlu and all my other colleagues, which I could not give their name here, for their support during the completion process of this thesis.

I would be remiss without mentioning my friends from METU Subaqua Society. I am grateful for their participation to my experiment, and their easing the process of finding participant for me.

Finally, for their supports and providing the motivation, a special thanks as well to Berrin Doğusoy, Nergis Ayşe Gürel and Selda Eren.

To each of the above, I extend my deepest appreciation.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

LIST OF ABBREVATIONS

ANOVA: Anaysis of Variance

DCT: Discrete Comparison Task

METU: Middle East Technical University

SD: Standart Deviation

STM: Short Term Memory

TAT: Trend Assessment Task

CHAPTER 1

INTRODUCTION

1.1 Purpose of the Study

Graphs are widely used means of visual communication and they are part of both daily life and more professional areas because they are a very efficient way of representing and conveying relation between variables.

There are lots of studies that investigate graph comprehension in terms of different aspects. They generally consist of investigating the effects of graph type by keeping the task constant, the effects of the task by keeping graph type constant, or the combination of these effects. The results of these studies share the conclusion that the given tasks or readers' aims affect the comprehension. Besides, different judgment tasks, for example discrete comparison or trend assessment tasks are better achieved by different types of graphs (Casner 1991; Cleveland 1985a, 1990; Gillan and Lewis 1994; Hollands and Spence 1992; Hollands and Spence, 1998; Lohse 1993; Shah and Carpenter 1995; Simkin and Hastie 1987; Pinker, 1990; Tan and Bensbasat, 1990, 1993). Most of these studies deal with the comprehension of simple graph designs like bar and line graphs in a linear graph structure that is based on a Cartesian coordinate system. On the other hand, pie charts, one of the other studied type of graphs, are based on a polar coordinate system and they are used to present proportions rather than giving absolute amounts (Ratwani and Trafton, 2008, Renshaw et. al, 2004).

In addition to the matching of task and graph type, appropriateness of event type as represented by the graph with the graph type and task should have a role in graph comprehension. While the effects of task and graph type have been well- studied, the effects of event type, in other words the effects of the characteristics of the concepts, needs more investigation.

The tasks generally used in graph perception and comprehension studies can be categorized into two categories: discrimination and trend assessment. Discrimination tasks contain the comparison of two or more data points in the graph. On the other hand, trend assessment is based on the extraction of the general trend (rising, falling etc.) represented by the graph. However, trend assessment studies predominantly deal with the concept conveyed by change-of-state verbs (like increase or decrease) and whether the graph design is convenient to convey this information.

In this study, I am investigating the effect of cyclic concepts of changes of states during graph comprehension by defining them as a sub-task of trend assessment. The winter season which consists of December, January and February is an example of a cyclic concept. It reoccurs each year. To demonstrate, in order to construct information presented by a linear graph about what change happens in winter, the reader should firstly read the data which is presented separately at the two extremity of the x-axis (January and February are at the left side, and December is at the right side of the graph), then integrate the information represented by this points. It can be concluded that the information about a particular season, in other words, seasonal data in a year is not captured directly from the linear graph. In order to provide an immediate comprehension of cyclic concepts like season (winter, summer) or time of day (night, morning) in a graph, these related data should be presented together. This feature is called “proximity” and is one of the Gestalt principles (Zacks and Tversky, 1999).

Cyclic concepts (events recurring in a day or year) and their representations play a role in many scientific areas and are also part of daily life. For example, the observation of the number of species in a specific region/place according to months over the years is one of the most basic research areas in ecological studies and is used to obtain information about seasonal change. This kind of data cannot be obtained from the data of one year only; it is based on the cumulative data of all years which were taken into account. Another example can be given from meteorology. The cumulative data of the amount of rainfall according to months gives us seasonal information from which many other areas like agriculture and the construction industry can benefit. In addition to change according to months, the change in the time of day (night, morning) is also another cyclic concept. For example, circadian biological parameters in the human body, light permeability (ecological parameter) change in the human body, or wind speed (meteorological parameter), are examples of these cyclic concepts. Although cyclic concepts are widely used, as illustrated above, they are generally represented by linear graphs. However, linear representations do not have the right properties to highlight their inherent cyclicty (see Figure 1-1).

In order to investigate this phenomenon, a novel graph type was designed by taking most relevant features of the Cartesian coordinate system and the Polar coordinate system to represent these types of concepts ideally. Briefly, in a Cartesian system (Figure 1-1a), two different scales are represented with two axes: the x-axis and the y- axis. On the other hand, the polar system (Figure 1-1b), is based on angles, and distance from origin. The data is shown on lines, and the respective value of this data is shown by distance from origin. Besides, each line corresponds to a particular angle.

Linearity of the Cartesian coordinate system as represented by the y-axis value and the sphericity of the polar coordinate system that provides circularity have been combined in order to design a novel round graph (Figure 1-1c). Consequently, most related entities such as months which constructs winter season are presented together in this graph for extraction of cyclic content. Therefore, one of the investigated areas in this thesis is whether the circularity of the graph design affects the comprehension of cyclic events. On a more abstract level, I investigate the effect of isomorphism between content and form: cyclic events are isomorphic with round graphical representations whereas linear events are isomorphic with linear graphs.

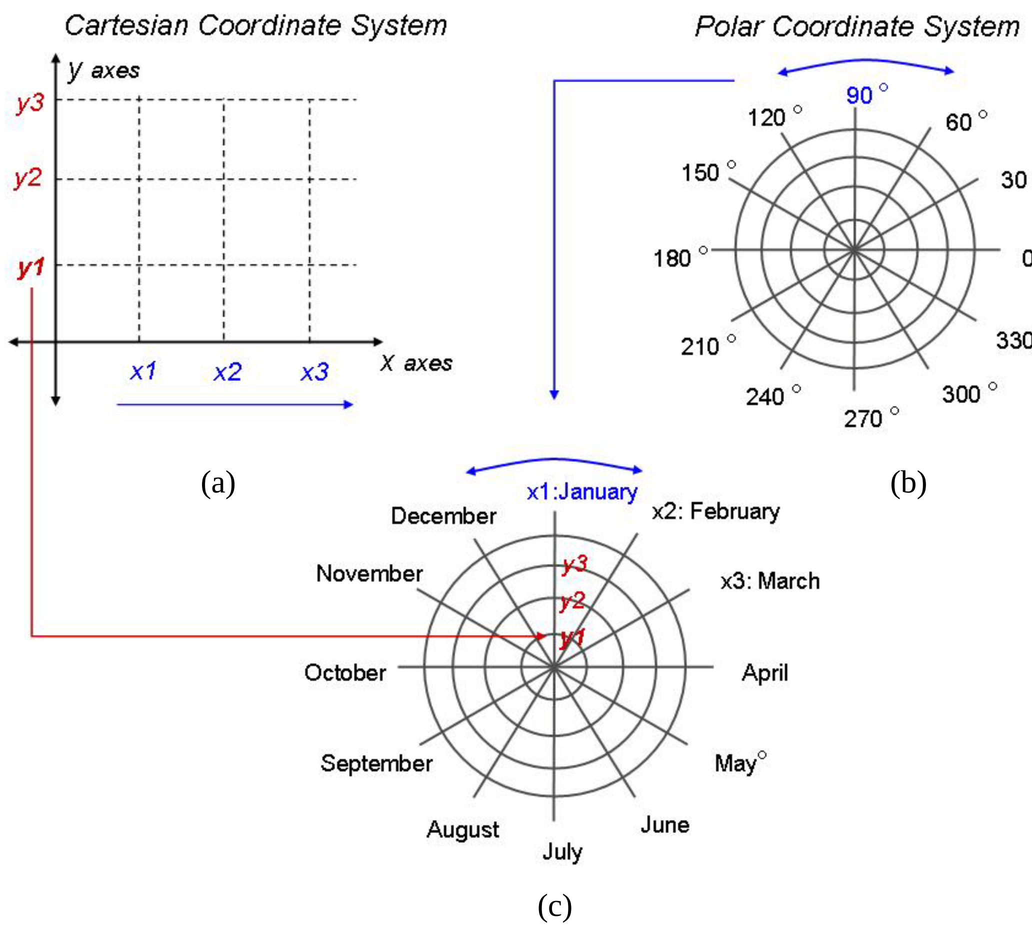


Figure 1-1: Round Graph Structure with relevant features imported from Cartesian and Polar Coordinate Systems

Additionally, for a task, such as assessing the change of a parameter according to years, the participant should need to compare the start and end values. The novel round graph positions start and end points on a cycle, and brings them nearer for comparison. This thesis also aims to explore whether it is easier to achieve this type of comparison in a round graph than in a linear graph.

Like in a usual graph representation, this round graph has also a co- ordinate system and respective labels, a data region with graphical components, such as points, lines or bars.

As a method to collect data on subjects' graph perception and comprehension, the investigation of eye movements such as scan paths, fixations and durations provides a powerful tool for observing bottom-up (perceptual) and top-down (conceptual) high level cognitive processes. Besides, they are tightly linked with moment-to-moment goals and sub-tasks. Therefore, eye movements of graph readers were analyzed in all the experiments in this research with an eye tracker (see section 2.4).

1.2 Research Questions

As introduced above, I have four main assumptions that guide my research questions. The first two of them are related to graph perception and comprehension. First, eye movements are affected by the perceptual properties or features of visualization. All graph types have different types of perceptual features and the differences between perceiving and comprehending graphs will be evaluated by the eye tracking method. Second, properties of visual elements also affect higher level cognitive process. As will be discussed below in more detail, line graphs are efficient tools for trend perception, bar graphs are more suitable for comparing discrete information. Also, an area graph can be helpful in representing the coverage of related variables. The other two are assumptions that are based on eye tracking as the main methodology of this thesis. The first is an eye/mind hypothesis (Just and Carpenter, 1976), which states that, when looking at a visual display and completing a task, the location of a fixation indicates the area of interest. The second assumption is that the duration of fixations and the pattern of eye movements (scanpath) in general are dependent upon how easy or difficult the display is to process (Renshaw et al, 2004).

This thesis explores effects of event type (concepts represented by the graph) in graph comprehension with three graph types (line, bar and area) and two graph designs (linear and round) by means of two different task styles (trend assessment and discrete comparison). My overall research question is “Do types of graphs, designs of graphs, types of events and types of tasks affect comparison strategies during graph comprehension?”

I have five hypotheses. The first one is that different types of graphs (area, line, bar) are comprehended by using different comparison strategies. The second one is that comparison strategies are also affected by the semantic properties of variables. For example, while months can be comprehended cyclically, years can be considered linearly. These characteristics of a variable affect the comprehension. The third one is that graph design (linear, round) affects the comprehension strategies. The fourth one is that graph design and event type interact, that is, isomorphism between graph design and event type facilitates comprehension and finally, the graphs are comprehended differently according to a given task (such as trend assessment or discrete comparison). Here is a brief summary of my hypotheses:

H1: Graph type affects comparison strategies.
H2: Event type affects comparison strategies.
H3: Graph design affects comparison strategies.
H4: Graph design and event type interact.
H5: Task type affects comparison strategies.

As an outcome of this research, differences are expected to be found in comparison strategies which are used for extracting information from different types of graphs such as line, bar and area graphs. In addition, it is also expected that graph design (linear or round graphs) will have effects on the understanding of the events (trend or cyclic events).

1.3 Organization of the Thesis

In the following, I will present a literature review on graphs and their general properties (section 2.1), theories and studies of graph perception and comprehension (sections 2.2 and 2.3), eye tracking technologies in cognitive science (section 2.4), and tasks and paradigms in graph comprehension studies (section 2.5).In Chapter 3, I will present the methodology of the present study. The results will be presented in Chapter 4 and I will discuss and draw some major conclusions and give an outlook on possible future work in Chapter 5.

CHAPTER 2

LITERATURE REVIEW

2.1 Graphs and Graph Properties

Graphs are widespread both in the daily life among non-graphic specialists and in most of the professional areas due to their efficiency in presenting and allowing the extraction of quantitative information and relationships between two or more variables. Graphs are very commonly used tools in the representation of data and they can provide detailed information clearly and concisely (Fischer, 2000; Kosslyn 1989*;* Larkin and Simon, 1987, Trickett et al., under review).

Graphs can be used for extracting a single piece of information from the graph, comparing two or more pieces of information, or they are also frequently used in determining trends (increasing, decreasing). Additionally, graphs are used for extracting information that may not even be explicitly represented in the graph (Trickett et al., under review).

In order to represent information for the different aims given above, different graph designs are being used. Most popular graphs are based on three basic designs: line graph, bar chart or pie chart. Different graph designs differ in pointing out specific features about the data. (Renshaw et al., 2004; Ratwani and Trafton, 2008)

Like the comprehension of all kinds of information, graph comprehension is also dependent on the given task or the readers' aim. According to Pinker (1990), "different types of graphs are not easier or more difficult in general, but are easier or more difficult depending on the particular class of information that is to be extracted". Lots of studies, whose common findings on graph perception and operations indicate that different graph types are better suited for different judgment tasks, support Pinker's view (Casner 1991; Cleveland 1985a, 1990; Gillan and Lewis 1994; Hollands and Spence 1992; Lohse 1993; Shah and Carpenter 1995; Simkin and Hastie 1987; Tan and Bensbasat, 1990, 1993; Hollands and Spence, 1998). In other words, given the correct design, readers can perceive relationships between variables and proportions of change, depict and extrapolate trends beyond the given information, and compare the variables presented in the graph more efficiently and effectively (Renshaw et. al, 2004). This topic will be detailed in Section 2.3.

Graphs are considered as one type of symbolic diagrams, which are widely used means of communication in scientific and technical areas. Iconic diagrams and schematic diagrams are other types of symbolic diagrams. Although there are fundamental differences between certain aspects of diagrammatic and textual representations, among these types of symbolic diagrams, graphs and charts are more formalized visualizations by language-like conventions (Schmidt-Weigand, 2006). Therefore, they are more useful for making linguistic analyses (Hegarty et al., 1991). Like textual representations, graphs also express relations and properties of objects (Gurr, 1999).

While looking at a graph or reading a sentence, internal mental representations of the content presented by these stimuli are constructed and then these textual or visual representations can be understood. This means that construction of this mental representation is dependent on the task given and on salient features of graphs/sentences which makes the extraction easier or more difficult (Schnotz, 2002).

Studies about the way of representing information with different means and their comprehensibility indicate two important concepts: informational and computational equivalence. Being informationally equivalent corresponds to having the same content and a set of information expressed in two different representations. For example, if the textual and visual representations of a statement “X is bigger than Y” have the same variables and same relations, then they are considered as informationally equivalent. Additionally, in the situation when they are informationally equivalent, in order to be called computationally equivalent, the retrieval of the information in one representation should be equally easy as the retrieval from the other representation. For the given example above, the statement represented by the graph should be retrieved as easily as the same statement from the textual representation. Furthermore, in order to compare the comprehensibility of two or more different visual representations - graph designs or graph structures - their informational and computational equivalence should be also considered (Larkin and Simon, 1987; Palmer, 1978; Schnotz, 2002).

Typically, a statistical graph consists of a co-ordinate system with two main axes and their respective labels, a data region which contains graphical components, such as points, lines or bars (Fischer, 2000). By these properties, graphs have syntactic, semantic and pragmatic levels like language (Kosslyn, 1989; Schnotz, 2002). Kosslyn states that a syntactic analysis focuses on properties of the lines and regions themselves; they are not interpreted in terms of what they convey or refer to. In addition, the semantic analysis focuses on the meanings of the configurations of lines, what they demonstrate (e.g. axes labels, etc.). The semantic analysis can be considered as the literal reading of each of the components of a chart or graph and the literal meaning that arises from the relationship between these components. And finally, he defines the pragmatic analysis as meaningful symbols conveying information above and beyond the direct semantic interpretation of the symbols. In addition, pragmatic considerations govern the relationship between the information in a display and the readers' purposes and needs. For example, Zacks and Tversky (1999) showed that when subjects see bar graphs, they describe discrete contrasts in the data; when they see line graphs, they describe trends. As a result, they concluded that both the graph type and the conceptual domain conveyed by the task affect viewers' descriptions.

On the syntactic level, in order to function effectively, the graph should be lucid, sound, and laconic. Lucid means that some single object or relation in the graph represents one single object or relation in the represented text rather than more.

Sound means that a representation permits a valid, well-formed and complete construction of a corresponding text, that is, every (relevant) object in the text appears in the graph, too. Laconic means that distinct objects in the graph refer to distinct objects in the represented world (Renshaw et al., 2004). This property has also support from one of the Gestalt principles, namely proximity. This principle states that entities placed in close proximity to one another are assumed to be related and a task is executed more successfully when the type of task is compatible with the perception of the information sources relevant to the task. This phenomenon is a simple and powerful way of emphasizing the relationship between data entities (Renshaw et al., 2004; Wickens and Carswell, 1995). Moreover, this property also provides directness to the graph, which increases the potential for semantically direct interpretation.

In addition to these, another factor that affects the graph comprehension is short term/working memory and long term memory constraints. Kosslyn (1989) states that since short term memory/working memory has a limited capacity, this constraint affects our ability both to integrate syntactic information and to hold semantic information in mind during graph comprehension. Thus, the complexity of a graph will be major factor in determining its comprehensibility. Moreover, long term memory has also some major constraints, most importantly the person's domain knowledge. The way of interpretation of a graph, both at the level of semantics and pragmatics, depends on which stored information is most closely associated with the way the stimulus properties of a graph are categorized.

In addition to perceptual properties of graphs, task completion on graphs also depends on the nature of a task in terms of short term/working memory limitations. For example, if the task involves integration operations, then performance is better when the graph design incorporates features that maximize the integration and extraction of information. This is achieved through perceptual features such as spatial proximity, similarity of color, shape and size provided that the end result is compatible with the task. Wickens and Carswell (1995) suggest that the reason for this is that such design features promote parallel processing and/or assist in the viewer's integration task. This reduces demands on working memory and subsequently enhances task performance. On the other hand, if the task requires that individual entities be processed separately, then their arrangement should be best organized to allow this through perceptual separation. Lohse (1993) also reaches a similar conclusion by stating that graph comprehension is facilitated if the desired information is represented explicitly in the graph since STM can hold three chunks and lasts for about seven seconds.

2.2 Graph Perception and Comprehension Theories

There is a large literature about graph perception and comprehension. Most studies involve both areas since they are considerably intertwined.

Most of the old literature on graph perception is based on the “incremental estimation model” which is a theoretically good model on perceptual discrimination of data shown in graphs (Petrusic, 1992; Vickers, 1980). However, since the graph usage is not limited to only discrimination of two or more quantities, but also may involve discrimination of proportions or percentages rather than absolute amounts, this model cannot properly account for all graph perception tasks (Hollands and Spence, 2001).

According to Cleveland (1985), several perceptual features are involved in the perception of graphs (e.g. length, angle, area) that differ in terms of their accuracy. The graph reader chooses the first property available from this set. However, according to Hollands and Spence (2001), observers sample from the set of available perceptual features rather than choosing the most effective one and using it consistently. The hierarchy of these perceptual features proposed by Cleveland is listed w.r.t to their accuracy levels below (Table 1-1).

Table 1-1: Hierarchy of perceptual features for discriminating proportions in graphs (Cleveland, 1985)

Accuracy	*Perceptual feature*
Most accurate	Position along a common scale
	Position along identical, non-aligned scales
	Length
	Slope / Angle
	Area
Less Accurate	Volume
	Color hue –saturation - density

During the perceptual process, the graph reader may choose between the perceptual clues or features, however, the available perceptual features may vary across graph designs. In addition to availability, the effective usage of the available features may also vary among graphs (Hollands and Spence, 2001).

The results of Hollands and Spence's study show that participants needed longer with pie charts than with divided bars, because the perceptual features that are available from divided bars rank higher in the hierarchy given in Table-1 than those available from pies. From this result, they concluded that observers sample from the set of available perceptual features when they are needed during comprehension process.

There are many theories of graph comprehension in literature; however, the three stated below are the most comprehensive ones in terms of covering any type of graph, making predictions about performance, and explaining how people extract information from graph (Trickett et al., under review). Like graph perception, graph comprehension is also supported by a very broad literature. Trickett et al.'s study, (under review) provides detailed investigations about tasks and whether current general theories account for all tasks in a well organized format. These current theories are Freedman and Shah's "Construction Integration Theory" (Freedman and Shah, 2002; Shah and Shellhammer, 1999; Shah, et. al., 2000; Shah, 2002; Shah et. al., in press), Pinker's "Propositional Model" (Pinker, 1990), and Lohse's "UCIE model" (Lohse, 1993).

Freedman and Shah's Construction Integration Theory

Freedman and Shah base their interpretation of graph comprehension processes on the construction-integration model of text comprehension (Kintsch, 1988). During the construction phase, the graph reader notices the visual features of the graph, and iterations occur between the reading of the graph and the legend. During the integration phase, the visual features from the construction phase are comprehended using prior knowledge about graph and domain knowledge. This information is activated early in the construction phase, facilitating chunking of the visual elements. Domain knowledge refers to any mental representation of the content of the graph (Freedman & Shah, 2002). Domain knowledge serves making numerical relationships more apparent and facilitating inferential processes. Readers who lack domain and/or graph knowledge will be less accurate in their interpretations and will produce only surface level descriptions of the graph (Freedman & Shah, 2002**).** Freedman and Shah assume that working memory is limited, and as a result, graph readers iterate between the construction and integration phase until the information is comprehended.

Pinker's Propositional Model

Pinker provides a more detailed graph comprehension model. The graph reader first scans the graph (scan patterns are not specified), and based on its perceptual properties, constructs a visual array. The graph reader then forms a propositional representation of the graph from this visual array. The appropriate graph schema, which allows the reader to create a conceptual question, is triggered by the propositional representation. The conceptual question is the information the graph reader wants to extract from the graph. The conceptual message is the actual information extracted from the graph.

Four main processes operate on the propositional representation of the graph, as given below;

(1) a matching process which allows the reader to recognize a graph as being of a particular type,

(2) a message assembly process that translates the visual information into conceptual information,

(3) an interrogation process which is used when needed information is not present from the message assembly,

(4) Inferential processes such as being able to perform mathematical operations from the context (Pinker, 1990).

Lohse's UCIE theory

Lohse's UCIE (Understanding Cognitive Information Engineering) theory aims at making quantitative predictions about how long it will take a graph reader to extract specific information from a graph. It contains the following operations; comparing two units in memory (digits, colors, words, shapes, forms), interpolating on a linear scale, making saccades, and making a perceptual judgment (Lohse, 1993). When interpreting a graph, Lohse suggests that

(1) Early visual processes detect and encode visual features such as shape and color,

(2) Short term memory (STM) builds a visual description from the early visual processes

(3) Information in STM triggers an association to a memory trace in long term memory (LTM) which instantiates the graph schema. The graph schema directs the interpretation of the graph.

In order to evaluate these comprehension theories, Trickett et al. (under review) use three levels of graphs (simple, medium and complex) and three levels of questions (read-off, integration, inference). According to the complexity level of graph types, simple graphs have only one or two variables and require no domain expertise to be interpreted. Medium-complexity graphs typically represent more variables, or levels of variables, and some domain knowledge may be needed to fully interpret these graphs. Complex graphs represent many variables and frequently substantial domain knowledge and specialized graph-reading skill are required to interpret them successfully (like meteorological and scientific visualizations).

In addition to graph type, the complexity of the task is also categorized. One of these three question types mentioned in Trickett's study (under review) is read-off questions, which ask for one piece of explicitly represented information to be extracted from the graph. Integration questions ask for multiple points to be extracted, which may or may not be explicitly represented in the graph. Inference questions require the graph reader to go beyond the explicitly represented data and may also involve read-offs and integration in order to derive an answer. All three theories account for read-off and integration questions in simple graphs. Table 2-2 shows steps in answering the integration question in simple graphs.

Table 2-2: Integration Question type for simple graphs (taken from Trickett et al. (under review, p.26)

<table>
<tr><td colspan="3">Task: Describe the general trend in the cost of tuition for public colleges, 1980-2000</td></tr>
<tr><td>Pinker</td><td>Lohse</td><td>Freedman & Shah</td></tr>
<tr><td>1. Early visual processes construct all possible relationships among graph elements (according to Gestalt principles)

2. Build propositional representation of whole graph (all scan patterns inferred)

3. Activate line graph schema

4. Devise notation for conceptual question: V1 range = 1980-2000; V2 college type = private; V3 trend = “?”</td><td>1. Look at legend and identify public college

2. Scan to circle and discriminate circle

3. Scan to 1980, discriminate 1980, read 1980

4. Scan back to black circle on Legend

5. Scan to 2000, discriminate and read

6. Scan up to circle at 1980</td><td>1. Notice visual features of graph, scanning between legend, axis and lines

2. Knowledge of line graphs and domain knowledge activated

3. This allows realization that tuition has increased over the years, leading to expectations of cost

4. Relevant information is represented in the visual features</td></tr>
</table>

Table 2-2 (continued)		
5. Realize time is on x-axis (from graph schema) 6. Translate visual information into quantitative information (on x-axis look from 1980 to 2000) 7. Realize 1980 corresponds to $2000 and 2000 to $4000 8. Take difference between value for 2000 and 1980, realize it's positive, declare increasing trend	7. Scan over to circle at 2000 8. Interpolate on linear scale; make perceptual judgment	5. Pattern perception allows interpretation of trend as increasing

2.3 Graph Perception and Comprehension Studies

Addition to theoretical studies, there are lots of experimental studies on graph comprehension. Although categorizing graph comprehension studies is challenging, since they are intertwined, generally they can be investigated under three broad categories: task-dependency studies, comparison studies, and the analysis of the tasks themselves.

As mentioned in section 2.1, complexity of the task differs in comprehension and also complexity of the graph designs has a role in comprehension (Trickett et al., under review). This study shows that different strategies are being used in graph comprehension and these depend on the increasing complexity of the graph. In addition to the studies which show that the graph comprehension performance is task-dependent (Cleveland, 1985; Kosslyn, 1994; Shah & Hoeffner, 2002), there are lots of studies that investigate which information is more effectively comprehended given which graph type. The studies that compare bars and lines from that perspective, show that readers tend to describe discrete information (like exact values, maximum points, higher, lower, greater than, less than relationships) if the information is given with bar graphs. On the other hand, readers tend to describe trends (rising, falling, increasing, decreasing) with line graphs (Fischer et al., 2005; Meyer et al;, 1997; Zacks and Tversky, 1999; Kosslyn, 1993, American Psychological Association, 1994; Levy et al., 1996).

This difference in the interpretation of information from different graphs, seem to be based on principles of cognitive "naturalness" which is supported by the Gestalt principles underlying figural perception. According to these principles, bars are convenient for conveying categorical information and lines for ordinal or interval data. Because in bar graphs, each value is represented as a separate bar, and this points to separate entities or categories, whereas in line graphs, values are connected by a single line, and this points to all the values belonging to the same entity (Zacks and Tversky, 1999).

Summarizing so far, graph comprehension processes are based on an interaction of bottom-up (organizing) and top-down (the selection of task-relevant information) activation of cognitive schemata (Schnotz, 2002).

In addition to information type, matching the order of the words which are asked in the task, and the order of the labels in the graph has also an effect in graph comprehension. Fisher's study (2005) shows that verification times were faster when the spatial order of the statement matched that of the labels in the graph.

Fisher (2000) argues that adding perceptual features like depth to graphs is not an effective way, because his study shows the comprehension is affected by the experimental task used in this experiment and generally the experimental task requires comprehension of the quantitative relation between two variables instead of mere feature perception or discrimination.

2.4 Eye tracking in perception, cognition and graph comprehension

The usage of eye tracking tools in cognitive science to study the relationship between fixations and cognitive activity has started in the early 1970's since eye tracking technologies were very invasive prior to the 1970's (Jacob, 2003). Since then, a huge number of studies involving perception and action, cognition, and language, among others, has been conducted using this technology. In addition to the fact that eye movements are central to the visual system extremely fast, and metabolically cheap, they have a lower threshold for being triggered as compared to other motor movements. This makes eye tracking a very powerful and accurate tool to investigate cognition (Richardson et al., 2007).

During comprehension of a visual stimulus, both bottom-up and top-down processes happen at the same time. The eyes appear to be driven by both visual properties of the stimulus (bottom-up) and top-down effects of knowledge and expectations (Henderson, 2003). In the study of Richardson et al. (2007), it is proposed that there is continuity between perception and cognition, and eye movement patterns during cognitive activity show striking resemblance to those during the perception and manipulation of objects in the world. Richardson states that a “low-level” motor process, such as eye movements, can actually have a role in “high-level” cognitive processes.

Eye tracking relies on two main assumptions. The first one is that when looking at a visual display and completing a task, the location of a fixation indicates the area of interest, and this assumption is called the “eye/mind” hypothesis (Just and Carpenter, 1976). The second is that the duration of fixations and the pattern of eye movements (scan path) are indicators of how easy or difficult the display is to process (Renshaw et al., 2004).

In addition to the importance of location of fixation, another important concept that should be taken into account is visual angle and visual acuity. The visual field of the reader during a fixation can be divided into three regions. These are foveal, parafoveal, and peripheral regions. In the foveal region, that is the central 2° of vision, the acuity is very sharp and clear. When the distance from the fixation point increases, the acuity decreases. In the parafoveal region, which is between 2° and 5° of vision on both sides of fixation, the acuity is not as good as in the foveal region. And in the periphery that is also called extrafoveal region (the region beyond the parafovea), there is no acuity at all (Rayner, 1998). However, when the object in the parafoveal or peripheral region is clear enough in terms of clarity and understandability, the next fixation duration to that point becomes shorter, or it is also possible to be identified without a saccade (Pollatsek et al., 1984).

Wickens and Holland's study (2000) on graph comprehension shows that eye movements are also influenced by the nature of the task being executed. For example the scan path and eye movement values differ between free search and targeted search (Renshaw, 2004; Richardson et al., 2007).

Determining the ambiguity or difficulty of comprehension is also easily captured by an eye movement analysis. The study of Land and Hayhoe (1999) shows that eye movements are tightly linked with the moment-to-moment purpose of the reader and that they differ if there is ambiguity in the input or the variables in the input exceed short term memory capacity.

Eye movement features during graph comprehension such as fixation duration, gaze time, number of fixations (instances when the eye remains relatively still within a particular location), the occurrence of regressions (transitions between the areas of interests), and a number of variations on these measures can be used to investigate moment-by-moment cognitive processing of a graph by the reader in order to assess comprehension strategies and the effectiveness of the graph (Just & Carpenter, 1980; Rayner et al., 1989; Renshaw, 2004).).

Generally, longer fixations are considered as an indication of more difficult processing, and increased durations on a sentence or on a picture would indicate the participant's difficulty in interpreting it. (Jacob, 2003; Underwood, 2004)

According to BS in 2.4: According to the definitions of the mostly analyzed eye movements by Jacob (2003), the fixation is described as a relatively stable eye movement over some minimum duration (usually 100–200 ms), and with a velocity below some threshold (usually 15–100 degrees per second). Another common eye movement is gaze duration which is cumulative fixation durations and the average spatial location of a series of consecutive fixations within an area of interest. Additionally, number of fixations corresponds to the overall number of fixations and is considered as an indication of ineffective search (Goldberg & Kotval, 1998; Kotval & Goldberg, 1998). Furthermore, Fixation duration mean is another parameter used in eye movement analysis. Longer fixations (and perhaps even more so, longer gazes) are generally believed to be an indication of a participant's difficulty extracting information from a display (Fitts et al., 1950; Goldberg & Kotval, 1998). Moreover, another commonly made eye movement analysis is Scan path analysis which investigates the spatial arrangement of a sequence of fixations. It can also be a derived measure such as the number of transitions between areas of interest, which is an indicator of the efficiency of the arrangement of elements. Finally, the area of a display or visual stimulus that is of interest to the researcher is called Area of interest.

2.5 Tasks and Paradigms in Graph Comprehension Studies

In the investigation of graph comprehension strategies, different types of measurement tools are used. The most commonly used method is to ask the graph reader to give a verbal or written report about the graph and the relation represented by it. Another method, which is mostly focused on in investigations of the effect of a task on graph comprehension rather than in comparison studies and effects of graph types, is to ask readers to produce a graph according to a given verbal or written definition.

The Sentence graph verification paradigm is another commonly used method in cognitive tasks (Clark and Chase 1972; Bower and Clapper, 1989; Feeney et al., 2000; Feeney and Simon, 2000; Feeney and Webber, 2003; Fisher, 2000; Renshaw, 2004; Underwood, 2004). In this paradigm, the graph and sentence are presented concurrently and readers are asked to make decision whether the sentence is an accurate description of the graph as quickly as possible.

CHAPTER 3

METHODOLOGY

In the methodology chapter, I will begin with reporting the details about the pilot studies two of which were conducted in order to test the procedure and the completion time before starting with the final version of the experiment. Then, in sections 3.2 and 3.3, I will give details about the participants, materials and apparatus used in this thesis. The procedures of the two parts and the data collection will be presented in sections 3.4 and 3.5, respectively. Finally, I will elaborate on the analysis procedure.

3.1 Pilot Studies

The first pilot study which contains line graphs with one variable was performed with 8 subjects. Number of fixation, gaze time, fixation duration, reaction time, number of transitions between the elements of the graph (scan path), and recognition were measured as dependent variables. Preliminary analyses showed that graph type had no main effect on the understanding of the events. Trend events (e.g., change of temperature according to years) needed less gaze duration than cyclic events (e.g., change of temperature according to months on a seasonal scale), however, event type did not differ in the other dependent variables. For all dependent variables, interactions between graph type and event type were obtained. Cyclic events were less well recognized, were looked at longer and more often with more transitions in linear graphs as compared to trend events. However, trend and cyclic events were equally well interpretable in round graphs. After obtaining these results, according to participants’ comments, some corrections were made. Some of the sentences used in Pilot-1 that seemed to be open to interpretation have been eliminated and the visibility of the labels was changed (Alaçam, et al., 2009).

Then a second pilot study which contains all graph designs (line / bar /area) was performed with 6 subjects in order to obtain information about test completion time and the procedure before starting to conduct the final version of the experiment. In addition, evaluation forms which ask for preferences of participants about graph designs and concepts were added to the procedure. Additionally, the experiment was adjusted to be finished in approximately 45 min. The final procedure was tested with one participant, and then the main experiment was started. In the following, information about the sample, material and apparatus, procedure, variables, data collection and data analysis of the main study will be presented.

3.2 Participants

40 subjects (28 female, 12 male), undergraduate and graduate university students ranging between 20 and 33 years of age, participated in this study. The age distribution is given in Figure 3-1. In order to obtain demographic data and information about their prior knowledge about graph usage, a questionnaire was carried out (given in Table 3.1).

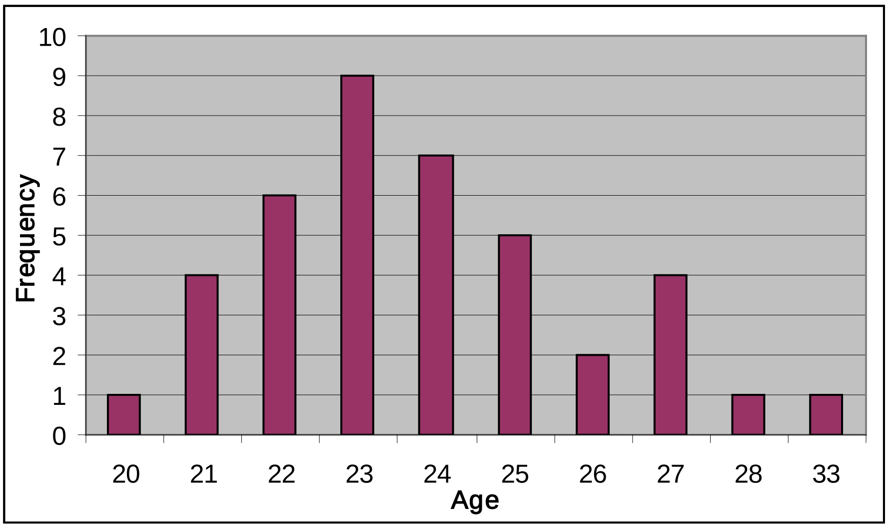


Figure 3-1: Age distribution of participants

Table 3-1: Demographic data of participants: number of subjects, according to various demographic parameters

		Linear	Round
Gender	Female	16	12
	Male	4	8
Major	Natural Science	15	18
	Social Science	5	2
Level of Education	Undergraduate	10	9
	Graduate	10	11
Statistic Course	Yes	11	12
	No	9	8

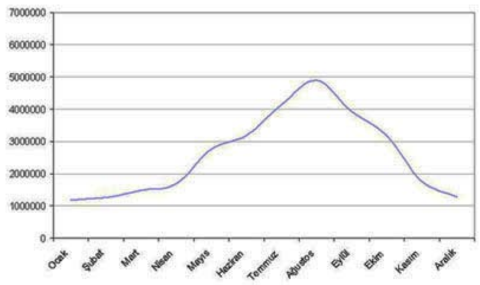
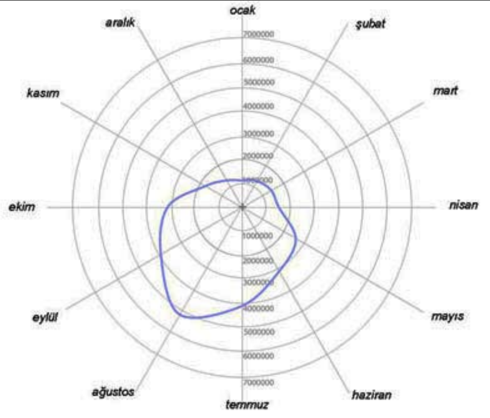
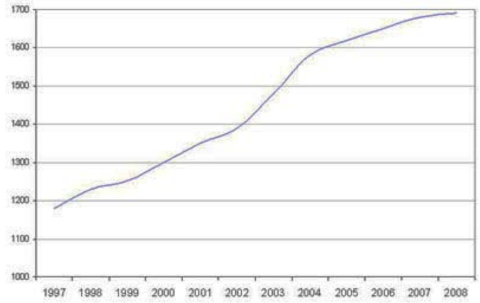
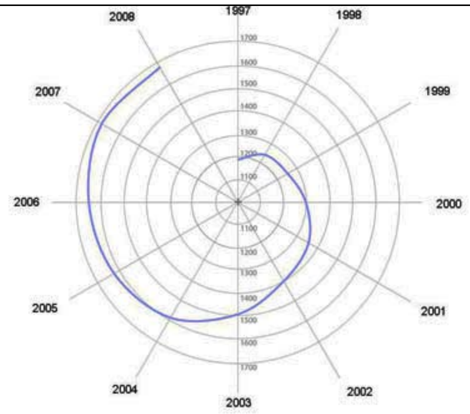
3.3 Materials and Apparatus

All experiments were conducted in METU's Computer Center, in the Human Computer Interaction Research and Application Laboratory. Participants' eye movements were collected by a Tobii 1750 Eye Tracker and analyzed with Tobii Studio software. The participants were seated at a distance of approximately 60 cm from eye tracker.

3.3.1 Graphs

Three types of simple graphs were used in this research: bar, line and area graphs. Graphs in this study have only one variable and require no domain expertise to interpret. In addition, two graph designs were used in this experiment: linear and round graphs (see the line graph samples in Table3-2).

Table 3-2: Sample line graphs for each graph design and event type

<table>
<tr><td></td><td>Linear Graph</td><td>Round Graph</td></tr>
<tr><td>Cyclic Event</td><td>
</td><td>
</td></tr>
<tr><td>Trend Event</td><td>
</td><td>
</td></tr>
<tr><td colspan="3">Translations of several months in Turkish to English:
Ocak: January , Şubat: February………. Aralık: December</td></tr>
</table>

A novel round graph type, similar in shape with polar graphs used in specific engineering areas, was designed for that purpose. All stimuli were created by using the combination of MS Excel, MS Power Point and Photoshop CS4 Applications. This novel round graph type can also be considered as a simple graph since it represents one variable and no domain knowledge is required in order to interpret it.

Furthermore, no domain knowledge about the contents represented by the graphs is needed to interpret these graphs. Additionally, the round graphs used in this experiment are informationally equivalent to linear graphs since they were presented with same contents and a set of information.

3.3.2 Tasks- Questions - Sentences

The comprehension of two different types of events was investigated in this experiment; these are trend events (such as change in the number of students according to years) and cyclic events (such as change in temperature according to seasons). Event type is a within subject variable. Depending on the feature of the content of the event presented by the graph, there are four patterns: rising, falling, v-shape (falling and then rising), and non-v shape (rising and then falling). However, these patterns are not variables. Different patterns were chosen in order to allow for a wider generalization of the results from the effect of the event type variable.

The contents of the first group, namely “trend events”, consist of events about the change of temperature, electric consumption, amount of erosion, air pollution, number of monk seals in years. The contents of the second group, namely “cyclic events”, consist of events about change in rainfall (in months), light permeability (in hours of a day), wind speed, metabolism speed, number of species (in months), number of tourists (in months), and the amount of zooplankton (in hours of a day).

3.3.2.1 Part 1: Recollection Evaluation

In order to measure the recognition of information represented by the graphs, after viewing the graph without a time limitation, the participants were asked either to describe or to draw the relationship between variables and the quantitative/numerical information represented by the graph.

For the “describe” question which is asked to assess the verbal information extracted from the graph, the participants are asked to describe the relationship with the values in written format. The time labels were always given in the description on the sheet, since there were only three types of time line (hours, months, years) in the experiments.

For the “drawing” question, the empty graph with reference lines and time labels was given and it was asked to draw the relationship with the values to assess the visual information extracted from the graph.

3.3.2.2 Part-2: Judgment Task

In this part of the experiment, the sentence-graph verification paradigm (see section 2.5) was used. In this paradigm, the participants see the graph and sentence concurrently and are required to decide whether the sentence is an accurate description of the graph (Feeney et al., 2000).

The sentences in this experiment (see Table 3-3 for the samples) ask for more than one piece of explicitly represented information even when the data presented together. In order to make decision about concept for change of state (about change in seasonal scale as cyclic event, amount of erosion by the year as linear event), the participants have to look at and process more than one point and their respective fields on the graph. Therefore the sentences used in this thesis were accepted as integration question. In additions to these read off questions, judgment sentences about cyclic events require integration since they ask the participant to read off multiple data points and then integrate that information using some kind of mental operation about the cyclicity of the event. Moreover, sentences which ask for “higher/lower” judgments were also used.

In order to decide about cyclic events (asked by sentences 1 and 2), the participant needs to extract cyclic concept information in addition to a trend assessment. Moreover, in order to assess the change according to years (asked by the sentences 1 and 2 for trend events), the participant has to make a comparison with the trend assessment due to the nature of the tasks. Sentences 3 and 4 for both cyclic and trend events are asked to make a decision about and a discrimination between two points. The sentence formats are given in Table- 3-3 below.

Table 3-3: The sentence formats for each task and event type

Cyclic Event (Cyc)	Trend Assessment (TAT)	S1: Eymir Gölünde gece vaktinde zooplankton sayısı artar. *(In Lake Eymir, the amount of zooplankton increases at night)*
		S2: Eymir Gölünde öğle vaktinde plankton sayısı azalır. *(In Lake Eymir, the amount of zooplankton decreases at noon)*
	Discrete Comparison (DCT)	S3: Eymir Gölünde saat 7:00’daki zooplankton sayısı, 13:00’dakinden fazladır. *(In Lake Eymir, the amount of zooplankton at 7:00 am is higher than at 1:00 pm.)*
		S4: Eymir Gölünde saat 19:00’daki zooplankton sayısı, 3:00’dakinden azdır. *(In Lake Eymir, the amount of zooplankton at 7:00 pm is lower than at 3:00 am.)*
Trend Event (Tr)	Trend Assessment (TAT)	S1: Akdeniz’de gözlemlenen Akdeniz Foku sayısı, yıllar geçtikçe azalmıştır. *(The number of monachus monkseals has increased over the years.)*
		S2: Akdeniz’de gözlemlenen Akdeniz Foku sayısı, son yıllarda azalmıştır. *(The number of monachus monkseals has increased in the last years.)*
	Discrete Comparison (DCT)	S3: Akdeniz’de 2001 yılında gözlemlenen Akdeniz Foku sayısı, 2006 yılındakinden fazladır. *(The number of monachus monkseals observed in 2001 is higher than in 2006)*
		S4: Akdeniz’de 2004 yılında gözlemlenen Akdeniz Foku sayısı, 2001 yılındakinden azdır. *(The number of monachus monkseals observed in 2004 is lower than in 2001)*

3.3.3 Graph- Sentence Stimuli

For Part-2 of the experiment, 480 graph-sentence stimuli (Graph type (2) × Graph design (3) × Event Type (2) × Content (5) × Sentence (4) × True/false (2)) were created. For each graph type group, 4 randomized versions consisting of 43 graphs (including all types of variables) were picked up randomly. 6 graphs in the familiarization part and 37 graphs in main experiment were presented to the participant. During the test, for each version, the order of the graphs in the main experiment was automatically changed in each run by Tobii Studio.

3.3.4 Questionnaire

After the participant completed Part-1 of the experiment, subjects were asked to fill the Graph evaluation form (Appendix –D for Linear Graph Group and Appendix – E for Round Graph Group) in order to obtain their opinion about each graph (line /bar / area). This form contains questions about

- the easiest graph in terms of remembering the labels,
- the easiest graph in terms of remembering the pattern,
- the easiest graph in terms of understanding/ grasping the relationship between the time and other variable,
- familiarity with the graph design (linear or round), and whether it was easy to become familiarized with it,
- open ended question about their general opinions on these graphs.

After Part-2 of the experiment, a Concept Evaluation form (Appendix-F) which asks the participant to fill the appropriate time zones which best describes the concept (e.g., winter and night) used in the experiment according to their opinion, had to be filled in. This evaluation form aimed at obtaining each participant's opinion about the concept. This was necessary because subjects may vary in their individual estimation of, e.g., what months constitute winter or how many hours the night lasts, even though they may have similar divisions of time.

3.3.5 Independent and Dependent Variables

There are four types of independent variables. The first one is graph design (linear or round graph), which is a between-subject variable. The second one is graph type (line, bar, area), which is a within-subject variable. The third independent variable is the type of events (linear and cyclic events), which is a within-subject variable. The last one is task type (discrete comparison and trend assessment), which is a within-subject variable again. All tasks were presented with all types of graph in order to study the interaction of task style, event type and graph type. The variable trees for each part were given in Appendix-R.

Recognition, gaze time, number of fixations, number of transitions between text and graph, and fixation duration were measured as dependent variables. In the experimental design of part-2, gaze time is an equivalent of reaction time which means how much time the task needs to be processed.

3.4 Procedure

The experiment was conducted in a mixed between/within- subject format. Each between-subject group includes 20 participants. The participants were randomly distributed into two groups. The first group, named "*Linear graph group*", evaluated the graphs in the linear format while the second group, named "*Round graph group*", evaluated the graphs in the round format. Additionally, one of the versions of each graph type condition (line, bar, area) was also randomly assigned to the participants.

Before the experiment, a consent form was signed by each participant. A calibration with 9 dots on the eye tracker was made. The experiment was not started until the participant made 9 successful fixations for calibration.

In order to eliminate random and redundant fixation on graphs and sentences, a fixation image (located at the bottom line in a centered position) was presented right before each graph. The participants were asked to fixate on the center of the image and then press any key while fixating on it.

Since the round graphs are an unusual graph design, for each group, a familiarization phase which contained samples from each graph type, depending on the group to which they belonged, was conducted. That is, subjects in both the linear and round graph design condition were presented with pictures of line, bar, and area graph types.

All participants took part in Part- 1 and Part-2 of the experiment, respectively. The average completion time for the entire experiment, including Part -1 and Part -2, was 35 minutes.

Part-1 of the experiment started with a brief information about the study, and then the instructions were given. Afterwards, familiarization trials that contain one describing and one drawing task for different types of graph (bar, line or area) were given. Subsequently, the main experiment which contains 6 tasks with different contents (with all graph design comprising three describing and three drawing tasks) was conducted. Each task started with the information screen which gives information what the graph is about. Then the graph was presented, and the participants pressed a key when they finished the observation. Subjects were asked to either describe or draw the relationship. After completion of all tasks in Part-1 of the experiment, the graph evaluation form was filled by the participant.

Then Part-2 of the experiment was given. This part also started with some brief information about study, and then the instructions were given. Subsequently, the familiarization trials that contain 6 graph-sentence stimuli (from each graph type and event type with different types of sentences) were given. Part-1 was supposed to provide familiarity to reading the graphs in Part-2. The main experiment in Part-2 contains 37 graph-sentence stimuli. The sentence given in a stimulus can be a true or wrong description of the information given in the graph, and the participants were asked to judge the accuracy of the sentence according to graph. There was no time limit but they were asked to give a response as quickly as possible. After all stimuli were presented and the main experiment ended, the concept evaluation form was filled by each participant.

3.5 Data Collection

To summarize, the data in this study collected by the various tools, is given below.

- Age, Sex, Departmental Information by the Demographic Data Sheet.

- Preferences on graph designs in terms of ease of understanding, ease of recollection and familiarity by the Graph Evaluation Form in Part-1.
- Subjects' comprehension of the time concepts in terms of time zones by the Concept Evaluation Form in Part-2.
- Eye movements (gaze duration, number of fixations, number of transitions between text and graph, and fixation duration) by the eye tracker.
- Recollected data from graphs in terms of verbal and visual description obtained by the answer sheets.

The experiment was administered with the permission of METU Ethical Committee.

3.6 Data Analysis

3.6.1 Analysis of the Recollection of Data (Part-1)

The answer sheets filled by the participants were evaluated according to their task type. The results of the drawing task were evaluated with three criteria. The first criterion is accuracy of the value. Each graph has 12 labels on the timeline, therefore the graph area was divided into grids, by drawing reference lines for the x and y axis. The fit of the graph elements (bar, area and line) drawn by participants was compared to the original data. If participants' drawings for a particular square (for example the y value corresponding to January) was accurate compared to the original graph, it was graded with +1. This process was repeated for each square which has a value in the graph. A second criterion is the recollection of the y-axis label. This criterion was evaluated with yes or no. The third criterion evaluated the consistency in the pattern between the subject's drawing and the original graph.

For the evaluation of the verbal description task, in addition to the value, the y-axis label and pattern recollection and the usage of keywords was also assessed. The keywords, i.e., the words which the participants tended to use to describe the information in the graph, were categorized under three aspects. These are discrete keywords (like minimum and maximum points), trend keywords (increasing, decreasing, falling etc.) and conceptual keywords which express the event represented in the graph (like “during the last years”, or “in winter”).

Furthermore, the participants did not know whether the next task was a drawing task or a describing task. Therefore, observation times of each graph were evaluated and tested with a three Way ANOVA (Graph Type * Graph Design * Event).

3.6.2 Eye movement analysis of the Sentence-Graph Verification Task (Part-2)

For the eye movement analysis for Part-2, graph area, sentence area and all stimulus screens were defined separately as Areas of Interest (AOI). For each graph viewed by each participant, gaze time, fixation count, observation length (reaction time) and observation count values for each sentence of each stimulus were taken from Tobii Studio and exported to MS. Excel. Then fixation duration was calculated and also the accuracy of the judgment was evaluated. Only values from accurate judgments were calculated. Afterwards, the stimuli were categorized w.r.t. their event type (trend / cyclic), sentence number, and graph type (line/ bar / area). Means were calculated for each sentence for each dependent variable. Each dependent variable was analyzed separately in SPSS with a mixed design ANOVA. Four-way ANOVAs with Graph type (3) x Graph design (2) x Event type (2) x Task type (2) were conducted).

Another analysis was conducted in order to investigate the Cyclic Event comprehension comparing Trend Assessment Task-1 (about winter and night) and Trend Assessment Task-2 (about summer and noon) sentences (see Table 3-3 for the examples of each task type). This analysis was also tested with a mixed design ANOVA (Three-way ANOVA).

Furthermore, the effect of the coherence of word order between the sentence and the graph was investigated with a four-way (3 (Graph type: Area, Line, Bar) * 2 (Graph design: Round, Linear) x 2 task (DCT-1, DCT-2) x 2 event (type (cyclic, trend)) mixed ANOVA.

3.6.3 Number of Errors in the Sentence-Graph Verification Task

The number of the errors committed by the participants in the decision task (trend assessment task) of the cyclic event between the graph design groups was evaluated with mixed ANOVAs. The first analysis to compare the linear and the round graph groups was a 2 (Event: cyclic, trend) x 2 (Graph Design: linear and round) mixed ANOVA. After re-grouping the participants according to their looking pattern and self-reports, a second analysis (one-way ANOVA) was conducted to compare the three graph groups (one-sided linear, two-sided linear and round)

3.6.4 *Scan path Analysis*

Gaze patterns for each task in each type of event and graph design were analyzed separately by the Tobii Studio Coding Schema. The general scan path of the participants was formed by analyzing the order of the transitions between the graph elements according to task type (trend assessment / discrete comparison), event type (cyclic / trend) and graph design (round /linear).

CHAPTER 4

RESULTS

In this chapter, I will present the results of the analysis of part-1 which was conducted to evaluate the Recollection of Data from graphs in terms of verbal and visual descriptions obtained by the answer sheets. Afterwards, in Section 4.2, the results of the analysis of Part-2 which was mainly based on eye tracking data (gaze length, number of fixations, and number of transitions between text and graph, and fixation duration) will be presented. in addition to the analysis of these eye movements, the analysis of scan paths will be presented in Section 4.3.

4.1 Part 1 – Analysis of Recollection of Data

4.1.1 Results of the Questionnaire

In the questionnaire that was given after the participants had completed Part-1 of the experiment, they were asked about the ease of memory of the values, the pattern, and the relation expressed in the graph. The results show that participants in both the linear and round graph condition rated bar graphs higher than linear and area graphs on the scale of remembering the values. On the other hand, area and line graphs were given similar scores.

For the recognition of patterns, while the participants in the linear graph group prefer the line graph, the round group prefers the area graph. Unlike for the recollection of the values, bar graphs were rated lower for the recollection of patterns.

The participants' preferences about the relation recognition are very similar to the value recognition. Bar graphs are favorites in remembering the relation, while area and line graphs have low ratings (Figure 4-1).

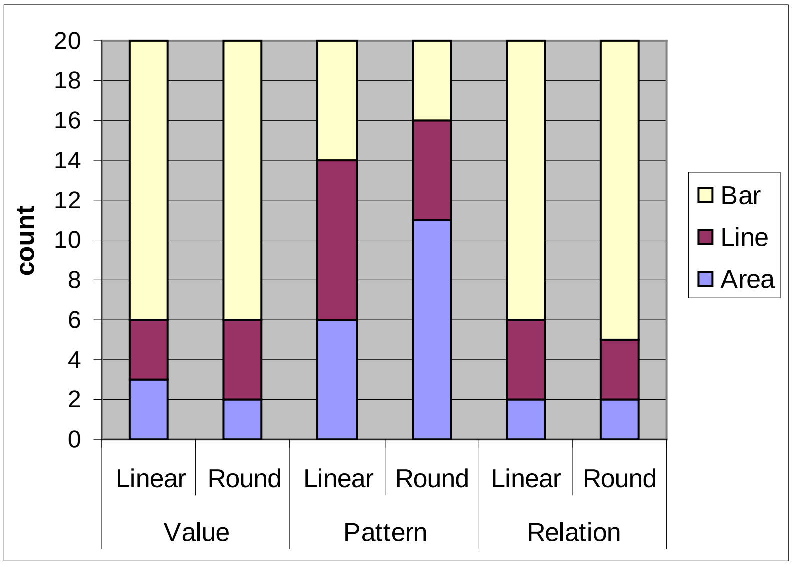


Figure 4-1: Easiness scores of participants in remembering the value, pattern and relation presented in the graphs according to graph type and graph design

The familiarity levels of participants to the graph designs were also evaluated. As expected, while the majority of participants in the linear group were familiar with the graph to a medium or high level, participants in the round graph group were not familiar with the round graph design. Participants who responded to this question with medium or high scores indicated that they put these round graphs into the same category with pie charts or polar graphs with which they are familiar (Figure 4-2).

In addition to collecting information about easiness of interpretation of and familiarity with the two graph designs, easiness of getting familiar with the graph designs was also assessed. Only the participants who reported not to be familiar with the graph design answered this question. 65% of the participants in the linear graph group reported being familiar with the graph on a medium or high level, therefore they skipped the question; 30% of them reported that the level of becoming familiar with the linear graph was medium, and one participant reported that it was easy to become familiar with linear graph.

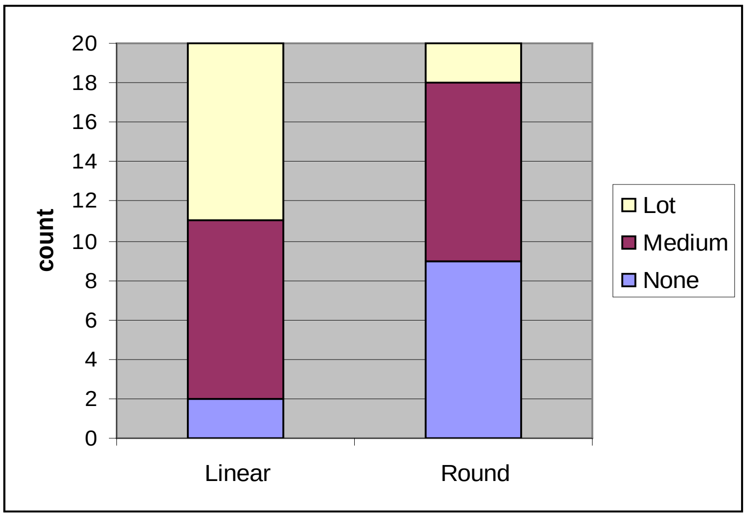


Figure 4-2: Familiarity scores of participants for linear and round graphs

For the participants in the round graph group, only 15 % of the participants were already familiar with this graph design. While one participant reported that getting familiar to the round graph was difficult, 35% of them gave medium scores, and 45% of them reported that it was easy to become familiar to them (Table 4-1).

Table 4-1: Participants' evaluation of graphs according to their level of easiness of familiarization

		Frequency	Percentage
NA (already familiar)	Linear	13	65
	Round	3	15
Difficult	Linear	0	0
	Round	1	5
Medium	Linear	6	30
	Round	7	35
Easy	Linear	1	5
	Round	9	45
Total	Linear	20	100
	Round	20	100

4.1.2 Observation Length and Recollection of Data

In the evaluation of the recollection part, 6 stimuli, that contained stimulus from each of the combination of the event type (cyclic and trend) and the design type (area, bar line), were given to the participants. The comparison of the observation length according to the graph design, the graph type and the event type were presented in the following part. The two types of tasks (the drawing task and the description task) given after observation were used for the recollection of data. Each participant took 3 stimuli for the each task. These 3 stimuli for each of the task type contained stimulus from each of the event type and each of the design type, however, did not contain stimuli from each of the combination of the event type and design type. Since there was not enough data to investigate the interaction between the event type and the graph type for each the task type, event and graph type effects were analyzed with different analyses for the recollection of data.

4.1.2.1 Observation Length

The observation length of graphs was tested with a three way mixed design ANOVA (Graph design (2) x Event type (2) x Graph type (3)).

The results of the ANOVA show that all effects were non-significant. The observation length of the graphs that represent trend and cyclic events are the same in general ($F(1, 37) = .036$, $p>.05$). Also, event type does not interact with graph design type ($F(1, 37) =.177$, $p>.05$). Additionally, there is no significant difference in the observation length of the graphs represented by different graph types ($F(2, 74) = .913$, $p>.05$). There is no significant interaction between graph type and graph design ($F(2, 74) =.626$, $p>.05$). Moreover, cyclic and trend events are observed in approximately the same time in different graph type types ($F(2, 74) = 1.493$, $p>.05$). Finally, graph design does not affect observation time ($F(1, 37) =.190$, $p>.05$).

In order to investigate the relationship between observation length, recollected value and number of keyword categories used in the description, bivariate correlations were applied. The Pearson coefficient indicates that there was a positive relationship between the amount of recollected values and the observation length (r=.363, p<.05). On the other hand, although the relationship between the amount of recollected values and number of keyword categories is negative, it is not significant (r=-.202, p>.05).

4.1.2.2 The Drawing Task

Effects of Event type

There is no significant main effect of event in the scores of the drawing tasks (F (1, 37) = .795, p>.05). Recollected data from cyclic events are similar to trend events. Also there is no significant interaction between event type and graph design (F (1, 37) =.065, p>.05), see (Figure 4-3)).

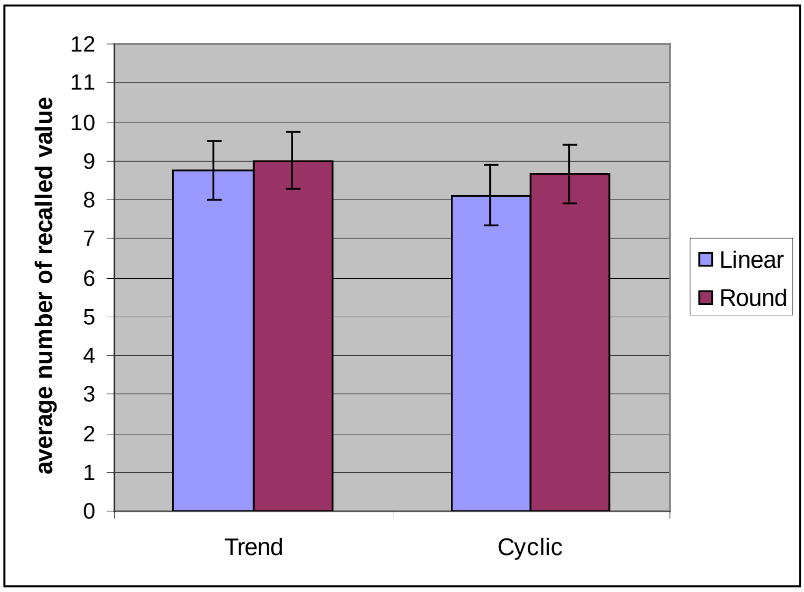


Figure 4-3: Average number of recollected values represented in the graphs, according to event type and graph design

Effects of Graph type

There was no significant main effect of graph type in the scores of the drawing tasks (F (1, 37) = 2.055, p>.05), indicating that all graph type types (area, line and bar graphs) have similar scores. Also there was no significant interaction between event type and graph type (F (1, 37) =.616, p>.05), see (Figure 4-4)).

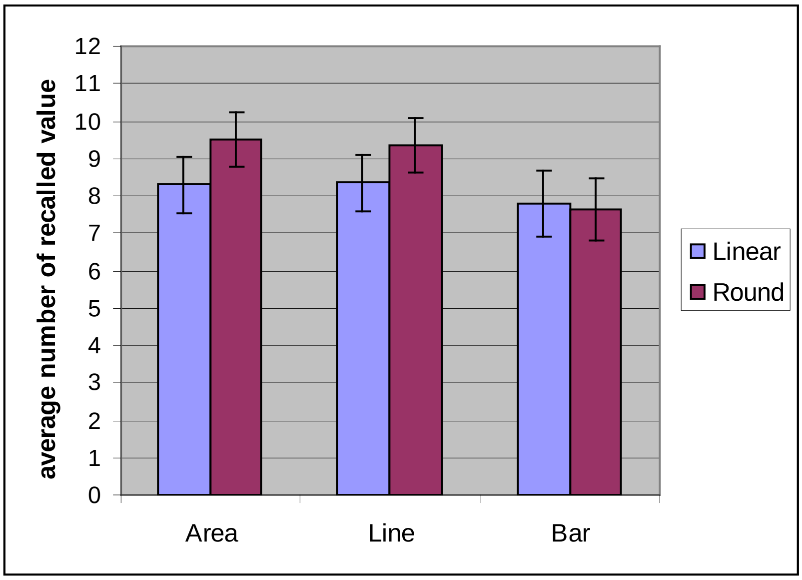


Figure 4-4: Average number of recollected values represented in the graph, according to graph type and graph design

Effects of Graph Design

There was no significant main effects of the graph design (F (1, 37) =.195, p>.05). This means that the recollected data from the drawing task did not differ for the linear and the round graphs, see Figure 4-5.

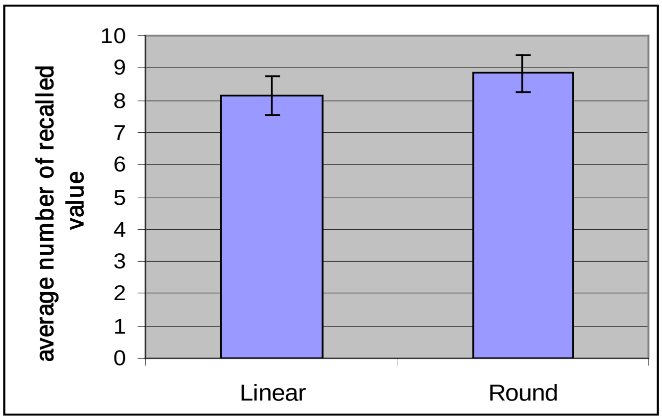


Figure 4-5: Average number of recollected values represented in the graph, according to graph design

4.1.2.3 Verbal Description of the Relations in Written Format

Effects of Event Type

Since the data for event type appeared to be significantly non-normal, non-parametric statistical procedures were used to compare means between cyclic and trend events. Additionally, in order to compare means between graph designs, the Wilcoxon W-test has been applied.

The result of Wilcoxon tests on value recollection showed that there is no difference in the recollection of the values of the graph elements for participants in the linear group (z=-1.865, p>.05, r=-0.41) and the round group (z=-1.446, p>.05, r=-0.32). There is also no significant difference between graph designs (z=-.081, p>.05, r=-0.01), see (Figure 4-6).

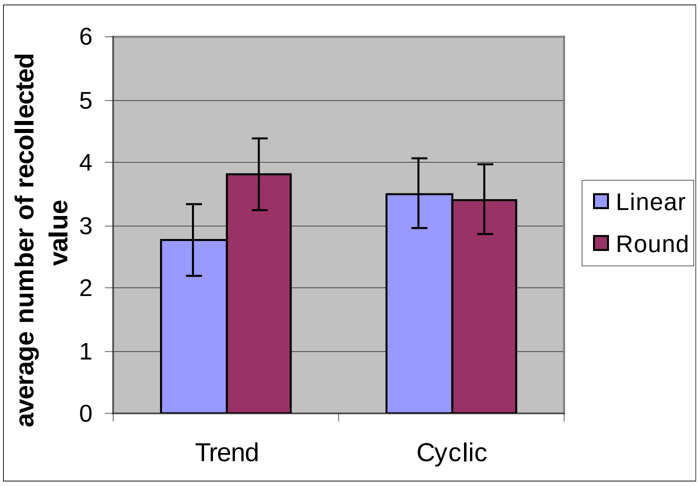


Figure 4-6: Average score of recollected values represented in the graph, according to event type and graph design

The analysis of the frequency in the usage of keywords indicated that in the linear graph group, discrete keywords (Maximum, minimum etc.) were used in cyclic events more often than in trend events (z=-2.309, p<.05, r=-0.51). On the other hand, the difference between event types in the round group was not significant (z=-1.897, p>.05, r=-0.42). Furthermore, there was no overall significant difference between graph designs (z=-1.448, p>.05, r=-0.43), although there seem to be much more discrete keywords in the cyclic event as compared to the trend event, see (Figure 4-7).

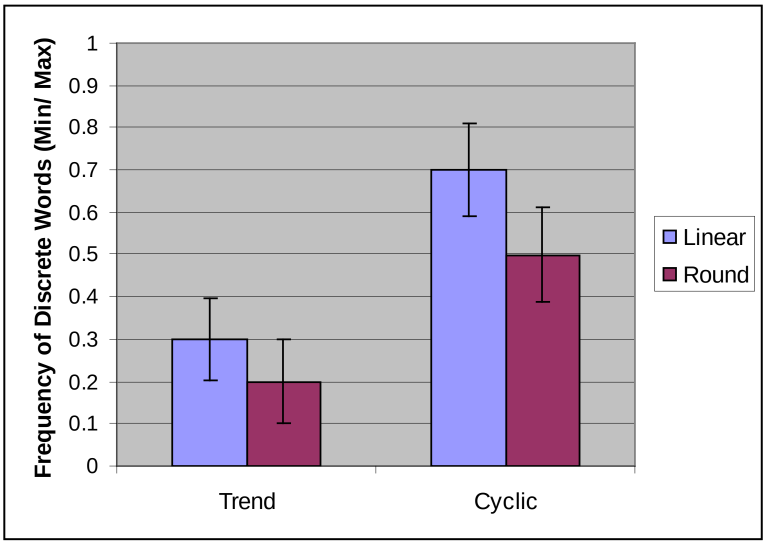


Figure 4-7: Frequency of Discrete keywords according to Event type and Graph design

Unlike discrete keyword, trend keywords were used in trend event more than in cyclic events in the linear graph group. (z=-2.236, p<.05, r=-0.50). However, again, there is no significant difference in the round group (z=.000, p>.05). In addition, there is no overall significant difference between graph designs (z=-.330, p>.05, r=-0.05), see (Figure 4-8).

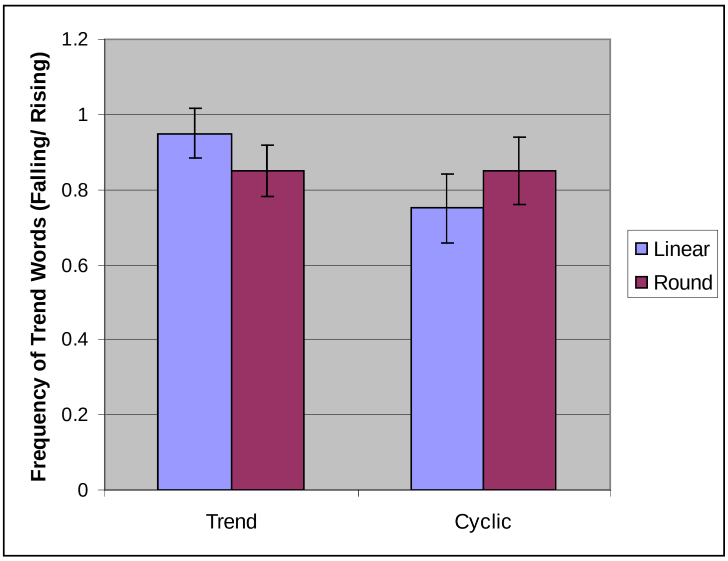


Figure 4-8: Frequency of Trend Words according to Event type and Graph design

Furthermore, conceptual words (“in winter”, “in last years” etc.) were used in cyclic events more often than in trend events ($z=-2.449$, $p<.05$, $r=-0.54$) in round graph, while there was no difference in the linear group. ($z=-.707$, $p>.05$, $r=-0.45$). Besides, the difference between graph designs was not significant ($z=-.154$, $p>.05$, $r=-0.02$), see (Figure 4-9). In general, the number of conceptual words was very low.

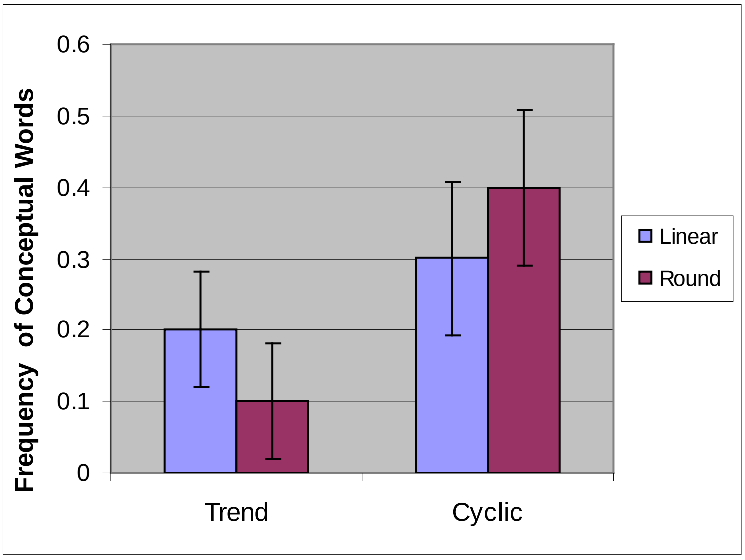


Figure 4-9: Frequency of Conceptual Words according to Event type and Graph design

The overall analysis of number of keyword categories used in description of the participants showed that in round graphs, the cyclic events were described by using more categories in terms of keyword than the trend events ($z=-2.676$, $p<.05$, $r=-0.59$). However, there was no difference between event types in linear graphs ($z=-.894$, $p>.05$, $r=-0.20$). Additionally, the difference between graph designs was not significant ($z=-.375$, $p>.05$, $r=-0.06$), see (Figure 4-10).

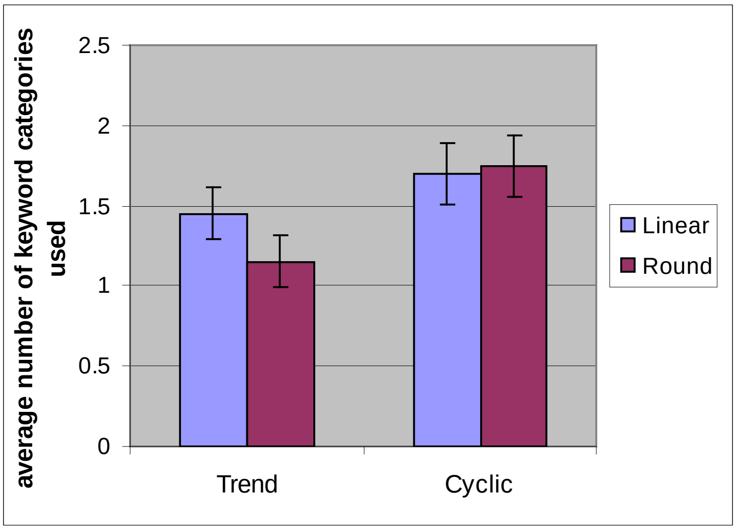


Figure 4-10: Average number of keyword categories used according to Event type and Graph design

Effects of Graph Type

The data for graph type was also significantly non-normal, therefore non-parametric statistical procedures (Friedman Test) was used to compare means between area, line and bar graphs. Additionally, means between graph designs were compared by using Wilcoxon W-tests.

All of the test results were reported as non-significant for graph type. The Friedman test results indicated that there was no difference between graph types in terms of number of recollected value in both linear graph (χ^2 (2) =1.677, $p>.05$) and round graph type ($\chi^2(2)=.160$, $p>.05$). The overall difference between graph designs was not significant either ($z=-.312$, $p>.05$, $r=-0.05$).

Likewise, discrete keyword usage in different graph types was not different from each other in both linear graph (χ^2 (2) =2.462, $p>.05$) and round graph type (χ^2 (2) =2.400, $p>.05$). There was no significant difference between graph designs either ($z=-1.672$, $p>.05$, $r=-0.26$). Additionally, graph types did not differ in terms of usage of trend words in both linear graph (χ^2 (2) =2.000, $p>.05$) and round graph designs (χ^2 (2) =4.333, $p>.05$). There was no significant difference between graph designs either ($z=-.995$, $p>.05$, $r=-0.16$). Like trend and discrete keywords, there was no significant difference in conceptual word usage between graph types in both linear graph (χ^2 (2) =1.143, p>.05) and round graph designs (χ^2 (2) =.250, p>.05). The overall difference between graph designs was not significant (z=-.929, p>.05, r=-0.05), see (Figure 4-11).

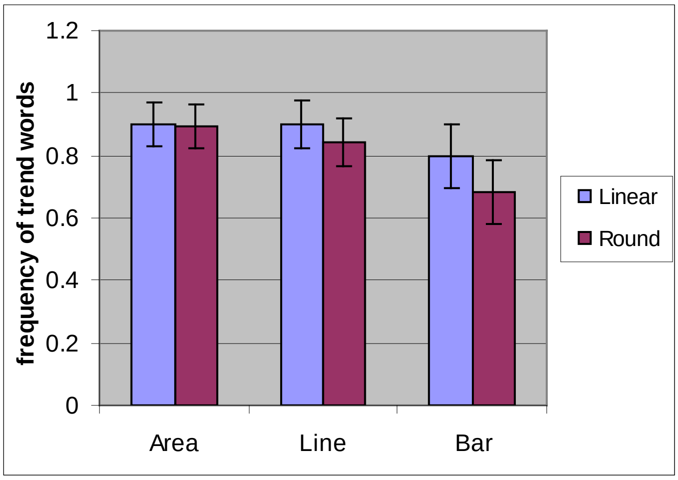


Figure 4-11: Frequency of Trend Words according to Graph type and Graph design

The Friedman test that was conducted to compare means of overall usage of keywords also revealed that there was no significant difference between graph types in both linear graph (χ^2 (2) =.778, p>.05) and round graph designs (χ^2 (2) =.703, p>.05). There was no significant difference between graph designs either (z=-.786, p>.05, r=-0.12), see (Figure 4-12).

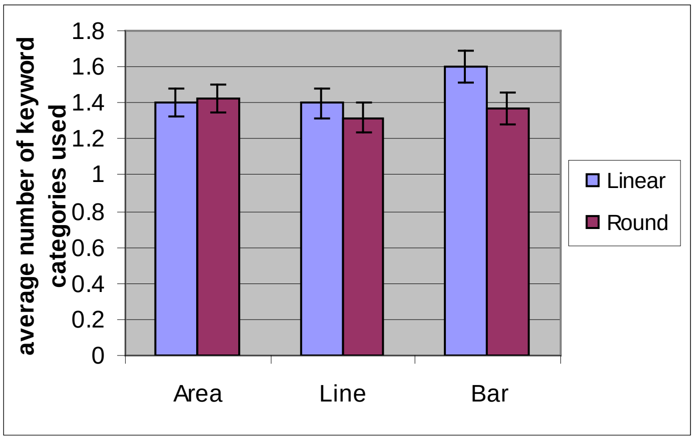


Figure 4-12: Average number of keyword categories used according to Graph type and Graph design

The Drawing vs. the Describing Task

For the "drawing" question, the empty graph with reference lines and time labels was given and the participants were asked to draw the relationship with the values to assess the visual information extracted from the graph. On the other hand, for the "describe" question which was asked to assess the verbal information extracted from the graph, the participants were asked to describe the relationship with the values in written format. Comparing the recalled values between the drawing and the describing task, there appears a significant difference. The number of values which was evaluated from the results of the drawing task was significantly higher than that of the verbal description task ($F(1, 35) = 68.708$, $r = .81$, $p < .05$).

4.2 PART-2 Judgment Task Results

4.2.1 Analysis 1: General Analysis

For the statistical analysis of the eye tracking data, 2 (Event type: Cyclic, Trend) x 2 (Task: Trend Assessment, Discrete Comparison) x 3 (Graph type: Area, Line, Bar) x 2 (graph design: linear, round) mixed ANOVAs were applied to compare the gaze time, number of fixation, number of transition and mean fixation duration between the two groups (Graph design: Linear, Round).

A table of F statistics and effect size values is given for each effect looking across for each dependent variable in Appendix – I.

Furthermore, the descriptive statistics for each dependent variable are also given in the Appendix Section (J: Gaze time, K: Fixation count, L: Fixation duration and M: Number of Transition).

The partial eta squared (η_p^2) values which were less than .10 were not reported in the result section of the analysis.

Gaze time

Gaze time means the sum of the individual fixation lengths, and increase in the score of this variable indicates increase in the processing time of the task. Gaze time scores showed that there was no significant main effect of graph design, indicating that gaze time for linear and round graphs were in general the same ($F(1,38)=1.448$, $\eta_p^2=0.36$, $p>.05$).

On the other hand, there was a high significant main effect of event type (cyclic/trend) $F(1, 38) =39.058$, $\eta_p^2=0.51$, $p<.001$. Pairwise comparisons revealed that making decisions about cyclic events took much longer than about linear events. On the other hand, there was no significant interaction between event type and graph design ($F(1, 38) =2.720$, $\eta_p^2=.07$, $p>.05$). This indicates that the gaze time for event types, namely the cyclic and trend events, did not differ for the different graph designs (linear, round), see Figure 4-13.

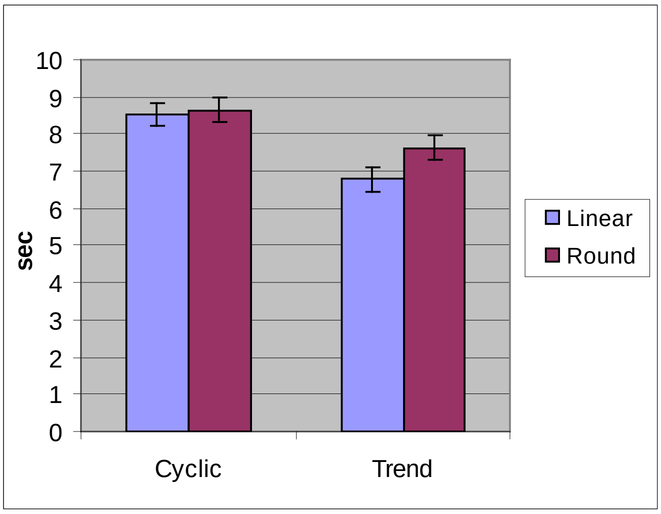


Figure 4-13: Gaze time for different event types and graph designs

There was a significant main effect of graph type ($F(2, 76) =3.158$, $\eta_p^2=.08$, $p<.05$). This indicates that gaze time on area, line and bar graphs were different from each other. The contrast revealed that gaze time on bar graphs was significantly shorter than on line graphs ($F(1, 38) = 8.543$, $\eta_p^2=.18$, $p<.05$), but there was no significant difference between bar and area graphs ($F(1, 38) = 3.409$, $p>.05$). Additionally, there was no significant interaction between graph type and graph design (F (1, 38) = 2.946, p>.05), see (Figure 4-14).

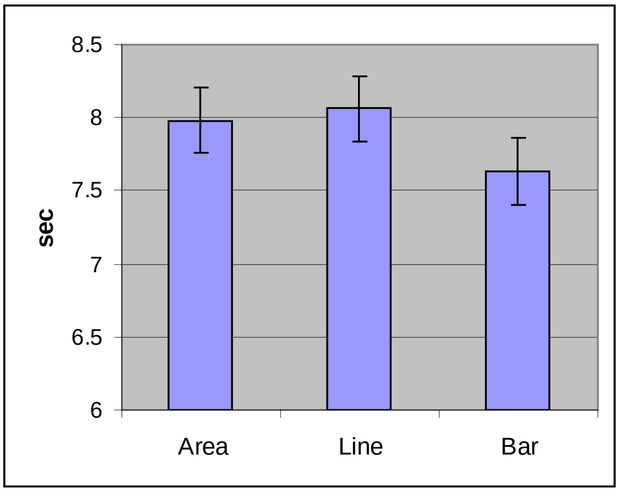


Figure 4-14: Gaze time according to graph types

There was also a highly significant main effect of task (trend assessment/ discrete comprehension) with a great effect size (F (1, 38) =260.368, η_p^2=.87, p<.001). The comprehension of trend assessment tasks needed less time than discrete comprehension tasks. However, the interaction between task and graph design was not significant (F (1, 38) =4.050, η_p^2=.10, p>.05), see (Figure 4-15).

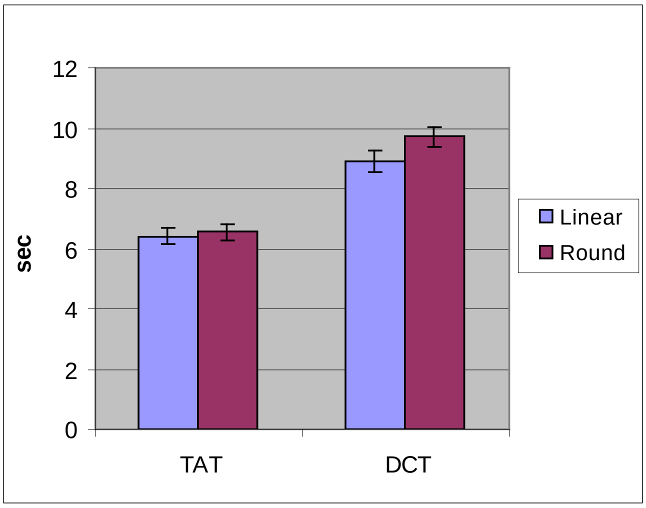


Figure 4-15: Gaze time in task types according to graph design type

There was no significant interaction effect between event type and task type ($F(1, 38) = 3.402$, $\eta_p^2 = .08$, $p > .05$). This indicates that the gaze time for different tasks did not differ in cyclic and trend events, see (Figure 4-16).

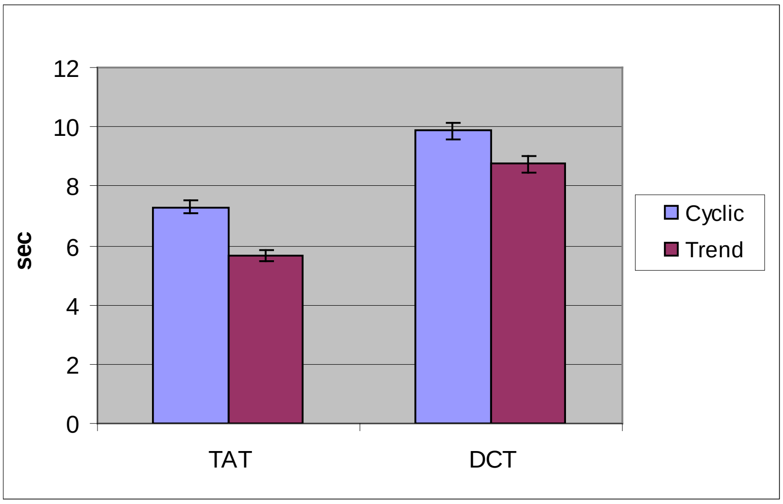


Figure 4-16: Gaze time for different event types for different event types

The interaction between event and graph was not significant. This reveals that gaze time of different type of events did not differ in graph type $F(2, 76) = .284$, $p > .05$.

There was no three-way interaction of event, task and graph design ($F(1, 38) = .084$, $p > .05$). Figure 4-17 shows the gaze time scores for these combinations.

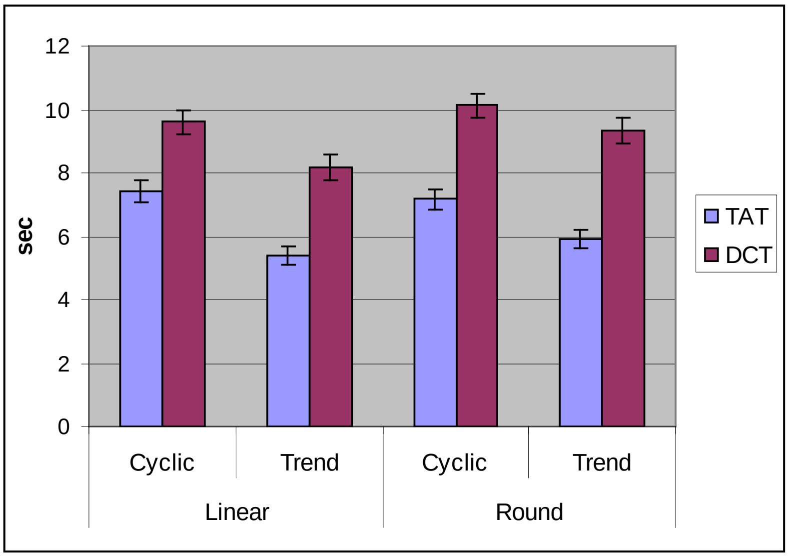


Figure 4-17: Gaze time for combinations of task, event, and graph design

Finally, the four-way interaction between graph type, task, event type, and graph design was not significant (F (2, 76) =3.115, η_p^2=.08, p<.05). Figure 4-18 illustrates the overall results of the gaze time variable.

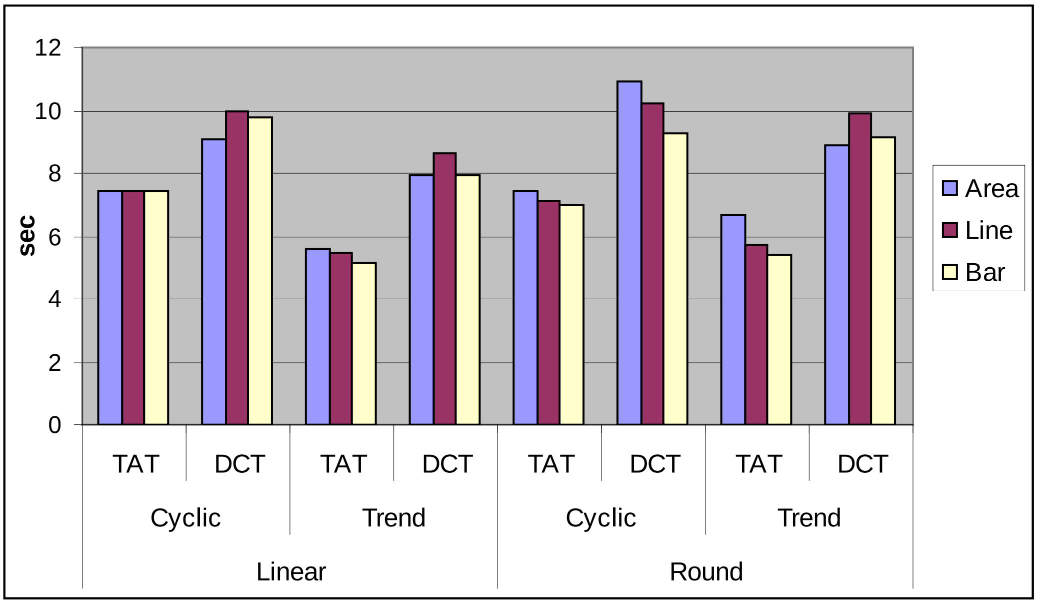


Figure 4-18: Gaze time according to the combinations of task, event type, graph type, and graph design.

Fixation count

The overall number of fixations is considered as an indication of ineffective search. The ANOVA analysis on fixation count indicates that there was no significant effect of graph design (F (1, 38) =2.708, η_p^2=.07, p>.05). That means that the number of fixations on linear and round graphs was the same in general. Furthermore, there was no significant main effect of graph type (bar/line/area) (F (2, 76) =1.399, p>.05). On the other hand, there was a significant main effect of event type (cyclic/trend) (F (1, 38) =23.962, η_p^2=.39, p<.05). Assessing cyclic events needed more fixations than assessing linear events. Furthermore, there was also a significant interaction effect between event type and graph design (F (1, 38) =4.700, η_p^2=.11, p<.05). This indicates that the fixation count for different event types; cyclic events and trend events, differed in the two graph designs. To break down this interaction, contrasts were performed comparing each type of event across each graph design. While the comprehension of cyclic events was the same between the two graph designs, trend events were comprehended much more easily in linear graphs than in round graphs and remarkably easier than in cyclic events; see (Figure 4-19).

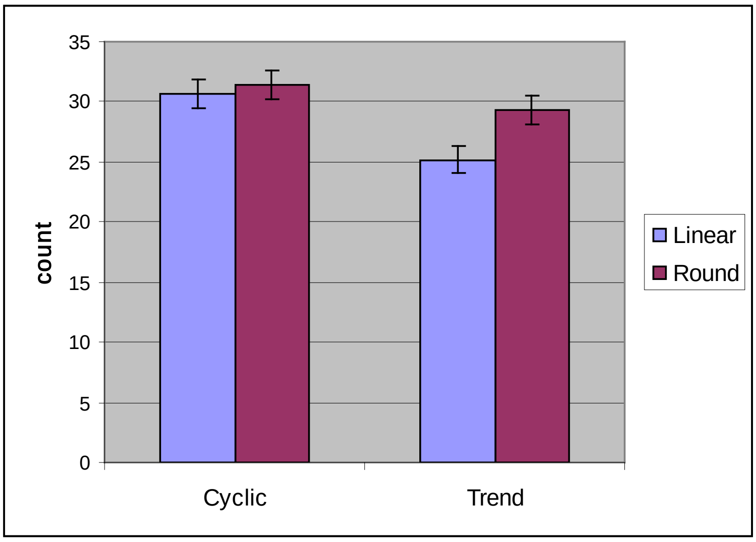


Figure 4-19: Fixation count for cyclic and trend events according to graph design

Another significant main effect with a high effect size was observed for the task variable (F (1, 38) =264.512, η_p^2=.87, p<.05). Trend assessment tasks were performed with fewer fixations than discrete comparison tasks. Furthermore, there was also a significant interaction effect between task and graph design (F (1, 38) =6.816, η_p^2=.15, p<.05). This indicates that the fixation count for different task types; namely the trend assessment task (TAT) and the discrete comparison task (DCT), differed for graph design. Pairwise comparisons revealed that the difference between the two tasks in round graphs is more pronounced than that in linear graphs (Figure 4-20).

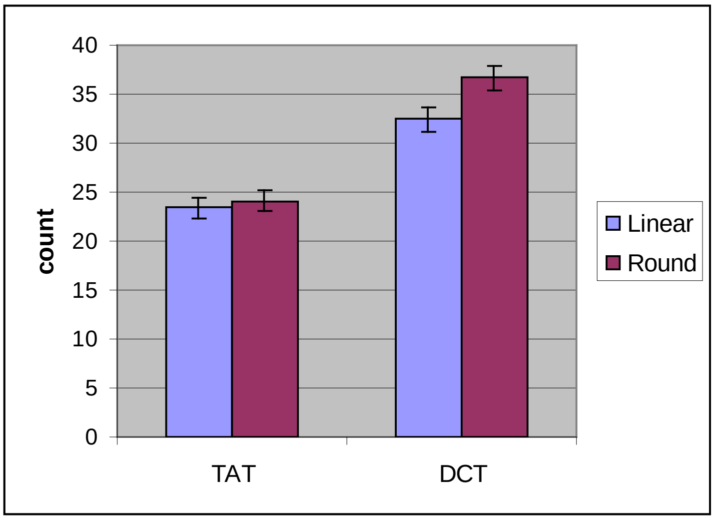


Figure 4-20: Fixation count in different tasks according to graph design

Moreover, there was a significant interaction effect between event type and task ($F_{(1, 38)} = 8.105$, $\eta_p^2 = .18$, $p < .05$). This indicates that the needed number of fixations for different tasks differed in cyclic and trend events. To break down this interaction, contrasts were performed comparing each type of task across cyclic and trend events. While discrete comparison tasks were performed with nearly the same number of fixations, there were fewer fixations in trend assessment tasks in trend events than in cyclic events (Figure 4-21).

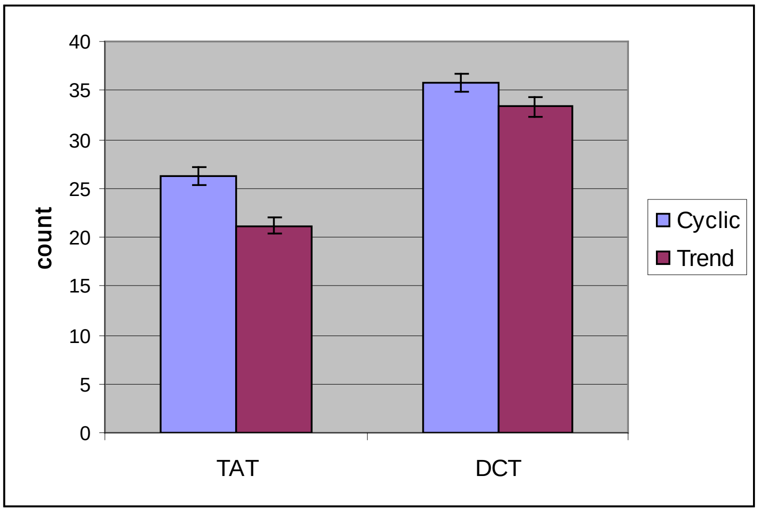


Figure 4-21: Fixation count for different event types and different task types

The four-way interaction between graph type, task, event type, and graph design was not significant ($F(2, 76) = 2.260$, $\eta_p^2=.06$, $p>.05$). Figure 4-22 illustrates the overall fixation count for all independent variables. The other two-way and three-way interactions were not significant (their test statistics can be seen in Appendix I).

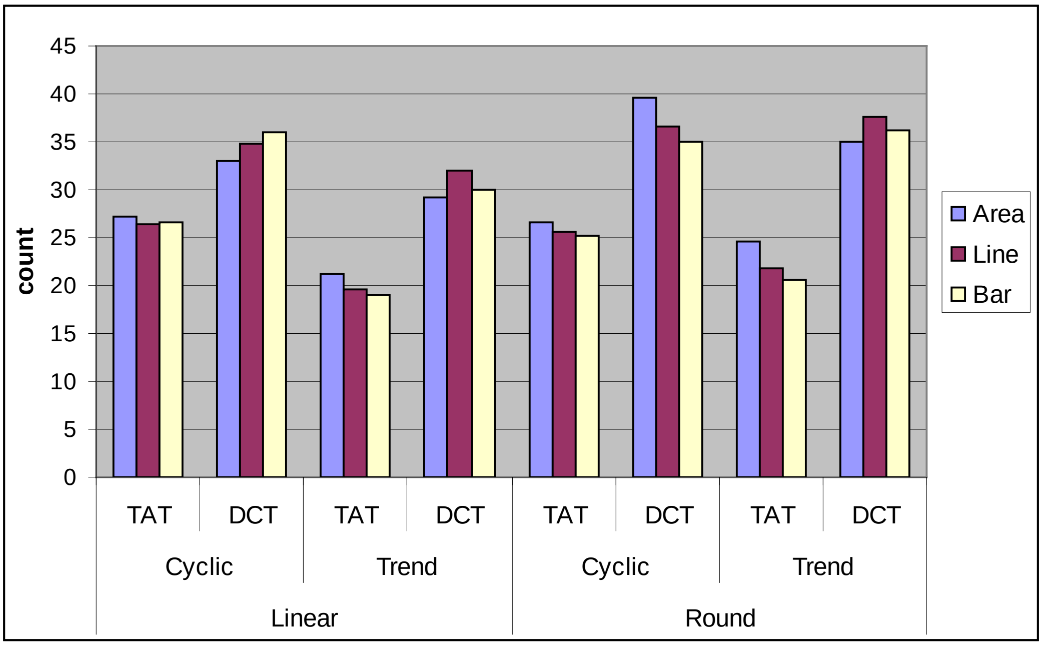


Figure 4-22: Fixation count in the combinations of task, event type, graph type, and graph design

Fixation Duration

Longer fixations are generally considered to be an indication of a participant's difficulty extracting information from a display. Similar to gaze time and fixation count, there was no significant effect of graph design ($F(1, 38) =.392$, $p>.05$). Furthermore, there was no significant main effect of event type (cyclic/trend) ($F(1, 38) =2.185$, $p>.05$). This indicates that fixation duration is the same in general for both cyclic and linear events. Unlike for gaze time, there was no significant main effect of graph type (area/bar/line) ($F(2, 76) = 2.695$, $p>.05$). On the other hand, there was a significant main effect of task (trend assessment/ discrete comprehension) ($F(1, 38) =8.359$, $\eta_p^2=.18$ $p<.05$). Fixation duration in trend assessment task was longer than in the discrete comprehension tasks (Figure 4-23).

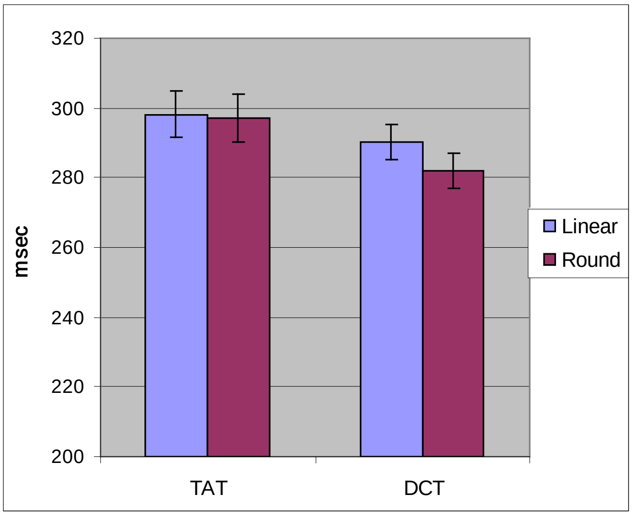


Figure 4-23: Fixation Duration in Trend Assessment and Discrete Comparison Tasks according to graph design

Two-way interactions between event type and graph design (F (1, 38) = .905), task and graph design (F (1, 38) = .797), graph type and graph design (F (2, 76) = .829) were all insignificant. However, the interaction between task and graph was significant (F (2, 76) = 3.772, p<.05). Pairwise comparisons indicate that, while fixation durations in line graphs did not differ according to task type, fixation duration in trend assessment tasks was longer than fixation duration in the discrete comparison task with area and bar graphs (Figure 4-24).

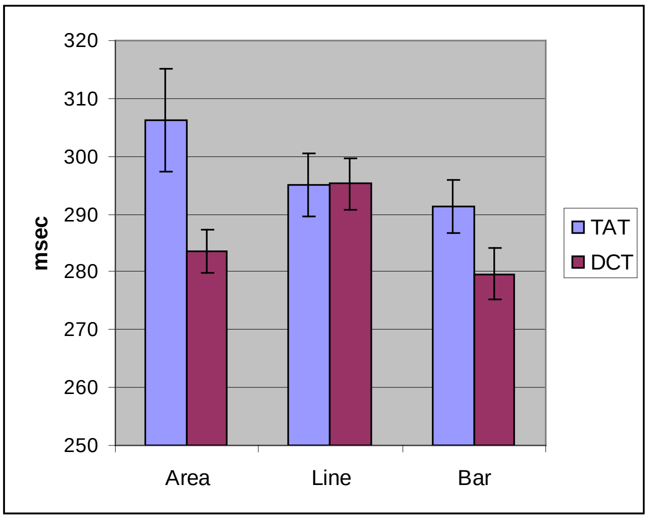


Figure 4-24: Fixation Duration in Trend Assessment and Discrete Comparison Tasks for each type of graph

Additionally, the four-way interaction between graph type, task, event type, and graph design was not significant (F (2, 76) = 1.168, p>.05). Figure 4-25 illustrates the overall fixation count for all independent variables. The other two-way and three-way interactions for fixation count were insignificant. Their test statistics are given in the Appendix I.

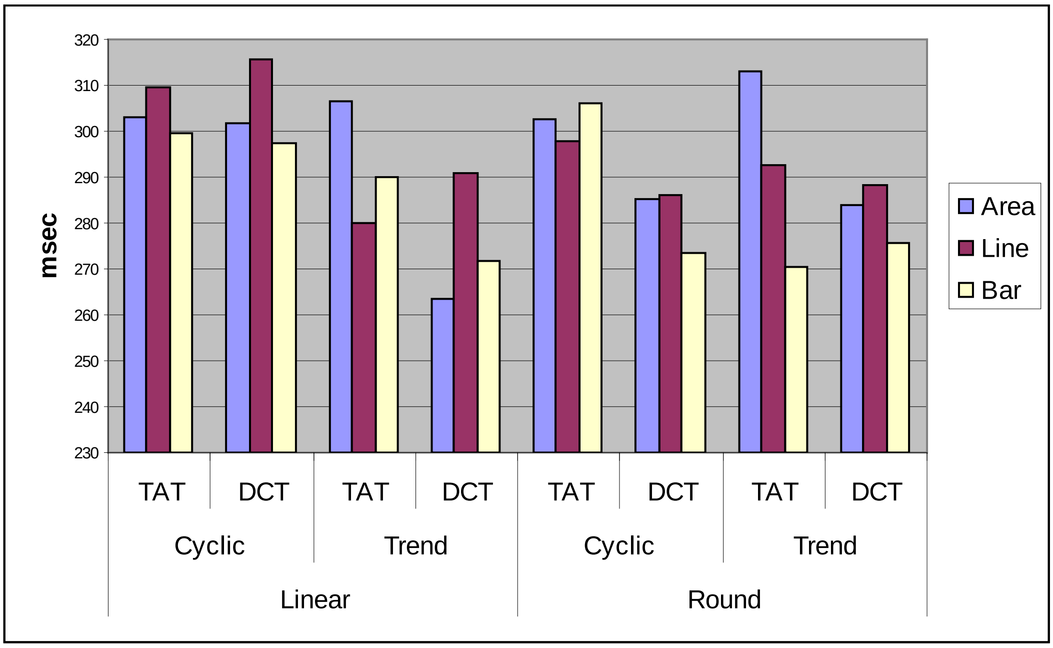


Figure 4-25: Fixation duration in the combinations of task, event type, graph type, and graph design

Number of Transitions

Increase in the number of transition between the components of the display indicates the difficulty in the processing of the information. As for the previous three dependent variables, there was no significant effect of graph design for number of transition either (F (1, 38) <0.21). On the other hand, there was a significant main effect of event type (cyclic/trend) (F (1, 38) =9.982, η_p^2=.21, p<.05). Pairwise comparisons revealed that cyclic events needed more transitions between graph elements than linear events. Furthermore, there was a significant main effect of task (F (1, 38) =25.612, η_p^2=.40, p<.05). Contrasts revealed that the number of transition in discrete comparisons is higher than in trend assessments. There was no significant main effect of graph type (bar/line/area) (F (2, 76) =1.927, p<.05). This means that the number of transition is the same in general for all types of graphs.

Similar to fixation duration, the two-way interactions between event type and graph design ($F(1, 38) = .951$), task and graph design ($F(1, 38) = .207$), graph type and graph design ($F(2, 76) = .562$) were not significant for the number of transitions.

However, there was a highly significant interaction effect between event type and task ($F(1, 38) = 8.068$, $\eta_p^2 = .17$, $p < .05$). This indicates that the number of transitions needed for the two tasks differed in cyclic and trend events. To break down this interaction, contrasts were performed comparing each type of task across cyclic and trend events. While discrete comparison tasks were performed with nearly the same number of transitions in cyclic and trend events, the number of transition needed in TAT tasks in trend events was lower than in cyclic event and lower than in DCT tasks in both event types (Figure 4-26).

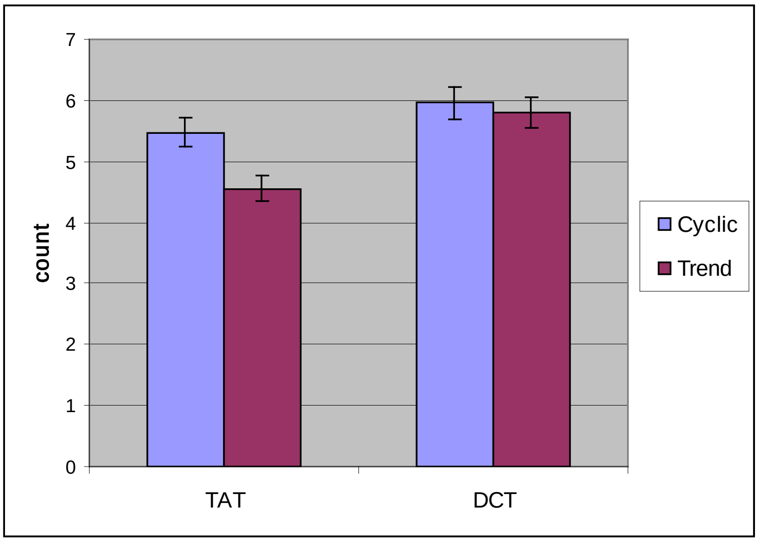


Figure 4-26: Observed number of transitions in different tasks according to event type

Number of Errors in the Decision Task

For the statistical analysis, a 2 (Event: Cyclic, Trend) x 2 (graph design: linear, round) mixed ANOVA was applied to compare the number of the errors in the decision task between the two groups (Graph design: linear, round). Their test statistics are given below (Table 4-2).

Table 4-2: Descriptive Statistics for the number of the error in the decision task according to event type and graph design

Event	Graph Type	Mean	Std. Deviation	N
Cyclic	Linear	2.35	1.17	17
	Round	1.80	0.77	20
	Total	2.05	1.00	37
Trend	Linear	1.06	0.97	17
	Round	0.85	0.99	20
	Total	0.95	0.97	37

The results showed that there was no significant effect of graph design (F (1, 34) =2.342, η_p^2=.07, p>.05). On the other hand, there was a highly significant main effect of event type (cyclic/ trend) with a large effect size (F (1, 34) =27.572, η_p^2=.44, p<.05). Pairwise comparisons revealed that the participants made more errors in the decision task of the cyclic events than in that of trend events. Furthermore, the interaction between event type and graph design was not significant (F (2, 34) =.651, p>.05). This indicates that the difference between the number of errors in cyclic events and trend events in the linear graph design was not different from that in the round graph design, see Figure 4-27.

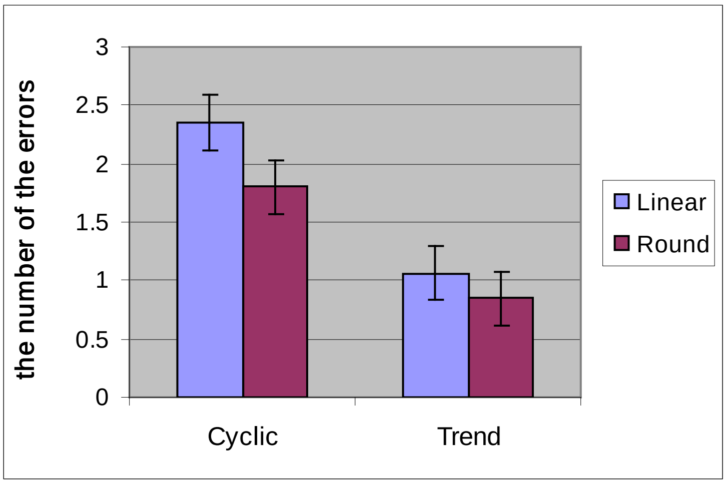
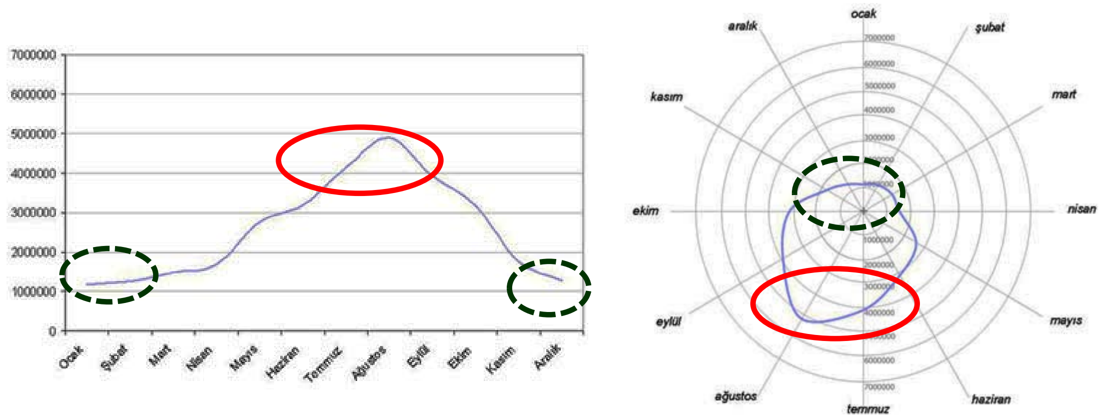


Figure 4-27: Mean number of the errors in the decision task according to Graph Design and Event Type

4.2.2 Analysis 2: The comprehension of Cyclic Concepts

After the general analysis on the data for all dependent and independent variables that aimed at the elucidation of the relationships between different types of tasks, events, graphs and graph design, a more detailed analysis was conducted in order to better understand the comprehension of cyclic concepts.

In this analysis, the score of two tasks in trend assessment category were compared. These were TAT-1 (about winter and night) and TAT-2 (about summer and noon) sentences for cyclic events (see Table 3-3 in Section 3.3.2.2 for the samples of each sentence). This analysis investigates how the comprehension of concepts is affected by their location in the graph. In round graphs, there should not be much difference since there is not much difference in their representation in terms of spatial properties. On the other hand, in linear graphs, while the information about summer or noon is located adjacently, the winter or night data points are distributed at both edges of the graph (see Figure 4-28).



(a) Linear Graph (b) Round Graph

TAT-1 (about winter and night)

TAT-2 (about summer and noon)

Figure 4-28: The location of information about cyclic concepts (e.g. winter and summer) in linear and round graph

For the statistical analysis, a 2 (Task: Trend Assessment Task-1, Trend Assessment Task-2) x 3 (Graph Type: Area, Line, Bar) x 2 (Graph Design: linear, round) mixed

ANOVA was applied to compare the gaze time and number of fixations between the two groups (Graph design: Linear, Round).

Since the previous analysis (see Appendix-I for the results) showed that Gaze time and Fixation Count are more robust dependent variables, in this analysis, these two were used. Descriptive statistics for the following statistical tests are given in Appendix N.

Gaze time

There was no significant effect of graph design ($F(1, 29) = 1.369$, $p > .05$). There was no significant main effect of task (trend assessment-1/ trend assessment-2) ($F(1, 29) = .042$, $p > .05$). Additionally, there was no significant interaction between task and graph design ($F(1, 29) = .996$, $p > .05$). This indicates that the difference between gaze times on TAT-1 and TAT-2 tasks in the linear graph design is not different from that in the round graph design, see Figure 4-29. There was no significant main effect of graph type (area/bar/line) ($F(2, 58) = .015$, $p > .05$). Additionally, there was no significant interaction between graph type and graph design ($F(2, 58) = .682$, $p > .05$).

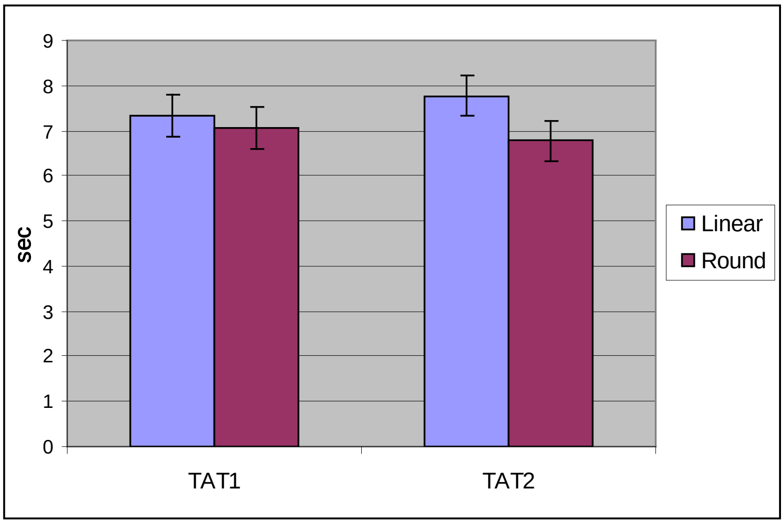


Figure 4-29: Gaze time on different tasks according to graph design

The two way interaction between task and graph type (area, line, bar) ($F(2, 58) = 1.898$, $p > .05$) and the three way interaction between task, graph type and graph design ($F(2, 58) = .985$, $p > .05$) were not significant either.

Fixation Count

The analysis on fixation count indicates that there was no significant effect of graph design ($F(1, 29) = 1.160$, $p > .05$). There was no significant main effect of task (trend assessment-1/ trend assessment-2) ($F(1, 29) = .132$, $p > .05$) either. Additionally, there was no significant interaction between tasks and graph design ($F(1, 29) = .185$, $p > .05$) (Figure 4-27). This means that the difference between number of fixations on TAT-1 and TAT-2 tasks in the linear graph design was not different from that in the round graph design. There was no significant main effect of graph type (area/bar/line) ($F(2, 58) = .019$, $p > .05$). Additionally, there was no significant interaction between graph type and graph design ($F(2, 58) = .374$, $p > .05$) (Figure 4-30).

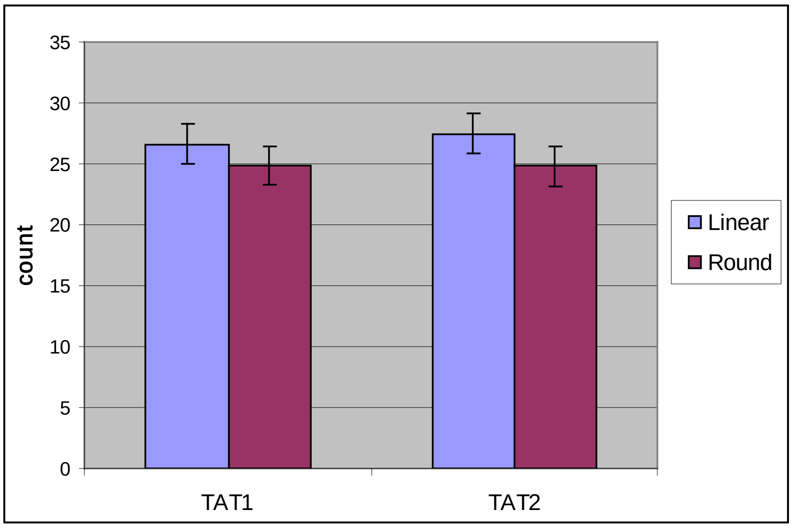


Figure 4-30: Fixation Count on different task types according to graph design

4.2.3 Analysis 3 : Trend Assessment Task-1 (Edge information) in the Linear Graph

It is assumed that in order to make decision about winter or night (presented at the edges of the axes separately) in a linear graph the participant should look at both left and right sides of the timeline. Therefore these tasks should be completed with longer gaze time and more fixation counts. However the previous analyses indicated that there are no significant differences in both the comparison of linear and round graph designs and the comparison of TAT-1 and TAT-2 revealed no differences, even in the linear graph although the data representation in TAT-1 (distributed) is different from that in TAT-2 (combined). For that reason, the gaze patterns of the participants in linear graph design group were investigated individually.

This additional analysis examined whether the participants in the linear graph design group look at both related areas when the task asks for edge information. These results were combined with the results of concept evaluation forms which were given after completion of Part-2. This evaluation form aimed at obtaining each participant's opinion about the concept by asking them to fill the appropriate time zones which best describes the concept (e.g., winter and night) used in the experiment according to their opinion.
According to the questionnaire results the participants were divided into two categories (see Figure 4-30). While some participants associated these concepts with the months or hours that are located in just one side of the graph (for example; defining winter as November and December instead of December, January and February), some of the participants associated these concepts with the entities represented at both sides of the linear timeline.

A scan path analysis of the linear graphs also showed that participants in the linear group can be grouped into two categories. The first category is called "one-sided linear group". The scores of the participants who looked at just one side of the graph in order to make a decision about data presented at both sides of the graph were evaluated under this category. The other participants belonged to the "two-sided linear group". They made two-sided reports in the questionnaire. However, also this group was divided into two subgroups because some participants did not look at both sides of the graph although their reports in the questionnaire contained months or hours from both sides of the timeline in the linear graph (see Figure 4-31).

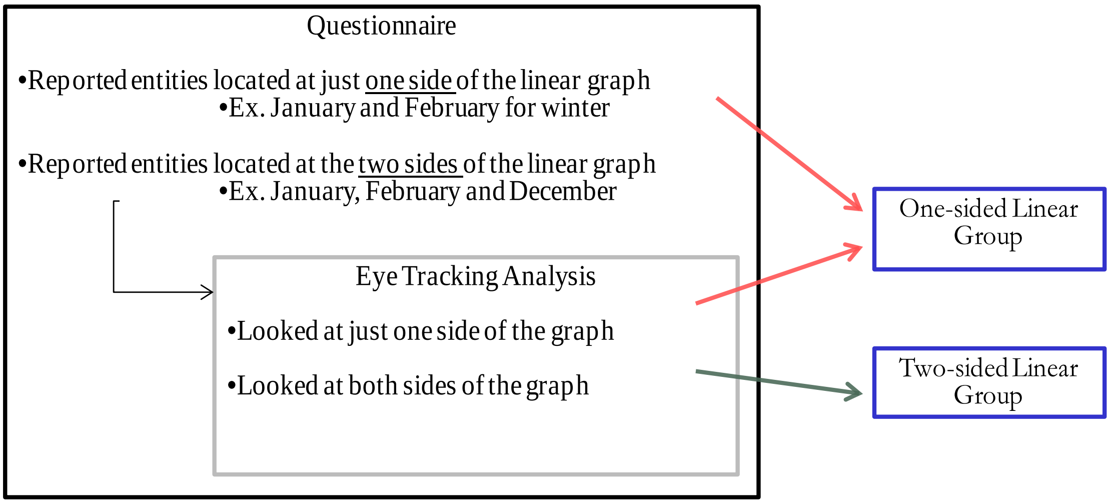


Figure 4-31: Linear graph type groups according to participants' questionnaire results and eye tracking analysis

Table 4-3 shows the results of the questionnaire (Q) and the eye movement (EM) analysis. In Part-2 of the experiment, each participant was presented 5 graph sentence stimuli about winter and night. The reports of 7 participants of the linear group showed that their winter or night concepts contained months or hours presented together in the linear graph. Therefore looking at both sides of the graph was not expected for them. There were 6 participants who made two-sided reports but looked at just one side of the graph in the experiment for the winter stimuli, and 8 for the night stimuli. The results of these subjects plus the ones from the one-side report group constituted the "one-side linear group". On the other hand, there were 7 participants who made two-sided reports and also looked at both sides for the winter stimuli, and 5 for the night stimuli. Their results were evaluated under "two-side linear group".

Table 4-3: The results of questionnaire (Q) and eye movement (EM) analysis for the Cyclic Event Analysis

Reporting Type	Reported (Q)	Looked (EM)	Not looked (EM)
Concept: Winter			
2-sides (ex. December / February)	13	7	6
1-side (ex. November/December)	7		
Concept :Night			
2-sides (ex. 23:00 - 3:00)	13	5	8
1-side (ex. 21:00 -23:00)	7		

4.2.4 Analysis 4: Investigation of Cyclic Events -2

After dividing the participants into three categories (one-sided linear, two-sided linear and round), the analysis with trend assessment task 1 (about winter and night) in the cyclic event was repeated. For the statistical analysis, a two-way ANOVA (Graph Group (3) x Graph type (3)) was applied to compare the gaze time and number of fixation between the three graph groups. In order to compare the difference between these three group post hoc tests were applied. Since the number of samples for each group was not equal, Hochberg's GT2 test for gaze time (since population variances were not significantly different) and the Games-Howell test for fixation count (since population variances differed) were chosen. The descriptive statistics is given in Appendix O.

Gaze time

There was a significant effect of graph group ($F(2, 34) = 10.092$, $\eta_p^2 = .35$, $p < .001$). Post hoc comparisons using Hochberg's GT2 test indicated that the mean score for the one-sided linear group ($M = 5.962$, $SD = 1.314$) was highly significantly different from the two-sided linear group ($M = 9.165$, $SD = 2.730$). Additionally, the mean score for the round group ($M = 7.237$, $SD = 2.249$) was highly significantly different from the two-sided linear group, see Figure 4-32. On the other hand, the difference between one sided linear group and round graph group was not significant. This result indicates that subjects in the one-sided linear group as well as in the round graph group tended to look at one area only and could gather the information relatively quickly there whereas subjects in the two-sided linear group tended to look at two separate areas at the edges of the linear graph and needed more time to gather the relevant information.

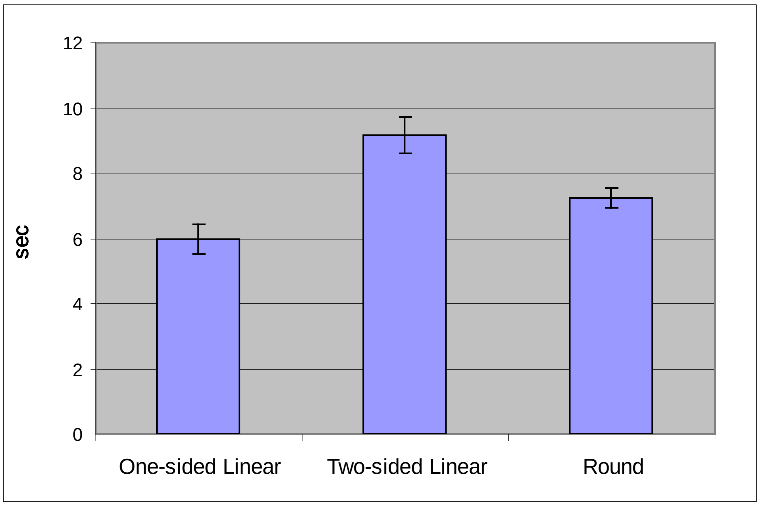


Figure 4-32: Gaze time in each graph group

On the other hand, there was no significant main effect of graph type (area/bar/line) ($F(2, 68) = .691$, $p>.05$). This indicates that the gaze time of graphs that have different graph types were the same in general. Furthermore, there was no significant interaction between graph type and graph design group ($F(4, 68) =1.263$, $\eta_p^2=.10$, $p>.05$) (Figure 4-33).

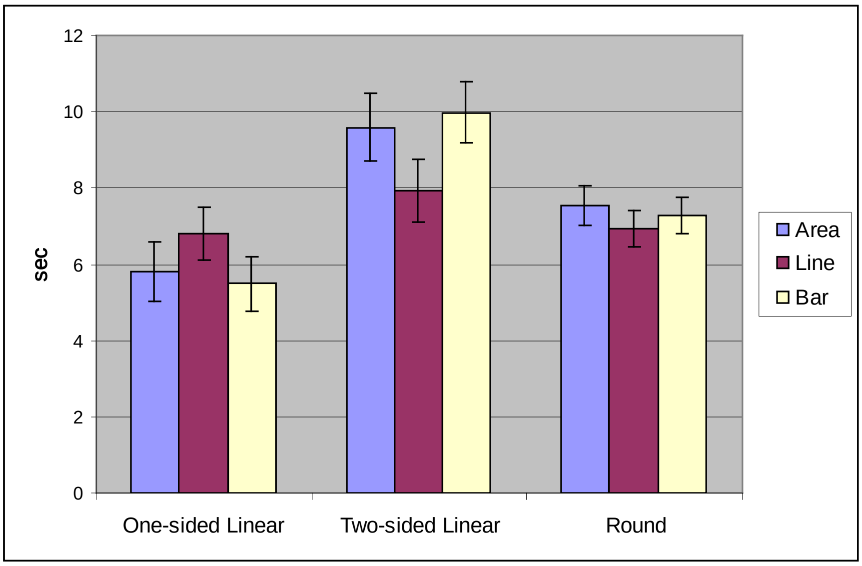


Figure 4-33: Gaze time for each graph groups according to graph type

Fixation Count

There was a significant effect of graph group (F (2, 34) =10.220, η_p^2=.40, p<.001). Post hoc comparisons using the Games-Howell test indicated that the mean score for the one-sided linear group (M = 21.650, SD =5.783) was significantly different from the two-sided linear group (M = 33.976, SD = 11.344). Additionally, the mean score for the round group (M = 25.775, SD =7.041) was significantly different from the two-sided linear group, see Figure 4-34. However, the one-sided linear group did not significantly differ from the round graph group.

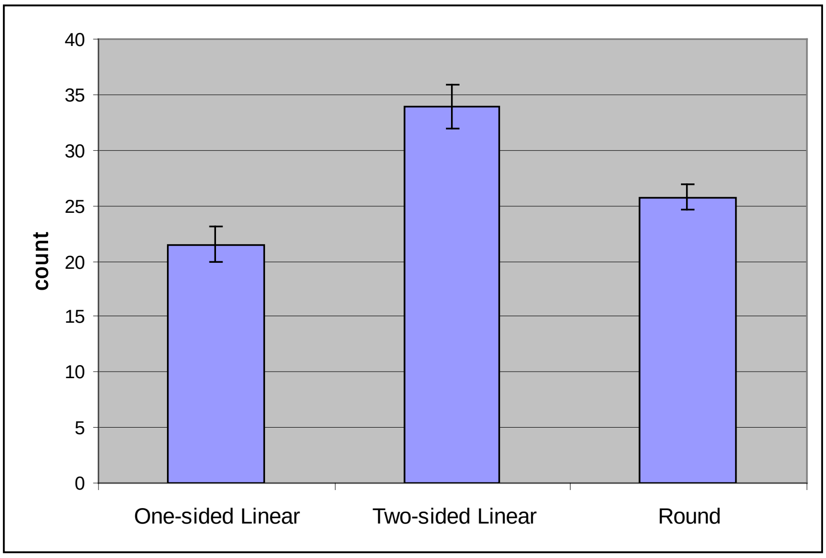


Figure 4-34: Gaze time for each between subject groups

There was no significant main effect of graph type (area/bar/line) (F (2, 68) = 1.140, p>.05). Additionally, there was no significant interaction between graph type and graph group (F (4, 68) =2.585, η_p^2= .13, p>.05) (Figure 4-35).

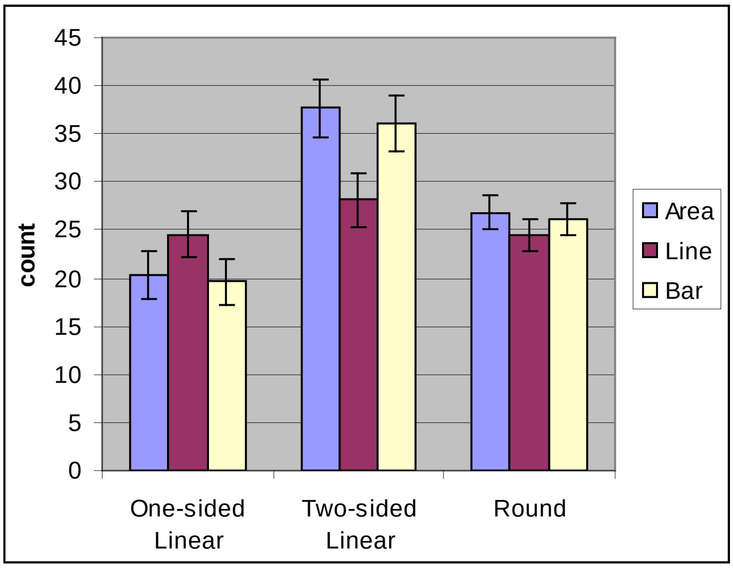


Figure 4-35: Fixation Count according to graph type and graph group

Number of the Errors in the Decision Task

In order to compare the number of the errors committed by the participants in the decision task (trend assessment task) of the cyclic event between the three graph groups (one-sided linear, two-sided linear and round), a one-way ANOVA and planned contrast were applied. Their test statistics are given below (Table 4-4).

Table 4-4: Descriptive Statistics for the number of the errors in the decision task

Graph Group	Mean	Std. Deviation	N
One-Sided	1,55	1,13	9
Two-Sided	1,00	0,75	8
Round	0,55	0,88	20
Total	0,89	1099	37

The results indicated that there was a significant main effect of graph group (F (2, 34) =3.727, η_p^2=.18, p<.05). Planned contrasts (Helmert contrast) that compares the first group (one-sided) vs. the other two groups (two-sided, round) revealed that the one-sided linear group (M = 1.55, SD =1.25) made significantly more errors in the judgment of the cyclic events than the two-sided linear group (M = 1.00, SD =.75) and the round group (M = .55, SD =.88). On the other hand, the contrast between the two-sided linear group and the round group was not significant. Post hoc comparisons using the Bonferroni (with directional hypotheses) test also indicated that the errors score of one-sided linear group was significantly higher than that of the round graph group. Although, the difference in the error score were not significant between one-sided linear and two-sided linear groups and between two-sided linear and the round groups (see Figure 4- 36).

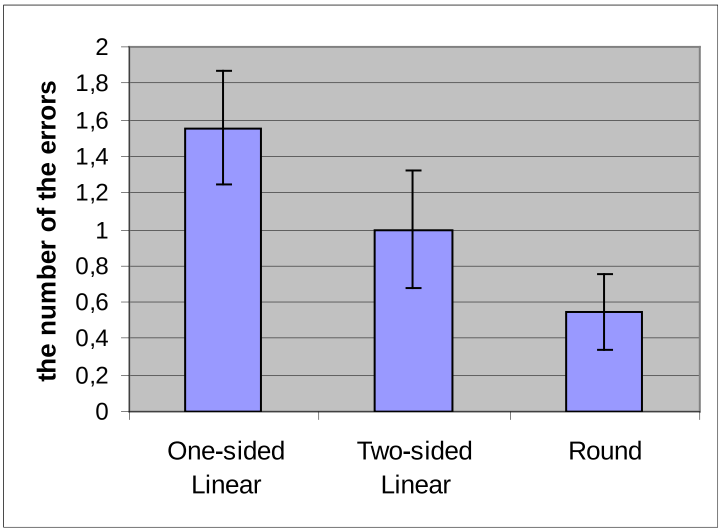
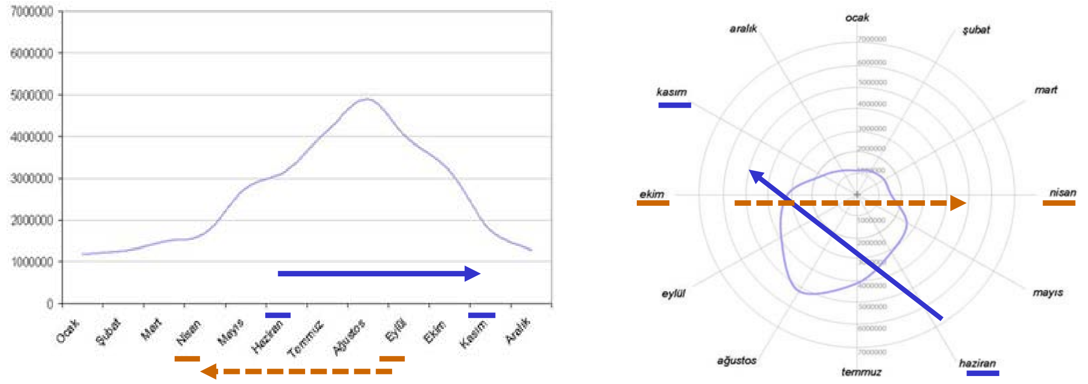


Figure 4-36: Mean number of errors committed by the participants in the decision task (of trend assessment) according to Graph Group

4.2.5 Analysis 5: Word Order Effect in Discrete Comparison Tasks

It is assumed that the comprehension of the graph is easier if the order of the stimuli is coherent with the order of the data presented in the graph. In order to investigate this effect in both of the graph design, the scores of the discrete comparison tasks, namely DCT-1 (congruous) and DCT-2 (incongruous) sentences, for each of the event type, the graph type, and the graph design were evaluated with additional analysis.

In linear graph, while the order of the words in the DCT-1 sentence was congruous with that in graph, it was incongruous in the DCT-2 (see Figure 4-37). On the other hand, since the labels in the circular timeline of the round graphs was located in clockwise order, while the DCT-1 sentence has an incongruous order, the order of the DCT-2 sentence is congruous with that in the graph.



DCT1 :
The number of tourists coming to Turkey in June is higher than in December

DCT2 :
The number of tourists coming to Turkey in September is higher than in April.

Figure 4-37: Congruous and incongruous sentences according to the task type and the graph design

Gaze time and fixation count were analyzed with a four-way (3 (Graph type: Area, Line, Bar) * 2 (Graph design: Round, Linear) x 2 task (DCT-1, DCT-2) x 2 event type (cyclic, trend)) mixed ANOVA. The descriptive statistics are given in Appendix P.

Gaze time

The ANOVA on the gaze time indicated that there was no significant effect of graph design F (1, 27) =.505, p>.05. This means that gaze times in the linear and the round graph design were the same in general. However, there was a significant main effect of event type (F (1, 27) =11.419, η_p^2=.27, p<.05).

There was no significant main effect of task (discrete comprehension-1 / discrete comprehension-2) (F (1, 27) =.55, p>.05). There was a significant interaction between tasks and graph design (F (1, 27) =5.874, η_p^2=.15, p<.05). In the linear graph, DCT-1 took less time since the word order in the graph was coherent with the sentence. On the other hand, in round graphs, DCT-2 takes less time because this word order in the round graph is also coherent with the sentence, because the labels are arranged in clock-wise order (Figure 4-38).

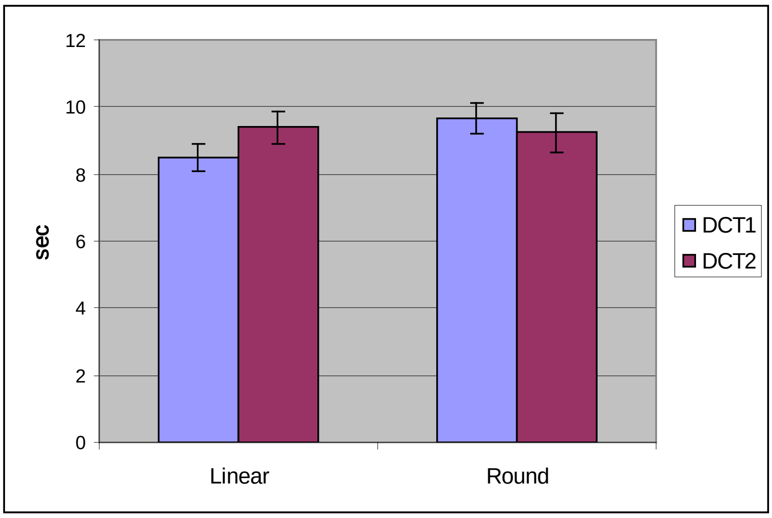


Figure 4-38: Gaze time for graph design according to discrete comparison tasks

There was no significant main effect of the graph type (area, bar and line) ($F (2, 54) =.1.270$, $p>.05$). Again, there was no significant interaction between the graph type and the graph design ($F (2, 54) =1.217$, $p>.05$). All of the other two-way interactions between the event type and the task type ($F (1, 27) =.000$, $p>.05$), between the event type and the graph type ($F (2, 54) =1.576$, $p>.05$) and between the task type and the graph type ($F (2, 54) =.596$, $p>.05$ were insignificant. The three-way interactions between the event type, the task type and the graph design ($F (1, 27) =.13$, $p>.05$) , between the event type, the graph type and the graph design ($F (2, 54) =2.552$, $p>.05$), between the graph type, the task type and the graph design ($F (2, 54) =.909$, $p>.05$) and between the event type, the task type and the graph type ($F (2, 54) =2.733$, $\eta_p^2=.10$, $p>.05$) were also insignificant. Lastly, the four-way interaction between the event type, the task type, the graph type and the graph design ($F (2, 54) =.174$, $p>.05$) was not significant.

Fixation Count

Like for gaze time, there was no significant effect of graph design for fixation count ($F (1, 27) =2.786$, $\eta_p^2=.10$, $p>.05$). There was a significant main effect of event type ($F (1, 27) =5.109$, $\eta_p^2=.16$, $p<.05$). The interaction between the event type and the graph design was not significant ($F (1, 27) = 1.377$, $p>.05$). There was no significant main effect of task (DCT-1, DCT-2) ($F (1, 27) = .599$, $p>.05$). On the other hand, there was a significant interaction between the task type and the graph design (F (1,

27) = 4.454, η_p^2=.14, p<.05). In the linear graph, DCT-1 tasks can be performed with fewer fixations since the word order in the graph is coherent with the sentence. On the other hand, in the round graph, DCT-2 was comprehended with fewer fixations. Furthermore, graph type had no effects on fixation count either (F (2, 54) = 1.001, r=.13, p>.05) (Figure 4-39).

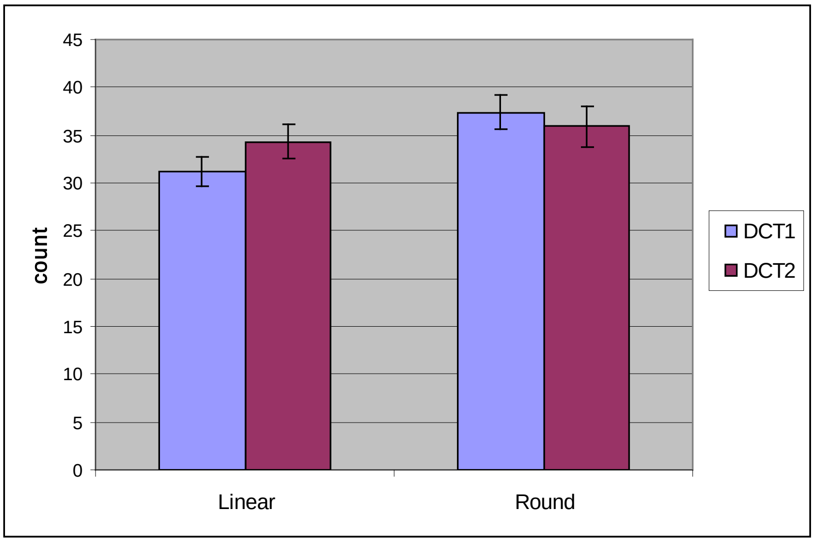


Figure 4-39: Fixation Count of the discrete comparison tasks according to graph design

Furthermore, all of the other two-way interactions between the event type and the task type (F (1, 27) =.473, p>.05), between the event type and the graph type (F (2, 54) =2.057, p>.05) and between the task type and the graph type (F (2, 54) =1.064, p>.05 were insignificant. The three-way interactions between the event type, the task type and the graph design (F (1, 27) =.19, p>.05) , between the event type, the graph type and the graph design (F (2, 54) =1.465, p>.05), between the graph type, the task type and the graph design (F (2, 54) =1.122, r=.14, p>.05) and between the event type, the task type and the graph type (F (2, 54) =1.833, p>.05) were also insignificant. Lastly, the four-way interaction between the event type, the task type, the graph type and the graph design (F (2, 54) =1.459, p>.05) was not significant.

4.3 Scan Path Analysis

The scan paths of all graphs in Part-2 of the experiment were analyzed according to task type (TAT, DCT), event type (cyclic, trend) and graph designs (linear, round).

The graphs were analyzed individually by using Tobii Studio Coding Schema to obtain information about the general way of visual investigation of graphs. The gaze plots for each combination presented in this section were picked from data which represents the general distribution/tendency ideally. In gaze plot representations, while circles correspond to fixation points, thin lines correspond to saccades. Moreover, bigger circles mean longer looking times to those particular points. Since increasing data points in gaze plot representations decreases understandability of the scan path, gaze plots were divided and numbered in order to make tracing easier.

4.3.1 The Trend Assessment Task

4.2.1.1 Cyclic Event:

As mentioned before (in Section 3.3.2.2), there are two different types of trend assessment tasks in this experiment. One of them (TAT-1) was conducted in order to get information about the target words which are presented separately in the linear graph, such as winter season or night. The second type (TAT-2) is about summer or noon which is presented adjacently in the linear graph. In the round graph, TAT-1 target words are also presented adjacently like TAT-2 target words. Sentence examples from both task types for a cyclic event are given below.

TAT-1: In Lake Eymir, the amount of zooplankton increases at night

TAT-2: In Lake Eymir, the amount of zooplankton decreases at noon.

Trend Assessment Task 1

The overview of the scan path analysis showed that the participants in the linear graph design group started with reading the sentence (sometimes twice), then their gaze tended to look at the center of the timeline or directly to the left or right side of the timeline which contained the target words (“hours” for night; “months” for winter) (Figure 4-37a). Then the related graph area and the middle area were looked at (Figure 4-37a). Before looking at the second target word in the timeline and its related area on the graph, the participants tended to return to the sentence, mostly to the action verb (“increases”; “decreases”) which contains the trend information (Figure 4-40b). Next, in order to check the decision, all informative data points were looked at again and finally the task was finished by checking the sentence (Figure 4-40c-d).

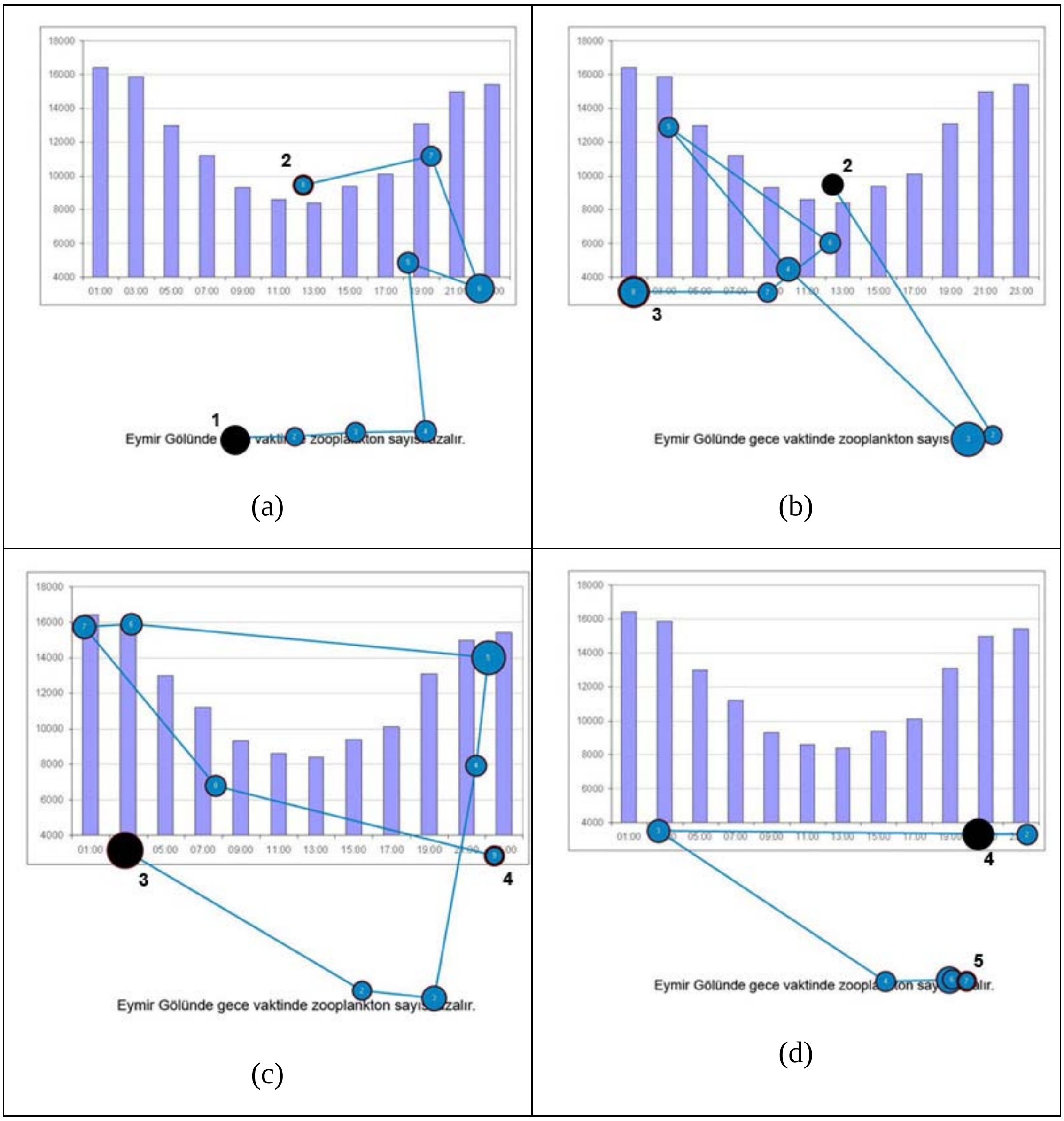


Figure 4-40: Linear Graph – Cyclic Event – Trend Assessment Task (TAT 1- Bar graph)

In the round graph, the sentence was read, then the target words (“hours” for night; “months” for winter) and their related area in the graph or the opposite concepts of targets (“hours” for noon; “months” for summer) were looked at (Figure 4-41a). Between fixating on the target concept and its graphical counterpart, or after gathering information about both of them, the sentence is read again, and then task ends (Figure 4-41b).

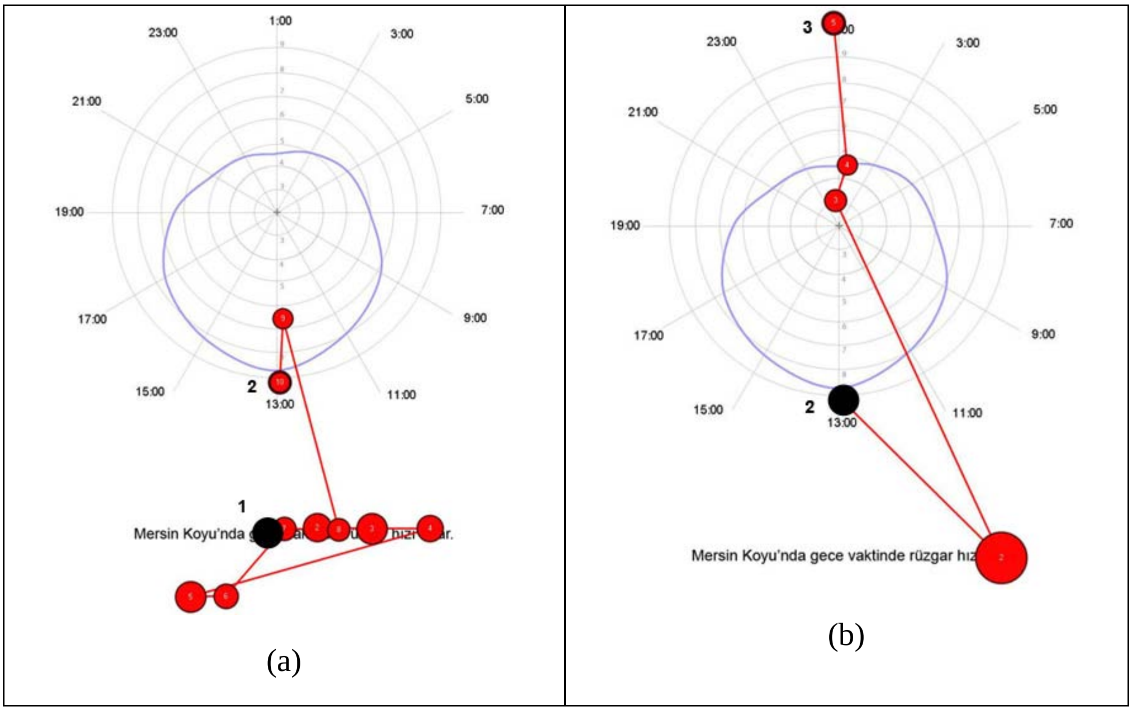


Figure 4-41: Round Graph – Cyclic Event – Trend Assessment Task (TAT 1- Line graph)

The Trend Assessment Task 2

In the linear graph design, the sentence is read first. Most of the participants had a tendency to look at the beginning of the graph rather than at the end point to make a trend judgment. Then the target words for the related concept in the timeline were attended. Then the sentence was read again. The graph area and, most of the time, the y –axis label are read. In order to check the decision, final fixations were made on the timeline and the sentence (Figure 4-42a, b, and c).

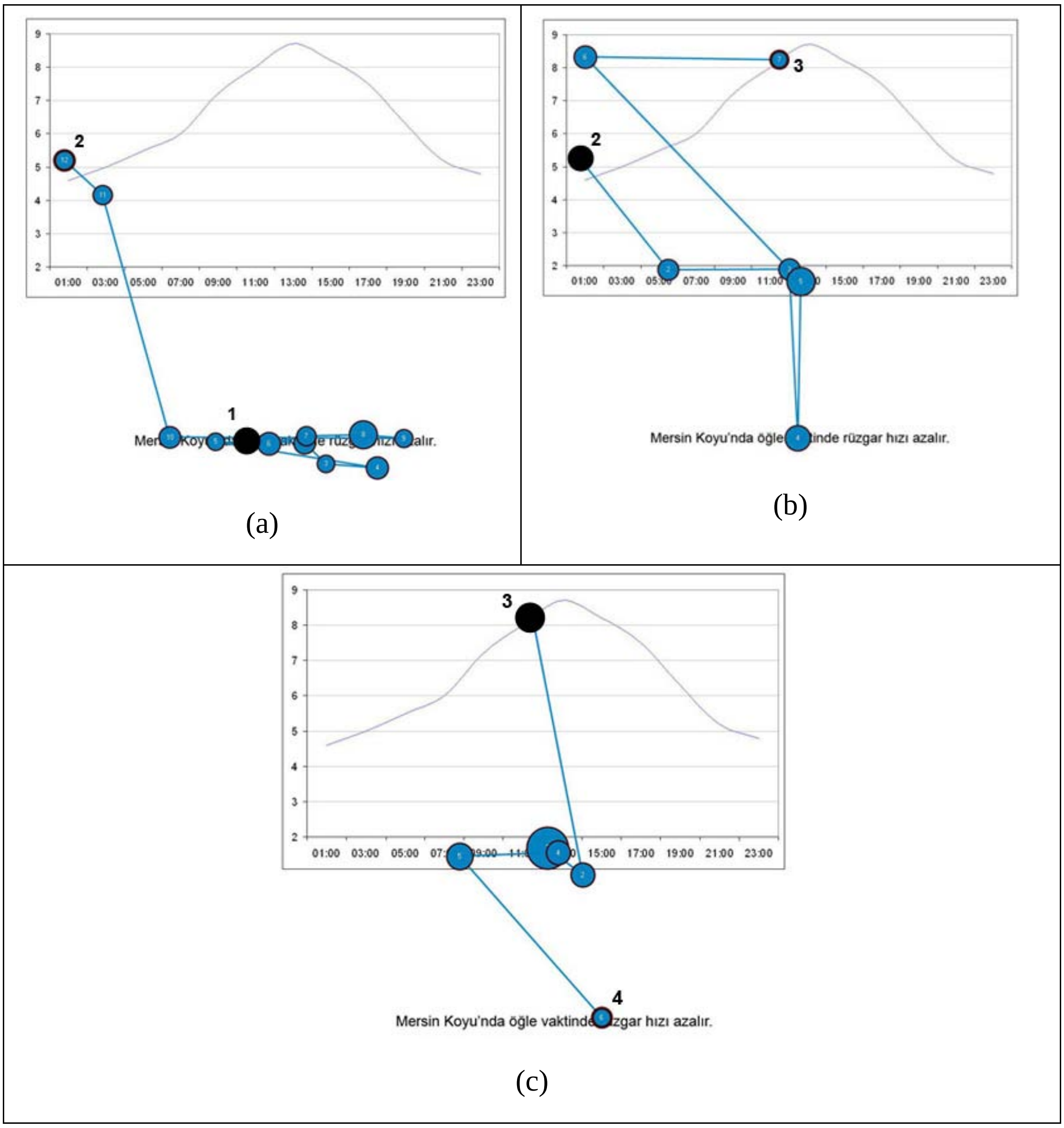


Figure 4-42: Linear Graph – Cyclic Event – Trend Assessment Task (TAT 2- Line graph)

In the round graph design, the scan path for TAT -2 for cyclic events is very similar to that of TAT-1. The sentence was read, target words and their respective areas in the graph was visited. The order of visiting target areas for target words or their visual counterpart was not strict. Finally, the sentence was read again and the task ended (Figure 4-43a, b).

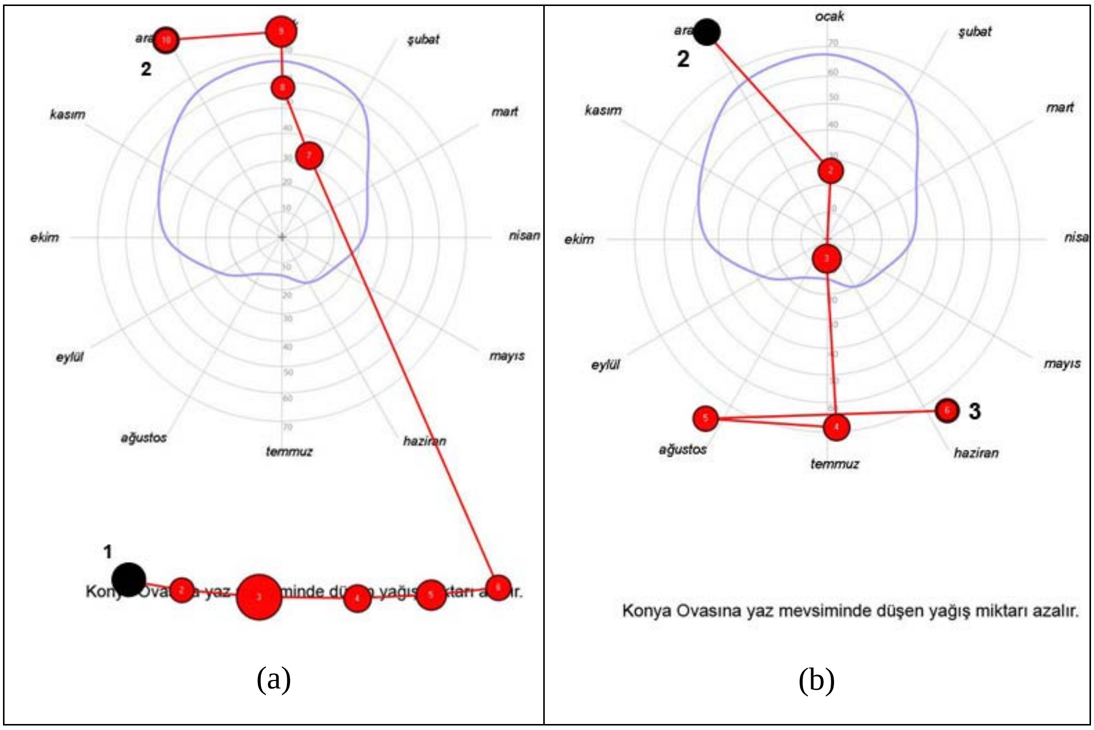


Figure 4-43: Round Graph – Cyclic Event – Trend Assessment Task (TAT 2- Line graph)

4.2.1.2 Trend Event:

There are also two different types of trend assessment tasks for trend events. TAT-1 for trend events asks for the evaluation of change as time passes. TAT-2 is about the change in the last years. Sentence examples from both task types for the trend event are given below.

TAT-1: The number of monachus monkseal has increased over the years.
TAT-2: The number of monachus monkseal has increased in the last years.

Trend Assessment Task 1

In the linear graph, when they are evaluating TAT-1 for the trend event, participants read the sentence, then went to the last two or three years in the timeline and their corresponding values in the graph, then they sometimes returned to the sentence and following the trend by making a left eye movement across the data (increasing or decreasing) (Figure 4-44a). Usually, they looked at the y-axis label, and went to the sentence to check the judgment that they had made (Figure 4-44b).

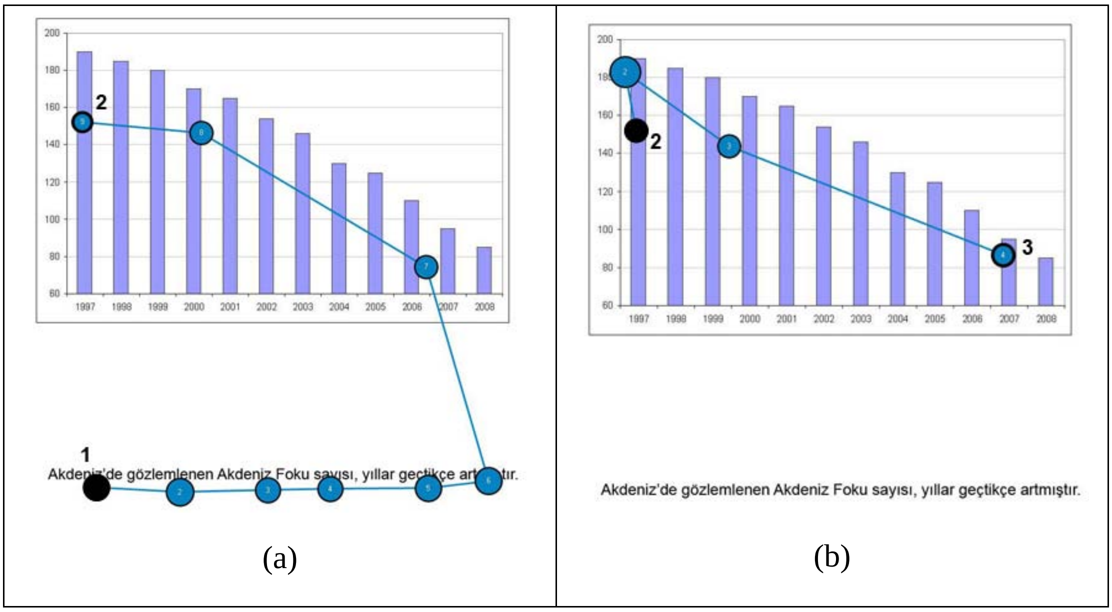


Figure 4-44: Linear Graph – Trend Event – Trend Assessment Task (TAT 1- Bar graph)

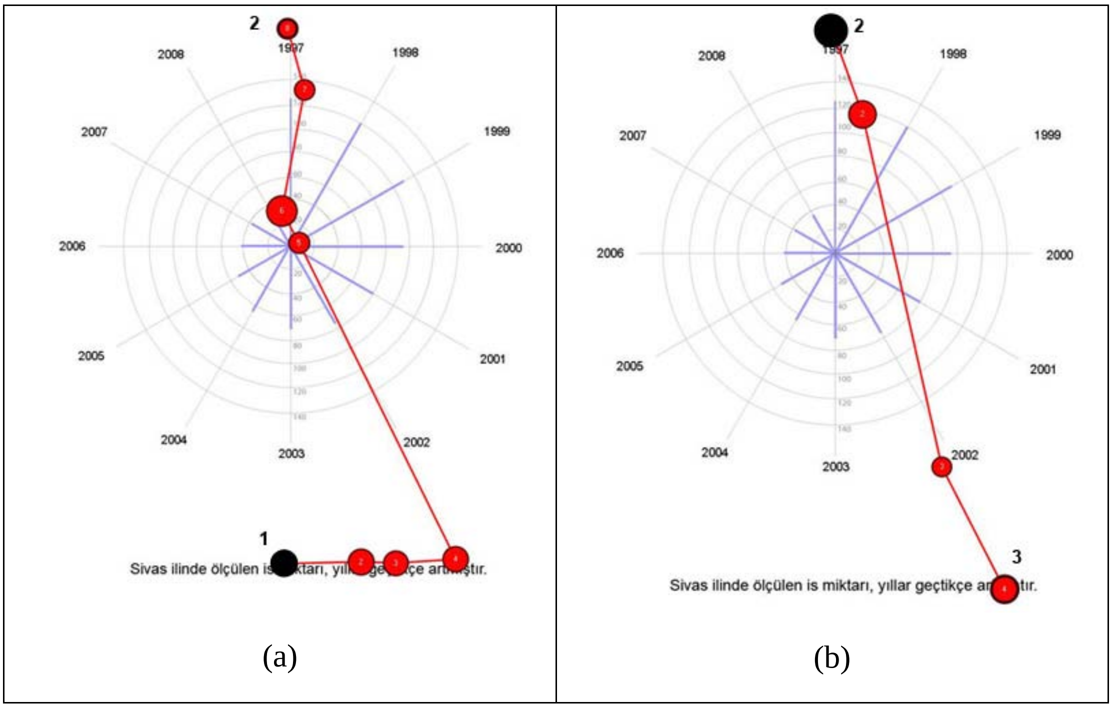


Figure 4-45: Round Graph – Trend Event – Trend Assessment Task (TAT 1- Bar graph)

After reading the sentence, participants in the round graph design group firstly looked at the center of the graph and read the data that belonged to the latest years (Figure 4-45a). Next, the top most point of the graph, where the end and start points of the timeline are located, were visited (Figure 4-45a). After making several fixations in this area, participants returned to the sentence and finished the task (Figure 4-45b).

Trend Assessment Task 2

TAT-2 scanpaths of trend event are very similar to TAT-1 scanpaths. Firstly, the sentence was read, the gaze went to the timeline and the graph area which contain targets (Figure 4-46a). Afterwards, a left eye movement parallel to the trend of the graph was done. Lastly, the timeline and sentence were checked (Figure 4-46b).

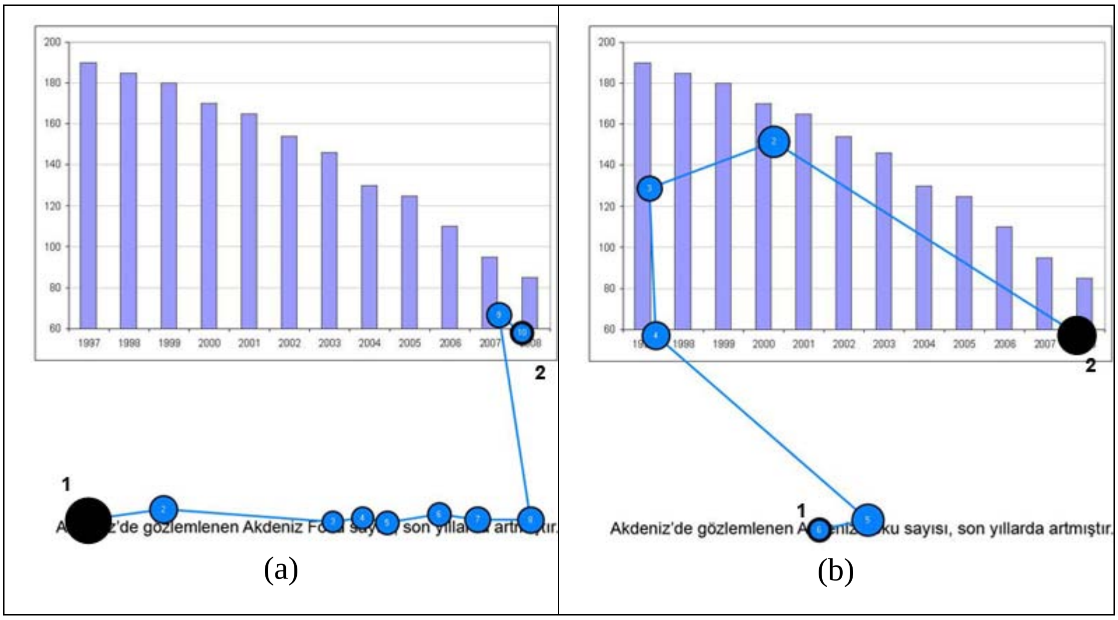


Figure 4-46: Linear Graph – Trend Event – Trend Assessment Task (TAT-2- Bar graph)

Gaze patterns of TAT-1 and TAT-2 tasks are also similar in the round graph. The center of the graph area was looked at after reading the sentence. The target labels in the timeline and graph area which contain start and end points of the timeline adjacently were attended before the task ended with a sentence check (Figure 4-47).

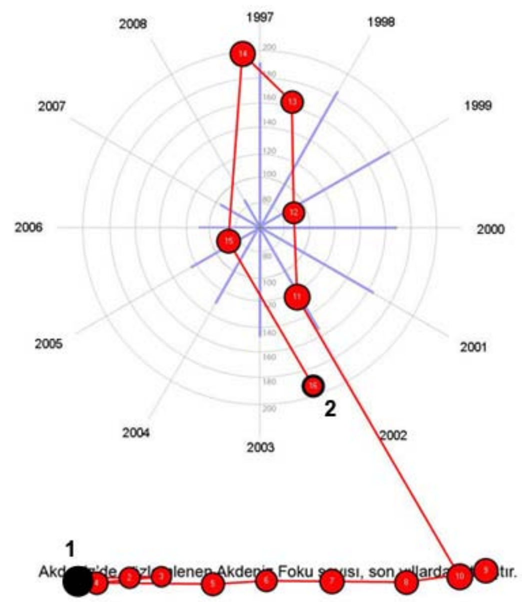


Figure 4-47: Round Graph – Trend Event – Trend Assessment Task (TAT-2 - Bar graph)

4.3.2 The Discrete Comparison Task

The Discrete comparison task has also two categories for each event type. In DCT-1, the order of the target words in sentence is congruent with the order of that information in the graph for the linear timeline. The target words are presented incongruously, i.e., in reverse order in DCT-2.

4.2.1.1 Cyclic Event:

The sentence examples from both task types in discrete comparison for cyclic event are given below.

DCT-1: In Lake Eymir, the amount of zooplankton at 7am is higher than at 1pm.
DCT-2: In Lake Eymir, the amount of zooplankton at 6pm is lower than at 3 am.

Discrete Comparison Task 1

The gaze pattern of the linear group participants showed that first the target and its related area in the graph were attended, after the sentence had been read (Figure 4-

45a). Generally, participants had a tendency to return to the sentence to look at the second target word (Figure 4-48b). Then, target words in the timeline and their respective areas in the graph or labels were visited (Figure 4-48b, c). Lastly, all informative data points were checked again before the task ended (Figure 4-48d, e).

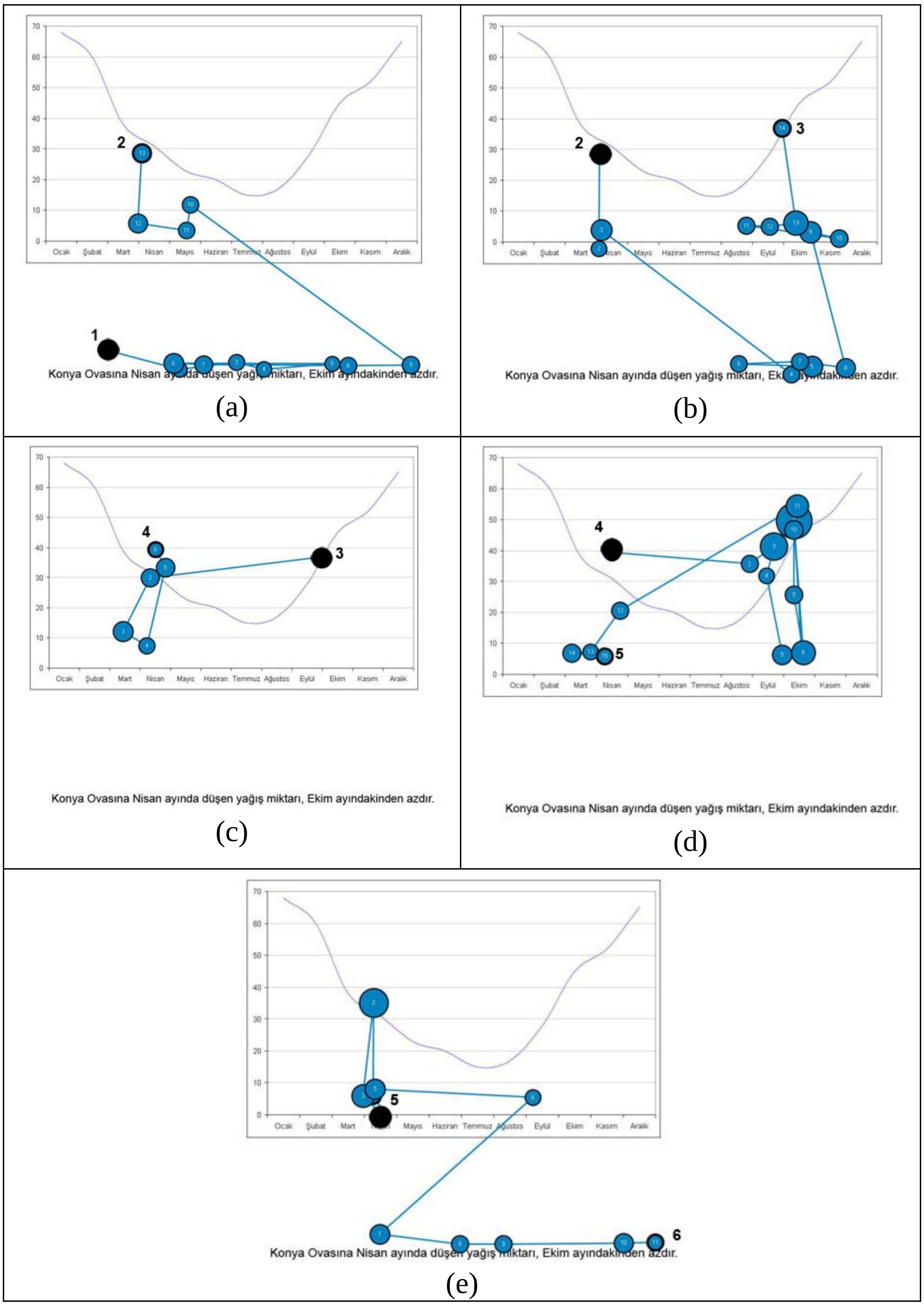


Figure 4-48: Linear Graph – Cyclic Event – Discrete Comparison Task (DCT-1 – Line Graph)

Discrete comparisons in the round graphs which represent cyclic events also started with reading the sentence (Figure 4-49a). After the first target word and its respective graph area were attended, the second target word in the timeline and its respective graph area were visited (Figure 4-49a, b). Between these, participants sometimes returned to the sentence to check target words. Several fixations to compare values of the target words were made, and then the task ended with a sentence check (Figure 4-49c).

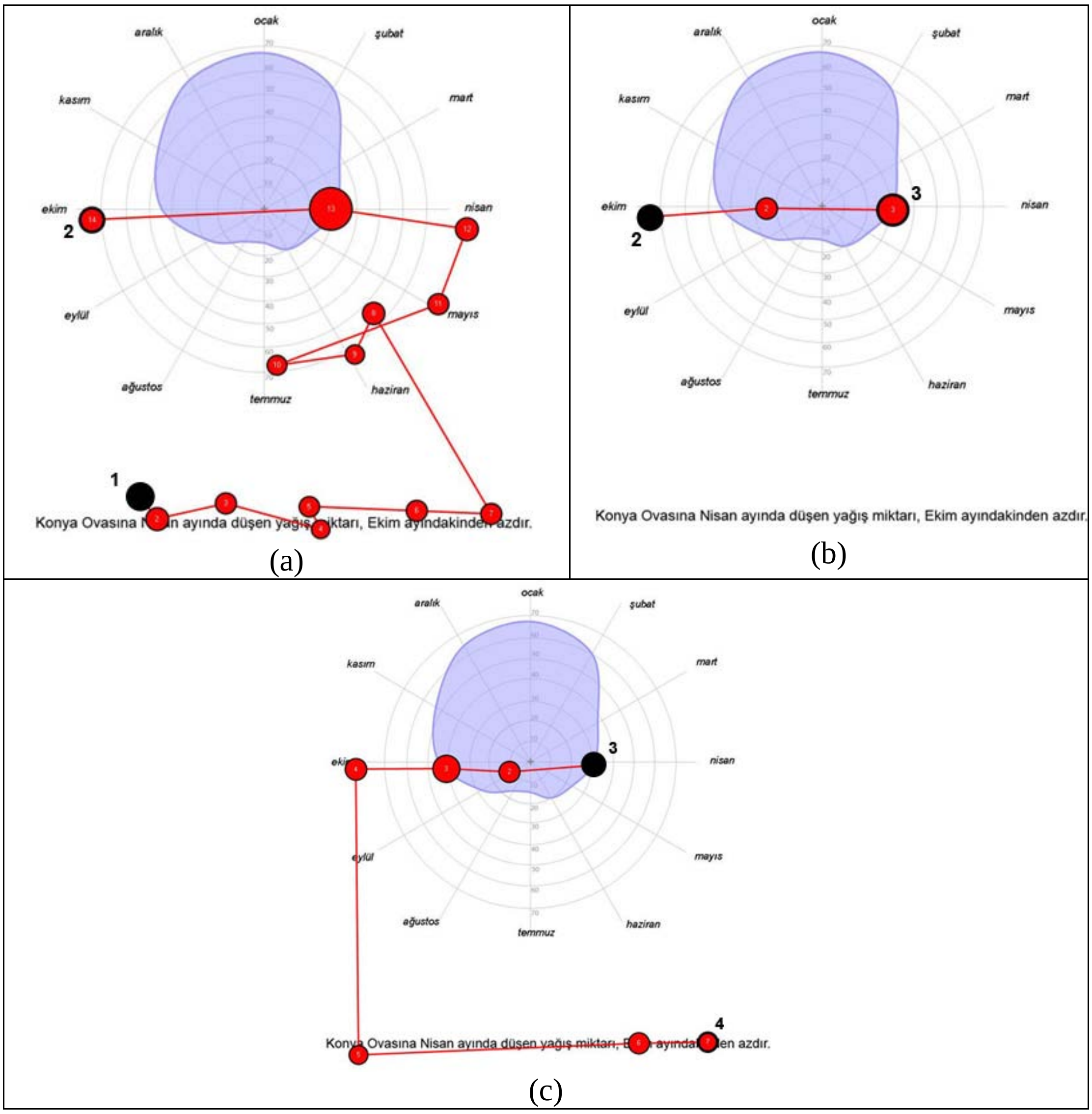


Figure 4-49: Round Graph – Cyclic Event – Discrete Comparison Task (DCT-1 – Area Graph)

Discrete Comparison Task 2

The scanpath analysis for DCT-2 in cyclic events represented with linear graphs showed that the participants started to read the graph with the sentence, and then the first target word in timeline and its respective field on the graph was visited (Figure 4-50a). . The second target word and its area were then looked at (Figure 4-50b). Then several fixations on this area could be made before ending the task (Figure 4-50c).

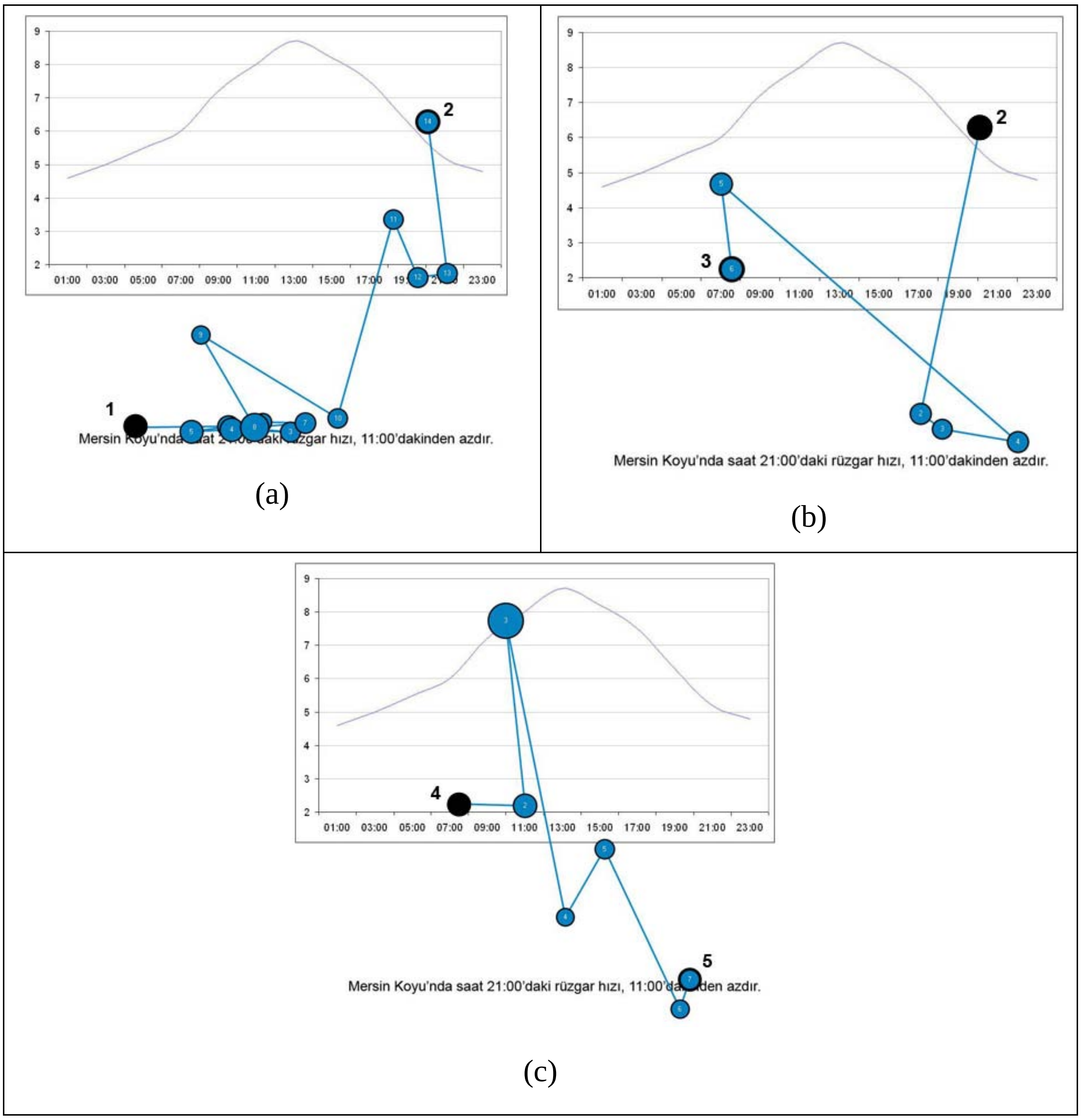


Figure 4-50: Linear Graph – Cyclic Event – Discrete Comparison Task (DCT-2 – Line Graph)

DCT-2 for cyclic events in round graphs was very similar to DCT-1. Firstly, the sentence was read and the target word in time line and its graph value were looked at (Figure 4-51a). Then the participants usually tended to return to the sentence and look at the second target word (Figure 4-51b). Then, after looking up its value in the graph, the decision was made (Figure 4-51b).

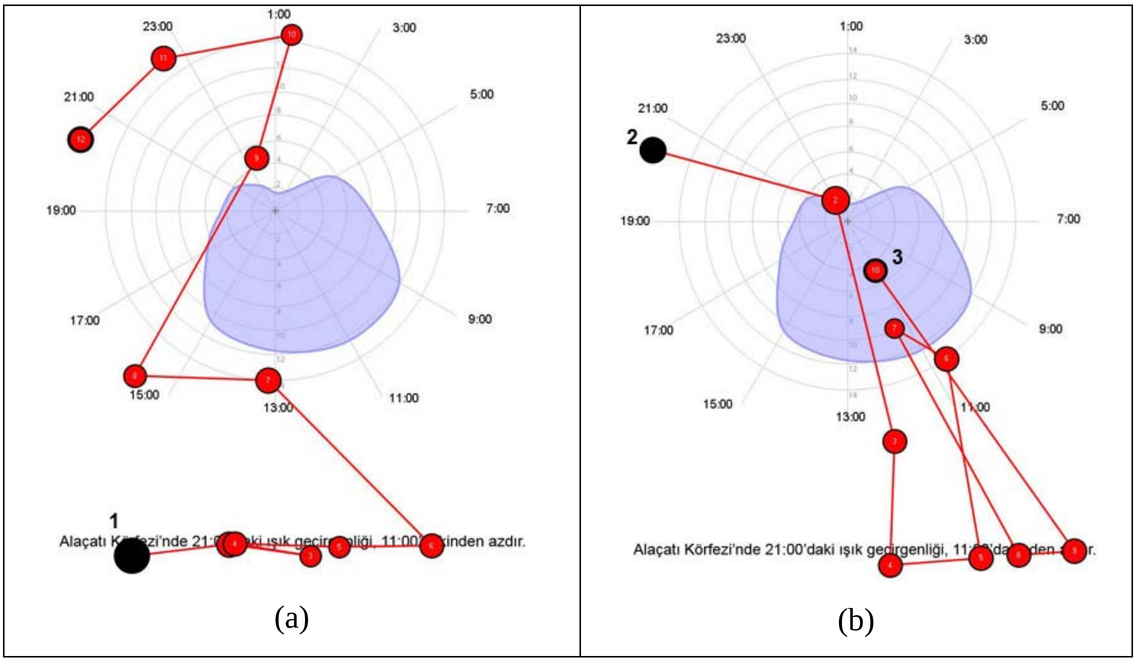


Figure 4-51: Round Graph – Cyclic Event – Discrete Comparison Task (DCT-2 – Area Graph)

4.2.2.2 Trend Events:

Sentence examples from both task types in discrete comparisons for trend events are given below.

DCT-1: The number of monachus monkseal observed in 2001 is higher than in 2006

DCT-2: The number of monachus monkseal observed in 2004 is lower than in 2001

Discrete Comparison Task 1

In the linear graph design, after reading the sentence, participant gaze went to the target word in the timeline and its related area in the graph (Figure 4-52a). Then participants tended to look at the sentence again for the second target word (Figure 4-52a). Next, the second target word and its value in the graph were attended (Figure 4-

52b). Then the gaze sometimes made a left movement to check the trend and participants read off the y- axis value (Figure 4-52c). Afterwards, the task was finished by checking the sentence and making a key press for the decision (Figure 4-52c).

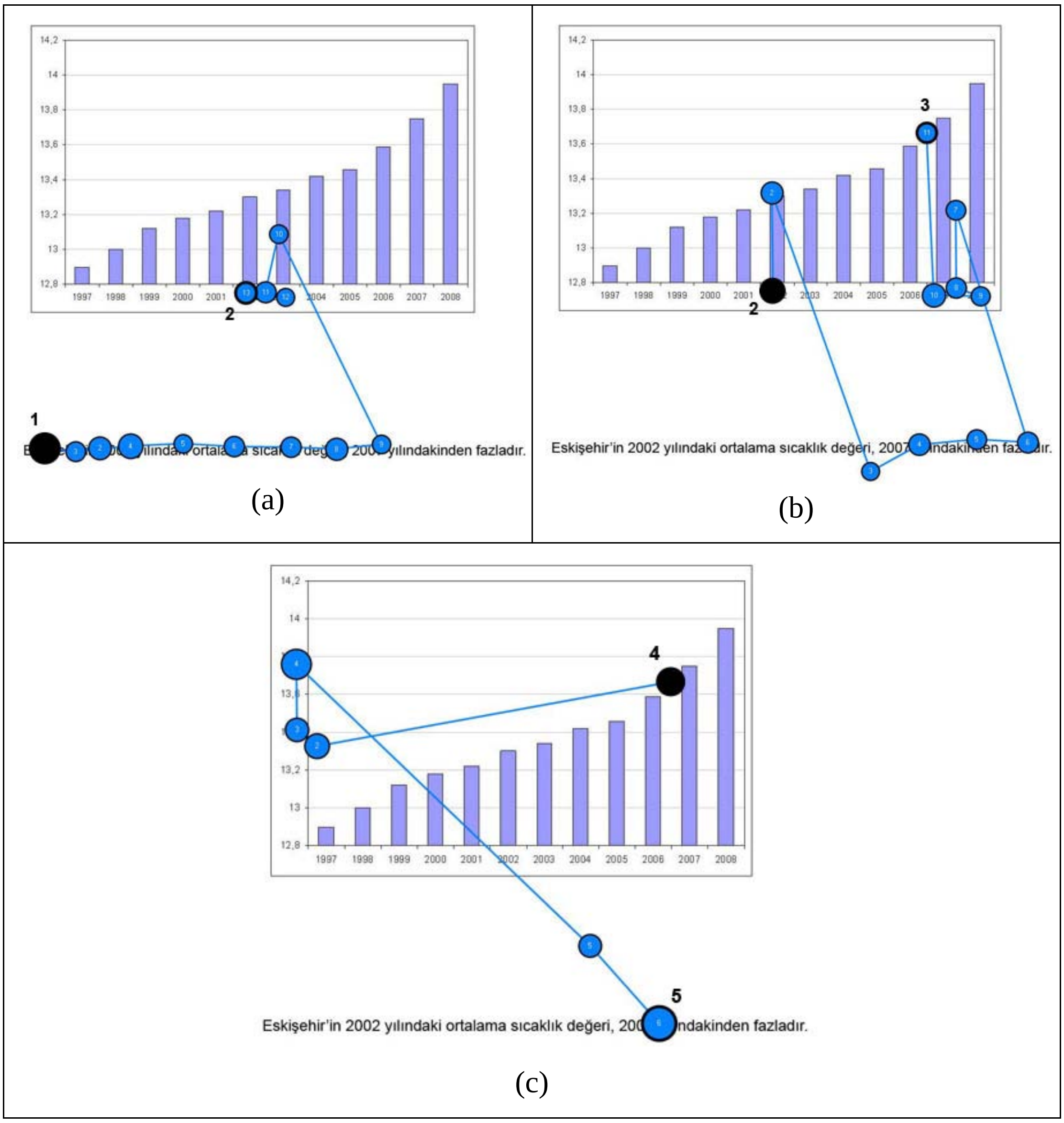


Figure 4-52: Linear Graph – Trend Event – Discrete Comparison Task (DCT-1 –Bar Graph)

The scanpath of DCT-1 in the trend event represented in the round graph design indicated that after reading sentence, gaze went to center of the graph area, and then the first target word was searched (Figure 4-53a). After finding target in the timeline and its respective value in the graph, usually the sentence was read again (Figure 4-

53b). Then the second target word and its value were also read (Figure 4-53b). Finally, after making several fixations on this area, the participant returned to the sentence to check the decision before the task was finished (Figure 4-53c).

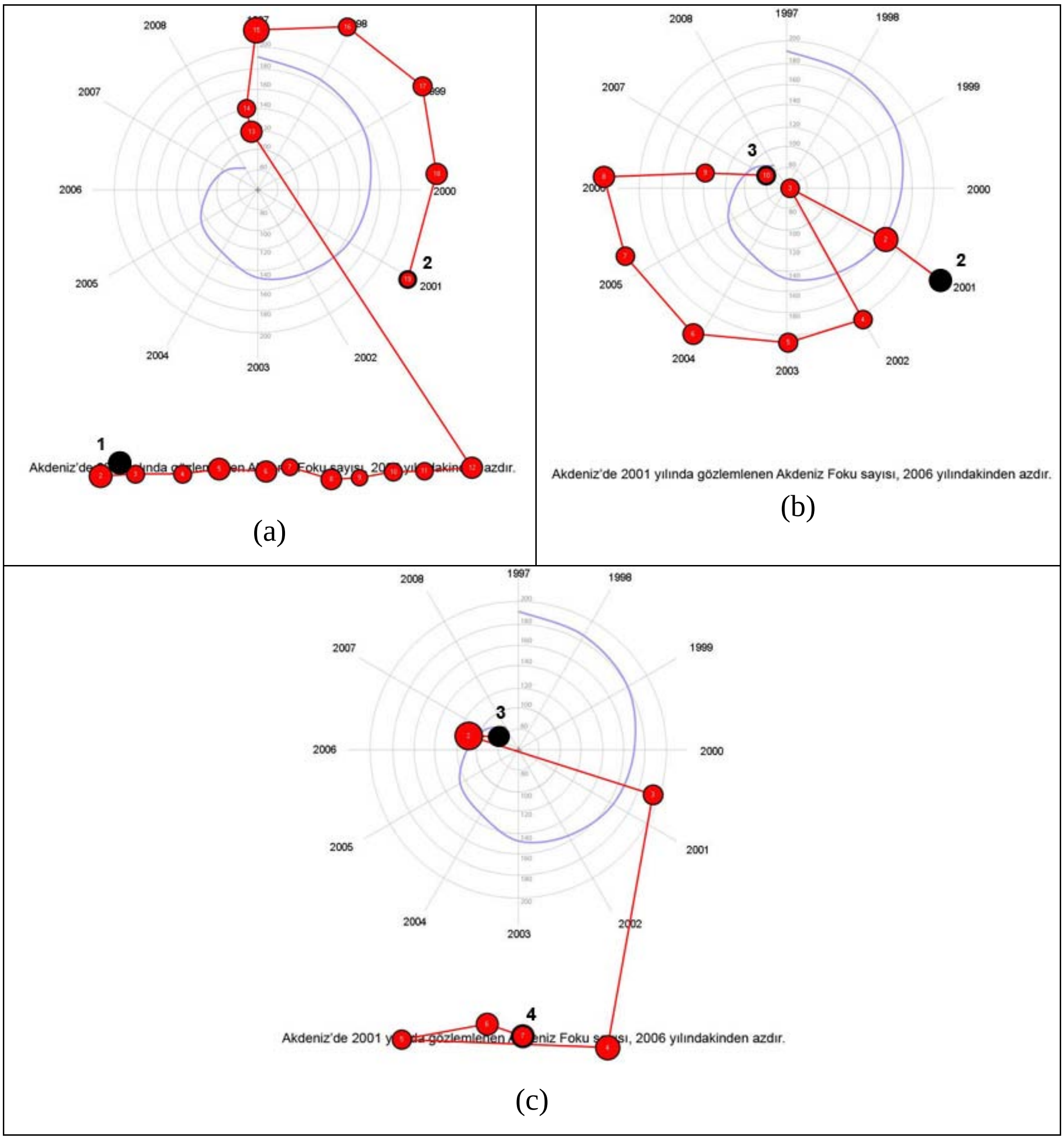


Figure 4-53: Round Graph – Trend Event – Discrete Comparison Task (DCT-1 – Line Graph)

Discrete Comparison Task 2

In linear graph, scan paths of DCT-2 were very similar to those of DCT 1. The sentence was read, first target and graph area were attended (Figure 4-54a). Then the second target in the timeline and graph area were visited (Figure 4-54a). Finally, several fixations were made on the timeline and the sentence before pressing a key for the decision (Figure 4-54b).

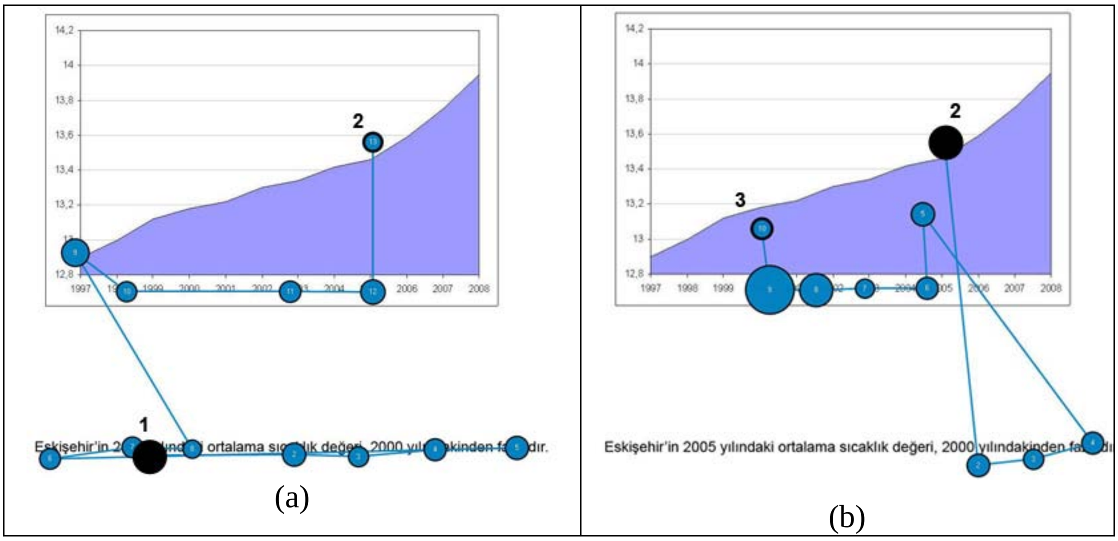


Figure 4-54: Linear Graph – Trend Event – Discrete Comparison Task (DCT-2 – Area Graph)

In the round graph design, DCT-2 was also similar to DCT-1. The center of the graph area was attended first (Figure 4-55a). Then the target word was searched (Figure 4-55b). After looking at its respective area in the graph, the second target word was searched (Figure 4-55b). Finally, its value was also read off and several fixations were made between these values before the task ended (Figure 4-55c).

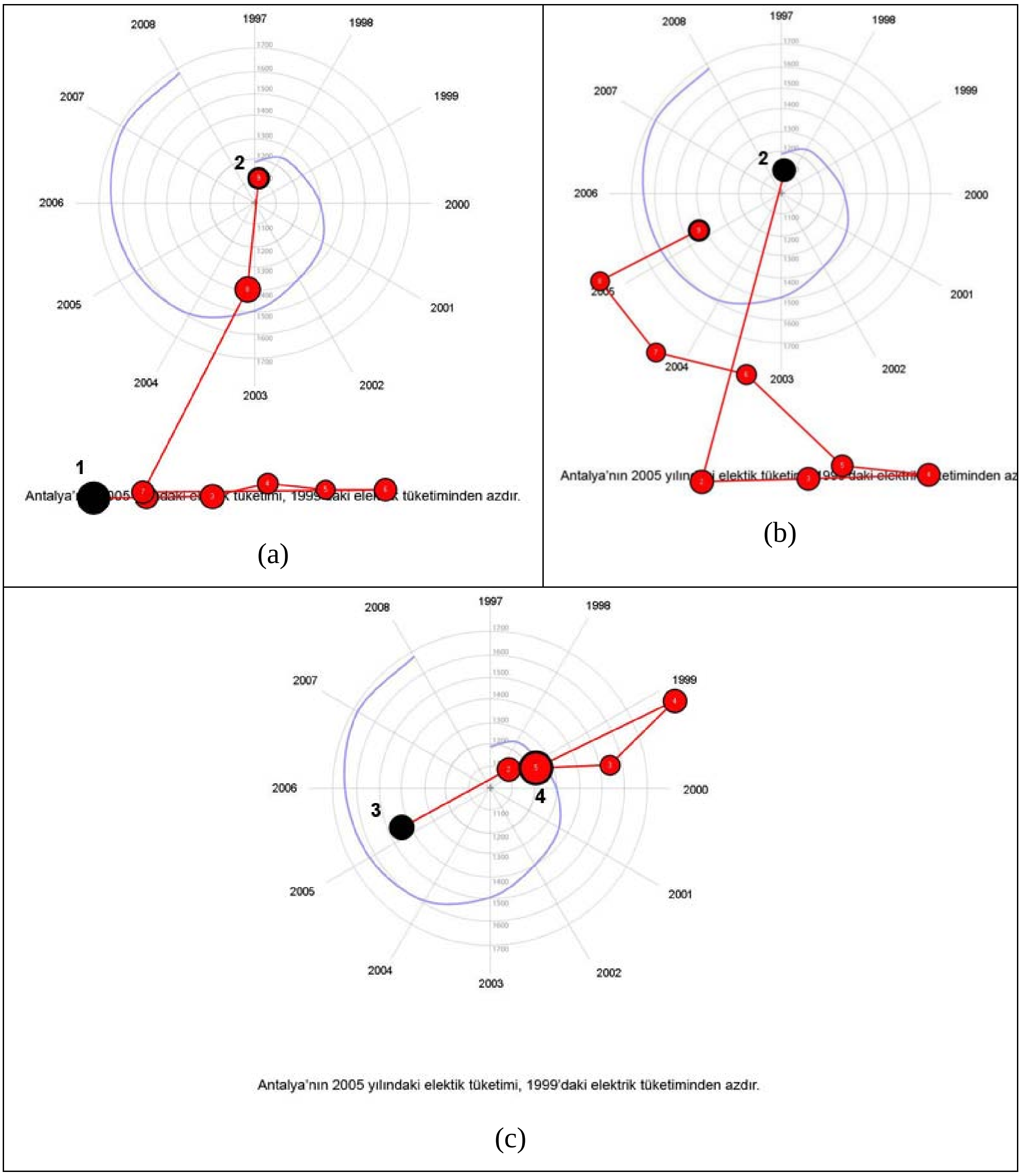


Figure 4-55: Round Graph – Trend Event – Discrete Comparison Task (DCT-2 – Line Graph)

CHAPTER 5

DISCUSSION AND CONCLUSION

5.1 Summary of the Methodology

In this study, comparison strategies in judgments on cyclic events for different tasks, graph types, and graph designs were evaluated by using two means of data collection. The first set of data is based on a questionnaire aiming at obtaining information about recollected data from different graph types in terms of value and keyword preferences. The second, main, set of data is based on the eye tracking tool providing information about subjects' gaze time, fixation count, fixation duration, and number of transition between graph elements. Scan paths, another set of data also obtained by the eye tracker, provided additional opportunity to observe gaze order during graph comprehension. The eye tracking results were supported by the questionnaire results, e.g., by obtaining the participants' opinions on the most easy graph in terms of recollection of information and on preferences on graph types, concepts, and entities which constitute these concepts.

5.2 Part 1 – Recollection of Data

The questionnaire results indicated that the bar graphs are rated higher compared to line and area graphs in terms of remembering values and relations. Since they are discrete objects and each bar corresponds to a single label in the timeline in a salient way, this feature may help participants in remembering the values and relations. On the other hand, patterns of change are continuous in nature like graph components in area and line graphs; therefore they were clearly preferred over bar graphs in terms of pattern recognition. The scores of line and area graphs are very close to each other because they have more features in common.

In addition to exploring the task performance of both groups in terms of the variables gaze time (corresponds to response time), fixation count, fixation duration, and number of transitions, participants' evaluation of the easiness to become familiar with the novel round graph design was requested right after completing Part-1 of the study. The results show that although the round graph is novel, the information that it conveys can be easily grasped.

Although the round graphs were novel to the participants, they were as good as linear graphs in terms of remembering the values and general pattern. The recollection of values, patterns, and y-axis values in both the drawing and the verbal description task did not differ for graph design, event and graph types at all. The overall results of the recollection evaluation also showed that event type, graph design or graph type have no effects on observation time. However, the results of the recollection evaluation in the description task indicate that participants focus on the value when their observation time increases rather than on the relation presented in the graph and they tend to make descriptions without using relational information.

Keyword usage was found to be affected by event type but not by graph design and graph type. While there is no difference in the usage of discrete and trend keywords between cyclic and trend events in the round graph, event type affects the usage of these keywords for describing relations in linear graphs. Discrete keywords (minimum, maximum) are preferred more in cyclic events while trend keywords are preferred more in trend events. Trends are characterized by 1 minimum and 1 maximum point (see Figure 5-1) which are readily expressed by trend keywords (increase, decrease) that involve such min-max relations intrinsically. Therefore they are more preferred in trend events in linear graphs. On the other hand, cyclic events in linear graphs have 3 informative points (2 max. and 1 min. or 1 max. and 2 min.), and the accurate description of these events needs reporting of these three points. Because of this feature, describing more than two points may have increased the use of discrete words. On the other hand, in round graphs, both cyclic and trend events have two informative data points (1 max. and 1 min.). Therefore, there is no difference between event types and also trend keywords are more preferred than discrete keywords in round graphs. This finding also suggests that discrete keywords like minimum and maximum are preferred more when there are more than two points that need to be described. The illustration of the number of salient points according to event type and graph design type is given in Figure 5-1.

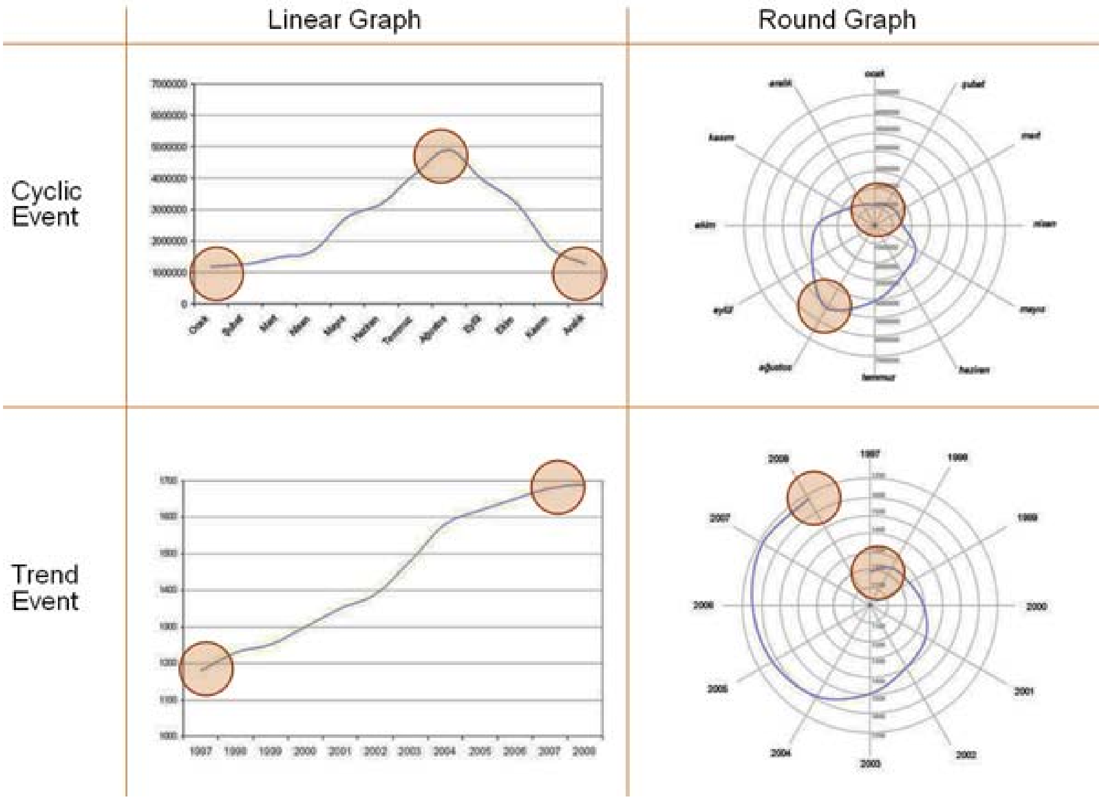


Figure 5-1: Number of salient points according to event type and graph design type

However, we found that in round graphs, describing the relation with conceptual key words (“in winter”, “during the last year” etc.) is preferred more in cyclic events than in trend events. Event type, though, does not affect the usage of these keywords in linear graphs. For example, the keyword “last year”, which is a conceptual word for a trend event, involves an end point which is quietly salient in the linear graph. On the other hand, the sphericity of the round graph may hinder the participant in grasping the end point from the graph. Therefore the occurrence of these keywords for trend events in round graphs may be low compared to linear graphs. Nevertheless, the same feature (sphericity) facilitates the occurrence of conceptual words for cyclic events in round graphs more than in linear graphs (see section 4.1.2.3). Additionally, describing the relation by using more than one keyword category is preferred in cyclic events as compared to trend events in round graphs, while there is no difference in linear graphs because of the effects of preference on the conceptual words mentioned above.

Contrary to event type, graph design does not affect the recollection of data in terms of all dependent variables: values and keyword usage.

5.3 Part-2: Judgment Task Performance

An overall four-way ANOVA had been conducted with all independent variables, namely graph type (area, line and bar), graph design (round and linear), event type (cyclic and trend), and task type (trend assessment and discrete comparison). The general results indicate that round and linear graphs that are informationally equivalent are also computationally equivalent. This is because there is no difference in any of the dependent variables: gaze time, fixation count, observation count and fixation duration (see Appendix I for the F statistics and effect sizes for all dependent variables and for all effects).

On the other hand, there is very clear and consistent effect of event type in most dependent variables with large effect sizes indicating that trend events are easier to comprehend than cyclic events, regardless of graph design (linear and round). The only significant interaction was observed for event and graph type for fixation count. This result indicates that while cyclic events are comprehended similarly in both graph types, trend events are processed with fewer fixations in linear graphs than in round graphs. However, this difference did not appear in the other dependent variables: gaze time, fixation duration, and number of transition. Furthermore, the mean scores show that while cyclic and trend events are comprehended equally well in round graphs, there is a consistent difference between these events in linear graphs. This can be also be explained by appealing to the number of informative data points in the graph mentioned in section 5.2. However, the detailed analysis of cyclic concept comprehension showed that linear graphs may misguide the interpretation of cyclic concepts. I will come back to this issue later.

The results of the ANOVA also showed a very clear and consistent effect of task type in all dependent variables with very large effect sizes. The trend assessment tasks are completed more easily and quickly as compared to discrete comparison tasks regardless of graph design. Furthermore, the interaction between event type and graph type is significant for the fixation count, indicating that comprehension of discrete entities needs more fixations in round graphs than in linear graphs. This may be due to the aligned positions of entities neatly provided by the linear timeline. Linearity is one of the factors which help participants to make comparisons with fewer fixations between two or more discrete entities. However, trend assessment tasks are completed with about the same number of fixations in both graph designs. There is no difference either in task performance in terms of gaze time, fixation duration, and number of transition in round and linear graphs.

The interactions between event type and task type observed in the fixation count and in the number of transition between the elements of the graph also indicate that while discrete comparison tasks for both event types can be completed with approximately the same number of fixation and transition between elements, trend assessment in cyclic events needs more fixations and transitions than in trend events.
An effect of graph type in comprehension of graphs was only observed for gaze time, though with a small effect size, indicating that task performance is affected by graph type (area, line or bar). However, the other three dependent variables (gaze time, fixation count, and number of transitions) that are more robust in the comparison of scores in this experiment reveal no differences between graph types. Additionally, the performance is not different for different graph designs (round/linear). This means that, overall, all graph types and both graph designs are equally convenient in conveying information for the tasks and events used in this experiment.

The analysis of task performance for cyclic concepts with respect to the two dependent variables gaze time and fixation count revealed that there is no difference between the two trend assessment tasks in the linear graph group. One of the tasks asks for distributed data from both edges of the graph, while the other task asks for adjacent data. Therefore there should be a difference between the two tasks. The lack of difference between the two tasks made me analyze the graphs in the cyclic events of the linear group individually for each participant. This analysis revealed a very interesting result (see Section 4-2-3). It showed that some of the linear graph readers just looked at one side of the graph although they had reported in the concept evaluation form that the relevant concept involved entities presented at both sides of the timeline. This result suggests that linear graphs representing cyclical events may either misguide the interpretation of the graph since the event that they represent is not coherent with the graph's features or may lead to a truncated interpretation that considers only partial evidence from one of the two sides of the graph.

For the second analysis subjects were re-categorized into three groups (one-sided linear, two-sided linear, round) according to their looking behavior and the self-report in the questionnaire. The second analysis revealed that cyclic events are comprehended in less fixation time (i.e., less reaction time), and with fewer fixations in the round graph than in the two-sided linear group, while task performance is the same for the round graph group and the one-sided linear graph group that might have been misguided in the evaluation of cyclic concepts, though. The analysis of the number or errors committed by the participants in the decision task also supported the idea that linear graph representing cyclical events may misguide the interpretation of the cyclic event, since the results indicated that one-sided linear group made significantly more error in the judgment of the cyclic concepts than the two-sided linear group and the round group (see Section 4-2-4).

There are three possible explanations of this result. First, the cyclicity of the event concept has been provided perceptually by the feature of sphericity in the round graph. The property of "being laconic", one of the criteria that describes the graphs' effectivity, means that objects that are close to each other in the graph, are also close to each other in the represented world. This essential feature is not provided in the linear graph in the case of cyclic events; however, in the round graphs. The isomorphism between the conceptual and the perceptual proximity may facilitate the comprehension of cyclic concepts and increase a direct semantic interpretation. Second, in addition to the isomorphism between content (the event type) and form (the graph design) on the conceptual level, the spatial proximity may also have a very important role in graph comprehension. In the foveal and even in the parafoveal field, the amount of data which contains relative information for the task is higher in the round graph than in the linear graph since the elements which construct an event concept in the round graph are presented neatly together, ready for extracting cyclic information. Third, in linear graphs, the representation of a cyclic event consists of at least three informative data points, as aforementioned, while the round graph is able to represent the cyclic information with just two informative points. Therefore, the comprehension can be also affected positively by the scarcity of data points needed for making a decision in the task.

The final evaluation of cyclic event (with previous analysis) and trend events (in general analysis) for both graph design groups shows that grasping trend information in cyclic events can be completed less effortfully in round graph. On the other hand, the trend events are comprehended in linear and round graph equally (see the Figure 4-17 and Figure 4-20 for the comparison of trend assessment tasks for trend event for both graph design group).

To summarize, the results of this study suggest that grasping trend information in cyclic events can be achieved with less effort in round graphs. This result is not trivial at all, given the fact that participants were not familiar with the round graph design and were confronted with them in this experiment for the first time. All eye tracking data results demonstrate that – despite informational equivalence between linear and round graphs – the latter are computationally superior to the former in the interpretation of cyclical concepts.

The results of another analysis conducted on the scores of the discrete comparison tasks support the literature. When the word order in the sentence is coherent with that in the graph, comprehension is completed with shorter gaze time and fewer fixations. The effect of word order is observed in the round graph although the direction is opposite of the linear timeline since the labeling around the circular timeline is clockwise (see Section 4-2-5).

5.4 Scan Path Analysis

In addition to the numerical data used in the general eye movement analysis using gaze time, fixation count, fixation duration, and transition between graph elements, the scan path analysis gave us the opportunity to observe the gaze sequence of the participants (see Section 4-3). The scan path analysis is able to show the interaction betweens graph design, event type and task type by means of a visual representation.

The scan path analysis clearly indicates the advantage of the round graph in the comprehension of cyclic concepts. Whereas entities are presented separately in linear graphs, they are represented adjacently in round graphs (Figure 4-37 and Figure 4-38). The reason why subjects in the linear group fixate longer and more often can be revealed by inspecting their gaze sequences. While the participants of the linear group look at the label and its respective field in the graph for the entities that appeared at both sides of the graph, it is enough to look at one area which comprises all relevant information in the round graph. This separation between the entities that constitute the concept in the linear graph also causes participants of this group to check the sentence more often as compared to participants of the round graph group. For the same reason participants of the linear graph group also display a higher number of transitions between graph elements. Friedman and Shah's Construction Integration theory explains this increase in terms of working memory capacity. Multiple iterations between construction and integration phases, which can be observed as an increase in the number of transitions between sentence and graph, are a sign of exceeding working memory capacity. In addition to the inappropriate representation of the event, task complexity also increases iteration between graph elements. When the task is getting complex, the number of transitions increases. This also explains the higher fixation count in discrete comparison tasks as compared to trend assessment tasks, and in cyclic events as compared to trend events.

In contrast to trend assessment tasks for cyclical concepts, trend events are processed in quite the same way in both graph designs in terms of gaze sequence and the huge difference observed in cyclic concept comprehension between graph designs does not appear. The scan path analysis also revealed similar patterns between graph types for discrete events in both cyclic and trend events represented in both graph designs. This also explains the similarity in task performance in general, and supports their being informationally and computationally equivalent, except in cyclic concept comprehension.

Furthermore, the result of the scan path analysis is consistent with the conclusion in Trickett's study (under review) on graph comprehension theories. All three major theories (Pinker' Propositional theory, Freedman and Shah' Construction Integration theory, and Lohse's UCIE theory) can account for all integration questions in this experiment in simple graphs although one of the graph designs is novel for the participants.

Although linear and round graphs have similar perceptual features in terms of length and area, they have different scales (linear vs. circular), which do or do not violate the alignment. Alignment, according to Clevelend (1985) is the most accurate perceptual feature in the hierarchy of perceptual features (Table 1-1). According to this hierarchy, while the linear scale can be considered as aligned, the circular scale should have a disadvantage of alignment between their entities, therefore the round graph should be hard to interpret. However, the results showed that there was no significant difference between linear and round graphs in terms of task performance. The results of the experiment in this thesis are also consistent with Hollands and Spence's conclusion (2001) that had revised Cleveland's hierarchy by indicating that the reader chooses a sample from the set of perceptual features rather than choosing the most accurate or most effective one. The advantage of round graph in cyclic events may be due to their being laconic. By this property, they may facilitate conceptual features which may overweight other features like length, and area, ranks higher in hierarchy, makes the comprehension of cyclic events easier.

5.5 Conclusion

I shall now summarize the results in terms of my hypotheses. My first hypothesis which stated that graph type (area/line/bar) affects comparison strategies was rejected. On the other hand, my second hypothesis that indicated that event type affects comparison strategies was accepted. Trend events were achieved with less effort compared to cyclic events. My third hypothesis which stated that the graph design affects comprehension was partially accepted, since the only difference between the two graph designs was observed in the comprehension of cyclic concepts. On the other hand, discrete comparison tasks in cyclic events and both of the tasks in trend events were equally interpreted. My fourth hypothesis which stated that graph design and event type interact was also partially accepted, only difference was observed in fixation count parameter, indicating while the comprehension of cyclic events was the same between the two graph designs, trend events were comprehended much more easily in linear graphs than in round graphs. The last hypothesis that stated that task type affects comparison strategies was also accepted. The results showed that discrimination tasks needed more effort to be comprehended compared to trend assessment tasks.

Although the literature which investigated the task dependency in graph comprehension about readers' preferences on choosing words in describing the relation presented in the graph shows that bar graphs are preferred to describe discrete information (maximum and minimum points) and line graphs to describe trends (increasing and decreasing), the effect of design in task performance in this study was only observed in the participant' preferences on graph type. However, consistent results were observed in preferences on keyword usage in the description task for different event types. Participants tend to describe cyclic events with discrete keywords and trend events with trend keywords in the linear graph design. In the round graph design, however, the difference between events is eliminated, by representing cyclic events with two salient points like trend events.

The results gathered from all data collection tools converged on the same conclusion which is that cyclic events have another dimension (cyclicity) as compared to trend events which renders its comprehension more difficult. Linear graphs, however, are inefficient representations for making this semantic information explicit and conveying this dimension. This semantic characteristic of cyclic events can effectively be accommodated by the sphericity of the round graph. Furthermore, the round graph also provides spatial proximity by displaying the relevant information within the same visual field, thereby decreasing the number of salient points to be processed. This helps the participant in the completion of the task without exceeding the capacity of short term /working memory.

However, all other tasks for both cyclic and trend events, except trend assessment task for cyclic events, are processed with approximately the same effort in both graph design groups. This result is highly remarkable when the fact is taken into account that participants were unfamiliar with the round type of graph.

Overall, these results indicate that participants find it hard to interpret a cyclic event in a linear graph but not a linear event in a round graph. This may mean that a less complex event (trend event) can readily be interpreted in a graph design that is more complex than necessary, that is, in a round graph. However, if there is a more complex event (cyclic event) it cannot be very well interpreted in a too simple graph (linear graph) that misses the crucial feature of the event, namely the cyclicity.

In conclusion, this study aimed at contributing to the existing literature insights on effects of event types and new representation alternatives to existing graph design types by pointing out the importance of compatibility between event features and graph features, and effects of the semantic relationship between these on graph comprehension. In a more cognitive perspective, the study addresses the issues of isomorphism between content and form and complexity (of event types and graph designs). In the present study the positive effect of isomorphism was most clearly observable in the case of cyclic events represented in round graphs. The results show that such isomorphism facilitates graph comprehension, even if subjects are unfamiliar with the round graph design prior to the experiment. However, trend events were not necessarily understood better or faster in linear as compared to round graph designs. Here, an asymmetric inclusion relation seems to hold: a complex graph design (round graph) may accommodate the interpretation of simple (trend) as well as complex (cyclic) event types similarly well; however, a simple graph design may only accommodate the interpretation of a simple (trend) event type but fail to accommodate a (too) complex (cyclic) event type. In the light of our positive results on the new round graph design, it might not seem to far-fetched to propose common graphical soft-ware to include this graph design into their inventory. As this study shows very clearly, a cyclical graph design is not only informationally equivalent to a linear graph design but also computationally more effective. Moreover, on the user-side, subjects readily accept this novel graph design and are easily familiarized with it.

5.6 Limitations and Future Studies

Part-1 of the experiment which was conducted in order to evaluate the recollected data might be repeated without time limitation on the observation length. In the current experiment, participants were allowed to observe the graphs as long as they wanted. A time limitation may force them to focus on the most relevant features of the graph and thus help identify the most salient feature of the graph that has a role on immediate comprehension. Furthermore, this would give us the opportunity to investigate whether this feature has a relationship with the event types that are presented in the graph.

Eye tracking data is very robust data; therefore the sample size of the part that is based on the eye tracking method is sufficient to make generalizations. On the other hand, the evaluation of the recollected data is based on the answer sheet filled by the participants. This kind of data is more subjective as compared to the eye tracking data; therefore, in order to make any generalization more valid, the experiment could be repeated by increasing the sample size of the recollection evaluation part.

Furthermore, each participant finished both parts (the recollection evaluation and the judgment task) of the experiment with all combinations of event, task, and graph type. In the end the experiment took about 45 min. However, the data of the recollection evaluation meant to investigate the interaction between event type (cyclic and trend) and graph type (area, bar and line) is insufficient to make a comparison. Since the results of the drawing task did not show any difference the test could be conducted with a verbal description task in written format with more diversity in terms of graph type but by omitting the drawing part.

The investigation of cyclic events in trend assessment tasks between graph design types informed us about the relation between event type and graph design type. The discrete comparison tasks used in this experiment for both cyclic and trend events requested a discrimination of two entities on the timeline. In order to investigate the conceptual effect on graph comprehension in a discrete comparison task, a further study could be carried out. This study might explore the interaction between event and graph design type by comparing task performance in entity-based and concept-based discrimination tasks. An example for an entity-based stimulus sentence might be: “In Lake Eymir, the amount of zooplankton at 7:00 pm is lower than at 3:00 am.”; an example for a concept-based stimulus sentence might be: “In Lake Eymir, the amount of zooplankton in the evening is lower than in the morning.”). This study might inform us about whether even type and task type interact and whether event type (cyclic) overweights effects of task, while cyclicity has an equal advantage in the discrimination of cyclic concepts in both graph design types, during comprehension of entity-based comparisons.

Last, the effect of familiarity with the novel round graph design type might be worth-while exploring. Increasing familiarity, as induced by repeated exposure to round graphs, might yield even stronger facilitatory effects of round graphs on the interpretation of cyclic events than already observed in the present experiment. The novelty of the round graph seems like an excellent basis for exploring learning effects in the domain of graph comprehension.

REFERENCES

Alaçam, Ö., Hohenberger, A. E. and Çağıltay, K. .(2009) .Comparison Strategies of Different Event Types in Different Graph Types. In *Proceedings of* 15th European Conference on Eye Movements, Southampton, UK, 2009. University of Southampton.

American Psychological Association. (1994). *Publication Manual of the American Psychological Association.* Washington, DC: Author. Clark, H. H. (1996). Using language, (pp. xi, 432). Cambridge England; New York: Cambridge University Press.

Bower, G. H. and Clapper, J. P. (1989). Experimental Methods in Cognitive Science. In Posner, M. I. (Ed.), Foundations of Cognitive Science (pp. 245-300). MIT Press 1993

Casner, S. M. (1991). A task- analytic approach to the automated design of graphic presentations. *ACM Transactions on Graphics*, 10, 111-151.

Clark, H. H., and Chase, W. G. (1972). On the process of comparing sentences against pictures. *Cognitive Psychology*, 3, 472–517.

Cleveland, W. S. (1985). The elements of graphing data. Monterey, CA:Wadcworth

Cleveland, W. S. (1990). A model for graphical perception. In proceedings of the Section on Statistical Graphcis. Alexandria, VA : American Statistical Association, pp 1-32

Feeney, A., Hola, A. K. W., Liversedge, S. P., Findlay, J. M., and Metcalf, R. (2000). How people extract information from graphs: Evidence from a sentence–graph verification paradigm. In M. Anderson, P. Cheng, and V. Haarslev (Eds.), Theory and application of diagrams: First international conference, Diagrams 2000 (pp.149–161). Berlin: Springer-Verlag.

Feeney, A. and Webber, L. J. (2003). Analogical representation and graph comprehension. *Lecture Notes in Computer Science*, 2733, 212-221.

Fischer, M. H. (2000). Do Irrelevant Depth Cues Affect the Comprehension of Bar Graphs? , *Applied Cognitive Psychology*, 14,151-162

Fischer, M. H., Dewulf, N., Hill, R. L. (2005). Designing Bar Graphs: Orientation Matters, *Applied Cognitive Psychology*,19: 953–962.

Fitts, P. M., Jones, R. E., & Milton, J. L. (1950). Eye movements of aircraft pilots during instrument- landing approaches. *Aeronautical Engineering Review* 9(2), 24–29.

Freedman, E. G., & Shah, P. (2002). Toward a Model of Knowledge-Based Graph Comprehension. Paper presented at Diagrammatic Representation and Inference, Second International Conference, Diagrams 2002, Callaway Gardens, GA, USA.

Gillan, D. J. and Lewis, R. (1994). A componential model of human interaction with graphs: 1. Linear regression modeling. Human Factors, 36, 419-440.

Goldberg, J. H., & Kotval, X. P. (1998). Eye movement-based evaluation of the computer interface. In: S. K. Kumar (ed.), *Advances in Occupational Ergonomics and Safety* (pp. 529–532). Amsterdam: I SO Press.

Gurr, C.A. (1999).Effective diagrammatic communication: Syntactic, semantic and pragmatic issues. *Journal of Visual Languages and Computing*, 10(4):317 -342.

Hegarty, M., Carpenter, P. A., Just, M. A. (1991). Diagrams in the comprehension of scientific texts, *Handbook of reading research*, 641-668.

Henderson, J. M. (2003). Human gaze control in real-world scene perception. *Trends in Cognitive Sciences*, 7, 498-504.

Hollands J. G. and Spence I. (1992). Judgments of change and proportion in graphical perception. Human Factors, 34,312-334.

Hollands J. G. and Spence I. (1998). Judging proportion with graphs: The summation Model*, Applied Cognitive Psychology, 12, 173-190.*

Hollands J. G. and Spence I. (2001), The discrimination of graphical elements, *Applied Cognitive Psychology,* 15, 413-431.

Jacob, R.J.K., and Karn, K.S., Eye tracking in human-computer interaction and usability research: Ready to deliver the promises (Section Commentary), in The Mind's Eye: Cognitive and Applied Aspects of Eye Movement Research, J. Hyona, Radach, R., and Deubel, H., Editor 2003, Elsevier Science. Amsterdam. p. 573-605.

Just, M.A., Carpenter, P.A., 1976. Eye fixations and cognitive processes. *Cognitive Psychology*. 8, 441–480.

Just, M. A., & Carpenter, P. A. (1980). A theory of reading: From eye fixations to comprehension. *Psychological Review*, 87(4), 329-354.

Kieras, D. , 2007, http://www.eecs.umich.edu/~kieras/goms.html

Kosslyn, S. M., (1989), Understanding Charts and Graphs, *Applied Cognitive Psychology*, 3, 185-226.

Kosslyn, S. (1993). *Elements of graph design.* NewYork: W.H. Freeman and Company.

Kotval, X. P., & Goldberg, J. H. (1998). Eye movements and interface components grouping: An evaluation method. In: Proceedings of the 42nd Annual Meeting of the Human Factors and Ergonomics Society (pp. 486–490). Santa Monica: Human Factors and Ergonomics Society.

Larkin, J. H., & Simon, H. A. (1987). Why a diagram is (sometimes) worth 10,000 words. *Cognitive Science*, 4, 317-345.

Levy, E., Zacks, J., Tversky, B., & Schiano, D. (1996, April 13-18, 1996). *Gratuitous graphics? Putting preferences in perspective.* Paper presented at the ACM conference on human factors in computing systems, Vancouver.

Lohse, G. L. (1993). A cognitive model for understanding graphical perception. *Human-computer Interaction,* 8, 353-388.

Macdonald-Ross, M. (1977). How numbers are shown: a review of research on the presentation of quantitative data in texts. *Audio-Visual Communication Review*, 25,359-407.

Martin, C. R., and Thompson, D. R. 2000. Design and Analysis of Clinical Nursing Research Studies, Routledge, London. (p. 69).

Meyer, J., Shinar, D., and Leiser, D. (1997). Multiple factors that determine performance with tables and graphs. *Human Factors*, 39/2, 268–286.

Palmer, S. E. (1978). Fundamental aspects of cognitive representation. In Rosch, E., and Lloyd, B. B. (eds.), *Cognition and Categorization*, Erlbaum, Hillsdale, NJ, pp. 259–303.

Petrusic W. M. (1992). Semantic congruity effects and theories of the comparison process. *Journal of Experimental Psychology: Human Perception and Performance.* 18 , 962-986.

Pinker, S. (1990). A theory of graph comprehension. In R. Freedle (Ed.), Artificial intelligence and the future of testing. (pp.73-126). Hillsdale, NJ: Erlbaum.

Pollatsek, A., Rayner, K., & Collins, W. E. (1984). Integrating pictorial information across eye movements. Journal of Experimental Psychology: General, 113, 426-442.

Ratwani, R.M. & Trafton, J.G. (2008). Shedding light on the graph schema: Perceptual features vs. invariant structure. *Psychonomic Bulletin & Review*. 15(4), 757-762.

Rayner, K. (1998). Eye movements in reading and information processing: 20 years of research. Psychological Bulletin, 124(3), 372-422.

Rayner, K., Sereno, S. C., Morris, R. K., Schmauder, A. R., & et al. (1989). Eye movements and on-line language comprehension processes. *Language and Cognitive Processes*. 4(3-4), 21-49

Renshaw, J., Finlay, J., Tyfa, D. and Ward, R.(2004). Understanding visual influence in graph design through temporal and spatial eye movement characteristics. Interacting with Computers, 16:557—578.

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

Schnotz, W. (2002). Commentary - Towards an integrated view of learning from text and visual displays. *Educational Psychology Review,* 14(1), 101-120.

Schmidt-Weigand F. (2006). Dynamic visualizations in multimedia learning: The influence of verbal explanations on visual attention, cognitive load and learning outcome.

Shah, P. and Carpenter, P. A. (1995). Conceptual limitations in comprehending line graphs. *Journal of experimental Psychology: General*, 124, 43-61.

Shah, P. (2002). Graph comprehension: The role of format, content, and individual difference. In M. Anderson, B. M. M. & P. Olivier (Eds.), Diagrammatic Representation and Reasoning. London, New York: Springer Verlag.

Shah, P., Freedman, E. G., & Vekiri, I. (in press). The comprehension of graphs. In P. Shah & A.Miyake (Eds.), Handbook of Visuospatial Cognition. New York: Cambridge University Press.

Shah, P., Freedman, E. G., & Vekiri, I. (2005). The comprehension of Quantitative Information in Graphical Displays. In P. Shah & A.Miyake (Eds.), The Cambridge Handbook of Visuospatial Thinking. New York: Cambridge University Press.

Shah, P., & Shellhammer, D. (1999). The role of domain knowledge and graph reading skills in graph comprehension. Paper presented at the Society of Applied Research in Memory and Cognition, Boulder, CO

Shah, P., & Hoeffner, J. (2002). Review of graph comprehension research: implications for instruction. *Educational Psychology Review*, 14/1, 47–69.

Simkin, D. and Hastie, R. (1987). An information processing analysis of graph perception. *Journal of the American Statistical Association,* 82, 454-465.
Spence I. (1990). Visual Psychophysics of Simple Graphical Elements, *Journal of Experimental Psychology: Human Perception and Performance.* 16(4):683-92.

Spence, I. and Lewandowsky, S. (1991). Displaying proportions and percentages. *Applied Cognitive Psychology.* 5 : 61-77.

Tan, J. K. H. and Benbasat, I. (1990). Processing of graphical information: decomposition taxonomy to match data extraction tasks and graphical representations. *Information systems Research*, 1, 416-439.

Tan, J. K. H. and Benbasat, I. (1993). The effectiveness of graphical presentation for information extraction: a cumulative experimental approach. *Decision Science*, 24, 167-191.

Trickett S. B., Ratwani, R. M. and Trafton, J. G, (under review). Real-world graph comprehension: High-level questions, complex graphs, and spatial cognition.

Underwood,G., Jebbett,L., and Roberts, K (2004). Inspecting pictures for information to verify a sentence: Eye movements in general encoding and in focused. The Quarterly Journal Of Experimental Psychology, *57A* (1), 165–182

Vickers D. (1980) Discrimination. In Reaction Times, Welford AT (Ed.). Academic Press: London; pp. 25-72.

Wickens, C.D., Carswell, C.M., 1995. The proximity compatibility principle: its psychological foundations and its relevance to display design. *Human Factors* 37, 473–494.

Wickens, C.D., Holland, J.G., 2000, 3rd ed., Engineering Psychology and Human Performance, Longman, London.

Zacks J., Tversky B.(1999) , Bars and Lines: A Study of Graphic Communication, Memory and Cognition (27), p. 1073-9.

APPENDICES

APPENDIX A: Demographic Information Form (In Turkish)

KATILIMCI VERİLERİ

Katılımcı No:_________

Yaş: ______

Cinsiyet: Kadın: ___ Erkek ___

Hangi elinizle yazıyorsunuz: Sağ: ___ Sol: ___

Anadiliniz: __________

Eğitim: Lisans öğrencisi: __ Lisansüstü ya da Doktora: __ Diğer (belirtiniz): ___

Öğrenim görmekte olduğunuz ya da mezun olduğunuz bölüm: ____________

Öğreniminiz süresince İstatistik Bilgisi içeren ders aldınız mı? ____________

APPENDIX B: Gönüllü Katılım Formu (In Turkish)

Bu çalışma, ODTÜ Bilişsel Bilimler Bölümünde, Bilişsel Bilimler Anabilim dalında Öğretim Görevlisi Annette Hohenberger ve Bilgisayar ve Öğretim Teknolojileri Bölümü Öğretim Görevlisi Kürşat Çağıltay danışmanlığında Yüksek Lisans Öğrencisi Özge Alaçam tarafından yüksek lisans tezi kapsamında grafik algılamasında karşılaştırma stratejileri'ni incelemek amacıyla yürütülmektedir.

Çalışmanın amacı, grafiklerde kullanılan tüm öğelerin grafiğin kavranmasında role sahip olduğu ve farklı grafik tiplerinin farklı karşılaştırma stratejileri kullanılarak algılandığını göstermektir. Grafik öğelerin biçimsel özelliklerinin yanı sıra, ifade edilmek istenen verinin anlamsal özelliklerinin de grafiğin algılanmasında önemli bir rolü olduğu gösterilmeye çalışılmaktadır.

Bu çalışma süresince kullanıcıların grafiklerle olan etkileşimi göz izleme cihazı tarafından kaydedilecektir. Uygulama öncesi kullanıcıların yaş/cinsiyet/bölüm/sınıf bilgilerini edinmemizi sağlayacak bir anket verilmektedir. Yapılacak çalışma 80 öğrenciye uygulanacak ve bütün çalışmalar İnsan Bilgisayar Etkileşim Araştırma ve Uygulama Laboratuarında gerçekleştirilecektir.

Bilgileriniz tamamıyla gizli tutulacak ve sadece araştırmacılar tarafından değerlendirilecektir, elde edilen bilgiler yüksek lisans tezi kapsamında ve bilimsel yayımlarda kullanılacaktır. Uygulama sırasında sorulardan ya da herhangi başka bir nedenden ötürü kendinizi rahatsız hissederseniz cevaplama işini yarıda bırakıp çıkmakta serbestsiniz. Böyle bir durumda uygulamayı yürüten kişiye, uygulamayı ya da soruları tamamlamadığınızı söylemek yeterli olacaktır. Uygulama sonunda, bu çalışmayla ilgili sorularınız cevaplanacaktır. Bu çalışmaya katıldığınız için şimdiden teşekkür ederiz.

Çalışma hakkında daha fazla bilgi almak için İnsan Bilgisayar Etkileşimi Araştırma ve Uygulama Laboratuvarı Sorumlusu Özge Alaçam ile (Oda: 118; Tel: 210 3357; E-posta: ozge@metu.edu.tr) iletişim kurabilirsiniz.

Bu çalışmaya tamamen gönüllü olarak katılıyorum ve istediğim zaman yarıda kesip çıkabileceğimi biliyorum. Verdiğim bilgilerin bilimsel amaçlı yayımlarda

kullanılmasını kabul ediyorum. (Formu doldurup imzaladıktan sonra uygulayıcıya geri veriniz).

İsim Soyad	Tarih	İmza
	----/----/-----	

APPENDIX C: Katılım Sonrası Bilgi Formu (In Turkish)

Bu çalışma, ODTÜ Bilişsel Bilimler Bölümünde, Bilişsel Bilimler Anabilim dalında Öğretim Görevlisi Annette Hohenberger ve Bilgisayar ve Öğretim Teknolojileri Bölümü Öğretim Görevlisi Kürşat Çağıltay danışmanlığında Yüksek Lisans Öğrencisi Özge Alaçam tarafından yüksek lisans tezi kapsamında grafik algılamasında karşılaştırma stratejileri'ni incelemek amacıyla yürütülmektedir.

Çalışmanın amacı, grafiklerde kullanılan tüm öğelerin grafiğin kavranmasında role sahip olduğu ve farklı grafik tiplerinin farklı karşılaştırma stratejileri kullanılarak algılandığını göstermektir. Grafik öğelerin biçimsel özelliklerinin yanı sıra, ifade edilmek istenen verinin anlamsal özelliklerinin de grafiğin algılanmasında önemli bir rolü olduğu gösterilmeye çalışılmaktadır.

Bu çalışma süresince kullanıcıların grafiklerle olan etkileşimi göz izleme cihazı tarafından kaydedilecektir. Uygulama öncesi kullanıcıların yaş/cinsiyet/bölüm/sınıf bilgilerini girmeleri istenecek bir anket verilmektedir. Yapılacak çalışma 80 öğrenciye uygulanacak ve bütün çalışmalar İnsan Bilgisayar Etkileşim Araştırma ve Uygulama Laboratuarında gerçekleştirilecektir.

Bu çalışma kapsamında göz izleme hareketlerini kullanarak kişilerin grafik algılama ve karşılaştırma stratejilerine yönelik bilgi edinilmesi amaçlanmaktadır. Kişilerin kavrama süreci içerisinde hangi noktalara daha fazla odaklandıkları, karşılaştırma yaparken nasıl bir yol izlediklerini, grafik kavrama sürecinde zorlandıkları noktaların tespit edilmesi ve grafiklerde karşılaştırma sürecinin ayrıntılı olarak ortaya konulması amaçlanmaktadır. Bu çalışmadan alınacak ilk verilerin Haziran 2009 sonunda elde edilmesi amaçlanmaktadır. Bilgileriniz tamamıyla gizli tutulacak ve sadece araştırmacılar tarafından değerlendirilecektir, elde edilen bilgiler yüksek lisans tezi kapsamında ve bilimsel yayınlarda kullanılacaktır.

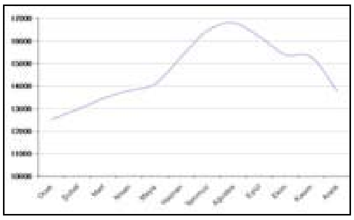
Çalışma hakkında daha fazla bilgi almak için İnsan Bilgisayar Etkileşimi Araştırma ve Uygulama Laboratuvarı Sorumlusu Özge Alaçam (Oda: 118; Tel: 210 3357; E-posta: ozge@metu.edu.tr) ile iletişim kurabilirsiniz.

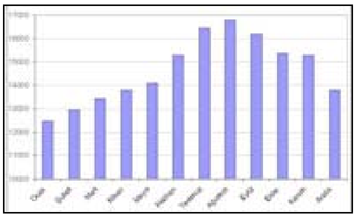
Bu çalışmaya katıldığınız için teşekkür ederiz.

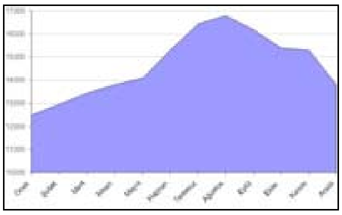
APPENDIX D: Linear Graph Evaluation Form for Part 1 Experiments (In Turkish)

PART-1 Lineer Grafik Değerlendirme Formu **Katılımcı No:**

Bu çalışmada lineer grafik üzerinde üç farklı grafik türü kullanılmıştır. Bunlar çizgi, sütün ve alan grafikleridir.


Çizgi Grafik


Sütun Grafik


Alan Grafik

Lütfen size göre en uygun olan cevabı X ile işaretleyiniz.

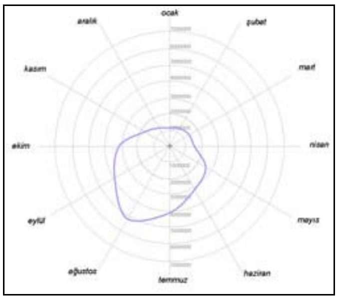
	Çizgi	**Sütun**	**Alan**
Hangi grafik türünde değerleri hatırlamak daha kolaydı?			
Hangi grafik türünde biçimi (pattern) hatırlamak kolaydı?			
Hangi grafik türünde, iki değişken arasındaki ilişki daha kolay anlaşılırdı?			
	Hiç	**Orta**	**Çok**
Bu grafik türüne (lineer grafik) ne kadar aşikarsınız?			
	Hayır	**Orta**	**Evet**
Eğer alışkın değilseniz, grafik türüne kolaylıkla alıştınız mı?			

Lütfen bu deneyde kullanılan grafikler ve bilgilerin/ilişkinin ifade edilişi ile ilgili genel düşüncülerinizi belirtiniz.

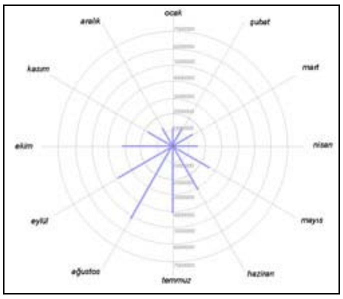
APPENDIX E: Round Graph Evaluation Form for Part 1 Experiments (In Turkish)

PART-1 Dairesel Grafik Değerlendirme Formu **Katılımcı No:**

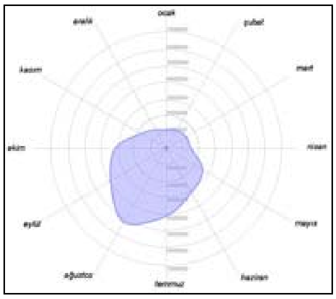
Bu çalışmada dairesel grafik üzerinde üç farklı grafik türü kullanılmıştır. Bunlar çizgi, sütün ve alan grafikleridir.



Çizgi Grafik



Sütun Grafik



Alan Grafik

Lütfen size göre en uygun olan cevabı X ile işaretleyiniz.

	Çizgi	**Sütun**	**Alan**
Hangi grafik türünde değerleri hatırlamak daha kolaydı?			
Hangi grafik türünde biçimi (pattern) hatırlamak kolaydı?			
Hangi grafik türünde, iki değişken arasındaki ilişki daha kolay anlaşılırdı?			
	Hiç	**Orta**	**Çok**
Bu grafik türüne (dairesel grafik) ne kadar aşikarsınız?			
	Hayır	**Orta**	**Evet**
Eğer alışkın değilseniz, grafik türüne kolaylıkla alıştınız mı?			

Lütfen bu deneyde kullanılan grafikler ve bilgilerin/ilişkinin ifade edilişi ile ilgili genel düşüncülerinizi belirtiniz.

____

____

APPENDIX F: Cyclic Concept Evaluation Form for Part 2 Experiments (In Turkish)

PART 2 –Kavram Değerlendirmesi **Katılımcı No:**

Bu formda, aylar ve gün saatleri ile ilgili iki ölçek verilmiştir. Belirli bir zaman dilimini ifade etmek için kullanılan bu terimler ile ilgili olarak, sizin için bu terimi en iyi karşılayan zaman dilimlerini ölçekten seçerek ilgili terimin karşısına yazınız.

Günler için;

01:00	03:00	05:00	07:00	09:00	11:00
13:00	15:00	17:00	19:00	21:00	23:00

Aylar için;

Ocak	Şubat	Mart	Nisan	Mayıs	Haziran
Temmuz	Ağustos	Eylül	Ekim	Kasım	Aralık

Gece : ______________________________

Öğle : ______________________________

Akşam : ______________________________

Yıl Sonu : ______________________________

Kış Mevsimi : ______________________________

Yaz Mevsimi : ______________________________

APPENDIX G: Descriptive Statistics of the Drawing Task

Dependent Variable: Observation Length (msec)

Event	Graph	Graph design	Mean	Std. Deviation	N
Cyclic	Area	Linear	55071	26723	19
		Round	48274	24557	20
		Total	51585	25527	39
	Line	Linear	58102	25271	19
		Round	54701	29243	20
		Total	56358	27075	39
	Bar	Linear	61067	35640	19
		Round	59012	31832	20
		Total	60013	33308	39
Trend	Area	Linear	59840	24298	19
		Round	53348	28425	20
		Total	56510	26353	39
	Line	Linear	57525	27705	19
		Round	58672	27026	20
		Total	58113	27003	39
	Bar	Linear	55358	30782	19
		Round	53968	19917	20
		Total	54645	25449	39

Appendix H: Descriptive Statistics for the Written Description Task

Independent Variables: Event Type and graph design

Means	Event	Graph design	Mean	Std. Deviation	N
Number of Recollected value	Cyclic	Linear	3,5	2,200478	20
		Round	3,8	3,096688	20
	Trend	Linear	2,775	1,888156	20
		Round	3,8	3,096688	20
Number of Discrete Keyword	Cyclic	Linear	0,7	0,470162	20
		Round	0,5	0,512989	20
	Trend	Linear	0,3	0,470162	20
		Round	0,2	0,410391	20
Number of Trend Keyword	Cyclic	Linear	0,7	0,470162	20
		Round	0,85	0,366348	20
	Trend	Linear	0,95	0,223607	20
		Round	0,85	0,366348	20
Number of Conceptual Keyword	Cyclic	Linear	0,3	0,470162	20
		Round	0,4	0,502625	20
	Trend	Linear	0,2	0,410391	20
		Round	0,1	0,307794	20
Number of Keyword Categories used	Cyclic	Linear	1,7	0,978721	20
		Round	1,75	0,71635	20
	Trend	Linear	1,45	0,686333	20
		Round	1,15	0,74516	20

Independent Variables: Graph Type and graph design

Means	Graph	Graph design	Mean	Std. Deviation	N
Number of Recollected value	Area	Linear	3,368	1,422	19
		Round	3,737	3,364	19
	Line	Linear	2,947	1,810	19
		Round	3,632	3,041	19
	Bar	Linear	3,368	2,608	19
		Round	3,474	3,133	19
Number of Discrete Keyword	Area	Linear	0,400	0,503	19
		Round	0,316	0,478	19
	Line	Linear	0,400	0,503	19
		Round	0,211	0,419	19
	Bar	Linear	0,600	0,503	19
		Round	0,421	0,507	19
Number of Trend Keyword	Area	Linear	0,900	0,308	19
		Round	0,895	0,315	19
	Line	Linear	0,900	0,308	19
		Round	0,842	0,375	19
	Bar	Linear	0,800	0,410	19
		Round	0,684	0,478	19

Number of Conceptual Keyword	Area	Linear	0,100	0,308	19
		Round	0,211	0,419	19
	Line	Linear	0,100	0,308	19
		Round	0,263	0,452	19
	Bar	Linear	0,200	0,410	19
		Round	0,263	0,452	19
Number of Keyword Categories used	Area	Linear	1,400	0,754	20
		Round	1,421	0,769	20
	Line	Linear	1,400	0,681	20
		Round	1,316	0,820	20
	Bar	Linear	1,600	0,940	20
		Round	1,368	0,895	20

Appendix I: A table indicating the F- Statistics and Effect Size values for main and interaction effects for each independent and dependent variables in the General Analysis (section 4.2.1).

(Graph=short for graph type; event= short for event type)

	Gaze time	***Fixation Count***	***Fixation Duration***	***Number of Transition***
Graph design	F (1, 38) =1.448, η_p^2=.04, *ns.*	F (1, 38) = 2.708, η_p^2=.07, *ns.*	F (1, 38) = .392, η_p^2=.01, *ns.*	F (1, 38) = .021, η_p^2=.001, *ns.*
Event (large effect size)	F (1, 38) =39.058, η_p^2=.51, *s.*	F (1, 38) = 23.962, η_p^2=.39 *s.*	F (1, 38) = 2.185, *ns.*	F (1, 38) = 9.982, η_p^2=.21, *s.*
Task (large effect size)	F (1, 38) =260.368, η_p^2=.87, s	F (1, 38) = 264.512, η_p^2=.87, *s.*	F (1, 38) = 8.359, η_p^2=.18, *s.*	F (1, 38) =25 .612, η_p^2=.40, *s.*
Graph Type	F (2, 76) =3.158, η_p^2=.07, *s.*	F (2, 76) = 1.399, *ns.*	F (2, 76) = 2.696, η_p^2=.06, ns.	F (2, 76) = 1.927, *ns.*
Graph design*Event	F (1, 38) =2.720, ns.	F (1, 38) = 4.700, η_p^2=.11, *s.*	F (1, 38) = .905, *ns.*	F (1, 38) = .951, *ns.*
Graph design*Task	F (1, 38) =4.050, η_p^2=.10, ns.	F (1, 38) = 6.816, η_p^2=.15, *s.*	F (1, 38) = .797, ns.	F (1, 38) = .207, *ns.*
Event*Task	F (1, 38) =3.402, ns.	F (1, 38) = 8.105, η_p^2=.18, *s.*	F (1, 38) = .212, *ns.*	F (1, 38) = 8.068, η_p^2=.17, *s.*

Graph design*Event*Task	F (1, 38) = .084, *ns.*	F (1, 38) = .008, *ns.*	F (1, 38) = 5.189, η_p^2=.12, *s*	F (1, 38) = .975, *ns.*
Graph design*Graph type	F (2, 76) =2.946, n*s.*	F (2, 76) = 1.821, *ns.*	F (2, 76) = .829, *ns.*	F (2, 76) = .562, *ns.*
Event*Graph type	F (2, 76) = .284, *ns.*	F (2, 76) = .626, *ns.*	F (2, 76) = .849, *ns.*	F (2, 76) = 1.403, *ns.*
Graph design*Event*Graph type	F (2, 76) = 1.532, *ns.*	F (2, 76) = 1.152, *ns.*	F (2, 76) = 1.178, *ns.*	F (2, 76) = .447, *ns.*
Task*Graph type	F (2, 76) = 2.044, *ns.*	F (2, 76) = 1.889, *ns.*	F (2, 76) = 3.771, η_p^2=.09, *s.*	F (2, 76) = 1.066, *ns.*
Graph design*Task*Graph type	F (2, 76) = .171, *ns.*	F (2, 76) = .344, *ns.*	F (2, 76) = .510, *ns.*	F (2, 76) = .115, *ns.*
Event*Task*Graph type	F (2, 76) = 2.487, n*s.*	F (2, 76) = 2.653, *ns.*	F (2, 76) = 3.013, *ns.*	F (2, 76) = .333, *ns.*
Graph design*Event*Task*Graph type	F (2, 76) = 3.115, *ns.*	F (2, 76) = 2.260, *ns.*	F (2, 76) = 1.168, *ns.*	F (2, 76) = .468, *ns.*

Appendix J: Descriptive Statistics For Gaze time in the General Analysis

Independent Variables: **Event type, task, graph type, graph design**

Event	Task	Graph Type	Graph design	Mean	Std. Deviation	N
Cyclic	TAT	Area	Linear	7.4262	2.758212	20
			Round	7.401647	2.298223	20
			Total	7.413924	2.50593	40
		Line	Linear	7.445494	1.624324	20
			Round	7.110114	1.997267	20
			Total	7.277804	1.804889	40
		Bar	Linear	7.451626	2.123504	20
			Round	6.972918	1.529033	20
			Total	7.212272	1.842439	40
	DCT	Area	Linear	9.079072	1.401308	20
			Round	10.90543	2.492887	20
			Total	9.99225	2.199887	40
		Line	Linear	9.971624	1.972918	20
			Round	10.20475	2.883902	20
			Total	10.08819	2.441732	40
		Bar	Linear	9.79983	2.093556	20
			Round	9.259166	1.888525	20
			Total	9.529498	1.986904	40
Trend	TAT	Area	Linear	5.589479	1.761395	20
			Round	6.644435	1.920989	20
			Total	6.116957	1.895953	40
		Line	Linear	5.450117	1.778166	20
			Round	5.740535	1.397207	20
			Total	5.595326	1.585274	40
		Bar	Linear	5.141289	2.002283	20
			Round	5.392544	1.564823	20
			Total	5.266916	1.778285	40
	DCT	Area	Linear	7.920365	1.713749	20
			Round	8.892569	1.847414	20
			Total	8.406467	1.82644	40
		Line	Linear	8.627514	2.134101	20
			Round	9.936207	2.59465	20
			Total	9.28186	2.436749	40
		Bar	Linear	7.945106	2.507947	20
			Round	9.132517	2.087234	20
			Total	8.538812	2.355463	40

Appendix K: Descriptive Statistics For Fixation Count in the General Analysis

Independent variables: **event type, task, graph type, graph design**

Event	Task	Graph Type	Graph design	Mean	Std. Deviation	N
Cyclic	TAT	Area	Linear	27.275	10.08513	20
			Round	26.6875	8.7899	20
			Total	26.98125	9.342381	40
		Line	Linear	26.475	6.119909	20
			Round	25.65	7.381699	20
			Total	26.0625	6.705755	40
		Bar	Linear	26.5375	8.156881	20
			Round	25.1375	4.345653	20
			Total	25.8375	6.489773	40
	DCT	Area	Linear	33.0625	4.733611	20
			Round	39.675	8.478727	20
			Total	36.36875	7.559802	40
		Line	Linear	34.7625	7.828119	20
			Round	36.5625	8.123178	20
			Total	35.6625	7.926657	40
		Bar	Linear	36	8.390941	20
			Round	34.9875	6.985641	20
			Total	35.49375	7.637938	40
Trend	TAT	Area	Linear	21.2625	7.010084	20
			Round	24.6	6.150738	20
			Total	22.93125	6.725141	40
		Line	Linear	19.7	6.501822	20
			Round	21.9	5.022581	20
			Total	20.8	5.841716	40
		Bar	Linear	19.0125	7.925008	20
			Round	20.5375	5.468375	20
			Total	19.775	6.764775	40
	DCT	Area	Linear	29.2	6.266347	20
			Round	34.9875	7.149933	20
			Total	32.09375	7.25424	40
		Line	Linear	31.9875	8.489729	20
			Round	37.6625	8.988763	20
			Total	34.825	9.095857	40
		Bar	Linear	29.9625	9.428496	20
			Round	36.275	7.813997	20
			Total	33.11875	9.125371	40

Appendix L: Descriptive Statistics For Fixation Duration in the General Analysis

Independent variables: **Event type, task, graph type, graph design**

Event	Task	Graph Type	Graph design	Mean	Std. Deviation	N
Cyclic	TAT	Area	Linear	0.303127	0.059169	20
			Round	0.302483	0.049624	20
			Total	0.302805	0.053902	40
		Line	Linear	0.309577	0.055373	20
			Round	0.29799	0.035226	20
			Total	0.303784	0.046181	40
		Bar	Linear	0.299527	0.036577	20
			Round	0.305902	0.050841	20
			Total	0.302715	0.043834	40
	DCT	Area	Linear	0.301644	0.046344	20
			Round	0.285043	0.043879	20
			Total	0.293344	0.045333	40
		Line	Linear	0.315852	0.046995	20
			Round	0.286131	0.044026	20
			Total	0.300992	0.0474	40
		Bar	Linear	0.297451	0.036057	20
			Round	0.273473	0.039828	20
			Total	0.285462	0.039416	40
Trend	TAT	Area	Linear	0.306613	0.061457	20
			Round	0.313159	0.111569	20
			Total	0.309886	0.088968	40
		Line	Linear	0.280075	0.060971	20
			Round	0.292665	0.040346	20
			Total	0.28637	0.051427	40
		Bar	Linear	0.289886	0.045969	20
			Round	0.270343	0.050558	20
			Total	0.280114	0.04871	40
	DCT	Area	Linear	0.263544	0.028206	20
			Round	0.284126	0.048878	20
			Total	0.273835	0.040744	40
		Line	Linear	0.290843	0.039577	20
			Round	0.288164	0.034971	20
			Total	0.289503	0.036888	40
		Bar	Linear	0.271948	0.037616	20
			Round	0.275607	0.044293	20
			Total	0.273777	0.040602	40

Appendix M: Descriptive Statistics for Number of Transition in the General Analysis

Independent variables: Event type, task, graph type, graph design

Event	Task	Graph Type	Graph design	Mean	Std. Deviation	N
Cyclic	TAT	Area	Linear	5.725	2.650596	20
			Round	5.675	1.495828	20
			Total	5.7	2.124491	40
		Line	Linear	5.1875	2.004723	20
			Round	5.5	1.200329	20
			Total	5.34375	1.638565	40
		Bar	Linear	5.35	2.26152	20
			Round	5.4	1.326352	20
			Total	5.375	1.830125	40
	DCT	Area	Linear	6.0875	1.897497	20
			Round	5.9875	1.846895	20
			Total	6.0375	1.8489	40
		Line	Linear	6.1875	2.343629	20
			Round	5.6	2.028417	20
			Total	5.89375	2.183774	40
		Bar	Linear	6.075	2.090234	20
			Round	5.8	1.174286	20
			Total	5.9375	1.6792	40
Trend	TAT	Area	Linear	4.625	1.923709	20
			Round	4.975	1.551527	20
			Total	4.8	1.734085	40
		Line	Linear	4.4625	1.888809	20
			Round	4.725	1.019223	20
			Total	4.59375	1.503934	40
		Bar	Linear	4.3125	2.153447	20
			Round	4.2	1.204706	20
			Total	4.25625	1.723227	40
	DCT	Area	Linear	5.4625	2.101495	20
			Round	5.8375	1.136346	20
			Total	5.65	1.678293	40
		Line	Linear	5.85	2.347171	20
			Round	6.3625	1.812738	20
			Total	6.10625	2.086194	40
		Bar	Linear	5.6875	2.070906	20
			Round	5.6375	1.625202	20
			Total	5.6625	1.837597	40

Appendix N: Descriptive statistics for the Analysis of Cyclic Concept Comprehension

Dependent Variable: Gaze time

Independent variables: Event type, task, graph type, graph design

Task Type	Graph Type	Graph design	Mean	Std. Deviation	N
TAT 1	Area	Linear	7.11483	2.597657	15
		Round	7.482478	2.560175	16
		Total	7.304584	2.541872	31
	Bar	Linear	7.67525	2.951349	15
		Round	7.333299	2.302444	16
		Total	7.498759	2.597245	31
	Line	Linear	7.210483	2.082193	15
		Round	6.373521	2.303532	16
		Total	6.778502	2.203895	31
TAT 2	Area	Linear	7.524861	3.547512	15
		Round	6.718748	2.810418	16
		Total	7.108803	3.160672	31
	Bar	Linear	7.345779	2.417953	15
		Round	6.761805	2.321707	16
		Total	7.044373	2.347666	31
	Line	Linear	8.431621	2.136711	15
		Round	6.850748	2.346651	16
		Total	7.615686	2.351364	31

Dependent Variable: Fixation Count

Task Type	Graph Type	Graph design	Mean	Std. Deviation	N
TAT 1	Area	Linear	26.33333	9.764489	15
		Round	26.3125	6.923089	16
		Total	26.32258	8.274004	31
	Bar	Linear	26.33333	7.64308	15
		Round	22.375	7.421815	16
		Total	24.29032	7.67113	31
	Line	Linear	27.16667	9.538768	15
		Round	26	7.348469	16
		Total	26.56452	8.355379	31
TAT 2	Area	Linear	27.03333	12.58892	15
		Round	24.625	12.14976	16
		Total	25.79032	12.21732	31
	Bar	Linear	28.93333	7.367561	15
		Round	25.5	9.916317	16
		Total	27.16129	8.805668	31
	Line	Linear	26.46667	8.118468	15
		Round	24.34375	7.478232	16
		Total	25.37097	7.738397	31

Appendix O: Descriptive statistics for the Analysis of Cyclic Concept Comprehension -2 (for the three graph types and the three graph design)

Dependent Variable: Gaze time

Independent variables: graph type, graph group

Graph Type	Graph group	Mean	Std. Deviation	N
Area	One - sided Linear	5.828908	1.201225	9
	Two - sided Linear	9.594068	3.26097	8
	Round	7.517202	2.315813	20
	Total	7.453827	2.567073	37
Line	One - sided Linear	6.517356	1.564241	9
	Two - sided Linear	7.926332	1.771496	8
	Round	6.915431	2.466165	20
	Total	6.999095	2.140141	37
Bar	One - sided Linear	5.6486	1.27891	9
	Two - sided Linear	9.97408	2.884019	8
	Round	7.278084	2.130336	20
	Total	7.318142	2.538122	37

Dependent Variable: Fixation Count

Graph Type	Graph design	Mean	Std. Deviation	N
Area	One - sided Linear	20.35	5.462244	9
	Two - sided Linear	37.64286	13.36885	8
	Round	26.825	6.759545	20
	Total	27.12162	9.781222	37
Line	One - sided Linear	24.55	5.875798	9
	Two - sided Linear	28.14286	7.033931	8
	Round	24.4	8.113666	20
	Total	25.14865	7.333205	37
Bar	One - sided Linear	20.05	6.173645	9
	Two - sided Linear	36.14286	11.82763	8
	Round	26.1	6.616566	20
	Total	26.25676	9.332308	37

Appendix P: Descriptive statistics for the analysis of Word Order in Discrete Comparison Tasks

Dependent Variable: Gaze time

Independent variables: Event type, task, graph type, graph design

Event	Task	Graph Type	Graph design	Mean	Std. Deviation	N
Cyclic	DCT -1	Area	Linear	9,014	1,407	17
			Round	11,343	3,467	12
			Total	9,978	2,686	29
		Line	Linear	9,090	2,337	17
			Round	9,645	3,106	12
			Total	9,320	2,643	29
		Bar	Linear	9,300	3,049	17
			Round	9,286	2,100	12
			Total	9,294	2,654	29
	DCT -2	Area	Linear	9,037	2,144	17
			Round	9,854	3,050	12
			Total	9,375	2,539	29
		Line	Linear	10,687	2,825	17
			Round	9,490	3,971	12
			Total	10,192	3,334	29
		Bar	Linear	10,477	2,361	17
			Round	9,619	3,075	12
			Total	10,122	2,661	29
Trend	DCT -1	Area	Linear	7,392	1,574	17
			Round	8,822	2,155	12
			Total	7,984	1,937	29
		Line	Linear	8,219	1,404	17
			Round	10,038	3,761	12
			Total	8,972	2,741	29
		Bar	Linear	7,897	3,170	17
			Round	8,824	2,490	12
			Total	8,281	2,897	29
	DCT -2	Area	Linear	8,771	2,511	17
			Round	8,499	2,645	12
			Total	8,658	2,524	29
		Line	Linear	9,397	3,744	17
			Round	9,236	2,899	12
			Total	9,330	3,364	29
		Bar	Linear	7,937	2,924	17
			Round	8,778	3,080	12
			Total	8,285	2,965	29

Dependent Variable: Fixation count

Event	Task	Graph Type	Graph design	Mean	Std. Deviation	N
Cyclic	DCT -1	Area	Linear	34,0	5,8	17
			Round	40,8	10,1	12
			Total	36,8	8,4	29
		Line	Linear	31,2	7,3	17
			Round	36,4	9,5	12
			Total	33,3	8,5	29
		Bar	Linear	33,7	13,5	17
			Round	35,8	6,7	12
			Total	34,6	11,1	29
	DCT -2	Area	Linear	32,4	7,7	17
			Round	38,3	11,1	12
			Total	34,8	9,5	29
		Line	Linear	38,3	11,1	17
			Round	34,9	10,3	12
			Total	36,9	10,7	29
		Bar	Linear	39,1	7,1	17
			Round	37,6	11,4	12
			Total	38,5	9,0	29
Trend	DCT -1	Area	Linear	27,8	6,1	17
			Round	36,8	10,0	12
			Total	31,5	9,0	29
		Line	Linear	30,7	5,6	17
			Round	39,0	13,7	12
			Total	34,1	10,5	29
		Bar	Linear	29,8	12,3	17
			Round	35,6	9,7	12
			Total	32,2	11,5	29
	DCT -2	Area	Linear	32,0	8,2	17
			Round	33,2	10,3	12
			Total	32,5	9,0	29
		Line	Linear	34,5	14,3	17
			Round	35,9	10,3	12
			Total	35,1	12,6	29
		Bar	Linear	29,7	10,0	17
			Round	35,8	12,7	12
			Total	32,2	11,4	29

Appendix R: Variable Trees for the experiment:

BSV*: Between subject variable / WSV* : Within subject variable

Variable Tree for Part-1 (*The Recollection of Data*)

Graph Design Type (2) / (BSV*) :	Round Graph Linear Graph
Event Type (2) / (WSV*) :	Cyclic Trend
Graph Type (3) / (WSV*) :	Area Line Bar
Task Type (2) / (WSV*) :	Draw Describe

Variable Tree for Part-2 (*Judgment Task Performance)*

Graph Design Type (2) / (BSV*) :	Round Graph Linear Graph
Event Type (2) / (WSV*) :	Cyclic Trend
Graph Type (3) / (WSV*) :	Area Line Bar
Task Type (2) / (WSV*) :	Trend Assessment (TAT) Discrete Comparison (DCT)
Sentence Type :	TAT-1 and TAT-2 DCT-1 and DCT-2
TASK:	Respond: Yes or No

CURRICULUM VITAE

PERSONAL INFORMATION

Surname, Name: Alaçam, Özge
Nationality: Turkish (TC)
Date and Place of Birth: 19 July 1982, Ankara
Phone: +90 312 210 33 57
Fax: +90 312 210 33 03
E-mail: ozge@metu.edu.tr

EDUCATION

Degree	**Institution**	**Year of Graduation**
MS	METU Cognitive Science	
BS	METU Computer Education & Instructional Technology	2005

WORK EXPERIENCE

Year	**Place**	**Enrollment**
2006- Present	METU Computer Center	Research Assistant

PUBLICATIONS

- Alaçam, Ö., Hohenberger, A. E. and Çağıltay, K. .(2009) .Comparison Strategies of Different Event Types in Different Graph Types. In *Proceedings of 15th European Conference on Eye Movements*, Southampton, UK, 2009. University of Southampton.
- Alaçam, Ö., and Dalcı, M.(2009). A Usability Study of Web Maps with Eye Tracking Tool: The Effects of Iconic Representation of Information (Accepted by HCI International 2009, San Diego, USA, Published in Lecture Notes In Computer Science; Vol. 5610, Proceedings of the 13th International Conference on Human-Computer Interaction. Part I: New Trends, San Diego, CA).
- Acarturk,C., Habel, C., Cagiltay, K., and Alacam, O. (2008). Multimodal comprehension of language and graphics: Graphs with and without annotations. *Journal of Eye Movement Research*, 1(3):2, 1-15.
- Alaçam Ö., and Çağıltay, K.,(2008). Bilişsel Süreçlerin takibinde Teknoloji Kullanımı- Göz Hareketlerini İzleme Cihazları ve Kullanım Alanları, Akademik Bilişim 2008, Çanakkale Onsekiz Mart Üniversitesi, Çanakkale 30 Ocak - 1 Şubat 2008.

- Dalcı M., Alaçam Ö, O. Saatçioğlu Y., ve Erdal, F., Odtü Kütüphanesi Yeni Web Sayfasının Tasarımı ve Kullanılabilirlik Çalışması, Akademik Bilişim 2008, Çanakkale Onsekiz Mart Üniversitesi, Çanakkale 30 Ocak - 1 Şubat 2008
- Ersoy A. Fatih, ve Alaçam Ö., ODTÜ BİDB Kullanıcı Eğitimleri Web Sayfası, Akademik Bilişim 2008,Çanakkale Onsekiz Mart Üniversitesi, Çanakkale 30 Ocak - 1 Şubat 2008
- Alaçam Ö., and Çağıltay, K., Odtü Bidb İnsan Bilgisayar Etkileşimi Laboratuvarı ve Yürütülen Çalışmalar, Akademik Bilişim 2007, Dumlupınar Üniversitesi, Kütahya 31 Ocak-2 Şubat 2007

Publications about Underwater Ecology

- Eyyüboğlu, A., Kırbulut, D. Polat, E. , Gençoğlu, G., Ergin, M., Filiz, N., Alaçam, Ö., Eren, S., Altay,S. , and Raşitoğlu, Z. (2009). Kaş, Hidayet Koyu'nda Uluburun III Batığı Yapay Resifi Balık Faunasının İncelenmesi. 2009 Mavi Paylaşım / 7. Sualtı Sporları ve Bilimleri Toplantısı, Kocaeli Üniversitesi.
- Kırbulut, D.,Tuncalı, E. , Koban, E., Gençoğlu, G. Fidan, G. ,Süder, İ., Ergin, M. , Mercan, M. , Koşar, O., Alaçam, Ö., Sayın, P., Altay, S. and Oruç, S.(2008). Observations on Fish Assemblage on Uluburun III Artificial Reef 2007. 2nd International Symposium on "Underwater Research", DoğuAkdeniz Üniversitesi, Kıbrıs.
- Alaçam Ö., Altay S., Eren S., Ergin M., Gençoğlu G., Kırbulut D., Koban E., (2007). Küçükkuyu-Assos-Bababurun Bölgesinde *Posidonia oceanica* (L) Delile Yayılımının İncelenmesi. Sualtı Bilim ve Teknoloji Toplantısı, Koç Üniversitesi, SBT 2007, İstanbul.

RESEARCH INTERESTS

Graph comprehension, Eye Tracking Methodology, Human-Computer Interaction, Visual Perception

HOBBIES

Scuba Diving, Underwater Photography, Underwater Ecology