COMPREHENSIBILITY OF TRAFFIC SIGNS IN TURKEY: THE EFFECTS OF COGNITIVE SIGN DESIGN FEATURES AND USER FACTORS

by

Gizem ZENGİN

B.S., Industrial Engineering, Kadir Has University, 2009

Submitted to the Institute for Graduate Studies in
Science and Engineering in partial fulfillment of
the requirements for the degree of
Master of Science

Graduate Program in Industrial Engineering
Boğaziçi University
2016

ACKNOWLEDGEMENTS

It is a pleasure to share my warm gratitude and thank you to all the people that made this thesis possible. First of all, I would like to thank to my thesis supervisor, Assoc. Prof. Dr. Mahmut Ekşioğlu for his endless encouragement, support, precious guidance, and inexhaustible patience and consciousness during my study. I would like to thank Assist. Prof. Dr. Mustafa Gökçe Baydoğan and Assist. Prof. Dr. Alev Çavdar Sideris, as members of my thesis committee, for their valuable suggestions.

I would like to thank all the subjects who participated in the study voluntarily and spared their time for all the steps of survey patiently.

I would like to thank to TÜBİTAK for its financial support during my graduate education.

Last but not least, I would like to express my gratitude to my family and friends for their support, encouragement and assistance during my thesis and also throughout my life.

ABSTRACT

COMPREHENSION OF TURKISH TRAFFIC SIGNS: THE EFFECT OF COGNITIVE SIGN DESIGN FEATURES AND USER CHARACTERISTICS

The correct and easy comprehension of traffic signs is essential for the effective traffic control and safe driving. It is known that comprehension varies widely among different traffic signs symbols, road user characteristics and also countries. Therefore, the main objective of this study was to examine the comprehensibility of Turkish traffic signs, investigate the effect of cognitive sign design features and user characteristics on comprehension, and understand the interrelationship between them. 100 male and 101 female volunteers aged between 18 and 69 years with family roots from all seven regions of Turkey were the participates of the study. 60 Turkish traffics signs were guessed and cognitive sign design features; familiarity, concreteness, simplicity, meaningfulness and semantics closeness were rated by participants for the each sign. The effects of age, gender, marital status, educational background, active car driving experience and smoking habits were evaluated in the study. Through statistical analyses descriptive values of guessability score as an indicator of comprehension and factor effects were determined. Current traffic signs were not guessed well with the 39.2% guessability score by the population. More than 50% of the participants could not guess the meaning of the 35% traffic signs. 5 signs were comprehended as opposite of their true meaning. Guessability score is highly and positively affected by familiarity, meaningfulness and semantic closeness. Age was found a significant predictor of guessability score where moderately effected by active car driving and marital status. Comprehension decreased with age and increased with driving experience. Guessability performance was not related with the educational background but performance increased when educational level increased. Road works sign was found more familiar, concrete, simple, meaningful and semantically close, followed by warning type. This was the first study that evaluates the effect of both the sign design features and user characteristics on the sign comprehension. The results can serve as a reference in design of traffic signs not only in Turkey but also in the world.

ÖZET

KAVRAMSAL İŞARET TASARIM FAKTÖRLERİ VE KULLANICI ÖZELLİKLERİNE GÖRE TÜRK TRAFİK İŞARETLERİNİN ANLAŞILIRLIĞININ DEĞERLENDİRİLMESİ

Etkili trafik kontrolünün ve yolda güvenliğin sağlanması için trafik işaretlerinin yol kullanıcılarını tarafından kolay ve anlaşılabilir olması gerekir. Anlaşılmanın farklı trafik işareti sembollerine, yol kullanıcılarının karakteristik özelliklerine ve ülkelere göre değiştiği bilinmektedir. Bu nedenle, bu çalışmanın amacı, mevcut Türk trafik işaretlerinin ne kadar anlaşılabilir olduğunu bulmak, kavramsal işaret tasarım faktörleri ve kullanıcı özelliklerine göre anlaşılabilirliğin nasıl etkilendiğini ve aralarındaki ilişkiyi istatistiksel olarak hesaplamaktı. Aile kökenleri Türkiye'nin yedi farklı bölgesinden olan 100 erkek ve 101 kadın katılımcı çalışmaya katıldı. 60 trafik işaretini katılımcılar tarafından tahmin edildi ve her bir işaret beş kavramsal işaret tasarım faktörü; aşinalık, somutluk, basitlik, anlamlılık, anlatılmak isteneni karşılama faktörlerine göre puanlandırıldı. Kullanıcı karakterleri olarak yaş, cinsiyet, medeni durum, eğitim durumu, aktif araç kullanımı ve sigara kullanımı araştırıldı. İstatistiksel analiz yapılarak işaretin anlamını tahmin skoru ve faktörlerin etkilerini tespit edildi. Sonuçlar mevcut trafik işaretlerinin %39.2'lik skor ile popülasyon tarafından yeterince iyi anlaşılamadığını ortaya koydu. Katılımcıların %50'den fazlası, trafik işaretlerinden %35'ini tahmin bile edemedi. Ayrıca, trafik işaretlerinden beş tanesinin verilmek istenen anlama zıt anlamda anlaşıldığı tespit edildi. Aşinalık, anlamlılık ve anlatılmak istenini karşılama skoru yüksek işaretlerin daha iyi anlaşıldığı görüldü. Tahmin etme performansı yaş artıkça düşmekte, aktif sürüç deneyimi artıkça artmaktadır. Eğitim durumunun tahmin etme performansı ile ilişkili olmadığı fakat eğitim seviyesi artıkça yükseldiği görülmüştür. En aşina olunan, somut, basit, anlamlı ve verilmek istenen anlamla uygun grup Yol Yapım grubu işaretleridir. Bu çalışma ile sadece Türkiye'de değil dünya çapında bir ilk başarılarak, trafik işaretlerinin anlaşılırlığı işaret dizayn faktörlerini ve kullanıcı karakterlerinin ile birlikte araştırılmıştır. Sonuçlar trafik işaretlerinin geliştirilmesi dizaynı ve anlaşılırlığı konusunda referans olarak kullanılabilecektir.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

LIST OF SYMBOLS

$\bar{x}$	Sample mean
H_0	Null hypothesis
H_1	Alternative hypothesis
R^2	Coefficient of determination
R^2_{adj}	Adjusted coefficient of determination
a	The percentage of relative accuracy desired
τ_i	The effect of the i^{th} treatment
β_j	The effect of the j^{th} block
ε_{ij}	Random error
δ_k	Effect of k^{th} level of age group factor
φ_l	Effect of l^{th} level of educational background group factor
ϑ_m	Effect of m^{th} level of active car usage group factor
∂_n	Effect of n^{th} level of marital status group factor
γ_p	Effect of p^{th} level of gender group factor
ε_{klm}	NID $(0,\sigma^2)$ random error component

LIST OF ACRONYMS/ABBREVIATIONS

ANOVA	Analysis of Variance
CV	Coefficient of Variation
EU	European
EA	European Agreement
GDH	General Directorate of Highway
HTSC	Highway Traffic Safety Committee
ISO	International Organization for Standardization
MS	Mean of squares
N/n	Sample size
NA	Not Applied
NM	Not Mentioned
NID	Normally and independently distributed
SGD	Security General Directorate
Std. Dev.	Standard deviation
SE	Standard error
SS	Sum of squares
VIF	Variance inflation factor
Yrs / Yr	Years / Year

1. INTRODUCTION

Road traffic accidents are not only one of the serious problems of countries but also a serious challenge that needs to be prevented for the public health and safety. Each year, as many as 50 million people are injured and approximately 1.2 million people die in traffic accidents in worldwide. The economic cost of those injuries is estimated as US$518 billion per year (Peden *et al*., 2004). Situation in Turkey is also worrisome. The number of traffic accidents with the fatality and injury in Turkey are approximately two to nine times higher than those in some European countries. From 2004 to 2014, the number of traffic accidents in Turkey has increased by 123% and in 2014 the human factors with the major ratio of 97.7%, seem to be the main responsible in occurrence of accidents (Turkish Statistical Institute, 2014). Moreover, based on the data, in 2014, almost 15% of faults linked to the driver in accidents derive from the failure to comply with the traffic signs (General Directorate of Security, 2014).

Traffic signs convey the message with the use of symbols/icons, words or combination of both in order to regulate, warn and guide road users to ensure the safe movement (Ng and Chan, 2008a). Traffic signs can be stated as the best known symbols since they are encountered daily in a traffic system. In general, using symbols in designs are preferred since they provide an international type of communication means that they overcome the barrier of culture and language and are more quickly recognized than their words equivalents (Ellis and Dewar, 1979; Bruyas *et al*., 1998). Although symbolic representation is a way of effective communication, the lack of a guideline that makes the meanings of symbols explicit in the same way as spoken or written communication, makes them to be interpreted in many different ways (McDougall *et al*., 1999). Previous studies have found that the different recognition and comprehension of traffic signs by users is one of the serious factors leading to traffic accidents (Retting *et al.*, 2003; Malaterre, 1990). Since, correct comprehension is very important, the design of traffic signs that should allow users to effectively comprehend and understand the correct information is very critical.

Comprehension is one of the criteria for an icon to be mentioned as effective. The visual and cognitive features of an icon have significant influence on icon effectiveness and thus comprehension. Many studied the effect of visual features like color, shape and size of icons on effective sign comprehension and found significant effect of it to understand the symbols (Ng and Chan, 2008a). McDougall *et al.* (1999) instead of considering visual features that the results of them are already obvious, focused on cognitive features as a central concern in sign and icon research and defined five cognitive features as criteria for measuring sign design; familiarity, concreteness, complexity, meaningfulness, and semantic distance. Familiarity is basically stands for the frequency with which icons had been encountered by users in their daily lives. McDougall *et al.* (1999) defined familiarity as a substantial determinant of usability. Ben-Bassat and Shinar (2006) found familiarity as one of the factors that contributed the most to comprehension. Complexity is about that the icons contain a lot of details; on the other hand icons are called as simple if they have a few details or elements. Bryne (1993) worked on symbol complexity on search performance and revealed that simplicity is the best in symbol design. Icons that depict objects, materials, or people that are obviously connected with real world is defined as concrete; those do not depict real objects are named as abstract. The performance of concrete icons by keeping extra details in minimum are found more effective than the abstract ones that convey message generally graphical features such as arrows, shapes (McDougall *et al.,* 1999). Meaningfulness is about whether the icon is perceived as meaningful or not by subjects. A strong relation between familiarity and meaningfulness was found. Familiar signs are more meaningful for the subjects. Semantic distance is regarded as the closeness of the relationship between what is depicted in an icon and the function it is intended to represent (Ng and Chan 2007a, b, 2011). Furthermore, in order to build a consistent response scales, in this study we used the definitions of sign design features proposed by Ng and Chan (2007a, b, 2011), and in this manner changed complexity to simplicity and semantic distance to semantic closeness, respectively. In the literature, positive and strong interrelation was found between familiarity, concreteness, meaningfulness and semantic closeness. Moreover, familiarity was found significantly correlated with meaningfulness and semantic closeness (Ng and Chan, 2007; McDougall *et al.*, 1999).

In our study, Turkish traffic signs will be evaluated by using those five cognitive design features and the intercorrelation between features will be investigated.

Usability has a significant role in consideration of product design. Usability is defined by The International Organization for Standardization (ISO) as the "extent to which a product can be used by specified users to achieve specified goals with effectiveness, efficiency, and satisfaction in a specified context of use" (ISO 9241-11, 1998). In relation to usability-based design, Jordan *et al.* (1991) identified three components that have an impact on the usability of products; those are guessability, learnability and experienced user performance. As a further step, two more components; namely system potential and re-usability are added to the model (Jordan, 1994). Guessability is associated with the cost (e.g. in terms of time on task or error made) of performing a task for the first time. If the guessability of a task is high than the measure of the cost is low. Learnability is defined as the cost of reaching the competent level of performance on a task while experienced user performance refers to the unchanging performance of user who has used product many times before. System potential is the optimum level of a performance that is possible to complete the task. Re-usability is the decrease in performance after an extended period of time that user has not used (Jordan, 1994). Each components of usability that are mentioned above, stand for a different part of learning curve. In terms of icon design, the first part of this learning curve which is guessability is the most crucial step since more guessable design means understanding meaning faster and with fewer errors. Furthermore, the guessability level of traffic signs is vital for users due to the fact that higher guessability will lead to fewer errors. We will be focusing on understanding the guessability level of Turkish traffic signs in this study.

In addition to all, user characteristics (e.g. experience, knowledge, culture, age and gender etc.) are expected to have an influence on usability since what is usable for one person will not necessarily be usable for another (Jordan, 1994). If a user has done a task before, it is obvious that he or she will find task easier as a second time. Ng and Chan (2007b) found that previous experience of visiting China was a crucial effect on guessing performance 120 Mainland China traffic signs. In a multinational study, Al Madani and Al-Janahi (2002a) showed that nationality has an influence on guessability whereas Ng and Chan (2007) suggested that the traffic signs are need to be designed as culture independent

in order to avoid culture bias. There are different results regarding the effect of age and gender in the literature. Al Madani and Al-Janahi (2002a) found no effect of age on sign comprehension and significant effect of gender as male drivers scored higher than females. On the other hand in another research the effect of age was found as significant by Al-Madani and Al-Janahi (2002b). Ng and Chan (2008) showed no significant effect of age on comprehension. In a recent cross-cultural study, sign comprehension of drivers from four different countries was compared and it was found that there are large differences among comprehension level of drivers according to the country and user factors such as age and experience etc. (Shinar *et al*., 2003). In our study, we will investigate also the effect of user factors on sign guessability and so on comprehension.

The traffic signs currently used in Turkey are based on the Highway Light and Road Signs agreement that have signed in Vienna on 8 November 1968 by the Economic Council of the European (General Directorate of Highway, 2004). However, there has not been any study conducted to investigate the comprehensibility of current signs by Turkish population considering user factors and compliance of them with ergonomic design principles. Therefore, to fill this gap, this study attempts to analyze the comprehensibility level of current traffic signs by the Turkish population. In this study, we will not only investigate the guessability score of current traffic signs but also analyze them considering ergonomic cognitive sign design features; familiarity, concreteness, complexity, meaningfulness, and semantic distance and the user factors like age, gender, region, educational background, driving experience, marital status etc. and reveal the relationships. In addition, we will investigate the effect of user factors on evaluation on five cognitive sign design features which there is not any study in the literature that focus on that analysis. Moreover, this study will be the first of its kind in Turkey and it is believed that the results will bring a different approach to design more user friendly traffic signs.

To do this, the thesis is organized as follows: Chapter 2 includes a brief summary of literature about usability of product design focusing more on guessability step, the effect of user factors on guessability and comprehensibility of signs considering the cognitive sign design features. In Chapter 3, the rationale behind the study and also the main objectives of the current study are presented. In Chapter 4 the methodology of the experiments that includes description of subjects, equipment used, procedures while conducting the tests

and statistical procedure followed to analyze the collected data. In Chapter 5, the results of the sign guessability scores of each signs, user factors' effect, cognitive sign design features scores and intercorrelation of them are given in detail. Moreover, the statistical analysis results of the study are presented in this chapter. In Chapter 6, discussions about the results of the current study presented and also comparison with the results of the other studies is presented in this chapter. In the last chapter, conclusion of current study and recommendation for future studies are presented.

1. LITERATURE REVIEW

1.1. History, development and importance of traffic signs

Traffic signs are designed to display information which conveys a particular message to the drivers and located along the roads. The main purpose of traffic signs is to give particular information, direction, warnings and guidance to drivers and pedestrians in order to control traffic, ensure safe driving and reduce accidents and fatalities (Ng and Chan, 2011; Ells and Dewar, 1979).

Before continuing with the importance of the design of traffic signs and the proper comprehension of them, it will be valuable to understand the history and the development of traffic signs. Focus on design, usage and standardization of road signs started with the proposal that was considered in a Congress of the International League of Touring Organizations in Paris, by 1900's. After the organization, between 1926 and 1949, European road sign system and Inter-American developped the first traffic signs according to the agreement which was done in that organization. Later on that, Vienna Convention on Road Traffic was organized in Vienna on 8 November 1968 in order to discuss international uniformity in rules managing the road signs, markings, signals and symbols in Europe and standarts are defined in Vienna Protocol. European Agreement Supplementing the Convention on Road Signs and Signals of 1968 had signed in Geneva on 1 May 1971 by the countries who accept that protocol (Kirmizioglu, 2010). The protocol describes specific symbols and colors for each category and also offers different sign and symbol alternatives that the countries can choose to use in only one of them. Also, making simple adaptations on symbols are allowed by not changing the basic characteristic. That protocol and agreement were an important step for not only standardization but also the development of road safety policies to reduce the traffic accidents and fatalities.

Traffic signs convey their intending messages in words and/or symbols. Even though symbols are highly used since they found as an effective way of communication and easily recognizable, they can also be interpreted in a different ways that can lead

misunderstanding (Ells and Dewar, 1979; McDougall *et al.*, 1999). Therefore, the design of traffic signs has a crucial importance for effective traffic control.

According to Canfield (1999), traffic signs are most effective when they comply with the following requirements: fulfill a need, command attention, convey a clear and simple message, command respect of the road users and give adequate time for proper response. Recent studies on traffic signs indicate that some of the signs fail to convey their messages as effectively as expected (Ng and Chan 2008a). Therefore, the traffic signs need to be designed with a high level of usability to minimize the risk of misunderstanding and support proper comprehension. Shinar *et al.* (2003) suggests that traffic signs that are designed consistent with the ergonomic guidelines are more likely to be understood by people.

Another substantial aspect influencing the effective traffic control is the proper understanding of traffic signs by the road users. Yannis *et al*. (2007) indicates that drivers who lack understanding of traffic signs have an increased risk of accident and injury involvement. Recent studies show that although traffic signs are very significant, they are not always understood properly by the drivers. In a multicountry study in which the sign comprehension level of five Arabic countries had been assessed, Al-Madani and Al-Janahi (2002a) found that only 56% of the 28 posted signs are comprehended correctly by the drivers. Moreover, Ben-Bassat and Shinar (2006) conducted a study with 23 male and 17 female licensed drivers with three to fifteen years active driving experience and found that comprehension varied among signs. Some of the 30 posted traffic signs were misunderstood by over 50% of the participants whereas at least 10% of the participants misinterpreted the sign to mean the opposite of its true meaning. Al-Yousifi (1999) also showed that the comprehension level of some traffic signs is misinterpreted. In another study, Shinar *et al*. (2003) evaluated the comprehension level of traffic signs in four different countries, which are Canada, Finland, Israel and Poland, and compared the comprehension levels of different traffic signs. They found that the comprehension level significantly differ among drivers of different countries and driver populations of these countries. They also observed that the signs consisted with the ergonomic guidelines are comprehended well. Ward *et al.* (2004) investigated the comprehension level of U.S drivers by using 100 international traffic signs and evaluate the effect of training on

comprehension. It was concluded that sign comprehension of U.S. drivers was very low only 17 out of 100 signs were comprehended in 85% accuracy.

To sum up, traffic signs have a significant role to ensure safe driving and effective traffic control. The misunderstanding of them is one of the factors of traffic accidents and injuries. There is a need of effective sign design for public health and safety.

1.2. History and development of traffic signs in Turkey

Traffic signs are used in Turkey for the traffic management and safety system, as in many countries. The rules in order to manage traffic safety are defined in "Road Traffic Law (2918)" in Turkey. This law defines the traffic rules, right and obligations, implementation and control of the rules, the responsible institutes and their authorization, responsibilities, working procedures and other provisions (Official Journal of the Republic of Turkey, 1983).

In Turkey, the main authority on making the last decision on the recommendation for traffic safety and the necessary organisation for implementation of them is the Highway Traffic Safety High Committee (in Turkish *Karayolu Trafik Güvenliği Yüksek Kurulu*) which is chaired by Prime Minister with the participation of 8 related Ministers. In addition, there is a support committee called Highway Traffic Safety Committee (HTSC) (in Turkish *Karayolu Trafik Güvenliği Kurulu*), which is chaired by Head of Turkish Security General Directorate (SGD) (in Turkish *Emniyet Genel Müdürlüğü*) and the role of this committe is to discuss the applicability of the recommendation proposed by the participants and make a latest decision on it. Last but not least, the General Directorate of Highway (GDH) is the main authority on determination of the standards of road signs and the publishment and controlling of them on highways in Turkey (Official Journal of the Republic of Turkey, 1983).

The traffic signs that are currently used in Turkey are based on the Highway Light and Road Signs agreement that have signed in Vienna on 8 November 1968 by the Economic Council of the European (General Directorate of Highway, 2004). In 1968, Turkey accepted to be a part of "Vienna Convention" and used the standard signs and

signals accepted with Vienna Convention, but did not mandatorily follow the further limitations accepted in European Agreement (EA).

In the European Agreement, which was signed on 1 May 1971 between the European countries accepted to be a part of Vienna Protocol 1968, some revisions are made on Vienna Convention (United Nations, 2008a, b; Kirmizioglu, 2010). Vienna Convention had been offered different alternatives for the design of traffic signs. For example, for the overtaking prohibited traffic sign, Vienna Convention accept to use any of the versions with or without a crossing red line whereas EA accepts sign do not have this red crossing line (See in Figure 2.1). In Turkey, GDH changed some of the sign in complience with European Agreement in 2004 and removed the red crossing line from the signs. In the Traffic Sign Manual, it is explained that traffic signs should be replaced with the new ones when they complete their lifetime and become unusable (General Directorate of Highway, 2004). Therefore, both the old and new versions of those traffic signs are on the roads together since 2004.

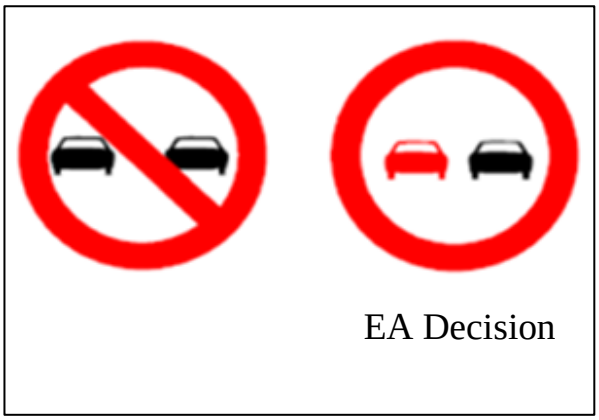


Figure 2.1. Vienna Convention Alternatives for “Overtaking Prohibited ” Traffic Sign (General Directorate of Highway, 2004)

2.2.1. Traffic sign types in Turkey

As indicated in GDH (2004), there are two types of marking system in Turkey as horizontal and vertical. Horizontal markings are the lines, arrows, writings and symbols which are drawn on the roads in order to guide road users. Vertical lines include traffic

signs and other control equipments which are placed outside of the road and convey messages about the road and road environment (See in Figure 2.2 and 2.3).

Figure 2.2. An Example of Horizontal Marking

Figure 2.3. An Example of Vertical Marking

Traffic signs in Turkey are classified according to their type and each type has different colors and shapes. There are three basic shapes of traffic signs: triangular shape for warning, circular shape for regulatory, rectangular shape for information. The color of triangular signs are always red since red is associated with danger. Blue and red colors are used with circle signs. Blue circles state mandatory information such as “straight and left only” or “turn right only”. Red rings in circle signs express an order and shows what you are not allowed to do, for example “no entry to the vehicles indicated” or “no right turn”. Blue, green, yellow, white and brown colors are used together with rectangle signs. Blue rectangle is used to give information except on highways. Green rectangle is used to give direction in primary routes in highways. Yellow rectangle is used to express construction,

maintenance and repair. White rectangle is used on non-primary routes. Brown rectangle is used in tourist areas (General Directorate of Highway, 2004).

As mentioned in GDH (2004), traffic signs in Turkey are grouped in six main categories as indicated below (Sample traffic signs of Turkey are given in Table 2.1):

- Warning Signs (T Group): The main purpose of these signs is to warn the road-users regarding a danger and the type of the danger on the road. In addition to danger warning, these signs also warn drivers to reduce their speed and drive more carefully. In Turkey, these signs are shown with an equilateral triangle with white background, thick red border and black symbols as the other countries.
- Regulatory Signs (TT Group): These signs inform road-users of restriction, prohibition and special obligations. In Turkey, these signs are shown with a circular shape with blue background or red ring around. Only "Stop" and "Yield" signs have different shape than circular.
- Informative Signs (B Group): The main purpose of that type of signs is to give useful information to the road-users. These signs are generally used in blue rectangular background with white symbols. Highway signs are used with green rectangular background whereas tourist area signs are used with brown rectangular background.
- Standing and Parking Signs (P Group): These signs show where vehicles can stand or park on the road. They have circle shape with red symbols or blue rectangular with white symbols and/or letters.
- Road Works and Maintenance Signs (YB Group): These signs may have orange or yellow background with an original shape. The aim of these signs is to give information of temporary situations such as road works or detours.
- Additional Panels (PL Group): These signs are used as an integrated part of traffic signs. They are placed under the traffic signs in necessary conditions in order to give more detailed information.

230 traffic signs (53 warning signs, 64 regulatory signs, 82 informational signs, 7 standing and parking signs, 12 road works and maintenance signs, 12 types panel signs) in total are currently used in Turkey (General Directorate of Highway, 2004).

2.2.2. Comparison of Turkish traffic signs with other countries

In many studies, the effect of cultural background on comprehension had been observed (Ng and Chan, 2007b; Summala, 1998; Shinar *et al*., 2003; Hawkins, 1998). As a result of cultural effect on comprehension, the design of traffic signs can differ in color and shape between countries. For example, in South Korea, warning signs are triangular shape with yellow background and prohibition signs are black or white background. In United States, yellow is used for warning signs, the use of red on signs are limited to stop, yield and prohibition signs. Rectangular traffic signs give information in Turkey, China, Germany, Spain, while they give prohibition and restriction for United States and New Zealand.

Since the globalization is increasing and one person is licensed in a place can drive in another place without further training, the standardization of traffic signs has a crucial importance in order to overcome misunderstanding.

2.3. Comprehension of traffic signs

In the literature there is limited number of studies conducted to investigate the comprehension of traffic signs. Some of the studies focused on to understand the effect of cognitive sign design features on comprehensibility of traffic signs. Those studies explored the effect of each sign design features on comprehension and the interrelation of those sign design features. Some of studies concentrated more on user factors such as age, gender, experience, education and tried to examine whether those factors had any effect on comprehension of traffic signs or not. In Turkey, there are only two studies worked on comprehension of traffic signs, which are the master thesis from different universities (Kirmizioglu, 2010; Yakut, 2006). Both of the studies are conducted to investigate the comprehension level of traffic signs considering user factors, none of them discussed the effect of cognitive sign design features.

Kirmizioglu (2010) studied to measure of the comprehension level of traffic signs designated as the effect of them on traffic accidents can be more than expected, with the support of a sub-committee under Road traffic Safety Board (in Turkish *Karayolu Trafik*

Güvenliği Kurulu). A pilot study was done in the Ankara, the capital of Turkey. The awareness of 39 traffic signs and the effect of user factors such as age, gender, educational background, occupation, driving license class, driving year, occupational driving requirements, the average inner-city and outer-city kilometers per year and the number of traffic fines taken in last five years on comprehension are investigated via a multiple-choice survey. Yakut (2006) analysed the the quessability of 20 traffic signs and the effect of user factors such as education, gender and type of driving licence on comprehension. Since our study will focus on to investigate comprehension level of current traffic signs in Turkey considering both the effect of cognitive sign features and the user factors on guessability score, literature review will be done in two parts in below. Some of the significant studies related with those issues will be summarized below are examined in detail before conducting the current study, even though they do not have exactly the same conditions as the current study has.

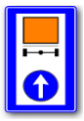
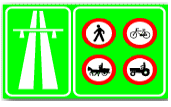
2.3.1. Effect of cognitive design features on comprehension

International Organization for Standardization (1998) defines usability as the effective, efficient and satisfactory use of a product in order to achieve a specific goal. Jordan *et al.* (1994) identified five components of usability as a part of usability–based design approach, which are, guessability, learnability, experienced user performance, system potential and re-usability. Each components of usability build the different parts of learning curve and the first step of this curve, meaning, guessability has the significant importance to understand faster and with a fewer error. Guessability is associated with the cost (e.g.in terms of time on task or error made) of performing a task for the first time. Higher guessability brings lower cost. Therefore, to have a higher guessability rate of traffic signs has a crucial importance in order to reduce the defects stem from lack of comprehension.

McDougall (1999) focused on to understand what makes a symbol easy to understand and conducted a study to provide rating norms for 239 symbols and icons that are selected from different sources (computers (n = 77), traffic and public information (n = 48), industrial (n = 69), and household goods (n =50)). The main aim of the study was to quantify symbol characteristics to contribute symbol usability. 100 female and male

university students participated the study. As a methodology, symbols are given as a 20 page booklet, 12 symbols in each page and for randomization, symbols are distributed in pages by using latin square design. A brief information regarding rating and the meaning of design features are included in the booklet. Participants rated each design features for

Table 2.1. Sample of different types of Traffic Signs currently used in Turkey (General Directorate of Highway, 2004)

Warning Signs		 Curve to Right	 Other Danger	 Dangerous Bend of Road to the Left
Regulatory Signs	Mandatory Signs	 Right Only	 End of Minimum Speed Limit	 Direction for Vehicles Carrying Dangerous Goods
	Prohibition Signs	 No Vehicles Taller than Indicated	 No Overtaking by Trucks	 End of Any Restrictions
	Priority Signs	 Stop	 Yield	 Priority Over Oncoming Vehicles
Informative Signs		 Hospital	 Highway begins	 Priority Road
		 Direction Sign	 Direction for Tourist Area	 Police
Standing and Parking Signs		 No Parking	 Parking	 No Stoping and Parking
Road Works and Maintanance Signs		 Road Works	 Traffic Lights	 Speed Limit

each sign in 5–point rating scale. To evaluate the validity of complexity ratings, complexity metrix developped by Garcia *et al.* (1994) is used.For meaningfulness part, after rating, participants were asked to state briefly what they thought the meaning of the symbol to be. In the study, it was suggested to consider cognitive design features rather than visual design features like shape, color, and size. Therefore, five symbol characteristics which are familiarity, concreteness, complexity, meaningfulness, and semantic distance are defined as of central concern. The interrelationship between those five symbol characteristics are also evaluated. Familiarity, as a substantial determinant of usability is defined as the frequency with which icons had been encountered by users in their daily lives. It was found that familiar items make a sense to the participants with the help of real life experience, even though they might not be correct and familiarity improves user performance. Icons that depict objects, materials, or people that are obviously connected with real world is defined as concrete; those do not depict real objects are named as abstract. The performance of concrete icons by keeping extra details in minimum are found more effective than the abstract ones that convey message generally graphical features such as arrows, shapes. Complexity is about that the icons contain a lot of details; on the other hand icons are called as simple if they have a few details or elements. In the study no strong relationship between concreteness and complexity is found and it is suggested that although concrete designs are thought easy to use since they include extra visual design, they can also be simple. Altogether, simplicity needs to be considered to produce effective symbols. Meaningfulness is about whether the icon is perceived as meaningful or not by subjects. An interchangeable and strong connection between familiarity and meaningfulness was found. Familiar signs are more meaningful for the subjects. In addition to this, a strong relation between concreteness and meaningfulness is also observed. Concrete symbols are found to be more meaningful for participants than abstract ones. Semantic distance is regarded as the closeness of the relationship between what is depicted in an icon and the function it is intended to represent. In our study, traffic signs of Turkey are also evaluated considering those cognitive design features.

Ng and Chan (2007a) examined the cognitive design features of 120 Mainland China traffic signs by considering the five cognitive design features determined by McDougall *et al.* (1999). In order to build consistent order of response scale, the terms ‘complexity’ and ‘semantic distance’ used in revised as ‘simplicity’ and ‘semantic closeness’, respectively in

this study. 19 male and 22 female, ages between 18 to 27 yrs (median=22,5 yrs), Hong Kong Chinese undergraduates who have never taken any driving tests or any driving licence and with no color deficiency participated to study. An adjustable chair, personal computer (1200 MHz microprocessor and 17-inch CRT monitor) as an apparatus and power point presentation were used to screen the traffic signs. A sign evaluation sheet to evalute traffic signs are evaluated and distributed. The Pesudo-Isochromatic Plates was used to evaluate the color deficiency problem of participants. Subjects were briefed about the meaning of features before the experiment and after practice trial with five signs, 120 Mainland China traffic signs presented randomly on the screen. Participants gave subjective ratings in 0-100 point scale for each sign. Subjects evaluated familiarity, concreteness, simplicity, meaningfulness at first, after that a verbal label is appeared under the sign and subjects were asked to rate sematic closeness. 1 minute rest was given after each 40 signs. Ng and Chan (2007a) give essentially the same information with McDougall *et al.* (1999). A strong correlation among the familiarity, meaninfulness and concreteness was found. Familiar signs were more meaningful for the participants. The more frequently encountered signs were evaluated with higher semantic closeness rating. It was found that concrete signs were more easy to understand for Chienese people. There was no correlation between the simplicity and other four features in the study.

Ng and Chan (2007b) investigated the usability of traffic signs with guessing and comprehension tasks conducting two experiments with two different groups of Hong Kong Chines. Different than the other studies, one sign design feature (Concreteness) and two user factors (Mainland China visit experience and non-local driving experience) were searched in this study to evaluate the effect of culturel issues on sign usability. In the first experiment, the participants were Hong Kong Chinese engineering undergraduates with no driving experience and driving licence and 120 signs were examined. Subjects guessed the meaning of each sign and gave 0-100 point scale subjective ratings for concreteness. In the second experiement, Hong Kong full driving licence holders participated the study and reported their non-local driving experience first, and then completed multiple-choice questions for evaluating their traffic sign comprehension for 120 traffic signs which 26 of them were similiar with the first experiment group. It is found that the guessing performance of participants who had visited the Mainland China was significantly better than the ones who had not visited. It is suggested that to use symbols which are culture

independent or multi-culturel might be beneficial to increase usability of signs across international markets since one icon can have different meanings in different cultures. Moreover, a positive effect of non-local driving experience on local sign comprehension was observed when the signs looked alike and conveyed the same meaning, but the comprehension is decreased when the meaning of those signs are different. Also, the quessing performance was found higher in concrete signs.

Ng and Chan (2011) evaluated the effect of different training methods (paired-associate learning, recall training, and recognition training) and the association between effectiveness of those training methods and five cognitive sign design features (familiarity, concreteness, simplicity, meaningfulness and semantic closeness) evaluation. 36 Hong Kong Chinese engineering undergraduates (29 males and 7 females) with no driving experience or licence participated the study. Five sign design features are evaluated in 0-100 point scale by participants. It was found that training did not have a serious effect on the comprehension of traffic signs. A strong relation was found between semantic closeness and before training and after training results. Traffic signs with higher semantic closeness scores tend to be easily guessed and interpreted. No strong effect of familiarity, concreteness, simplicity and meaninfulness on training was observed.

Ou and Liu (2012) conducted a study in Taiwanese and Vietnamese groups to investigate effect of five sign design features (familiarity, concreteness, simplicity, meaningfulness and semantic closeness) and three training conditions (before training, immediately following training, and one month after training) on the comprehension of traffic signs. Only the undergraduate students (30 Taiwanese and 30 Vietnamese with no driving licence or experience) participated the study and 65 traffic signs were used. In contrary to the Ng and Chan (2011) findings, significant effect of training on sign comprehension was observed in both groups. It was found that the effect of semantic closeness on comprehension is the most crucial one and followed by meaningfulness, concreteness, familiarity, and simplicity. Moreover, the long lasting effect of semantic closeness was observed. Higher score in semantic closeness means higher comprehension level and less decline of comprehension even after one month period.

Bryne (1993) is reserched the factors when searching for documents and defined the type of icon as the most important factor in icon search. It was concluded that icons should be simple and easy disciriminiable to be effective and simple icons significantly performs than complex icons in usability aspect.

Ben-Bassat and Shinar (2006) investigated the guessability of traffic signs and the effect of convenience to ergonomic principle in design. In the study five main ergonomic principles (spatial compatibility, conceptual compatibility, physical presentation, familiarity and standardization) were concerned. The physical and conceptual compatibility were defined as the correspondence between the sign and message it represents. Standardization was mentioned as the uniformity in the forms/colors/symbols/ directions in all traffic signs for representing a similar meaning. Familiarity was the frequency to encounter a sign in real life. Participants of the study were all industrial engineering and management undergraduate students (17 male, 23 female with average age 28.8 yrs.). Compatibility evaluation was made by 13 human factors and ergonomic professors and advanced graduate students, whereas standardization assessment was made by Ben-Bassat and Shinar. Thirteen traffic signs were evaluated in the study. Participants were asked to guess the meaning of presented sign at first and than evaluated the principles in 0-10 point scale. In order to evalute the guessing score +2, +1, -1 and -2 scores are used (+2 for fully correct answer, +1 for partially correct answer, -1 wrong answer, -2 for opposite of correct answer). Although large differences in comprehension among different signs was observed, signs were comprehended best when they comply with the defined ergonomic principles. It was very strange that two of the signs were answered as the opposite of their true meanings by more than %10 of the participants which was explained as dangerous. High correlation between compatibility level and comprehension was observed. Similiar to the findings of McDougall *et al.* (1999) and Ng and Chan (2007a, b), signs related with the reality, in other words concrete, and high familiarity ratio were found more comprehensible by participants.

Bruyas *et al*. (1998) conducted a study with 32 participants. Half of them were young (23 to 33 yrs old) and half of them were old (60 to 70 yrs old) to set up ergonomic recommendation for the optimisation of pictoral information design by considering age effect. Car, suitcase, car battery, cup, tunnel, cable car, plane, horn and electrical plug

objects with various versions of them (more or less complete due to the additon or the withdrawing of some of their components) were tested. And it was concluded that using non-important extra objects in the design makes difficulty in comprehension. Graphical representation should be easy to understand with no doubt and wide variety of the users need to be considered since the perception changes by age, cultural background and comprehension.

Isherwood *et al.* (2007) searched the relation of icon characteristics (concreteness, visual complexity, semantic distance and familiarity) with the speed and accuracy of icon identification by testing 40 icons that were selected between the icons from the study of McDougall *et al.* (1999). In the experiment, participants matched the icon with it's correct functional label and also rated the icon characteristics in 1 to 5 points scale using computer screen and a mouse. The most crucial finding of the study was that the importance of icon characteristics on comprehension changed with experience. Semantic closeness was found important at initial step while icon and it's function relationships were learned, but familiarity was found as crucial in later step since it has long lasting effect on access the long term memory representation. Concreteness had not a serious importance to identify icon.

Chan and Chan (2011) evaluated the relation of cognitive design features familiarity, simplicity, concreteness and meaninfulness with the comprehension score and found that the comprehension ability was higher for familiar, concrete, simple and meaningful signs.

Ng and Chan (2008b) examined the effect of cognitive design features on comprehension score of traffic signs. As a result of the study, familiarity was found as siginificantly correlated with comprehension score whereas no strong relationship were observed between concreteness, simplicity and meaningfulness and comprehension score.

To gather subjective ratings for cognitive design features, McDougall *et al.* (1999) used five-point scale whereas Ng and Chan (2007a, b) prefered 0-100 point scale. Gulta *et al.* (2001) showed that using higher number of scale points makes data easy to use in various kinds of statistical analysis. Preston and Colman (2004) mentioned that shorter

scales are found by people as easy to use, however longer scales are prefered by the people since they allow to express their feelings more adequately.

To measure icon concreteness Garcia *et al.* (1994) proposed a metric by adding closed figures, letters, open figures, special characters, horizantal lines, vertical lines, arcs, arrowhead. In the background of this metric, complex symbols are taken as more concrete. Higher metric value showed higher concreteness. However, McDougall *et al.* (1999) showed that there was not a significant relationship between complexity and concreteness; simple symbols could also be concrete. In the review study of Ng and Chan (2008a), subjective rating was found as the most efficient method to quantify cognitive design features.

Shinar and Vogelzang (2012) examined the comprehension speed and accuracy of 30 traffic signs varying in their of familiarity in three different condition (symbol-only, text-only, and symbol + text). The particiapants of the study were all university engineering students (26 male, 22 female average age 24,8 yrs with at least two years driving experience). Recommendation from this study were using symbols are beneficial only when the symbol is familiar to the driver, and even then, comprehension reaction time is slower for symbols than for text. Koyuncu and Amado (2008) also evaluated the effect of type of sign on reaction times and came up with the solution that the best reaction times were found with the traffic signs include both word+symbols.

Lin and Kreifeldt (1992) examined to define the appropriate design style for the images. 30 icons from different drawing packages were matched with the referents in the study by 59 male industrial design students (16 to 20 yrs). From the study, it was concluded that if the information or function was highly related with an object, pictoral icon should be used. Moreover, if a command is strongly related with a concept than abstract icons should be prefered to use.

2.3.2. Effect of user factors on comprehension

In order to evaluate the acceptable level of comprehension, guidelines defined the criterion percentage in a comprehension test to be considered corretly. The International

Organisation for Standardization (ISO 3864-3:2012) adviced that symbols must meet the at least 67% correctness criterion, while American National Standard Institute (ANSI Z535.3-2007) defined the correctness criterion as at least 85% for an acceptable comprehension test.

Al-Madani and Al-Janahi (2002) conducted a multi-national study to investigate the effect of personal characteristics experience, accident per experience ratio, age, marital status, sex type, nationality, educational background and monthly income on comprehension of traffic signs. 28 posted signs are evaluated by 9000 drivers in total from Bahrain, Kuwait, Oman, Qatar and United Arab Emirates. Age was grouped in 5 categories (16-24 yrs., 25- 34 yrs., 35-44 yrs., 45-55 yrs., 55+ yrs.). A questionnaire is designed which was involved with short-answers to define user characteristics and multiple-choice questions to evaluate the driver's comprehension. As a result, only 56% of the posted signs on an average were correctly comprehended by the population. Nationality was found as a significant indicator in comprehension. Western drivers performed better than other nationalities and it was concluded that difference stems from the better traffic education system in Western societies than Gulf Cooperation Council States. In terms of gender, the finding was that male drivers performed better than female drivers, however it was expected that the difference in results could be limited to that region. No significiant effect of age, marital status, experience and accident rates on comprehension wasobserved. However the effect of education and monthly income was found as significant. And also, the results showed that driver's personal characteristics were not associated with their accident involvement.

Lesch (2003) conducted an experimental study with an objective to observe the effect of three different training methods (verbal label, explanatory statement and accident scenario) on comprehension and memory. This evaluation was assessed for two participant groups according to age; young (18–35 yrs of age) and old (50–67 yrs of age). 31 symbols were tested and trained while 10 symbols were only tested. It was revealed that both younger and older participants have difficulty in comprehension of warning signs while the performance of older subjects were more poor. And also, the improvement with training on comprehension was observed.

Al-Madani (2000) studied the comprehension level of posted signs on safety related characteristics in three of the Gulf Cooperation Council (GCC) states, Bahrain, Qatar and United Arab Emirates (UAE) with over 6000 participants. Driving experience, accident involvement, experience per accident, citations received in the last 3 years on speed limit violations, and seat belt usage are evaluated in the study. A questionnarie includes short-answer and multiple-choice questions was used to gather the data of 28 posted signs. 47% of the participants answered the questionnarie. Different than the results of study of Al-Madani and Al-Janahi (2002a), it was observed that drivers with the high years of driving experience performed better than the ones had less experience. No significant influence was found with the understanding of the signs and accident involvement, experience per accident ratios, or speed citations. And it was also shown that the understanding of driver's increased when the seat belt usage was increased.

Ng and Chan (2007c) investigated the effect of prospective-users characteristics and the five cognitive design features on guessability of traffic signs in a study using 120 Mainland China traffic signs. Nine prospective-user characteristics which are cycling experience, sign information receiving experience, gender, Mainland China visit experience, vehicle ownership, intention of being driver, car game experience, attention to the traffic sign, traffic incident experience were examined in the study. The effect of cycling experience, experience with receiving sign information and gender on guessing score was not found significant. Contrary to the study of Al-Madani and Al-Janahi (2002a), the performance was found as the same for females and males. The subjects were all university students. Similar to the finding of Al-Madani (2000), effect of traffic incident experience was not significant on guessability. Only for the subjects who want to become driver or have car game experience, the effect of vehicle ownership was observed as significant. Subjects who pay attention to traffic signs in their daily life had better performance in guessing traffic signs than those who did not. In addition to the effect of characteristics on guessing score, the effect of five cognitive sign design features determined by McDougall *et al.* (1999) was also considered in the study. Although the guessability of traffic signs was found better when the signs were familiar, simple, concrete, meaningful, the contribution of each feature on guessability was different. Semantic closeness was concluded as the best predictor of guessing score, and followed by familiarity, meaningfulness, concreteness and simplicity.

Al-Madani and Al-Janahi (2000) conducted a study to determine the effect of age, gender, nationality, educational background and monthly income on comprehension of traffic signs. Similar to the other studies of Al-Madani, a questionnaire that involves short answers for user characteristics and one multiple choice question for each 28 traffic signs were used. Sample was gathered from five Arabian Gulf Countries and 4774 subjects with 53% response rate were participated to the study. Similar to the study of Al-Madani and Al-Janahi (2002a), the understanding of participants were found low, just over half of the signs presented were understood by them. The effect of age, gender, education and income on drivers' comprehension was found significant. In contrast to the study of Lesch (2003), it was found that the performance of older subjects were less than younger subjects. Al-Madani and Al-Janahi (2002a) asserted that comprehension was increased with the age. And also the performances of male drivers were found better than female drivers. Moreover, it was indicated that drivers with high education level understood signs better than those with lower education level. Marital status showed no significant effect on comprehension.

Ells and Dewar (1979) measured the reaction time and the speed of drivers to comprehend traffic signs in two-stage experiment. In the second stage of experiment, the speed of comprehension for symbolic and verbal signs was evaluated with 32 university students (20 male, 12 female) who had minimum two years experiences on operating a motor vehicle. At the end of the study, symbolic signs were found to be comprehended more quickly than verbal signs.

Hancock *et al.* (1999) searched the effect of age on recognition, perception and comprehension of symbolic warning signs with the participant of 865 subjects (467 male and 398 female, age from 17 to 91 yrs). The experiement had three sections for recognition, perception and comprehension. In comprehension part, 12 warning symbols which had high recognition rate and had been observed on a product or in the environment were tested. Participants answered the question whether they had seen those symbols before or not and guessed the meaning of the symbols. Results were rated as completely incorrect, partially correct, or completely correct. As a result of the study, a strong relation between the familiarity of symbol and comprehension of symbol's meaning were observed. The comprehension performance of younger adults were higher for the symbols that were

more familiar to them. Old age group perfomed poorly in comprehension of symbols. To sum up, the age-related differences were observed between younger and older group, younger adults understood the symbols better than older adults.

Kirmizioglu and Tuydes-Yaman (2011) evaluated the comprehensibility of traffic signs among urban drivers of Turkey in the city of Ankara considering some user factors (age, gender, educational background, occupation, driving license class, years of driving, occupational driving requirement, the average inner-city and outer-city kilometers per year, and the number of traffic fines in last 5 years). A total of 39 traffic signs were tested with 1478 drivers (1134 male and 327 female). A survey was prepared to gather information of user factors and the three questions to understand familiarity, comprehension, and knowledge of participants about required response. The same category description with Shinar *et al.* (2003) to evaluate comprehension score was used in the study, except that inccorect category was divided into two groups. The results of the study indicated that the drivers do not understand most of the traffic signs correctly. Only 12 out of the 39 traffic signs were correctly defined 70% or more participants. Not only crucial but also dangerous finding was that 5 out of the 39 signs were comprehended in opposite meanings rather than true meanings of them by more than 10% of the drivers. The recommendation from the study was that in order to increase the comprehension level for especially the signs comprehended as opposite meaning, education should be given to the drivers.

Chan *et al*. (2009) was conducted a study to understand the comprehension of Hong Kong Chinese and Korean people on American security safety symbols. In the study the effect of gender on comprehension score of 24 security symbols were measured. The main objective of the study was to observe the security safety symbols, designed for use by American and with non-American users from different regions. The differences in comprehension performance between local (Americans) and international users also determined in the study. In this manner, the comprehensibility score of appropriateness of design rating and personal characteristics were gathered from 81 Hong Kong Chinese and 60 Korean participants. This study showed that the comprehension score of two Asian population were lower than that for Americans. No significant relation of age and design appropriateness rating on comprehension were found.

Chan and Chan (2011) conducted a study to investigate the effect of 10 user factors on comprehension score. 91 participants (85 male, 7 female) with the age from 18 to 58 yrs old joined the questionnarie. No significant difference was found in comprehension performance amongst different age groups. Moreover, no significant effect of education level was also observed in the study. Only 1 out of 10 factors which was possesion of registrated safety officer was found to be a significant predictor of comprehension performance.

Ng and Chan (2008b) examined the relation of user factors and comprehension performance in a study in which 21 traffic signs were tested. To compare with the findings of Chan and Chan (2011), the results for age were similiar that no difference in comprehension performance amongst age group was observed however the results for education were different, it was found that participants with higher education level performed better in comprehension. Regarding years with licence factor, comprehension performance was found as better during the first year of holding a driving license and decreased with years of being a licensed driver. No significiant effect of actual driving experience, last time driving, driving frequency and non-local driving experience on comprehension was observed in the study.

2.3.3. Summary and Critics of findings

To summarize all studies and to be able to see their lacks, classification according to the factors searched, sample size, methodology and findings, summary tables were done (Table 2.2, Table 2.3, Table 2.4, and Table 2.5).

Table 2.4 and Table 2.5 show the sample type and materials used in each study. Several studies have conducted face to face or online questionnaires to gather the data of comprehension score, user factors and sign design feature ratings. To investigate the comprehension score, the answers of participants for the meaning of traffic signs were collected via multiple-choice questionnaire rather than the open-ended answer questionnaire in some of the studies. In multiple-choice test subjects were asked to select one of the answers that best expresses the meaning of the sign whereas in open-ended test subjects were asked to give the meaning of sign in their own words. In multiple-choice

questionnaire participants can be influenced by the distracting alternatives (wrong answers) or guess the meaning of the sign even though they do not know. Wolff and Wogalter (1998) suggested that the open-ended method is better than multiple - choice method to gather information on sign comprehension and guessing score.

In some of the studies the sample size was very limited. Only some of them (Al-Madani and Al-Janahi, 2002a, b; Al-Madani, 2000; Shinar *et al.*, 2003; Hancock *et al*., 1999; Yakut, 2006) which conducted multinational studies and/or used online questionnaire or multiple-choice questionnaire reached higher sample size except Kirmizioglu and Tuydes-Yaman (2011) which used open-ended answer questionnaire. Several studies that evaluated the relation of sign design features with comprehension without considering age and education factors, used generally university students (age from 19 to 25 yrs), prospective or experienced drivers as participants in the studies. Although, the findings of these studies were highly relevant to the industry and sign designers and the experiments were very successful in examining the cognitive design features on traffic signs, more extended sample type regarding age, education, experience is needed for the generalization of the results.

The user characteristics and sign design features were discussed in the studies. Some of the studies analyzed the effect of user factors or sign design features on comprehension whereas only a few of them evaluated the effect of both together on comprehension. The user factors that were considered in studies changed amongst the studies. Most of the studies discussed the effect of age, education, gender, driving experience on comprehension and a few of them evaluated the effect of the marital status, monthly income, occupation, accident ratio and other specific factors additionally. Moreover, the findings regarding the effect of user factors on comprehension were varied amongst studies. Different findings regarding age factor were investigated by studies. Some of them found significant difference in comprehension by age whereas some of them did not observe any significant effect of it on comprehension. The difference could stem from the fact that different age groups were used as an older and younger in the studies. The same results for the education and gender were also revealed. Findings regarding the effect of education and gender changed amongst studies and also nationalities, but generally males were performed better than females and the performance of participants with higher

educational level was higher. In order to evaluate the effect of user factors, it is crucial to consider the interrelationships of those factors due to the fact that two of the factors could be highly correlated and could change the perspective of results.

Some studies considered familiarity, concreteness, complexity/simplicity, meaningfulness and semantic distance/closeness as a cognitive sign design features (McDougall *et al.*, 1999; Ng and Chan, 2007a, c, 2008a, b, 2011; Ou and Liu, 2012; Liu and Ho, 2012; Isherwood *et al.*, 2007; Chan and Chan, 2011) where Ben-Bassat and Shinar (2006) used compatibility, familiarity and standardization. Although the name expressions were different the approach was similar except standardization. In our study, we used the features defined by McDougall *et al.* (1999).

In order to evaluate the comprehension score, several studies used percentage of correct answers in multi-choice answer questionnaire or scoring was done for the answer of correct, partially correct, wrong and opposite meaning in open-ended questionnaire. In order to make a more accurate evaluation, scoring for each type of answers may be beneficial for the researchers, since this method enables the researches to group not only partially correct answers separate from the fully correct ones but also the opposite comprehension separate from the wrong ones. For the evaluation of cognitive design features different point scales were used in the literatures (0 to100 points, 1 to 10 scales, 1 to 5 scales, 1 to 3 scales, 7 points Likert scale). Previous studies showed that longer scales are prefered by the people to express their feelings more accurately and also using higher number of scale points makes data easy to use in statistical analysis (Gupta *et al.*, 2001; Preston and Colman, 2004).

Some of the previous studies gave brief information before starting the experiments but not all of them made practice trials before starting the experiment. In order to make accurate subjective ratings, fully understanding the definition of each sign design factors is crucial. Practice trials make participants to understand the methodology and give more conscious answers. In addition to this, especially for the long duration experiments, giving a rest time is important in order to keep respondents' interest high and avoid fatigue. However, it is also very important to avoid making longer experiments. Some of the studies gave short rest times during the experiment and if there were more than one

experiment to conduct in different issues, giving some periods between experiments is beneficial.

In Turkey, there are only two studies conducted to investigate the comprehension score of current traffic signs. Yakut (2006) in master thesis investigated only the comprehension score by using 20 current traffic signs with multi-choice answer technique, where Kirmizioglu (2010) in master thesis and Kirmizioglu and Tuydes-Yaman (2011) evaluated the effect of user characteristic on comprehension performance together with comprehension score by using 39 traffic signs in a pilot study in Ankara. Although the results of the studies are beneficial to understand the comprehension of traffic signs in Turkey, more extended study design is needed for the generalization of the results.

There have been several previous studies designed to measure the comprehension score but really a few of them evaluated both the effect of user factors and sign design features together on comprehension with appropriate sample size and methodology. Therefore, it can be said that the studies performed on understanding the comprehension of traffic signs so far do not provide extended and generalized information regarding the effect of both user factors and cognitive sign design features.

Studies previously undertaken by the researchers mainly have the following results:

- Comprehension performance of traffic signs is low and large differences in comprehension level among different countries are observed.
- Ergonomically designed traffic signs are more understandable and guidelines will improve comprehension.
- Some signs were guessed as opposite of their true meanings which is very dangerous and could result with deadly outcomes on the road.
- Five symbol features significantly influence comprehension of symbols and performance.
- Contribution of five sign design features (familiarity, concreteness, complexity/simplicity, meaningfulness, semantic distance/closeness) on guessability is different. Results regarding the interrelation of five sign design features vary from study to study.

- Generally, familiarity and semantic closeness are found highly correlated with comprehension.
- Close relationship between concreteness and meaningfulness are found. Concrete signs are more meaningful.
- The results regarding the effect of age on comprehension vary from study to study, but generally older adults experienced greater difficulty in understanding particular symbols as compared to younger adults.
- Results of the effect of education on comprehension are inconsistent in literature. But generally the performance of highly educated participants was higher.
- Results of the effect of driving experience on comprehension are inconsistent in literature. Some studies founded no significant effect of experience whereas others founded that the comprehension level of highly experienced drivers higher than less experienced drivers.
- The findings regarding gender are different amongst studies but generally the performance of males is better than females.
- Comprehension performance is not related with the years of active driving

Table 2.2. Summary of studies in literature according to the purpose, factors investigated and sample type

Source	**Location**	**Purpose**	**User Factors**	**Cognitive Design Features**	**Signs/symbols analysed (size - type)**	**Sample (Nationality, Size & Gender, Driving experience)**	**Age Groups**	**Education Level Groups**	**Driving Experi-ence Groups**
Ells, Dewar, (1979)	Canada	•Reaction time •Comprehension score	NA	NA	8-Warning, 8-Regulatory	6M-6F, experienced.	NA	NA	NA
					7-Warning, 7-Regulatory	12M-12F, experienced.	NA	NA	NA
Lin, Kreifeldt, (1992)	USA	Define right image functions for icon design	NA	Concreteness	30-Icons	59M, NM	NA	NA	NA
Bryne, (1993)	Atlanta	•Factors involved in icon search	NA	Simplicity	15 subjects - 45 icon type	Atlanta, 45, NM	NA	NA	NA
Bruyas, Breton, Pauzie (1996)	France	•Comprehension score •Sign reaction time •Effect of age on recognition	Age	NA	91-Symbols	•16, NM •16, NM	23-33 yrs 60-70 yrs	NA	NA
McDougall Curry, Brujin, (1999)	United Kingdom	Evaluation of sign design features	NA	•Familiarity, •Concreteness •Complexity, •Meaningfulness •SemanticDistance	77-Computer 48-Traffic& public 69-Industrial 50-Houssehold goods symbols	100M-100F, NM	NA	NA	NA

*NM- Not mentioned, NA – Not Applied

Table 2.2. Summary of studies in literature according to the purpose, factors investigated and sample type (cont.)

Source	Location	Purpose	User Factors	Cognitive Design Features	Signs/symbols analysed (size - type)	Sample (Nationality, Size & Gender, Driving experience)	Age Groups	Education Level Groups	Driving Experi-ence Groups
Hancock, Rogers, Fisk, (1999)	USA	•Symbol perception, recognition and comprehension •Effect of age	Age	NA	12-Warning symbols	Atlanta,467M-398F, NM	17-35 yrs 35-54 yrs 55-64 yrs 65-91 yrs	NA	NA
Al Madani (2000)	Saudi Arabia	•Comprehension score •Effect of user factors	•Driving experience •Accident involvement •Experience per accident •Citiation in last 3 yrs •Age	NA	10-Warning 18-Regulatory	•Bahrayn, 970, NM •Quatar,1000, NM •UAE, 850, NM	< 24 yrs 24-34 yrs 35-44 yrs 45-54 yrs 54+ yrs	NA	0-5 yrs, 6-10 yrs, 11-20 yrs, 20+ yrs
Al-Madani, Al-Janahi, (2002a)	Saudi Arabia	•Comprehension score •Effect of user factors	•Experience •Accident per experience, •Age, •Marital status, • Gender, •Nationality, •Educational Background, • Monthly income	NA	18-Regulatority, 10-Warnings	• Kuwait,617M - 301F, experienced •Oman,715M - 314F, experienced •Oatar, 995M - 37F experienced •UAE, 808M - 56F, experienced • Bahrain,717M - 220F, experienced.	16-24 yrs 25-34 yrs 35-44 yrs 45-55 yrs 55+ yrs	• Below secondary, •Secondary to under diploma, •Diploma holders, •Higher studies	< 5 yrs, 5-10 yrs, 10-15 yrs, > 15 yrs

Table 2.2. Summary of studies in literature according to the purpose, factors investigated and sample type (cont.)

Source	Location	Purpose	User Factors	Cognitive Design Features	Signs/symbols analysed (size - type)	Sample (Nationality, Size & Gender, Driving experience)	Age Groups	Education Level Groups	Driving Experi-ence Groups
Al Madani, Al Janahi (2002b)	Saudi Arabia	• Comprehension score • Effect of user factors	• Age, •Marital Status, •Gender, •Nationality, •Educational Background, •Monthly income	NA	10-Warning 18-Regulatory	•Kuwait, 592M - 291F, licenced •Oman, 694M - 303F, licenced •Oatar, 990M - 37 F, licenced •UAE, 796M - 53F, licenced •Bahrain, 709M - 218F, licenced.	16-24 yrs 25-34 yrs 35-44 yrs 45+ yrs	•Below secondary, •Secondary & diploma, • B.Sc • Higher degrees	NA
Lesch, (2003)	USA	•Comprehension score •Effect of age •Effect of training	Age	NA	41-Warnings symbol	•14M - 32F, NM •14M - 33F, NM	18-35 yrs 50-67 yrs	NA	NA
Shinar, Dewar, Summala, Zakowska (2003)	Israel, Canada, Finland, Poland	•Comprehension score •Effect of user factors	•Driver type: novice drivers, college students, tourists, problem drivers, older drivers •Age •Country	NA	31-Mixed	4 different countries: Canada, Finland, Israel, Poland, 250 from each country, 25M - 25F for each driver type	Older 65+ yrs	University	>2 yrs, >10 yrs,
Ward, Wogalter, Mercer, (2004)	USA	Effect of training	NA	NA	100-Signs	Nort Carolina, 60M - 40F, licenced	NA	NA	NA

Table 2.2. Summary of studies in literature according to the purpose, factors investigated and sample type (cont.)

Source	Location	Purpose	User Factors	Cognitive Design Features	Signs/symbols analysed (size - type)	Sample (Nationality, Size & Gender, Driving experience)	Age Groups	Education Level Groups	Driving Experi-ence Groups
Ben-Bassat, Shinar, (2006)	Isreal	•Comprehension score •Effect of ergonomic design features	NA	•Competibility, •Familiarity, •Standization	30-Warning, Regulatory, Informational	Israel, 23M - 17F, experienced	NA	NA	NA
Yakut, Master Thesis, (2006)	Turkey	Comprehension score	• Education level, •Gender, •Driving licence type	NA	20-Mixed	Turkish,187M - 23F, NM	NA	•Primary school, •Secondary school, • High school, •University and higher	NA
Koyuncu, Amado, (2007)	Turkey	Effect of sign type	NA	NA	24-Warnings	•Turkish,19M - 20F, NM	NA	NA	NA
Isherwood, McDougall, Curry, (2007)	United Kingdom	Effect of icon design features in speed and accuracy of icon identification	NA	•Concreteness •Complexity •Semantic Distance •Familiarity	40-Icons	4M - 26F, NM	NA	NA	NA

Table 2.2. Summary of studies in literature according to the purpose, factors investigated and sample type (cont.)

Source	Location	Purpose	User Factors	Cognitive Design Features	Signs/symbols analysed (size - type)	Sample (Nationality, Size & Gender, Driving experience)	Age Groups	Education Level Groups	Driving Experience Groups
Ng, Chan, (2007a)	Hong Kong	Effect of sign design features	NA	•Familiarity, •Concreteness •Simplicity, •Meaningfulness •Semantic Closeness	36-Warnings, 30-Prohibition, 25-Mandatory, 13-Guide, 15-Tourist, 1-Road Works	Hongkong Chinese, 19M - 22F, non-experienced	NA	NA	NA
Ng, Chan, (2007b)	Hong Kong	•Effect of user factors •Relation of one sign design feature	•Chaina visit experience •Non-local driving experience	Concreteness NA	120-Signs	Hongkong Chinese, 19M - 22F,non-experienced	NA	NA	NA
					21-Signs	Hongkong, 100M - 9F, experienced	NA	NA	NA
Ng, Chan, (2007c)	Hong Kong	•Guessability score • Effect of sign design features •Effect of prospective user factors	•Potential to be driver •Car racing game experience •Vehicle ownership •Cycling experience •Experience with receiving sign info •Attention to traffic signs •Car accident/ cautioning experience •China visit experience •Gender	•Familiarity, •Concreteness •Simplicity, •Meaningfulnes •Semantic Closeness	36-Warnings, 30-Prohibition, 25-Mandatory, 13-Guide, 15-Tourist, 1-Road Works	Hongkong Chinese,19M - 22F, non-experienced.	NA	NA	NA

Table 2.2. Summary of studies in literature according to the purpose, factors investigated and sample type (cont.)

Source	**Location**	**Purpose**	**User Factors**	**Cognitive Design Features**	**Signs/symbols analysed (size - type)**	**Sample (Nationality, Size & Gender, Driving experience)**	**Age Groups**	**Education Level Groups**	**Driving Experi-ence Groups**
Ng, Chan, (2008a)	Hongkong	Literature review on coginitive and visual features	NA	•Familiarity, •Concreteness • Simplicity, •Meaningfulnes • Semantic Closeness	NA	NA	NA	NA	NA
Ng, Chan, (2008b)	Hong Kong	•Comprehension score •Effect of user factors • Effect of sign design features	•Yrs with license •Yrs active driving •Last time of driving •Driving frequency •Driving hours per trip •Non-local driving experience •Occupation •Education level •Gender and Age	•Familiarity •Concreteness •Simplicity •Meaningfulness	21-Mixed	Hong Kong, 100M - 9F, experienced.	18-27 yrs 28-37 yrs 38-57 yrs	•Below university, •University or above	< 1 yr, 1-2 yrs, 2-4 yrs, 4 + yrs
Chan, Han, Ng, Park, (2009)	South Korea	•Comprehension score •Effect of user factors •Effect of design appropriateness	•Gender, Age, •Education level, •Place of birth, •Length of time living in place of birth, •Experience with the test symbols	Design appropriateness rating	24-Safety symbol	* Hong Kong Chinese, 51M - 30F, NM *Korean, 52M - 8F, NM	< 21 yrs 21-30 yrs 31-40 yrs 41-50 yrs	•Secondary school, •Undergraduate/ College level, •Postgraduate or above	NA

Table 2.2. Summary of studies in literature according to the purpose, factors investigated and sample type (cont.)

Source	**Location**	**Purpose**	**User Factors**	**Cognitive Design Features**	**Signs/symbols analysed (size - type)**	**Sample (Nationality, Size & Gender, Driving experience)**	**Age Groups**	**Education Level Groups**	**Driving Experi-ence Groups**
Kirmizioglu, Master Thesis, (2010)	Turkey	•Comprehension score •Effect of user factors	•Age, •Gender, •Educational background, •Occupation, •Driving license class, •Yrs of driving, •Occupational driving requirement, •Avg. kilometers per yrs, Number of traffic fines in last 5 yrs.	Familiarity	39-Mixed	Turkish, 1134 M - 327 F, experienced.	18-25 yrs 26-35 yrs 36-45 yrs 46-60 yrs 61 + yrs	• High school, •University	NM
Ng, Chan, (2011)	Hong Kong	•Effect of training method •Effect of sign design features	NA	•Familiarity •Concreteness • Simplicity •Meaningfulness •Semantic Closeness	37- Low guessability signs	Hong Kong Chinese, 29M - 7F, non-experienced.	NA	NA	NA

Table 2.2. Summary of studies in literature according to the purpose, factors investigated and sample type (cont.)

Source	**Location**	**Purpose**	**User Factors**	**Cognitive Design Features**	**Signs/symbols analysed (size - type)**	**Sample (Nationality, Size & Gender, Driving experience)**	**Age Groups**	**Education Level Groups**	**Driving Experi-ence Groups**
Kirmizi-oglu, Tuydes-Yaman, (2011)	Turkey	•Comprehension score • Effect of user factors	•Age, •Gender, •Educational background, •Occupation, •Driving license class, •Yrs of driving, •Occupational driving requirement, •Avg. kilometers per yrs, • Number of traffic fines in last 5 yrs.	Familiarity	39-Mixed	Turkish, 1134M - 327F, experienced.	18-25 yrs 26-35 yrs 36-45 yrs 46-60 yrs 61+ yrs	•High school, •University	NM
Liu, Ho, (2012)	Taiwan	•Comprehension score •Effect of sign design features • Lenght of time spent to finish route and the errors made.	Age	•Familiarity, •Concreteness •Simplicity, •Meaningfulnes •Semantic Closeness	39-Directional symbols 39-Directional route scenario	•20M - 10F, NM •18M - 12F, NM	23-30 yrs 65-74 yrs	NA	NA

Table 2.2. Summary of studies in literature according to the purpose, factors investigated and sample type (cont.)

Source	**Location**	**Purpose**	**User Factors**	**Cognitive Design Features**	**Signs/symbols analysed (size - type)**	**Sample (Nationality, Size & Gender, Driving experience)**	**Age Groups**	**Education Level Groups**	**Driving Experi-ence Groups**
Ou, Liu, (2012)	Taiwan	•Comprehension score •Effect of sign design features •Effect of training	Country of origin	• Familiarity, •Concreteness •Simplicity, •Meaningfulness Semantic Closeness	24-Warning, 28-Prohibition, 9-Indicatory, 4-Auxiliary	•Taiwanese,15M - 15F, NM •Vietnamese, 15M -15F, NM	NA	NA	NA
Shinar, Vogelzang, (2012)	Israel	•Comprehension score •Effect of display conditions •Effect of familiarity on the correctness	NA	Familiarity	30-low familiarity signs	26M - 22F, experienced.	NA	NA	NA

Table 2.3. Summary of studies classified according to the factors investigated

		User Factors			
Comprehension score of traffic signs	**Cognitive design features**	**Age**	**Education level**	**Gender**	**Driving experience**
•Al-Madani and Al-Janahi (2002a) •Ng and Chan (2007 b) •Ou and Liu (2012), •Ben-Bassat and Shinar (2006) •Al-Madani (2000) •Ng and Chan (2007 c) • Al-Madani and Al-Janahi (2002b) •Ward *et al.* (2004) •Shinar and Vogelzang (2012) •Ng and Chan (2011) •Ells and Dewar (1979), •Shinar *et al.* (2003) • Kirmizioglu and Tuydes-Yaman (2011) •Ng and Chan (2011) •Ells and Dewar (1979) •Kirmizioglu and Tuydes-Yaman (2011), •Kirmizioglu (2010), •Yakut (2006), •Chan and Chan (2011), •Ng and Chan (2008 b)	•Ng and Chan (2007a) •Ng and Chan (2007b) •Ou and Liu (2012) •Ben-Bassat and Shinar (2006), • Ng and Chan (2007c) •Mc Dougall *et al.* (1999) •Ng and Chan (2011) •Shinar and Vogelzang (2012) •Liu and Ho (2012) •Lin and Kreifeldt (1992) • Kirmizioglu and Tuydes-Yaman (2011) • Chan *et al.* (2009) •Isherwood *et al.* (2007) •Chan and Chan (2011) • Ng and Chan (2008b)	•Al-Madani and Al-Janahi (2002a) • Lesch (2003) • Bruyas *et al.* (1996) • Al-Madani (2000) • Al-Madani and Al-Janahi (2002b) • Liu and Ho (2012) • Shinar *et al.* (2003) • Hancock *et al.* (1999) • Kirmizioglu and Tuydes-Yaman (2011) • Chan *et al.* (2009) • Kirmizioglu (2010) • Chan and Chan (2011) • Ng and Chan (2008b)	•Al-Madani and Al-Janahi (2002a) • Al-Madani and Al-Janahi (2002b) • Kirmizioglu and Tuydes-Yaman (2011) •Chan *et al.* (2009) •Yakut (2006) • Kirmizioglu (2010) • Chan and Chan (2011) • Ng and Chan (2008 b)	•Al-Madani and Al-Janahi (2002a) • Ng and Chan (2007c) • Al-Madani and Al-Janahi (2002b) • Kirmizioglu and Tuydes-Yaman (2011) •Chan *et al.* (2009) •Yakut (2006) • Kirmizioglu (2010) • Ng and Chan (2008b)	•Al-Madani and Al-Janahi (2002a) • Ng and Chan (2007b) • Al-Madani (2000) • Shinar *et al.* (2003) • Kirmizioglu and Tuydes-Yaman (2011) • Kirmizioglu (2010) • Ng and Chan (2008b)

Table 2.4. Summary of studies in literature according to the apparatus used, questionnaire method and results

Source	**Apparatus & Materials used**	**Questionnarie**	**Comprehension / Guessability evaluation**	**Cognitive Design Evaluation Scale**	**Practice Trial, Rest Time, Test Duration**	**Results**
Ells, Dewar (1979)	•Kodak Carousel projector •Shutter and drive	NA	NA	NA	•16 trials •NA •NM	•Mean reaction time for symbolic signs was less than that for verbal signs for both "yes" and "no" responses
	•Vision tunnel illuminated by two 100 Wbulbs.	NA	NA	NA	•I4 trials •NA •NM	• Existing symbolic traffic signs is understood more quickly than the verbal version of them
Lin, Kreifeldt (1992)	AutoCAD	Via computer programme	NA	NA	•NA •NA •NM	•A pictorial icon is the best design choice when information and function have strong ties •If a command is frequently used or strongly related with a common concept, then an abstract icon is a better design choice
Bryne, (1993)	Personal computer	NA	NA	NA	•3 trials • NA •NM	•Simple icons clearly outperform blank and complex icons • Icons must be simple and easily discriminiable for effective search •Complex icons performed worse than blank

*NM- Not mentioned, NA – Not Applied

Table 2.4. Summary of studies in literature according to the apparatus used, questionnaire method and results (cont.)

Source	**Apparatus & Materials used**	**Questionnarie**	**Comprehension / Guessability evaluation**	**Cognitive Design Evaluation Scale**	**Practice Trial, Rest Time, Test Duration**	**Results**
Bruyas, Breton, Pauzie (1996)	•Personal computer	•Open-ended comprehension •Reaction time	% of correct answers	NA	•NA •NA • NM	•Additional elements can disturb the quick understanding of the symbol •Very simple design is also very fare away from the real object •The non- important exstra objects make understanding difficult • The undestanding will be different according to the culture and belonging of the population
Mc Dougall Curry, Brujin (1999)	NA	Sign design features evaluation (face to face)	NA	1 to 5 points scale	•NA •NA •NM	•No significiant relation between complexity and concreteness. Concrete symbols can also be simple. •Close relationship between concreteness and meaningfulness. Concrete signs are more meaningful. •Familiarity and meaningfulness are interchangeable. For familiar items, participants could access meaning even if it is wrong. •Strong relations between concreteness, meaninfulness and familiarity
Hancock, Rogers, Fisk (1999)	NA	Perception, recognition and open-ended comprehension (Online)	2 independent raters evaluated the results as completely incorrect, partially correct, or completely correct	NA	•NA •NA •NM	•Younger adults understood the symbols better than older adults

Table 2.4. Summary of studies in literature according to the apparatus used, questionnaire method and results (cont.)

Source	**Apparatus & Materials used**	**Questionnarie**	**Comprehension / Guessability evaluation**	**Cognitive Design Evaluation Scale**	**Practice Trial, Rest Time, Test Duration**	**Results**
Al Madani (2000)	SPSS,1996	•User factors •Multiple choice comprehension (Online)	% of correct answers	NA	•NA •NA •NM	•High experienced drivers comprehension level is higher than less experienced •No significant relation is found between comprehension and accident involvement, age, experience per accident ratio or speed citiantions •Seat belt usage has significant effect on comprehension.When drivers' awareness of signs increases, seat belt usage increases also
Al Madani, Al Janahi, (2002a)	NA	•User factors •Multiple-choice comprehension (face to face)	% of correct answers	NA	•NA •NA • 30 months	•Education, monthly income and nationality were related with the comprehension •Male drivers scored higher than female •Age, marital status, experience and accident rates had no effect on drivers comprehension •Drivers' personal chracteristics are related with their understanding of signs not with their involvement of accident rates.

Table 2.4. Summary of studies in literature according to the apparatus used, questionnaire method and results (cont.)

Source	**Apparatus & Materials used**	**Questionnarie**	**Comprehension / Guessability evaluation**	**Cognitive Design Evaluation Scale**	**Practice Trial, Rest Time, Test Duration**	**Results**
Al Madani, Al Janahi (2002b)	NA	•User factors •Multiple-choice comprehension (face to face)	% of correct answers	NA	•NA •NA •30 months	•Age, gender, education and income played major roles in comprehension of signs, whereas marital status showed no significant effect. • Comprehension of signs generally increases with age. Drivers in younger age group (16-24 yrs) comprehend significantly less well than the older groups (35-44 yrs & over 44 yrs) •Comprehension of signs increases with level of education • Age and experience are highly correlated. Experience has no significant influence in improving comprehension of signs at the age of 45 or over •Male drivers understand traffic signs significantly better than female ones
Shinar, Dewar, Summala, Zakowska (2003)	Cardboards	•User factors •Open-ended sign comprehension (face to face)	Correct & complete (+2 point), Partially correct (+1 point), Incorrect (0 point), Opposite (-2 point)	NA	•NA •NA •NM	•Large difference in comprehension level among different countries is observed •Signs were comprehended best when they were consistent with ergonomic guidelines •Some of the signs were misinterpreted as opposite of their true meaning •Performance of elderly drivers was lower than other drivers, and had relatively more opposite responses than younger drivers

Table 2.4. Summary of studies in literature according to the apparatus used, questionnaire method and results (cont.)

Source	**Apparatus & Materials used**	**Questionnarie**	**Comprehension / Guessability evaluation**	**Cognitive Design Evaluation Scale**	**Practice Trial, Rest Time, Test Duration**	**Results**
Lesch, (2003)	•Personal Computer •E-Prime software program (presentation&record)	•User factors •Yes-No choice comprehension (online)	% of correct answers	NA	•3 trials •NA •2nd test- 1 weeks later from pre-test, 3rd test - 6/8 weeks later than pre-test	•Older participants demonstrated poor comprehension of warning symbols than younger participants •Age did not interact with test section or training condition •Olders have difficulty to understand warning signs, training is effective for both younger and older •Explanatory statement provide greater training benefit •Younger and olders tend to differ in terms of their perceptual and cognitive abilities.
Ward, Wogalter, Mercer (2004)	Cardboards	Open-ended comprehension	% of correct answers	NA	•NA •NA •NA	•The training helped decrease the open red circle critical confusions from 39% to 6% •U.S. drivers do not comprehend a relatively large number of the international road signs •Training of the sign meanings might counteract low comprehension and high critical confusions

Table 2.4. Summary of studies in literature according to the apparatus used, questionnaire method and results (cont.)

Source	**Apparatus & Materials used**	**Questionnarie**	**Comprehension / Guessability evaluation**	**Cognitive Design Evaluation Scale**	**Practice Trial, Rest Time, Test Duration**	**Results**
Ben-Bassat, Shinar (2006)	•Personal computer •Mouse & keyboard •Microsoft Visual Basic 5.0 programme	•User factors •Open-ended comprehension	Correct & accurate (+2 point), Partially correct (+1 point), Opposite (0 point), Wrong (-2 points)	•Design features in 1 to 10 scale •Experts evaluated compitability in 0-3 scale •Author evaluated standardization in 0-3 scale	•NA •NA •NM	•High compatibility has high correlation with comprehension •Familiarity and comprehensibility have significiant correlation alone on comprehension whereas standardization has not but they are more stronger together •Ergonomically designed traffic signs are more understandable •Standardization is very important in order to avoid culturel bias • Some signs guessed as opposite meaning which is very dangerous and results with deadly outcomes on the road
Yakut, Master Thesis (2006)	NA	•User factors •Multiple-choice comprehension (face to face)	5 points for each correct answer, Comprehension score: % of correct answers	NA	•NA •NA •NA	•Only 81% of the subjects comprehended traffic signs correctly.
Koyuncu, Amado (2007)	•Personal computer •Keyboard •Microsoft Visual Basic 5.0 programme	NA	NA	NA	•5 trials •NA •5 min	•The fastest response time were in written+symbolic stimuli • Duration of exposure to traffic sign affected response time • Location did not have any significant effect on priming.

Table 2.4. Summary of studies in literature according to the apparatus used, questionnaire method and results (cont.)

Source	**Apparatus & Materials used**	**Questionnarie**	**Comprehension / Guessability evaluation**	**Cognitive Design Evaluation Scale**	**Practice Trial, Rest Time, Test Duration**	**Results**
Isherwood, McDougall, Curry (2007)	•Personal computer •Mouse	Via computer programme	NA	1 to 5 point scale	•NA •NA •Apprx. 5 min	•Semantic distance is found important initial step where icon-function relationships are learned, familiarity is important later since it has lasting effects on access to long-term memory •Icon concreteness does not have significant effect on identification. Semantic distance and familiarity may be more important.
Ng, Chan, (2007a)	•The-Pesudo-Isochrometric Plates •Minolta Luminance Meter (LS-11) •Personal Computer •Mouse •Comfortable chair •Powerpoint	Sign design features evaluation (face to face)	NA	0 to 100 points scale	• 5 trials • 1-min rest after every 40 signs • 1 hour	•Need further study for experienced drivers •Meaningfulness is strongly correlated with familiarity •Significant relation with familiarity and semantic closeness observed •Concreteness is significantly associated with familiarity •Simplicity did not correlated with any of the factors significantly.

Table 2.4. Summary of studies in literature according to the apparatus used, questionnaire method and results (cont.)

Source	Apparatus & Materials used	Questionnarie	Comprehension / Guessability evaluation	Cognitive Design Evaluation Scale	Practice Trial, Rest Time, Test Duration	Results
Ng, Chan, (2007b)	NA	•User factors •Sign design feature evaluation	NA	0 to 100 points scale	•NA •NA •NM	•Concrete signs contributed to higher guessability scores than abstract ones •Concreteness influenced guessing performance
	NA	•Use factors •Multiple-choice comprehension (face to face)	% of correct answers	NA	•NA •NA •NM	•Usability in international markets can be increased by the use of culture independent icons •If culture specific graphics are used in a traffic sign, supplementary text may be needed • Non-local driving experience has a positive effect on local sign comprehension when the signs look alike and convey same meaning, otherwise it has not.
Ng, Chan (2007c)	• The-Pesudo-Isochrometric Plates • Minolta Luminance Meter (LS-11) •Personal Computer • Mouse •Comfortable chair •Powerpoint •Voice recorder (LG VP-998)	•User factors •Sign design features evaluation (face to face)	Fully correct (+2 points), Partially correct (+1 point), Incorrect (0 point)	0 to 100 points scale	•5 trial •1 min rest after 42 signs •1st test 40 min., evaluation test 1 hour (1 month later)	• Cycling experience, experience with receiving sign info, gender were not significiant predictors of quessing performance •Mainland China experience has an effect on guessing performance •Icons should be designed culture independent, or supported with text. Vehicle avaibility in familiarity has negative effect on quessing performance. Attention to traffic signs had significiant effect on quessing performance •Guessability of signs were better in signs that are familiar, concrete, simple, meaningful, associate with underlying concepts • Contribution of 5 sign design features on guessability is different. Semantic closeness is the best predictor .

Table 2.4. Summary of studies in literature according to the apparatus used, questionnaire method and results (cont.)

Source	**Apparatus & Materials used**	**Questionnarie**	**Comprehension / Guessability evaluation**	**Cognitive Design Evaluation Scale**	**Practice Trial, Rest Time, Test Duration**	**Results**
Ng, Chan (2008a)	NA	NA	NA	NA	•NA •NA •NA	•Color, shape and size of icons are important for icon effectiveness •Although strong relationship between familiarity, concreteness, meaningfulness and semantic closeness, no significiant interaction of simplicity •To evalute simplicity, subjective rating is the most comprehensive approach than metric and automated measurment •0-100 point scale is prefered since it makes data more amenable than 0-5 point scale
Ng, Chan (2008b)	NA	•User factors •Sign design features evaluation •Multiple-choice comprehension (face to face)	Correct (+1 point), Two of one answer correct (+0.5 point), No answer (+0.25 point), Incorrect (0 point)	0 to 100 points scale	•Pilot testing •NA •NM	•No difference of age is found on comprehension performance •University or above education performed better in sign comprehension •Comprehension performance is found better during the first year of holding a driving license and decreased with years of being a licensed driver •Familiarity is significantly correlated with comprehension score •Years of active driving, hours of driving and last time of driving do not have any relationship with comprehension performance

Table 2.4. Summary of studies in literature according to the apparatus used, questionnaire method and results (cont.)

Source	**Apparatus & Materials used**	**Questionnarie**	**Comprehension / Guessability evaluation**	**Cognitive Design Evaluation Scale**	**Practice Trial, Rest Time, Test Duration**	**Results**
Chan, Han, Ng, Park (2009)	NA	•User factors •Multiple-choice comprehension •Sign design feature evaluation (face to face)	Accurate (+1 point), Inaccurate(0 point)	7 points Likert scale	•NA •NA • NM	•Guessing performance was not related to gender •Guessability score did not correlate with design appropriateness ratings.
Kirmizioglu, Master Thesis (2010)	NA	•User factors •Open-ended Comprehension (face to face)	Correct(+1point), Partially correct (+0,5 point), No comment (0 point), Wrong (-0,5 point), Opposite (-1 point)	percentage of drivers reported to see them in traffic	•NA •NA •30 min	•Female drivers gave more correct and less wrong or opposite responses than male drivers • Highly educated drivers gave more successful responses than other groups •Public bus drivers gave the highest wrong responses •Drivers older than 46 yrs comprehended signs more correctly, aged between 18-25 yrs gave more wrong and not commented, 26-35 yrs drivers gave more opposite responses •Driving experience does not have significant effect on comprehension • Previous version of traffic signs is mostly comprehended correctly but new version (EU accession) is not learnt well • Drivers do not have enough knowledge about color and shape of a sign

Table 2.4. Summary of studies in literature according to the apparatus used, questionnaire method and results (cont.)

Source	**Apparatus & Materials used**	**Questionnarie**	**Comprehension / Guessability evaluation**	**Cognitive Design Evaluation Scale**	**Practice Trial, Rest Time, Test Duration**	**Results**
Kirmizioglu, Tuydes-Yaman (2011)	NA	•User factors •Open-ended Comprehension (face to face)	Correct(+1 point), Partially correct (+0,5 point), No comment (0 point), Wrong (-0,5 point), Opposite (-1 point)	Percentage of drivers reported to see them in traffic	•NA •NA •30 min	•The most of the signs were not known well by the drivers. •5 of the signs were oppositely comprehended by more than 10% of the drivers
Ng, Chan (2011)	• Pseudo-isochromatic plates •A Minolta luminance meter (Konica Minolta LS-11) •Personal computer •Voice recorder (LG VP-988) • Powerpoint • SPSS software	•Training evaluation •Sign design features evaluation (face to face)	NA	0 to 100 points scale	•2 trials •NA •2 experiments in 2 days	•The 5 sign features did not have a significant influence on training effectiveness but sign is perceived more meaningful after learned its intended meaning • Training improves the ability to comprehend traffic signs • Recall training is preferable to paired-associate learning and recognition training •The effect of semantic closeness in sign effectiveness and ease of interpretation is found to be long lasting

Table 2.4. Summary of studies in literature according to the apparatus used, questionnaire method and results (cont.)

Source	**Apparatus & Materials used**	**Questionnarie**	**Comprehension / Guessability evaluation**	**Cognitive Design Evaluation Scale**	**Practice Trial, Rest Time, Test Duration**	**Results**
Chan, Chan (2011)	NA	•User factors •Sign design features evaluation •Multiple-choice comprehension (face to face)	Incorrect(0 point), Partially correct (+1 point) , Correct answer (+2 points)	0 to 100 points scale	•Pilot testing •NA •30 min	•No difference was found in comprehension performance amongst different age groups •Education did not have significant effect on comprehension •Comprehension ability of participants was high for familiar signs • Concrete signs produced higher comprehensibility scores than abstract ones. •Simple signs led to higher comprehensibility score than complex signs. •The comprehensibility scores were high for meaningful signs
Shinar, Vogelzang , (2012)	•Personel computer •Keyboard	Open-ended Speed and accuracy of comprehension (face to face)	Correct (+2points) Partially correct (+1 point), Incorrect (0 point) Opposite (-2 points)	NA	•NA •NA •15 min	•Symbolic displays are beneficial only when the symbol is familiar to the driver •Text signs better comprehended, the comprehension time (based on the reaction time) was shorter for the text signs than corresponding symbolic version of them

Table 2.4. Summary of studies in literature according to the apparatus used, questionnaire method and results (cont.)

Source	**Apparatus & Materials used**	**Questionnarie**	**Comprehension / Guessability evaluation**	**Cognitive Design Evaluation Scale**	**Practice Trial, Rest Time, Test Duration**	**Results**
Liu, Ho (2012)	•Nikon •Adobe Photoshop •A micro computer •Flash software •PowerPoint	Sign design features evaluation (face to face)	Correct & complete (+2 points), Partially correct (+1 point), Incorrect (0 point)	0 to 100 points scale	•5 trials for sign features, •3 trials for simulation • NA • NM	•Older adults experienced greater difficulty in understanding some symbols as compared to younger adults. •Familiarity is the most highly correlated features with comprehension •Accuracy of semantic depiction is the best predictor of behavior in routes •Five symbol features significantly influence comprehension of symbols and performance
Ou, Liu, (2012)	•Ishiara color blindness test •Cardboards •SPSS software	Sign design features evaluation (face to face)	Correct (+1 point), Very close (+0,75 point), Close (+0,5 point), Opposite (0 point)	0 to 100 points scale	•5 trials •5 min. rest in the middle of experiment •1st test 60 min, 2nd test 20 min, 3rd test after 1 month 20 min	•Familiarity will improve comprehension •Subjects' culturel background is critical factor for sign recognition •Training can significantly support sign comprehension •Semantic closeness has significant correlation with comprehension, followed by meaningfulness, concreteness and familiarity and simplicity •Higher scores for 5 sign design features do not only lead the higher comprehension but also lower forgetfulness.

Table 2.5. Summary of studies classified according to the questionnaire type, design and evaluation method

Comprehension Questionnaire		**Comprehension Evaluation**			**Sign Design Evaluation**				**Questionnaire Design**	
Multiple-choice	**Open-Ended**	**% of correct answer**	**Scoring**	**Expert evaluation**	**0 to 5 points**	**0 to 10 points**	**1 to 7 points**	**0 to 100 points**	**Practice Trial**	**Rest Time**
•Al-Madani and Al-Janahi (2002a) •Ng and Chan (2007 b) •Al-Madani (1999) •Al-Madani and Al-Janahi (2002b) •Chan, Han, Ng, Park (2009) •Yakut (2006) •Chan and Chan (2011) •Ng and Chan (2008b) • Lesch (2003)	•Bruyas *et al*. (1996) •Ben-Bassat and Shinar (2006) •Ward *et al*. (2004) •Shinar and Vogelzang (2012) •Shinar *et al*. (2003) •Hancock *et al*. (1999) •Kirmizioglu and Tuydes-Yaman (2011) • Kirmizioglu (2010)	•Al-Madani and Al-Janahi (2002a) •Lesch (2003) •Ng and Chan (2007b) •Bruyas *et al*. (1996) •Al-Madani (2000) •Al-Madani and Al-Janahi (2002b) • Yakut (2006)	•Ou and Liu (2012) •Ben-Bassat and Shinar (2006) •Ng and Chan (2007c) •Shinar and Vogelzang (2012) •Liu and Ho (2012) •Shinar *et al*. (2003) •Kirmizioglu and Tuydes-Yaman (2011) •Chan *et al*. (2009) •Kirmizioglu (2010) •Chan and Chan (2011) •Ng and Chan (2008b)	•Hancock *et al*. (1999) • Ben-Bassat and Shinar (2006)	•Mc Dougall, *et al*. (1999) •Isherwood *et al*. (2007)	•Ben-Bassat and Shinar (2006)	•Chan *et al*. (2009)	•Ng and Chan (2007a) •Ng and Chan (2007b) •Ou and Liu (2012) •Ng and Chan (2007c) •Ng and Chan (2011) •Liu and Ho (2012) •Chan and Chan (2011) •Ng and Chan (2008b)	•Ng and Chan (2007a) •Lesch (2003) •Ou and Liu (2012) •Koyuncu and Amado (2007) •Bryne (1993) •Ng and Chan (2007c) •Ng and Chan (2011) •Liu and Ho (2012) •Ells and Dewar (1979) •Chan and Chan (2011) •Ng and Chan (2008b)	•Ng and Chan (2007a) •Ou and Liu (2012) •Ng and Chan (2007c)

3. RATIONALE AND OBJECTIVES OF THE STUDY

3.1. Rationale behind the study

Traffic signs are the main tools to regulate, warn and guide the drivers. Therefore, the correct comprehension of them is crucial for safe driving and effective control on the roads. In spite of their importance, they are not always very well understood by the drivers. The well designed traffic signs are necessary to make their messages more clearly for the drivers, which are essential to improve the comprehension and reduce the road traffic accidents and its consequences (i.e., fatalities, injuries and damages). One of the critical success factors in correct comprehension of traffic signs is to design them consistent with the ergonomic guidelines (Shinar *et al.*, 2003).

There are a few studies on comprehensibility of traffic signs in the literature. McDougall *et al.* (1999) emphasized the importance of cognitive design features on effective sign design and defined five features as criteria for measuring the effect of sign design. In addition to this, there have been a number of studies in different countries throughout the world, which had investigated the effect of user factors, such as gender, age, driving experience, marital status, on comprehension of traffic signs. The studies in the literature have investigated either only the effect of cognitive design features on traffic signs considering one specific user factor or only the effect of user factors on comprehension. None of them focused on both together. Moreover, in the literature there is no such a study that analyzes the effect of user factors on the evaluation of cognitive sign design features.

Since, the cultural differences among different nationalities lead to variation in comprehension, it is crucial to evaluate the information for the populations of different nationalities. However, there have not been any studies conducted in Turkey to evaluate the comprehensibility of current traffic signs by considering cognitive sign design features and the effect of user factors. Therefore, this study would be useful to understand the

comprehension levels of current traffic signs in Turkey and the effect of cognitive sign design features and user factors on it.

This study combines the mentioned factors and discusses both cognitive design features and user factors effect on comprehension to understand the comprehensibility of current Turkish traffic signs by Turkish population. Moreover, this study will be the first to evaluate the comprehensibility of Turkish traffic signs considering both cognitive design features and user factors in Turkey. Also, this study will be the first study in the literature worldwide that evaluates the effect of user factors on evaluation of sign design features in a single study. It is believed that the results of this study will be useful for the designer of road signs to design more user-friendly traffic signs that transmit clear message to drivers and support the safe driving.

3.2. Objectives of the study

Based on the rationale above, this study aims the following:

(i) Investigate comprehension level and guessability score of current Turkish traffic signs;
(ii) Investigate the effect of five cognitive sign design features: familiarity, concreteness, simplicity, meaningfulness, and semantic closeness on guessability score of traffic signs;
(iii) Investigate the effects of user factors such as age, gender, education, marital status, active driving experience on guessability score of traffic signs;
(iv) Investigate the effect of user factors on evaluation of cognitive sign design features;
(v) Determine the interrelationship between the user factors, cognitive sign design features and the guessability score of traffic signs.

4. METHODOLOGY

To gather the needed information to achieve the objectives of the study, an on-screen power point presentation of the traffic signs, a user factors questionnaire for gathering the participants' user factors, a guessability questionnaire to evaluate guessability score and a sign design evaluation sheet for participants to give the subjective ratings for five cognitive sign design features were developed. Details of the questionnaires are given in Appendix A.

4.1. Participants

100 male and 101 female volunteers, who were selected from the population of Turkey, voluntarily participated in this experiment. Almost all participants were recruited from İstanbul, the metropolitan city of Turkey. It is assumed that the population of İstanbul represents the general population of Turkey since its population is composed of people from all regions of Turkey.

In order to ensure a balanced distribution amongst the seven regions of Turkey, information of the birthplace, family origin city, mother's and father's birthplace was asked to the participants. The distribution family origins of the participants are shown Figure 4.1.

The minimum age to be able to get a driving license is 18 years in Turkey. Therefore, the recruitment in the study was done considering this age restriction. The age of the participants were between 18 and 78 years (average= 41,.9 yrs, median= 41 yrs). In order to analyze the effect of age on comprehension, participants were stratified into 5 age groups; 18-29 years, 30-39 years, 40-49 years, 50-59 years, 60-69 years. In this study, we wanted to include the 70-79 years age group, but they had a difficulty to understand the meaning of the sign design features. Therefore, that age group was excluded from the study. Participants were classified according to the educational background considering the last graduated/continuing school in 3 groups; primary/secondary school, high school,

university or higher degree. Both driver license holders and those without driver license were included in the study. Participants who had the driver license were asked to report their active car driving experience to evaluate the effect of experience on comprehension. Active car driving experience were evaluated in four categories; not an active driver, less than or equal to 1 year, more than 1 year and less than 5 years and equal to or more than 5 years. Participants reported their occupations also.

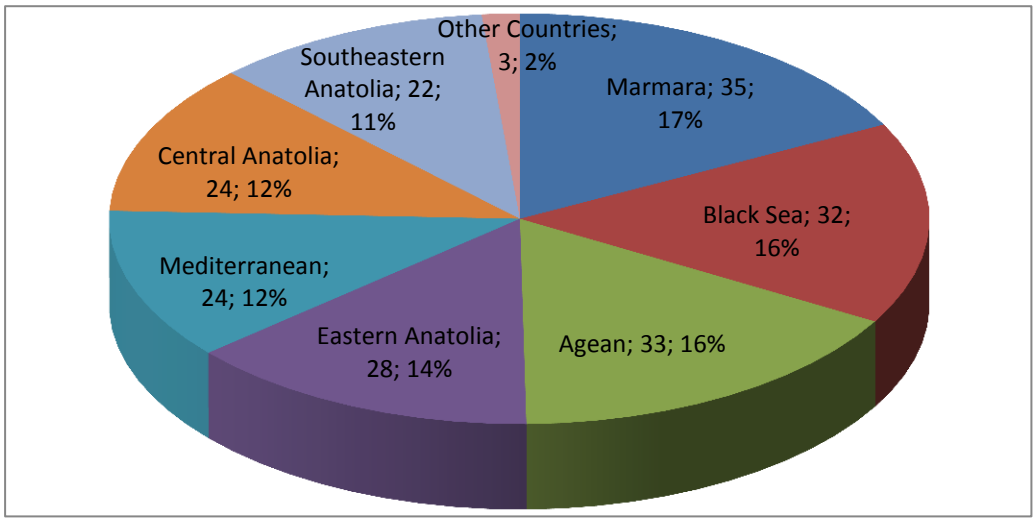


Figure 4.1. The distribution of family origins of the participants. (1st number in figure shows quantity of participants per region, 2nd number shows distribution in whole sample)

It is assumed that the behavior of drivers can change with the family responsibility. Moreover, it is know that smoking is highly damaging not only the physical health but also the performance of brain and attention. Therefore, to observe the effect of marital status and smoking habits on comprehension performance, participants were asked to report their status for those two factors. Marital status was evaluated in 3 categories; single, married, married with children. Smoking habits were reported in yes or no choices.

All participants had to be no color deficiency and be free from health disorders; therefore participants declared they were not suffering from color deficiency before starting the experiment. In addition to a color deficiency form (Color Deficiency Declaration Form – Appendix A) to discover the color deficiency situation of participants

were applied to the participants before the experiments. The participants, who had color deficiency were excluded from the experiments.

Number of participants, according to the gender, age, educational background, holding driver license or not, active car driving experience, marital status and smoking habits are shown in Table 4.1. The occupation distribution of participants is listed in Table 4.2.

Table 4.1. Participant categorization by gender, age, educational background, holding driver's license or not, active car driving experience, marital status and smoking habits

User Factors	**Levels**	**Number of Participants**	
		Male	**Female**
Gender	(Male, Female)	**100**	**101**
Age	18-29 years	20	21
	30-39 years	20	23
	40-49 years	20	20
	50-59 years	21	20
	60-69 years	19	17
Marital Status	Single	30	37
	Married	20	19
	Married with children	50	45
Educational Background	Primary / Secondary School	36	31
	High School	33	33
	University or higher degree	31	37
Driver Licence Holder	Yes	78	78
	No	22	23
Active Car Driving Experience	Not an active car driver	23	30
	Less than or equal to 1 year	25	28
	More than 1 year and less than 5 years	25	24
	More than or equal to 5 years	27	19
Smoking Habits	Yes (Smoker)	42	30
	No (Non-Smoker)	58	71

Table 4.2. Occupation distribution

Accountant (4)	Retired (Civil servant) (3)
Assistant (Research) (1)	Retired (Doctor) (1)
Bank Employee (6)	Retired (Lawyer) (1)
Business Manager (2)	Retired (Pharmasist) (1)
Civil Servant (13)	Retired (Teacher) (6)
Clothier (1)	Retiree (15)
Coiffeur (4)	Sales Representive (Drugs) (4)
Doctor (2)	Salesperson (2)
Driver (Taxi) (3)	Secretary (4)
Driver (Truck) (5)	Self-employed (7)
Engineer (2)	Seller of Furniture (4)
Engineer (Civil) (1)	Shopkeeper (3)
Engineer (Industrial) (1)	Specialist (Social Service) (1)
Engineer (Machine) (1)	Specialist (Customer Relationship) (1)
Engineer (Enviromental) (1)	Specialist (Finance) (1)
Graphic Designer (1)	Specialist (Human Research) (1)
Grocery (1)	Specialist (Information Technology) (1)
Guide (1)	Specialist (Market Research) (1)
Hostess (1)	Specialist (Marketing) (1)
Hotel Manager (1)	Specialist (Public Relations) (1)
House Painter (1)	Specialist (Sales) (1)
Housewife (29)	Student (12)
Insurance Agent (1)	Tailor (1)
Interior Architect (1)	Teacher (8)
Human Research Manager (1)	Technician (1)
Pastry seller (1)	Technician (Chemistry) (1)
Personnel Manager (1)	Technician (Electrics) (1)
Pharmasist (1)	Technician (Machine) (1)
Police (1)	Technician (Textile) (1)
Ready-Made Seller (2)	Tecnical Service Personnel (1)
Real Estate Dealer (1)	Translator (1)
Retired (Bank employer) (3)	Waiter (5)
	Worker (13)

4.2. Materials

The information about the materials and tools used to investigate the guessability score of traffic signs, the effect of user factors and the sign design features on guessability are presented in this section.

4.2.1. Apparatus

A laptop computer (HP Elite Book 820p) with 2.30 GHz processor and 12.5-inch HD monitor was used for the experiment. An adjustable chair was provided for the participant to complete experiment in a ergonomically comfortable position. To present the traffic signs Microsoft PowerPoint programme was used. Self-pased presentation slides of traffic signs were prepared in random sequence to minimize the order effect. A computer mouse was used by the experimenter to control the flow of the presentation.

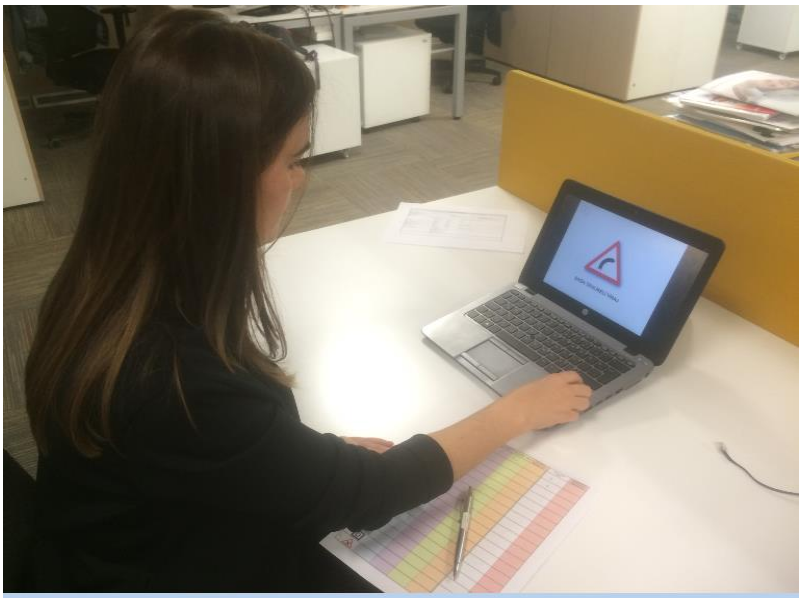


Figure 4.2. Scene from the experiment with experimental apparatus

4.2.2. Traffic Signs

Identifying the traffic signs to be used in the experiment was the first step to develop the survey materials. Similar to the method of Ng and Chan (2007a, b, c; 2008b, 2011),

two criteria were used to decide on the signs to be included in study: first, the signs should convey their messages with symbols only; second, they should not be used in accompaniment with other signs to give the message. 230 traffic signs (53 warning signs, 64 regulatory signs, 82 informational signs, 7 standing and parking signs, 12 road works and maintenance signs, 12 types panel signs) in total are currently used in Turkey (General Directorate of Highway, 2004). 146 (41 warning signs, 50 regulatory signs, 45 informational signs, 7 standing and parking signs, 3 road works and maintenance signs) of 230 Turkey traffic signs satisfy the above mentioned two criteria to be chosen. A long questionnaire is not preferred since it can decrease the attention and motivation of participants, so we carefully selected 60 of 146 traffic signs considering the criteria that sign conveys their message with symbol only and should not be used in accompaniment with other signs. 22 warning signs, 21 regulatory signs, 13 informational signs, 3 standing and parking signs, 1 road works and maintenance signs were used in the experiment that are shown in the Table 4.3.

4.2.3. User factors, guessability questionnaire and cognitive sign design features evaluation sheet

To gather the user factors, a questionnaire with seven questions was designed. Questionnaire began with the participant's information regarding name and surname, date of birth, birthplace, family origin city, mother's and father's birthplace and occupation. That information was used to ensure the balanced distribution of sample. The questionnaire continued with the user factors questions. The first question was about the age. Age was divided in 5 groups (18-29 yrs, 30-39 yrs, 40-49 yrs, 50-59 yrs, 60-69 yrs) and participants were asked to tick their age group and write their exact age in given area. In the second question, participants had to choose their gender type and tick "female" or "male" boxes. Question three was about marital status. Participants are required ticking "single", "married" and "married with children" boxes. In question four, participants were reported their smoking habits with the question "Do you smoke cigarettes?" They had to tick "yes" or "no" boxes. Fifth question was about educational background. Participants required ticking their last graduated or continuing degree amongst the boxes that were "primary or secondary school", "high school" and "university or higher degree". In question six,

participants were asked the question “Do you hold a driver license?” and they had to tick “yes” or “no” boxes.

In addition, they asked to write the type of driver license. In the last question, participants were asked about their active car driving experience. Participants had to tick “not an active driver”, “drive less than or equal to 1 year”, “drive more than 1 years and less than 5 years”, or “drive more than or equal to 5 years” boxes.

Traffic signs were presented on the computer screen using Microsoft PowerPoint program in random order. All the signs were fitted into 10 cm x10 cm squares without border and placed at the centre of the computer screen, at a viewing distance of about 60 cm (Ng and Chan, 2007a). In order to make easier to fill the questionnarie, the numbers that was given to each sign in “Guessability Questionnarie” were used in the presentation together with the sign. The verbal labes which was used when rating the semantic closeness were received from the General Directorate of Highway (2004).

Guessability questionnaire was prepared to evaluate their understanding of traffic signs. First of all, traffic signs (with the number given in guessability questionnaire) without verbal labels were presented on the screen using Microsoft Power Point program in random order and then participants were asked the guess of the meaning of traffic sign and write the meaning of the sign, in their own words, in the given text box area on guessability questionnaire. Those signs were displayed on sign design features evaluation sheet in color and in squares.

A cognitive sign design evaluation sheet was prepared for gathering the subjective ratings of participants. Participants were asked to give subjective ratings between 0 to100 points scale, similar to the methodology used by Ng and Chan (2007a, b, c; 2008b; 2011), (for familiarity (0 point = very unfamiliar, 100 points = very familiar), simplicity (0 point = very complex, 100 points= very simple), concreteness (0 point= very definitely abstract, 100 points= very definitely concrete), meaningfulness (0 point= very definitely meaningless, 100 points= very definitely meaningful), semantic closeness (0 point =very weakly related, 100 points= very strongly related). In subjective ratings 0 to 100 points scale was used since it was found more prefered by the users to express their feelings and

more effective to use in several kind of statistial anaysis. Gulta *et al.* (2001) showed that using higher number of scale points makes data easy to use in various kinds of statistical analysis. Preston and Colman (2004) mentioned that shorter scales are found easy to use by people, however longer scales are prefered by the people since they allow to express their feelings more adequately.

The user factors and guessability questionnarie, and sign design features evaluation sheet are shown in Appendix A.

Table 4.3. Turkey traffic signs used in the experiment (General Directorate of Highway, 2004)

No	Type	Traffic Sign	Meaning	No	Type	Traffic Sign	Meaning
1	W1		Dangareous curve to right	11	R11		No entry for motor vehicles except motorcycles
2	W2		School crossing	12	W14		Bridge begins
3	RM1		Road work	13	R12		No right turn
4	R1		No U-turn	14	R13		Mandatory direction for vehicles carrying dangerous goods
5	W3		Steep hill downwards	15	I6		Maintanance
6	W4		Uneven road	16	P2		Parking area
7	R2		Mandatory right direction	17	R14		No overtaking
8	W5		Traffic lights	18	R5		End of all restriction and prohibitions
9	R3		Motor vehicles prohibited	19	W15		Road narrows from right
10	W6		Pedestrian crossing ahead	20	W16		Drawbridge

*W=warning signs, R=regulatory signs, I=informational signs, P= standing and parking signs, RM= road works and maintenance signs

Table 4.3. Turkey traffic signs used in the experiment (General Directorate of Highway, 2004) (cont.)

No	Type	Traffic Sign	Meaning	No	Type	Traffic Sign	Meaning
21	W7		Slippery road	31	R16		No entrance fo vehicles carrying water-polluting cargo
22	R4		No entry	32	R17		No audible warning device
23	W8		Side winds	33	R18		End of pedestrians only
24	W9		Airport (low-flying aircraft)	34	I7		Hospital
25	R15		End of obligation to put on snow chains	35	W17		Road ends with sea or river
26	I1		Highway begins	36	R19		Closed to all vehicles in both directions
27	I2		Stop	37	W18		Risk of falling rocks
28	W10		Uncontrolled intersection	38	I8		Pedestrian crossing
29	W11		Entrance of priority road from left	39	R20		No handcarts
30	R6		Priority over oncoming vehicles	40	R21		No entry for the vehicles carrying hazardous cargo

Table 4.3. Turkey traffic signs used in the experiment (General Directorate of Highway, 2004) (cont.)

No	Type	Traffic Sign	Meaning	No	Type	Traffic Sign	Meaning
41	R7		End of overtaking restriction for trucks	51	W19		Danger
42	I3		Pedestrian underpass	52	I9		Do not swim
43	I4		No entry road junction	53	I10		Hotel
44	W12		Soft verge	54	I11		End of priority road
45	W13		Risk of ice	55	W20		Cycle route ahead
46	R8		No entry for vehicles carrying explosives	56	W21		Domestic animals
47	R9		Roundabout	57	P3		No standing or parking
48	P1		No parking	58	W22		Roundabout
49	I5		Tourism information	59	I12		Start road of motor vehicles
50	R10		Give way	60	I13		Youth camp

4.3. Testing procedure

Before starting the experiment, a pilot testing was conducted to estimate the questionnaries' completion time and to understand the appropriateness of the questions with the purpose of the study and whether they had correct wordings. Participants filled in the user factors sheet first, guessed the meaning of 60 traffic signs secondly and then gave ratings for the sign design features. Completion of the whole experiment in a one day took almost 2 hours and the attention of participant started to decrease towards the end of the experiment. Therefore, we decided to perform whole experiment in two different days with 1 day break was given between the experiments to minimize the effect of fatique.

201 participants joined the experiment volantarily. Each participant completed the experiment individually in two section on two different days. In the first day participants filled in user factors sheet and guessed the meaning of the trafic signs. Secondly, 1 day after the first experiment, participants gave ratings for each sign design features.

At the begining of the experiment, a brief description of the objectives and requirements of the study were given to each candidate. All participants filled the "Color Deficiency Declaration Form", and only the participants who reported had no color deficiency were included in the experiment. And then, the participants signed the "Personal Consent Form" to accept that he/she was voluntarily participated the experiment. This form gave a detailed explanation regarding the purpose and the procedure of the experiment. And also, it was included in the form that the information gathered during the experiment would be kept confidential. Both English and Turkish version of the "Color Deficiency Declaration Form" and "Personal Consent Form" were given in Appendix A.

During the questionnaire, in order to increase the reliability and validity of the data and also the confidence of the participants to give more proper responses, it is emphasized that the main purpose of the questionnaire was to understand the comprehensibility level of the current traffic signs by the Turkish population. This questionnaire was not an exam and not interested with the personal results or identity. Names and any other personal information would not be used in results. Participants were more willing to participate after hearing all this explanation.

4.3.1. First day: user factors and guessability questionnaire

The steps of the methodology in the first day were as follows:

- Participants filled the "User Factors Questionnaire" that took about in 5 minutes to fill out first.
- After that, each participant was briefed on the objectives of the guessability questionnarie and given verbal instructions at the beginning of the experiment.
- 60 Turkey traffic signs were presented in random order on the computer screen without verbal labels and than participants were required to quess the meaning of the each presented sign. After guessing, participants wrote the the meaning of the sign, in their own words, to the given area on "Guessability Questionnaire". To make easier to find the presented sign on "Guessability Questionnaire", sings were presented together with the numbers given in the "Guessability Questionnaire".
- There was no time pressure to complete the survey. No feedback was given during the experiment. To avoid fatigue, 1 minute break was given after 30 signs answered. Whole experiment took around 45 minutes.

4.3.2. Second day: sign design features evaluation questionnaire

Sign design features evaluation experiment was conducted 1 day after the first experiment.

- At the beginning, subjects were briefed with the rating instructions and the meanings of five sign design features; familiarity, concreteness, simplicity, meaningfulness and semantic closeness (Appendix A).
- Before starting the experiment, practice trial with 5 Turkey traffic signs, which were not included in the questionnaire, were done to ensure the methodology and the meaning of the design features were well understood by the participants.
- 60 Turkey traffic sign were randomly presented on computer screen. At the first step, participants were required to give subjective ratings, in 0 to 100 point scale, for familiarity, concreteness, simplicity and meaningfulness on "Sign Design Feature Evaluation Questionnaire" for each sign.

- After evaluation of four sign design features, the verbal label was given under the sign and participants gave the rating for the fifth feature, semantic closeness. And then both verbal label and sign disappeared.

The same process mentioned above was repeated until the ratings of all 60 traffic signs were completed. In order to avoid fatigue, 1 minute break was given after 30 signs were finished. There was no time pressure during the experiment. Whole experiment was completed about between 60 minutes to 100 minutes. No feedback was given during experiment.

4.3.3. Evaluation of the guessability questionnaire

Shinar *et al*. (2003), coded the guessability response of the each sign by using the four categories: Correct and complete (+2 points), partially correct (+1 point), incorrect (0 point), or opposite of the true meaning (-2 points). However, participants did not prefer to make comment for some of the signs since they could not guess the meaning of them, so left it empty. Therefore, incorrect category was changed with 2 other categories, "wrong answer" and "no comment". As a result, 5 scale coding was used to evaluate the guessability score similar to the Kirmizioglu and Tuydes-Yaman (2011) as follows:

- (+1) point for the completely correct answer
- (+0.5) point for the partially correct answer
- (0) point for the no comment
- (-0.5) point for the wrong answer
- (-1) point for the opposite answer

As an example regarding the coding, "no parking", "no waiting" etc. answers for the "no standing and parking" sign were coded as partially correct (+0.5). "Two way directions", "roundtrip" etc. answers for the "priority over oncoming vehicles" was coded as wrong (-0.5) since the given meanings were not opposite but wrong from its true meaning. The answer "handcart can be used" for the "no handcart" sign was coded as opposite (-1).

Both of the questionnaires were controlled and coded by only one person, the examiner, in order to keep standardization. It was also ensured that questionnaires were fully completed by the participants. Some of the participants could not join the second experiment because of the personal reasons. Those participants were excluded from the analysis. Only the responses of 6 participants were excluded from the study.

The flowcharts for both of the experiments were shown in Figure 4.3.

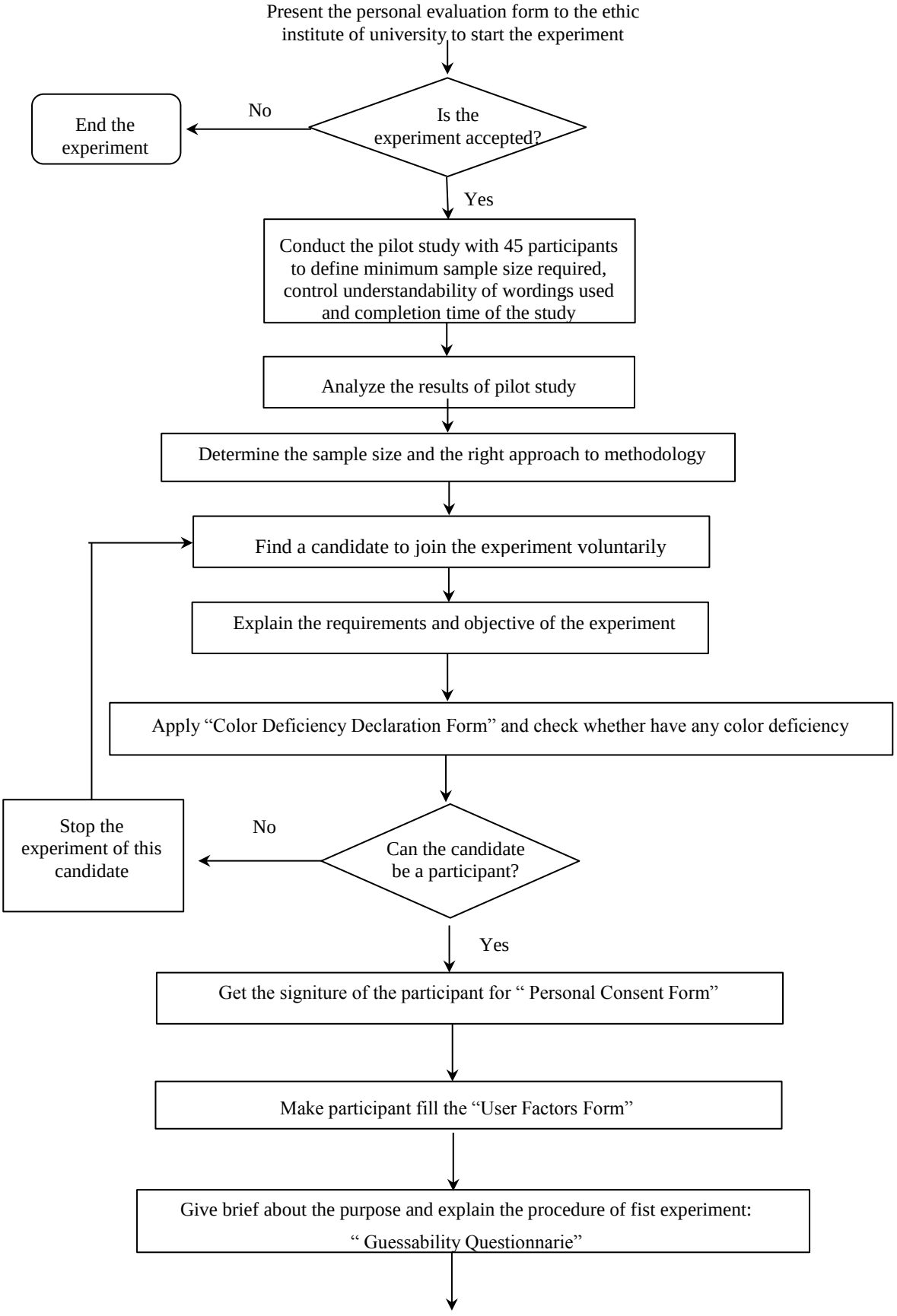


Figure 4.3. Flow Chart of the Experimental Procedure

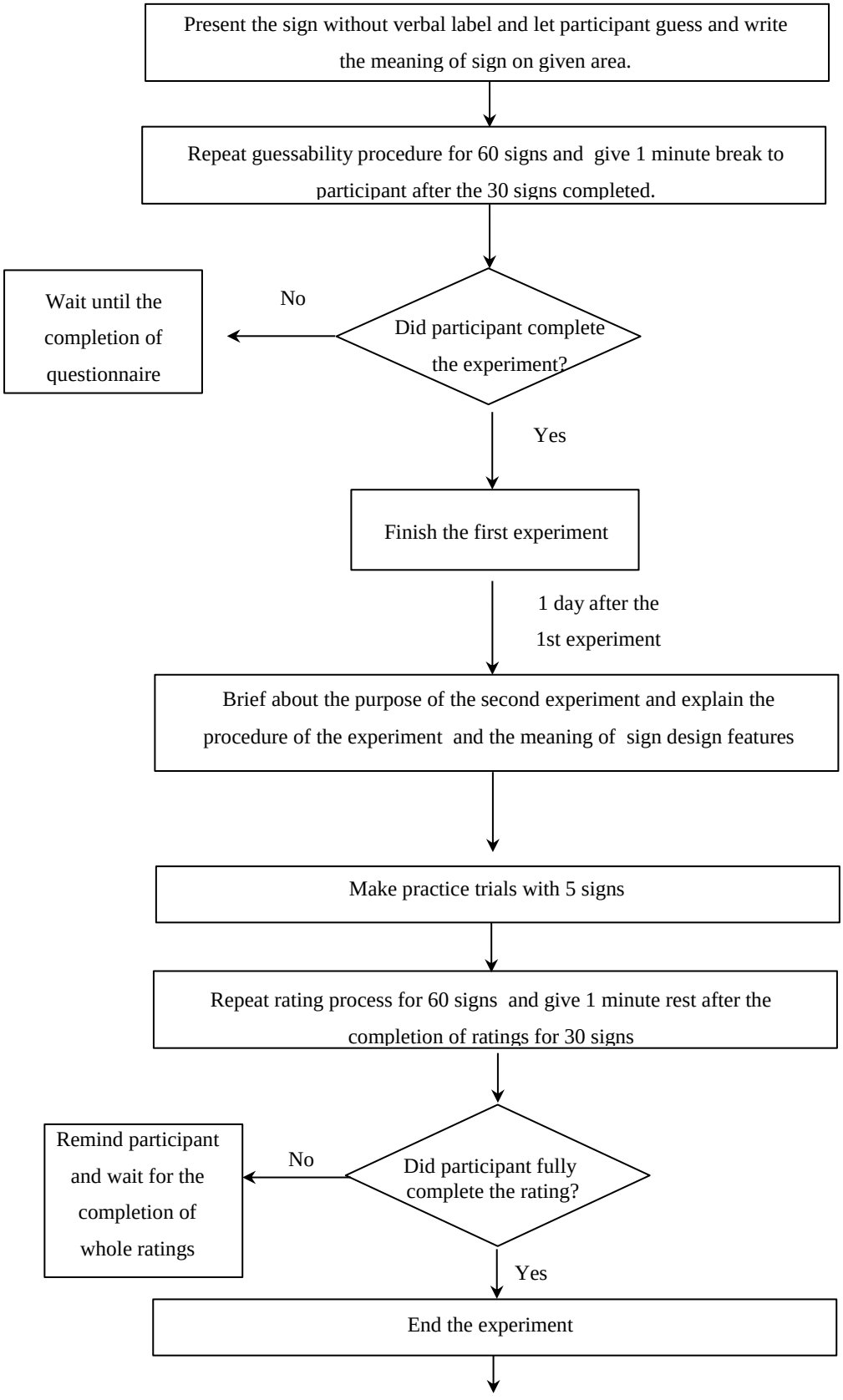


Figure 4.3. Flow Chart of the Experimental Procedure (cont.)

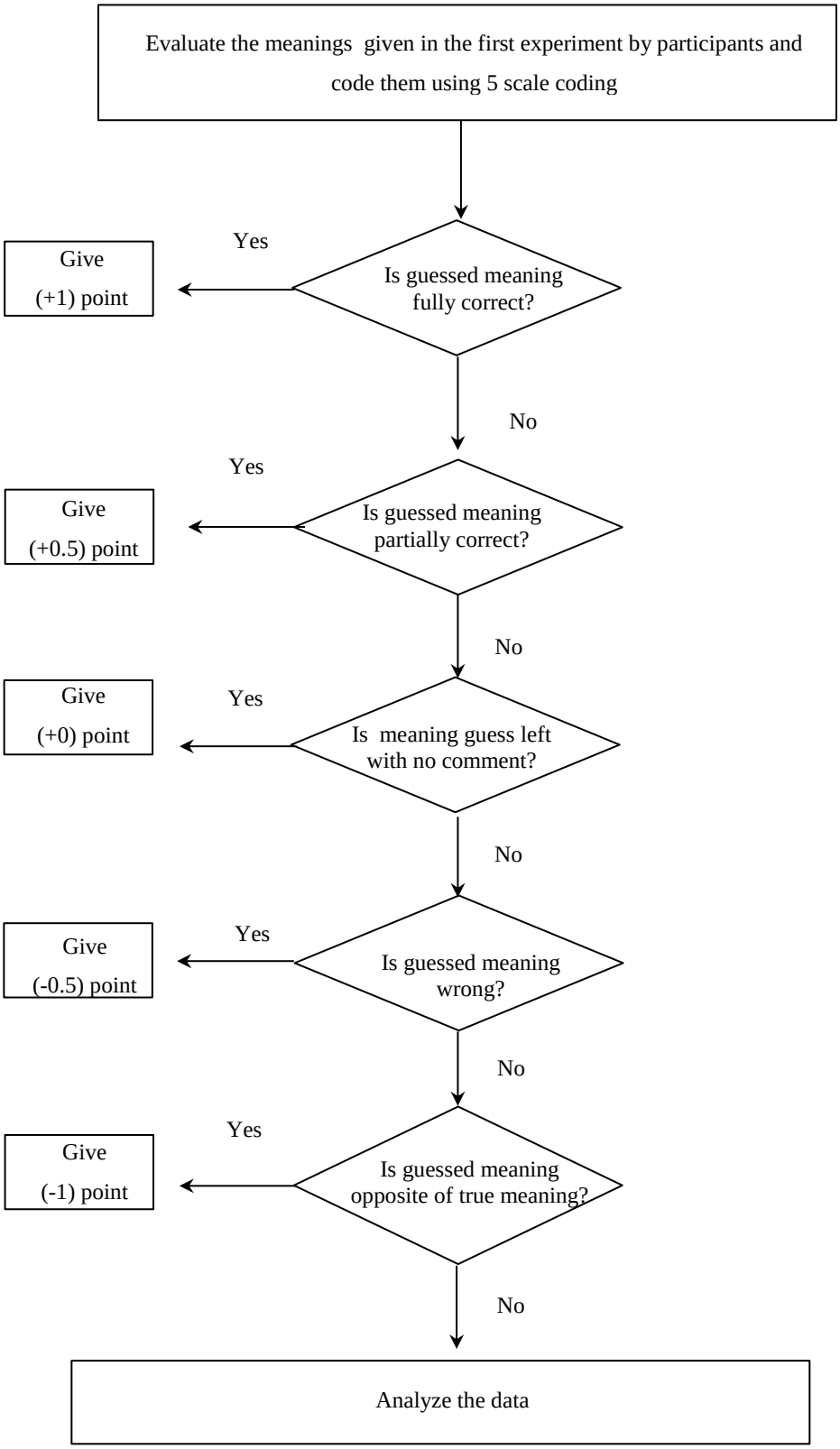


Figure 4.3. Flow Chart of the Experimental Procedure (cont.)

4.4. Experimental Design and Statistical Analysis

4.4.1. Experimental Variables

Three different experiments were conducted in this study: the effect of cognitive sign design features on guessability score, the effect of user factors on guessability score and the effect of user factors on cognitive sign design features. Levels of the factors can be seen in Table 4.4.

- Cognitive sign design features on guessability: The response (dependent) variable of the experiment is the guessability score which refers the accuracy level for guessing the meaning of a sign. Guessability score was calculated as the percentage of fully correct responses given by the participant for all sign. The design (independent) factors are the familiarity, concreteness, simplicity, meaningfulness and semantic closeness as cognitive sign design features. The score for design factors were calculated as the average score of all signs given by the participant for a design factor.
- User factors on guessability: The response (dependent) variable is the guessability score which refers the accuracy level for guessing the meaning of a sign. Guessability score was calculated as the percentage of fully correct responses given by the participant for all sign. The design (independent) factors are age group, marital status, educational background, active car driving experience and smoking habits.
- User factors on cognitive sign design features: There are five different response (dependent) variable; familiarity, concreteness, simplicity, meaningfulness and semantic closeness. The design (independent) factors are age group, marital status, educational background, active car driving experience, smoking habits and sign type as an additional independent variable included in analysis.

4.4.2. Experimental Conditions

60 traffic signs were included in the study and randomly presented to the participants. The experimental tasks were to guess and write the meaning of each traffic

signs and then give ratings of five sign design features for each sign. 201 participants were involved in the study. All responses were recorded in "Sign Design Features Evaluation Sheet" that can be seen in Appendix A. The number of recorded data points (test runs) was 201 x 60 x 6 = 72.360. Hence, the number of data point used in the analysis mentioned below:

(i) Cognitive sign design features on guessability: Guessability score of each sign was calculated as the % of fully correct responses (pointed as +1 point) given by the participants. The cognitive sign design feature of a sign was calculated as the average of responses given by the participants. Therefore each sign had one guessability score and one score for each cognitive sign design features. In total 60 data was analyzed.

(ii) User factors on guessability: Guessability score of each participant was calculated as the % of fully correct responses (pointed as +1 point) given for all signs and 201 participant joined the study. Therefore one % data was calculated per participant. In total 201 data was analyzed.

(iii) User factors on sign design features: 201 participant joined study and evaluation for 5 sign design features for each 60 traffic signs were done. Therefore the number of analyzed data point was 201 x 60 x 5 = 60.300.

4.4.3. Experimental Model

(i) To analyze the effect of sign design features on guessability score, multiple regression analysis was conducted. The model for regression analysis is:

$$y = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \cdots + \beta_k x_k + \varepsilon \quad (4.1)$$

where,

y : Response (Guessability score)

β_0 : Constant

$\beta_1, \ldots, \beta_5$: Regression coefficients for familiarity, concreteness, simplicity,

meaningfulness and semantic closeness

x_1, x_2 : Regression variables

ε : Error term (normally distributed)

The hypotheses of interest are:

$$H_0: \beta_1 = \beta_2 = \cdots = \beta_5 = 0 \quad (4.2)$$

$$H_1: \beta_k \neq 0 \; for \; at \; least \; one \; k \; (k = 1, \ldots, 5) \quad (4.3)$$

(ii) In order to analyze the effect of user factors on guessability score ANOVA was done. Completely randomized design was selected as an experimental model. The model for complete randomized design is:

$$y_{klmnp} = \mu + \delta_k + \varphi_l + \vartheta_m + \partial_n + \gamma_p + \varepsilon_{klmnp} \quad (4.4)$$

where,

y_{klmnp} : $klmnp^{th}$ Response (Guessability Score)

μ : The overall response mean

δ_k : Effect of k^{th} level of age group factor

φ_l : Effect of l^{th} level of educational background group factor

ϑ_m : Effect of m^{th} level of active car usage group factor

∂_n : Effect of n^{th} level of marital status group factor

γ_p : Effect of p^{th} level of gender group factor

ε_{klmnp} : Random error component NID $(0,\sigma^2)$

for

$k = 1, 2, 3, 4, 5$ (1: 18-29 yrs)

(2: 30-39 yrs)

(3: 40-49 yrs)

(4: 50-59 yrs)

(5: 60- 69 yrs)

$l = 1,2,3$ (1: Primary / Secondary School)

(2: High School)

(3: University or higher degree)

$m = 1, 2, 3, 4$ (1: Not an active car driver)

(2: Less than or equal to 1 year)

(3: More than 1 year and less than 5 years)

(4: More than or equal to 5 years)

$n = 1, 2, 3$ (1: Single)

(2: Married)

(3: Married with children)

$p = 1, 2$ (1: Male)

(2: Female)

The hypotheses of interest are:

$\mu_k = \mu + \delta_k$, for age k= 1,2,3,4,5

$H_0 = \mu_1 = \mu_2 = \ldots\ldots = \mu_5$

H_1 = At least one μ_k is dif.

$\mu_l = \mu + \varphi_l$, for educational background l = 1,2,3

$H_0 = \mu_1 = \mu_2 = \mu_3$

H_1 = At least one μ_l is dif.

$\mu_m = \mu + \vartheta_m$, for active car usage m= 1,2,3,4

$H_0 = \mu_1 = \mu_2 = \ldots\ldots = \mu_4$

H_1 = At least one μ_m is dif.

$\mu_n = \mu + \partial_n$, for marital status n= 1,2,3

$H_0 = \mu_1 = \mu_2 = \mu_3$

H_1 = At least one μ_n is dif.

$\mu_p = \mu + \gamma_p$, for gender p= 1,2

$H_0 = \mu_1 = \mu_2$

$H_1 = .\mu_1 \neq \mu_2$

Table 4.4. Design factors and levels of these factors

	Response Variable	Design Factors (Independent variable)	Number of Levels	Levels
Cognitive Sign Design Features on Guessability	Guessability Score	Cognitive Sign Design Features	5	(1) Familiarity (2) Concreteness (3) Simplicity (4) Meaningfulness (5) Semantic Closeness
User Factors on Guessability	Guessability Score	Age Group	5	(1) 18-29 yrs. (2) 30-39 yrs. (3) 40-49 yrs. (4) 50-59 yrs. (5) 60-69 yrs.
		Marital Status	3	(1) Single (2) Married (3) Married with Children
		Gender	2	(1) Male (2) Female
		Educational Background	3	(1) Primary / Secondary School (2) High School (3) University or higher degree
		Active Car Driving Experience	4	(1) Not an active car driver (2) Less than or equal to 1 yr (3)More than 1 yr and less than 5 yrs (4) More than or equal to 5 yrs
		Smoking Habit	2	(1) Yes (2) No

Table 4.4. Design factors and levels of these factors (cont.)

	Response Variables	Design Factors (Independent variable)	Number of Levels	Levels
User Factors on Cognitive Sign Design Features	(1) Familiarity (2) Concreteness (3) Simplicity (4) Meaningfulness (5) Semantic Closeness	Age Group	5	(1) 18-29 yrs. (2) 30-39 yrs. (3) 40-49 yrs. (4) 50-59 yrs. (5) 60-69 yrs.
		Marital Status	3	(1) Single (2) Married (3) Married w Children
		Educational Background	3	(1) Primary / Secondary School (2) High School (3) University or higher degree
		Gender	2	(1) Male (2) Female
		Active Car Driving Experience	4	(1) Not an active car driver (2) Less than or equal to 1 yr (3) More than 1 yr and less than 5 yrs (4) More than or equal to 5 yrs
		Smoking Habit	2	(1) Yes (2) No
		Sign Type	5	(1) Warning (2) Regulatory (3) Informational (4) Standing and Parking (5) Road works and maintenance

(iii) In order to analyze the effect of user factors on cognitive sign design features, ANOVA was done. There are five cognitive sign design features, so five different response variables and analysis. Separate ANOVA for each design feature was conducted. Completely randomized design was selected as an experimental model. The model for complete randomized design is:

$$y_{klmnpo} = \mu + \delta_k + \varphi_l + \vartheta_m + \partial_n + \gamma_p + \Omega_o + \varepsilon_{klmnpo} \quad (4.5)$$

where,

y_{klmnpo} : $klmnpo^{th}$ Response (familiarity, concreteness, simplicity, meaningfulness or Semantic Closeness)

μ : The overall response mean

δ_k : Effect of k^{th} level of age group factor

φ_l : Effect of l^{th} level of educational background group factor

ϑ_m : Effect of m^{th} level of active car usage group factor

∂_n : Effect of n^{th} level of marital status group factor

γ_p : Effect of p^{th} level of gender group factor

Ωo : Effect of o^{th} level of sign type group factor

ε_{klmnpo} : Random error component NID $(0,\sigma^2)$

for

$k = 1, 2, 3, 4, 5$ (1: 18-29 yrs)
(2: 30-39 yrs)
(3: 40-49 yrs)
(4: 50-59 yrs)
(5: 60-69 yrs)

$l = 1, 2, 3$ (1: Primary / Secondary School)
(2: High School)
(3: University or higher degree)

$m = 1, 2, 3, 4$ (1: Not an active car driver)
(2: Less than or equal to 1 year)
(3: More than 1 year and less than 5 years)
(4: More than or equal to 5 years)

$n = 1, 2, 3$ (1: Single)
(2: Married)

(3: Married with children)

$p = 1, 2$ (1: Male)

(2: Female)

$o = 1, 2, 3, 4, 5$ (1: Warning)

(2: Regulatory)

(3: Informational)

(4: Standing and Parking)

(5: Roadworks)

The hypotheses of interest are:

$\mu_k = \mu + \delta_k$, for age k= 1,2,3,4,5

$H_0 = \mu_1 = \mu_2 = \ldots.. = \mu_5$

H_1 = At least one μ_k is dif.

$\mu_l = \mu + \varphi_l$, for educational background l = 1,2,3

$H_0 = \mu_1 = \mu_2 = \mu_3$

H_1 = At least one μ_l is dif.

$\mu_m = \mu + \vartheta_m$, for active car usage m= 1,2,3,4

$H_0 = \mu_1 = \mu_2 = \ldots\ldots = \mu_4$

H_1 = At least one μ_m is dif.

$\mu_n = \mu + \partial_n$, for marital status n= 1,2,3

$H_0 = \mu_1 = \mu_2 = \mu_3$

H_1 = At least one μ_n is dif.

$\mu_p = \mu + \gamma_p$, for gender p= 1,2

$H_0 = \mu_1 = \mu_2$

$H_1 = .\mu_1 \neq \mu_2$

$\mu_o = \mu + \gamma_o$, for sign type o= 1,2,3,4,5

$H_0 = \mu_1 = \mu_2 = \ldots\ldots = \mu_5$

H_1 = At least one μ_m is dif.

4.4.4. Pilot study

Before starting main experiment, a pilot study was conducted with randomly selected 45 participants (23 female and 22 male). The categorization of participants can be seen in Table 4.5. The purposes of the pilot study were:

(i) Estimate the questionnaires' completion time and check whether the questions are clear and understandable enough,
(ii) Get familiar with experimental procedures and equipments,
(iii) Find statistical parameter values (mean and standard deviation) to calculate the required minimum sample size.

4.4.5. Sample Size Determination

Determining sample size is an important step of statistical analysis since it gives information of sample size to be able to make an accurate decision on the study. The sample size should be large enough to obtain accurate results and small enough to be collected in feasible time.

In order to estimate the sample size needed, the following method offered by ISO standards (ISO 15535:2006) for establishing anthropometric databases are applied in the study.

$$N = \left(\frac{1.96 \times CV}{a}\right)^2 \times 1.534^2 \quad (4.6)$$

1.96 : the critical (Z) value from a standard normal distribution for a 95% confidence interval
CV : the coefficient of variation
a : the percentage of relative accuracy desired

CV is calculated as the following:

$$CV = \frac{SD}{\bar{x}} \times 100 \tag{4.7}$$

where,

$\bar{x}$: Sample mean

SD : Sample standard deviation

In this study, to calculate the CV, the mean and standard deviation of the population are taken from the results of the pilot study. In pilot study, CV values vary for each sign type and sign feature type as can be seen in Table 4.6. Therefore, the highest CV value among sign types for each sign feature is chosen and considering relative accuracy to be 20%, required minimum sample sizes are calculated in Table 4.7.

Table 4.5. Pilot study participants categorization

User Factors	**Levels**	**Number of Participants**	
		Male	**Female**
Gender	(Male, Female)	**22**	**23**
Age	18-29 yrs	5	5
	30-39 yrs	5	5
	40-49 yrs	4	5
	50-59 yrs	4	4
	60-69 yrs	4	4
Marital Status	Single	4	10
	Married	3	5
	Married with children	15	8
Educational Background	Primary / Secondary School	11	7
	High School	8	7
	University or higher degree	3	9
Active Car Driving Experience	Not an active car driver	3	6
	Less than or equal to 1 yr	6	8
	More than 1 yr and less than 5 yrs	7	5
	More than or equal to 5 yrs	6	4
Smoking Habits	Yes (Smoker)	9	5
	No (Non-Smoker)	13	18

Table 4.6. Sample statistics of pilot study

Sign Type	Sign Feature	Mean	SD	CV	Min	Max
Warning (22)	Familiarity	60.5	34.4	56.8	16.4	92.3
	Concreteness	55.4	38.8	70.0	30.7	79.1
	Simplicity	71.1	27.8	39.1	26.5	93.2
	Meaningfulness	62.8	32.9	52.4	25.7	86.3
	Semantic Closeness	67.8	33.0	48.6	36.4	86.4
Regulatory (21)	Familiarity	53.2	36.4	68.4	14.8	82.9
	Concreteness	47.4	39.1	82.5	16.7	67.6
	Simplicity	65.9	31.0	47.0	33.2	92.4
	Meaningfulness	51.3	34.9	68.1	18.1	71.9
	Semantic Closeness	45.9	37.3	81.1	23.8	69.0
Informational (13)	Familiarity	48.0	39.0	81.3	13.8	92.3
	Concreteness	50.5	40.4	80.0	18.5	80.8
	Simplicity	63.8	31.0	48.5	33.8	88.5
	Meaningfulness	49.7	37.0	74.6	19.6	81.5
	Semantic Closeness	50.6	37.8	74.7	20.0	76.1
Standing and Parking (3)	Familiarity	78.2	28.1	35.9	0.0	100.0
	Concreteness	22.7	29.7	130.8	0.0	76.7
	Simplicity	75.5	22.7	30.0	33.7	100.0
	Meaningfulness	48.1	35.1	73.0	0.0	100.0
	Semantic Closeness	47.1	36.0	76.5	0.0	100.0
Road works and	Familiarity	71.0	NA	NA	NA	NA
Maintenance signs (1)	Concreteness	88.9	NA	NA	NA	NA
	Simplicity	77.7	NA	NA	NA	NA
	Meaningfulness	85.0	NA	NA	NA	NA
	Semantic Closeness	88.7	NA	NA	NA	NA

The largest value of the calculated sample sizes belongs to concreteness (327 subjects) and the second largest value corresponds to familiarity (150 subjects). Considering the data collection time, which completion of experiment for one subject takes almost 2 hours in total and 2 separate days are required for it, the calculated sample size for familiarity is taken as the minimum required sample size. However, a total of 201 subjects participated to study to increase the statistical power even further.

So the choice was;

$$N = \left(\frac{1.96 \times 81{,}35}{20}\right)^2 \times 1.534^2 = 149.561 = \boxed{150}$$

According to the selected sample size (201), the relative accuracy for each sign design feature differs (familiarity (17%), concreteness (28%), simplicity (10%), meaningfulness (16%), and semantic closeness (17%)).

Table 4.7. Minimum sample size for 95 % confidence and 20 % relative accuracy

Familiarity	$N = \left(\frac{1.96 \times 81.35}{20}\right)^2 \times 1.534^2 = 149.561 = 150$
Concreteness	$N = \left(\frac{1.96 \times 130.8}{20}\right)^2 \times 1.534^2 = 386.650 = 387$
Simplicity	$N = \left(\frac{1.96 \times 48.52}{20}\right)^2 \times 1.534^2 = 53.204 = 54$
Meaningfulness	$N = \left(\frac{1.96 \times 74.56}{20}\right)^2 \times 1.534^2 = 125.636 = 126$
Semantic Closeness	$N = \left(\frac{1.96 \times 81.12}{20}\right)^2 \times 1.534^2 = 148.716 = 149$

4.4.6. Statistical analysis

Statistical analysis was performed using Minitab 17.0. In the analysis, p-values ≤ 0.05 were accepted as significant and $0.05 <$p-values ≤ 0.1 were accepted as marginal.

For descriptive statistics: mean, standard deviation, coefficient of variation, range (min-max), scatterplots and correlation analysis of data was calculated. To determine if there were any sign with ratings very different from the other sign, the box plot was prepared. The signs with min and max scores were defined to make a comparison between signs with the lowest and highest guessability score.

For inferential statistics, regression analysis was used to estimate the relationship between guessability score and sign designs features (familiarity, concreteness, simplicity, meaningfulness, semantic closeness) and find prediction equations for guessability score. In addition to this, ANOVA was used to investigate the effects of user factors (age, gender,

marital status, educational background, active car driving experience, and smoking habits and sign type) on guessability score.

4.4.6.1. Regression Analysis. Regression analysis was conducted to understand the effect of five sign design features to predict the guessability score. Best subsets approach was performed in the statistical analysis. In the Best subsets approach, all possible models are compared and the best-fitting models are given by the system and this allows comparing by considering Mallow's Cp statistics and choosing the best.

According to Montgomery (2005), the general form of the multiple regression equation is as follows:

$$y = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \cdots + \beta_k x_k + \varepsilon \quad (4.8)$$

where,

y : Response (Guessability score)
β_0 : Constant
$\beta_1, \beta_2, \ldots$: Regression coefficients for independent variables,
x_1, x_2 : Regression variables,
ε : Error term (normally distributed)

To be able to perform regression analysis, the following assumptions must be satisfied. Therefore, assumptions mentioned in below were checked before starting the analysis (Montgomery, 2005):

- Normality of residuals: The residuals of the independent variables need to be normally distributed. This assumption was checked by using normal probability plots of residuals in Minitab 17.
- Homoscedasticity: The variances of the errors within groups must be equal.
- Linearity: Although the multiple regressions are not so much affected by the minor deviations, the relationship between the dependent and independent variables should

be linear. This assumption was checked by using scatter plot of dependent and independent variable.

- Multicollinearity: To be able to perform the regression analysis, multicollinearity should not exist. This can be checked by the variance inflation factor (VIF) criteria. VIF is recommended as smaller than 10.
- Autocorrelation: Residuals should not be correlated with each other. This assumption is checked by using Durbin-Watson statistic. Result of Durbin–Watson statistics should be between 1.5 and 2.5.

After checking the assumptions, regression models were performed. Moreover, the goodness of the model was checked by using the test for significance of regression to determine if there is a linear relationship between response and independent variables (Montgomery, 2005). The hypotheses are:

$$H_0: \beta_1 = \beta_2 = \cdots = \beta_p = 0 \tag{4.9}$$

$$H_0: \beta_k \neq 0 \; for \; at \; least \; one \; k \; (k = 1, \ldots, p) \tag{4.10}$$

Rejection of H_0 shows that, at least one of the independent variables contributes significantly to the model. Analysis of variance table can be used to check the significance of regression. If the p-value is less than α significance level, H_0 is rejected. Hence, at least one of the variables contributes significantly to the model (Montgomery, 2005).

In order to check goodness of fit of the regression model $R^2$and R^2_{adj} values were evaluated. If the unnecessary variables are added in to the model, the value of R^2_{adj} decreases. Therefore, the higher R^2_{adj} values indicate that response variable is determined by the independent variables better, another meaning, the fit of the model is better (Montgomery, 2005).

4.4.6.2. Analysis of Variance (ANOVA). The effect of user factors (age, gender, marital status, educational background, driver license holder, active car driving experience,

smoking habits and sign types) on guessability score were examined by using Analysis of variance (ANOVA), after checking the satisfaction of following ANOVA assumptions.

- Normality of residual assumption: The residuals of the model must follow normal distribution ($\varepsilon \sim$ NID Normal $(0, \sigma^2)$). This assumption was checked by using Anderson-Darling normality test. If the p value of the test result is higher than the 0,05, then the distribution of residuals are determined as normal. Moreover, another useful method to check normality is to construct the normal probability plot of residuals. If the plot looks like normal distribution centered at zero, so the assumption is satisfied. In our study, the distribution of guessability score and the residuals were not normal, therefore the box-cox transformation is applied with λ=0,5. After checking the transformed value, normality of residuals assumption was satisfied.
- Independence assumption: The correlation between residuals must be checked. There must not be any correlation between residuals. To check the correlation of residuals, plotting the residuals in time order is used. According to the Montgomery (2005), runs of positive and negative residuals violate the independence assumption, therefore proper randomization is essential for independence. In our study independence assumption was satisfied.
- Constant variance assumption: There must not be significant difference between the variances of response variables. In order to check this assumption, plot of residuals versus the fitted response was used and assumption was satisfied.

As a post-hoc analysis, Tukey's test was selected to use while making pairwise comparisons, the Tukey's method results in a narrower confidence limit, which is preferable (Toothaker, 1993). The hypotheses of this test are (Montgomery, 2005):

$$H_0: \mu_i = \mu_j$$
$$H_1: \mu_i \neq \mu_j \quad (4.11)$$

Where i and j are treatment levels $(i \neq j)$. Tukey's procedure makes use of the distribution of the studentized range statistic which is equal to (Montgomery, 2005)

$$q = \frac{\bar{y}_{max} - \bar{y}_{min}}{\sqrt{MS_E/n}} \quad (4.12)$$

Where $\bar{y}_{max}$ and $\bar{y}_{min}$ are the largest and the smallest sample means, MS_E is mean squares due to error and n is the sample size. Due to q value, T value of Tukey's test for unequal sample sizes can be calculated as (Montgomery, 2005):

$$T_\alpha = \frac{q_\alpha(a,f)}{\sqrt{2}} \sqrt{MS_E \left(\frac{1}{n_i} + \frac{1}{n_j}\right)} \quad (4.13)$$

Where $q_\alpha(a,f)$ is the upper α percentage points of studentized range statistics (q), f is the number of degrees of freedom associated with the MS_E, α is the number of groups will be compared, n_i and n_j are the sample sizes of the groups. Moreover, a set of $100(1-\alpha)$ confidence intervals for all pairs of means can be constructed.

5. RESULTS

5.1. Guessability Score

In this study the comprehension level of each sign based on the percentage of answers given in each defined categories (completely correct, partially correct, no comment, wrong answer and opposite answer) were calculated. Then, guessability score was defined by the percentage of completely correct answer (accurate answer) for a sign. The comprehensibility and guessability score of a sign can be defined with following equations mentioned below and calculation can be seen in Appendix B.

- Comprehensibility level: The percentage of the sample population that guess the given sign completely correct, partially correct, no comment ,wrong answer or opposite answer. For example, for the sign "Pedestrian crossing ahead" 95.5% of the participants gave completely correct answer, 3.5% of them gave partially correct answer, 1% of them left it without comment and nobody gave wrong or opposite answer.
- Guessability Score: The percentage of completely correct answer given (+1 point) by the sample population. Another word, the comprehensibility level of completely correct answers is accepted as guessability score. For example, the guessability score of "Pedestrian crossing ahead" sign is 95.5% which is equal to completely correct answer score.

In the beginning of the experiment, participants were asked to guess the meaning of the sign and write it in their own words on "Comprehension Questionnaire". The responses given by the participants were assessed by the experimenter considering the verbal labels received from the General Directorate of Highway (2004). An accurate response was the one that is completely correct (+1 point) and the given description points the correct meaning. The partially accurate answer (+0.5 point) was the one that partially correct answer which only a part of the correct meaning was given. No comments (0 point), wrong answer (-0.5 point) and the opposite answer (-1.0 point) were described as the inaccurate.

5.1.1. Guessability Score of sign categories

The descriptive statistics of guessability score for the five sign categories are shown in Table 5.1., and calculation equations are given in Appendix B. According to the results, the mean and the standard deviation of guessability score of all signs (60 signs) were 39.2% and 29.2%. Only the mean guessability score of warning signs (50.5%) and road works sign (82.6%) were greater than 50% whereas the mean guessability score of other categories were below 50%. Another calculation for guessability score was also calculated by considering the percentage of completely and partially correct answers together in order to see how the mean guessability score will change. The results were not very different from the guessability score in which only the percentage of completely correct answers had been taken. The mean and the standard deviation of partially accurate score of all signs were 55.6% and 33.7%. With the exception of warning signs (73.3%) and road works sign (96.5%), the mean score of sign categories were below than 50%.

In all sign category, with the exception of standing and parking signs, the variability was high and regulatory signs had the highest variability in guessability score (97.15%) as shown in Table 5.1.

The guessability score of signs were also analyzed and top ten highest and lowest guessability scored signs were shown in Table 5.2.

5.1.2. Comprehension level of traffic signs

The comprehension level of each sign based on the percentage of answers given in each defined category (completely correct, partially correct, no comment, wrong answer and opposite answer) are presented in Table 5.3., 5.4, 5.5, 5.6, 5.7. Comprehension level varied among the signs. Some of the signs were comprehended well (completely correct), some of them were partially comprehended (partially correct), some of them could not be comprehended (no comment) and some of them misunderstood (wrong answer and opposite answer). Therefore, signs were examined by grouping according to the comprehension level of participants as follows:

- Well known signs
- Partially known signs
- Sings left with no comment
- Misinterpreted sign
- Oppositely understood signs

Table 5.1. Descriptive Statistics of Guessability Score and Partially Accurate Score for Signs in the Five Sign Categories

	Guessability Score (%)				
	Mean	**Standart Deviation**	**Coefficiant of Variation**	**Minimum**	**Maximum**
Warning (22)	50.5	27.6	55%	2.5	95.5
Regulatory (21)	27.2	26.4	97%	0.5	88.1
Informational (13)	34.6	27.4	79%	1.0	95.0
Standing and Parking (3)	29.0	8.9	30%	17.9	39.8
Road Works (1)	82.6	NA	NA	NA	NA
Overall (60)	**39.2**	**26.2**	**66.7**	**1.0**	**95.0**

	Partially Accurate Score (%)				
	Mean	**Standart Deviation**	**Coefficiant of Variation**	**Minimum**	**Maximum**
Warning (22)	73.3	25.0	34%	10.4	99.0
Regulatory (21)	42.4	35.4	83%	1.5	93.5
Informational (13)	48.3	23.0	48%	0.0	95.0
Standing nd Parking (3)	37.3	26.5	71%	30.3	44.8
Road Works (1)	96.5	NA	NA	NA	NA
Overall (60)	**55.6**	**33.7**	**61%**	**0.00**	**95.5**

* The number of signs included in each category are shown in parenthesis.
NA=Not applicable

5.1.2.1. Well Known Signs. According to the analysis, some of the signs were comprehended highly correct, which means high percentage of the participants gave completely correct answer for them. Since the guessability score was associated with the percentage of completely correct answer, those well comprehended signs are the signs also with the highest guessability score.

12 traffic signs were comprehended correctly (+1 points) by the 70% or higher of the participants were included in that group. “Pedestrian Crossing Ahead” from the warning

Table 5.2. The top ten sign with the highest and lowest guessabiliy score (General Directorate of Highway, 2004)

	Sign Type	Sign	Correct Meaning	Guessability Score (%)
Sign with the highest score	W6		Pedestrian crossing ahead	95.5
	I-27		Pedestrian crossing	95.0
	W5		Traffic lights	89.5
	W-12		School crossing	88.1
	R12		No right turn	88.1
	R1		No U-turn	85.1
	RM1		Road work	82.6
	W1		Dangareous curve to right	80.1
	I7		Hospital	74.1
	W20		Cycle route ahead	73.6
Sign with the lowest score	R16		No entrance for vehicles carrying water-polluting cargo	0.5
	R21		No entrance for the vehicles carrying hazardous cargo	0.5
	I11		End of priority road	1.0
	R19		Closed to all vehicles in both directions	1.5
	R15		End of all restriction and prohibitions	2.0
	W14		Bridge begins	2.5
	R13		Mandatory direction for vehicles carrying dangerous Goods	2.5
	W10		Uncontrolled intersection	5.0
	R11		No entry for motor vehicles except Motorcycles	5.0
	R8		No entry for vehicles carrying explosives	5.5

sign category is the well-known sign by the 95.5% of the participants. “Pedestrian Crossing” from the informational sign category was known from by the 95.0% of the participants also. Those two similar signs were included in our study to observe the approach of participants to different categories and understand whether the difference of warning and information were recognized. Participants interpreted those two signs as same. “Traffic Lights”, “School Crossing” and “No Turn Right” were other well-known signs with respectively 89.5%, 88.1% and 88.1%. All of 12 well-known signs were presented in Table 5.3.

5.1.2.2. Partially Known Signs. The meanings of the some signs were known partially by 33% or higher of the participants. 6 signs were included in that category as can be seen in Table 5.4.

5.1.2.3. Signs Left with No Comment. More than 50% of the participants told that they do not know the meaning of some of the signs and left them without comment. This group includes 21 traffic signs when 50% or more score is considered. “End of priority road” was the sign that left without comment by 94.5% of the participants. “End of all restriction and prohibition”, “No entry for vehicle carrying explosives”, “No entry for vehicles carrying water-polluting cargo” and “Closed to all vehicle in both direction” had also highest score in no comment group by 88.1%, 87.6%, 86.6% and 86%. The score of other signs left without comment is presented in Table 5.5.

5.1.2.4. Misinterpreted Signs. Some of the signs were known as wrong by 20% or more of the participants (See Table 5.6.).“No entry for vehicle carrying explosives” sign was known as wrong by 42.8% of the participants. The most given answers for this sign were “Car can explode” and “explosion”. “Start road of motor vehicles” was the other sign that was known as wrong by 36.3% of the participants. “Parking” and “Car” were the most frequently given answer for that sign. “Youth camp” and “Priority over oncoming vehicles” were also the other misunderstood signs by 23.4% and 22.9% of the participants. “Home with garden” and “Tree to fall” answers were given instead of “Youth camp”. Moreover, “Round trip” answer was given for “Priority over oncoming vehicles” sign.

Table 5.3. Well known traffic signs (General Directorate of Highway, 2004)

Sign	Correct Meaning	Completely Correct (%)	Partially Correct (%)	No Comment (%)	Wrong Answer (%)	Opposite Answer (%)
W6 	Pedestrian crossing ahead	**95.5**	3.5	1.0	0.0	0.0
I27 	Pedestrian crossing	**95.0**	3.0	2.0	0.0	0.0
W5 	Traffic lights	**89.5**	9.5	1.0	0.0	0.0
W12 	School crossing	**88.1**	10.4	0.5	1.0	0.0
R12 	No right turn	**88.1**	2.5	9.0	0.5	0.0
R1 	No U-turn	**85.1**	6.0	5.0	4.0	0.0
RM1 	Road work	**82.6**	13.9	3.0	0.5	0.0
W1 	Dangareous curve to right	**80.1**	13.9	4.5	1.5	0.0
I7 	Hospital	**74.1**	1.0	24.4	0.5	0.0
W20 	Cycle route ahead	**73.6**	17.9	8.0	0.5	0.0
W21 	Domestic animals	**72.1**	26.9	1.0	0.0	0.0
W9 	Airport (Low-flying aircraft)	**70.6**	26.4	2.5	0.5	0.0

5.1.2.5. Oppositely understood signs. Some of the signs that the participants understood as the opposite of their true meanings included in that group. “No handcarts” sign was oppositely understood by 51.7% of the participants and “Handcarts can pass” or “Handcarts road” answers were given for this sign. “No entry for motor vehicle except motorcycles” was the other oppositely associated sign (46.8%) and “Car road” or “Motor Vehicle can pass” answers were given instead of its true meaning. For the “End of overtaking restriction for trucks (41.3%), “Overtaking prohibited for trucks” answer was given. “No overtaking” and “Priority over oncoming vehicles” were the other signs that were understood opposite of their true meanings by 23.9% and 6.0% of the participants. “Overtaking allowed” and “two line going road” answers were given for “No overtaking”

sign. Understanding the meaning of the sign as opposite of its true meaning can lead irreversible damage and even deadly outcomes on the roads.

Table 5.4. Partially known traffic signs (General Directorate of Highway, 2004)

Sign	Correct Meaning	Completely Correct (%)	Partially Correct (%)	No Comment (%)	Wrong Answer (%)	Opposite Answer (%)
R17 	No audible warning device	34.3	**58.7**	6.0	1.0	0.0
W17 	Road ends with sea or river	16.4	**48.8**	28.9	6.0	0.0
W3 	Steep hill downwards	12.9	**47.3**	30.8	9.0	0.0
R2 	Mandatory right direction	40.8	**40.8**	11.9	6.5	0.0
R9 	Roundabout	55.7	**37.8**	6.0	0.5	0.0
R15 	End of obligation to put on snow chains	24.9	**33.3**	30.0	11.5	0.5

Table 5.5. Signs left without comment (General Directorate of Highway, 2004)

Sign	Correct Meaning	Completely Correct (%)	Partially Correct (%)	No Comment (%)	Wrong Answer (%)	Opposite Answer (%)
I11 	End of priority road	1.0	2.0	**94.5**	3.0	0.0
R5 	End of all restriction and prohibitions	2.0	0.0	**88.1**	9.9	0.0
R8 	No entrance for vehicles carrying explosives	0.5	2.0	**87.6**	9.0	1.0
R16 	No entrance fo vehicles carrying water-polluting cargo	0.5	4.0	**86.6**	9.0	0.0
R19 	Closed to all vehicles in both directions	1.5	0.0	**85.1**	12.9	0.5
R13 	Mandatory direction for vehicles carrying dangerous goods	2.5	6.5	**83.6**	7.5	0.0
W14 	Bridge begins	2.5	9.0	**82.6**	6.0	0.0

Table 5.5. Signs left without comment (General Directorate of Highway, 2004) (cont.)

Sign	Correct Meaning	Completely Correct (%)	Partially Correct (%)	No Comment (%)	Wrong Answer (%)	Opposite Answer (%)
R10 	Give way	11.4	1.5	**78.6**	8.5	0.0
I5 	Tourism information	7.5	7.5	**78.1**	7.0	0.0
W10 	Uncontrolled intersection	5.0	5.5	**72.1**	17.4	0.0
P3 	No standing or parking	17.9	18.9	**55.7**	7.5	0.0
I4 	No entry road junction	17.9	25.4	**50.7**	6.0	0.0

Table 5.6. Misinterpreted signs (General Directorate of Highway, 2004)

Sign	Correct Meaning	Completely Correct (%)	Partially Correct (%)	No Comment (%)	Wrong Answer (%)	Opposite Answer (%)
R8 	No entry for vehicles carrying explosives	5.5	7.0	44.8	**42.8**	0.0
I12 	Start road of motor vehicles	8.5	12.4	41.3	**36.3**	1.5
I13 	Youth camp	21.4	13.4	40.8	**23.4**	0.0
R6 	Priority over oncoming vehicles	16.9	28.9	25.4	**22.9**	6.0

Table 5.7. Oppositely understood signs (General Directorate of Highway, 2004)

Sign	Correct Meaning	Completely Correct (%)	Partially Correct (%)	No Comment (%)	Wrong Answer (%)	Opposite Answer (%)
R20 	No handcarts	8.5	0.0	38.3	1.5	**51.7**
R11 	No entry for motor vehicles except Motorcycles	5.0	2.5	34.8	10.9	**46.8**
R7 	End of overtaking restriction for trucks	14.9	9.4	28.9	5.5	**41.3**
R14 	No overtaking	37.3	9.9	23.4	5.5	**23.9**
R6 	Priority over oncoming vehicles	16.9	28.9	25.4	22.9	**6.0**

5.2. Cognitive Sign Design Features

5.2.1. Descriptive statistics

In this part, Turkish traffic signs were examined in five categories of warning, regulatory, informational, standing and parking, and road works and maintenance. The ratings between 0-100 points were given for each sign design features, which are familiarity, simplicity, concreteness, meaningfulness and semantic closeness. The descriptive statistics of ratings for mentioned five categories are shown in Table 5.8., and calculation equations are shown in Appendix B. According to the results, mean rating for road works and warning signs were above the 50 for five sign design features. Only simplicity and meaningfulness were above 50 out of 100 for regulatory and standing & parking signs. With the exception of simplicity and semantic closeness, the ratings were below 50 for informational signs. It can be concluded that, warning and roadworks signs were found more familiar, simply, concrete, meaningful and semantically close by the Turkish population than the regulatory, informational and roadworks signs. As an overall evaluation, with the exception of concreteness, rating score of design features were above 50. Icons depict real objects, materials or people are defined as concrete. 48.8 score for concreteness shows that current traffic signs were found more abstract by the population. Moreover, current traffic signs were evaluated as moderately familiar, simply, meaningful and semantically close by subjects.

The first three signs with the highest and lowest ratings on familiarity, simplicity, concreteness, meaningfulness and semantic closeness were shown in Table 5.9. The sign R13 (No entry for vehicles carrying water-polluting cargo) was rated as the most unfamiliar sign with the rating 13.8, where the sign W6 (Pedestrian crossing ahead) received the highest familiarity rating. The sign W2 (School crossing) had the highest concreteness rating, it might be due to the fact that the depicted icon shows two students which is an image from real world. Moreover, the sign I11 (End of priority road) was rated as very abstract (5.6). For simplicity the sign W6 (Pedestrian crossing ahead) was rated as the most simple sign with rating 90.4, where the sign R13 (Mandatory direction for vehicles carrying dangerous goods) was found as the most complex sign (26.7). For meaningfulness and semantic closeness, same signs were rated as the highest and lowest.

The sign W6 (Pedestrian crossing ahead) received the highest ratings (91.4; 94.7) shows that it was found both meaningful and the depicted icon on this sign gives the real meaning well. And the sign I11 (End of priority road) received the lowest ratings (9.0; 7.8) that shows the sign I11 were not perceived as meaningful by the subjects and the relationship between depicted icon and function was not found strong enough.

Table 5.8. Descriptive statistics of Sign Design Features' ratings in five categories

Sign Type	**Sign Feature**	**Mean***	**SD**	**CV(%)**	**Min**	**Max**
Warning (22)	Familiarity	58.9	33.4	56.7	34.3	90.0
	Concreteness	53.9	36.5	67.7	13.0	89.8
	Simplicity	71.7	27.7	38.6	45.7	90.5
	Meaningfulness	64.6	32.6	50.4	19.5	91.2
	Semantic Closeness	70.0	32.5	46.5	14.8	94.0
Regulatory (21)	Familiarity	49.8	32.2	64.6	15.7	80.2
	Concreteness	43.6	34.4	78.9	9.8	79.6
	Simplicity	67.2	31.6	47.0	27.8	82.3
	Meaningfulness	52.1	32.1	61.6	19.2	79.4
	Semantic Closeness	46.3	34.1	73.6	11.5	86.5
Informational (13)	Familiarity	44.6	33.1	74.1	21.5	84.8
	Concreteness	49.8	35.8	72.0	6.5	84.3
	Simplicity	65.5	29.1	44.5	40.5	85.7
	Meaningfulness	50.0	33.0	66.1	8.9	84.5
	Semantic Closeness	50.8	33.1	65.2	8.5	89.5
Standing and Parking (3)	Familiarity	47.3	29.7	62.7	35.3	65.6
	Concreteness	31.1	27.7	89.0	22.3	48.7
	Simplicity	62.8	18.3	29.1	60.2	65.2
	Meaningfulness	54.4	23.0	42.2	45.2	72.0
	Semantic Closeness	38.4	24.8	64.7	33.1	55.9
Road works and maintenance (1)	Familiarity	73.6	NA	NA	NA	NA
	Concreteness	87.5	NA	NA	NA	NA
	Simplicity	82.0	NA	NA	NA	NA
	Meaningfulness	85.6	NA	NA	NA	NA
	Semantic Closeness	90.7	NA	NA	NA	NA
Overall (60)	**Familiarity**	**52.3**	**33.3**	**63.6**	**13.8**	**90.5**
	Concreteness	**48.8**	**36.1**	**73.8**	**5.6**	**90.2**
	Simplicity	**68.5**	**30.2**	**44.0**	**26.7**	**90.4**
	Meaningfulness	**56.9**	**32.9**	**57.7**	**9.0**	**91.4**
	Semantic Closeness	**56.3**	**35.0**	**62.1**	**2.2**	**94.7**

To observe whether there is a sign rated very different from others, the box plot of ratings for all signs on different sign design features were conducted as shown in table Figure 5.1. Only for simplicity, the sign R13 (Mandatory direction for vehicles carrying dangerous goods) was the only outliner with a much lower simplicity rating than the others.

In order to analyze whether the coefficient variation (CV) of sign features differ from each other's, the box plot of CV for ratings on each sign features for all sign was prepared. As shown in Figure 5.2, CV on each sign feature was different from sign to sign. The signs with extremely large CV on each sign features were shown in Table 5.10.

Table 5.9. First three signs with the highest and lowest ratings (General Directorate of Highway, 2004)

Sign Features	Signs with highest ratings	Signs with lowest ratings
Familiarity	W6  Pedestrian crossing ahead (90.5)	R13  No entrance fo vehicles carrying water-polluting cargo (13.8)
	W2  School crossing (88.4)	R21  No entrance for vehicles carrying hazardous cargo (16.9)
	W5  Traffic lights (88.2)	I13  Mandotory direction for vehicles carrying dangerous goods (19.8)
Concreteness	W2  School crossing (90.4)	I11  End of priority road (5.6)
	W6  Pedestrian crossing ahead (90.2)	R15  End of all restriction and prohibitions (10.3)
	W9  Airport (Low-flying aircraft) (87.2)	R10  Give way (10.8)
Simplicity	W6  Pedestrian crossing ahead (90.4)	R13  Mandatory direction for vehicles carrying dangerous goods (26.7)
	W5  Traffic lights (90.0)	R21  No Entry for the vehicles carrying hazardous cargo (34.7)
	W9  Airport (Low-flying aircraft) (89.3)	R16  No entrance fo vehicles carrying water-polluting cargo (35.6)
Meaningfulness & Semantic Closeness	W6  Pedestrian crossing ahead (91.4)	I11  End of priority road (9.0)
	W5  Traffic lights (91.0)	W14  Bridge begins (18.4)
	W2  School crossing (90.5)	R19  Closed to All Vehicles in Both Directions (18.7)

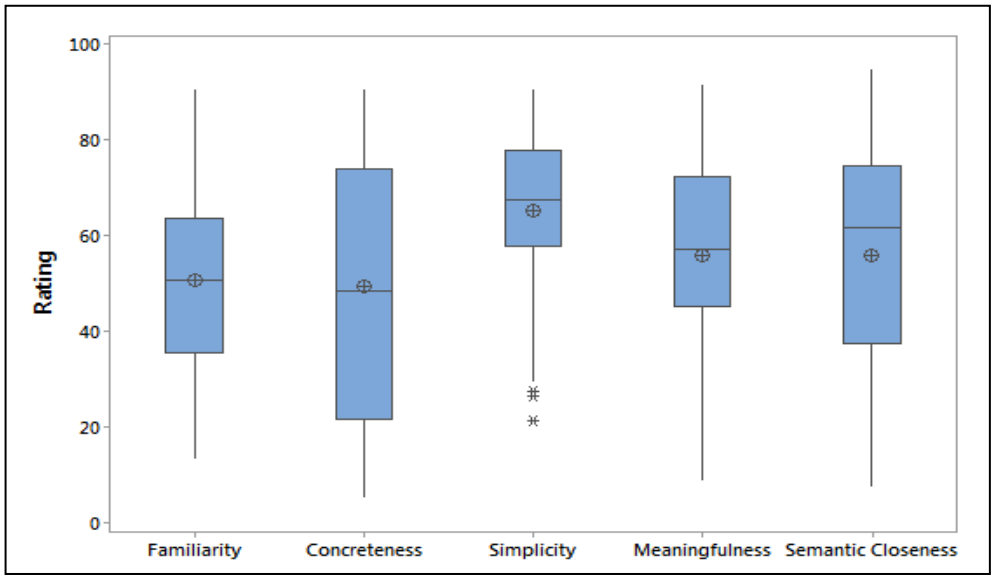


Figure 5.1. Box Plot of Ratings for All Signs on Sign Design Features

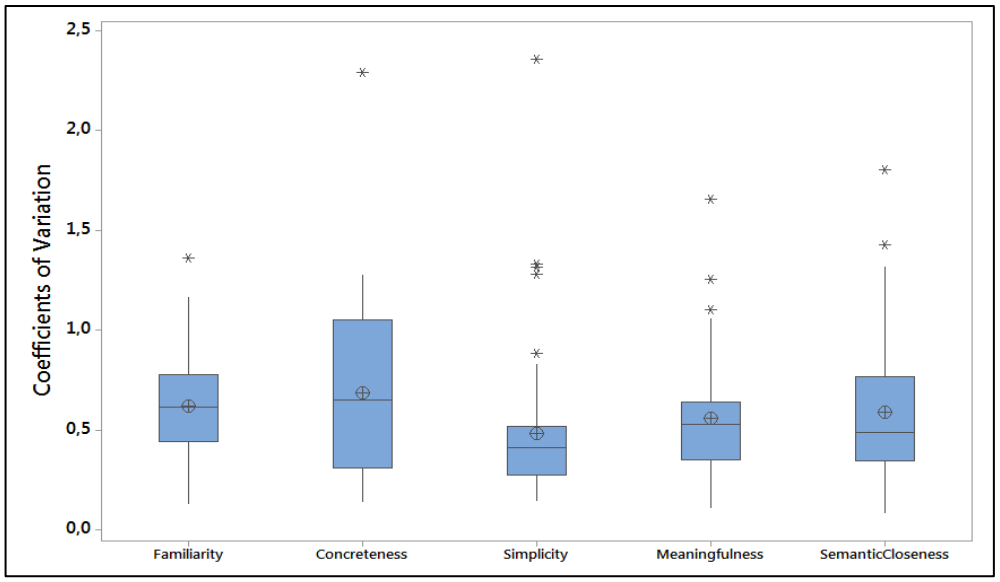


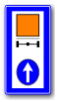
Figure 5.2. Box Plot of CV for All Signs on Sign Design Features

5.2.2. Interrelationships between Sign Design Features

The interrelationships between sign design features were analyzed. Except from concreteness, other four features were distributed normally (Anderson-Darling, $p>0.05$). Concreteness was also not far from the normality (Figure 5.3). The scatter plots (see in Figure 5.4) were also calculated and the relationships between sign features were observed

as linear. Since it is observed that the data fulfilled the assumptions, Pearson correlation were conducted in Table 5.11.

Table 5.10. Signs with extremely large CV in rating of sign features (General Directorate of Highway, 2004)

Sign Type	Sign Meaning	Familiarity	Concreteness	Simplicty	Meaningfulness	Semantic Closeness
	I11- End of priority road	127.8	226.8		158.9	181.9
	R16-No entrance for vehicles carrying water-polluting cargo			211.5		
	R21-No Entry for the vehicles carrying hazardous cargo			81.3		
	R13-Mandotory direction for vehicles carrying dangerous goods			85.8		
	I5-Turizm information				105.0	
	R19-Closed to All Vehicles in Both Directions				119.2	
	R10-Give way				104.8	
	W14-Bridge begins					142.4

As shown in Table 5.11., the highest correlation were found between meaningfulness and familiarity ($r = 0.812$, $n = 60$, $p < 0.001$) and between semantic closeness and meaningfulness ($r = 0.807$, $n = 60$, $p < 0.001$). Meaningfulness was also strongly correlated with concreteness ($r = 0.649$, $n = 60$, $p < 0.001$), and simplicity ($r = 0.656$, $n = 60$, $p < 0.001$). Semantic closeness was moderately associated with concreteness ($r = 0.523$, $n = 60$, $p < 0.001$) and simplicity ($r = 0.571$, $n = 60$, $p < 0.001$) and moderately correlated with familiarity ($r = 0.644$, $n = 60$, $p < 0.001$). The lowest correlation was found

between familiarity and concreteness (r = 0.227, n = 60, p < 0.081) and since the p value > 0.001 it can be concluded that there is not a significant relationship between concreteness and familiarity.

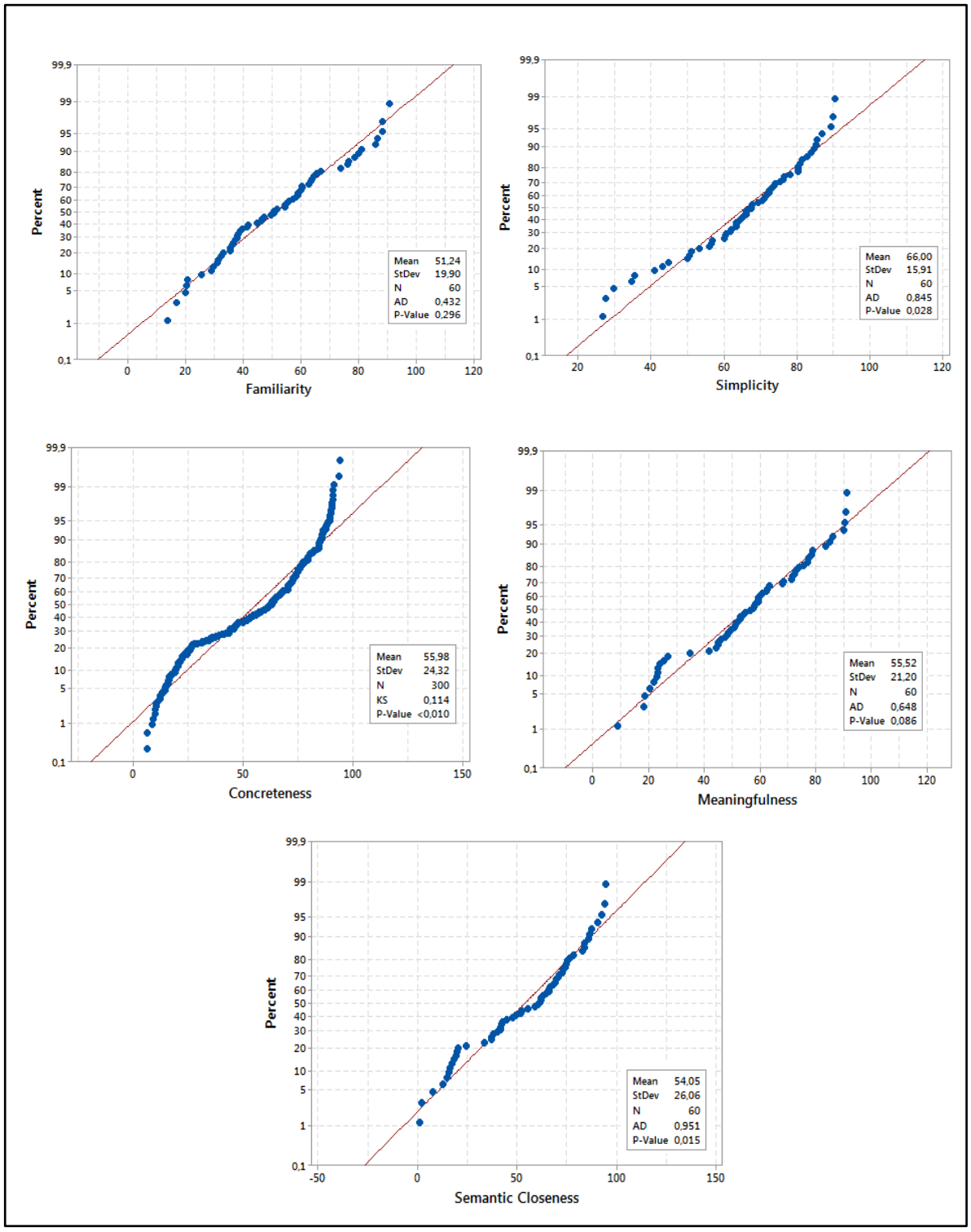


Figure 5.3. Normal Probability Plot of Sign Design Features

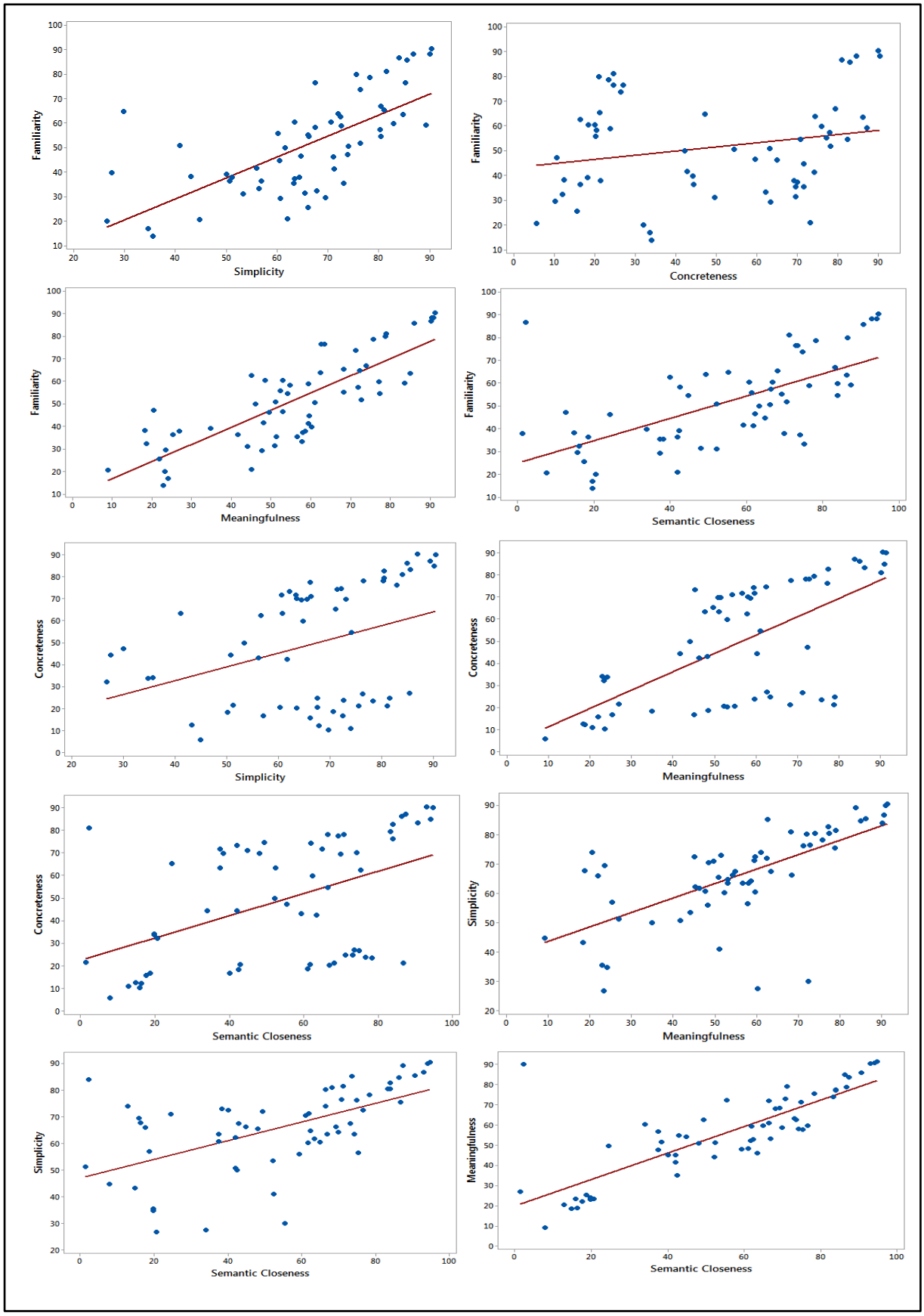


Figure 5.4. Scatter Graphs between two Sign Design Features

Precisely we can say that, sign design features are significantly and positively correlated with each other's, except from the relationship between familiarity and concreteness. There is no significant correlation between those two features.

Table 5.11. Pearson correlations of Sign Features

	Familiarity	Concreteness	Simplicity	Meaningfulness	Semantic Closeness
Concreteness	0,227* 0,081**				
Simplicity	0,686 <0,001	0,369 0,004			
Meaningfulness	0,812 <0,001	0,649 <0,001	0,656 <0,001		
Semantic Closeness	0,644 <0,001	0,523 <0,001	0,571 <0,001	0,807 <0,001	

*Pearson correlation coefficient; **p-value

5.2.3. Interrelationships between Guessability and Sign Design Features

In order to understand the relationship of guessability score with sign design features and check the hypothesis that the guessability score will be higher for the signs with higher familiarity, concreteness, meaningfulness, simplicity and semantic closeness, correlation analysis was done. According to the scatter plot, the relationship between guessability score and sign design features are almost linear. Pearson correlation analysis was performed.

As a result, except from concreteness, the other four sign design features were significantly and positively correlated with guessability score. Familiarity had the strongest correlation with guessability score ($r = 0.855$, $n = 60$, $p < 0.001$) whereas concreteness had the lowest correlation with guessability score ($r = 0.385$, $n = 60$, $p < 0.001$).

For further analysis, 15 signs with lowest familiarity score and 15 signs with the highest familiarity score were selected and Spearman correlation analysis was conducted. Sings with the highest familiarity score had high correlation with the guessability score ($r = 0.857$, $n = 15$, $p < 0.001$).

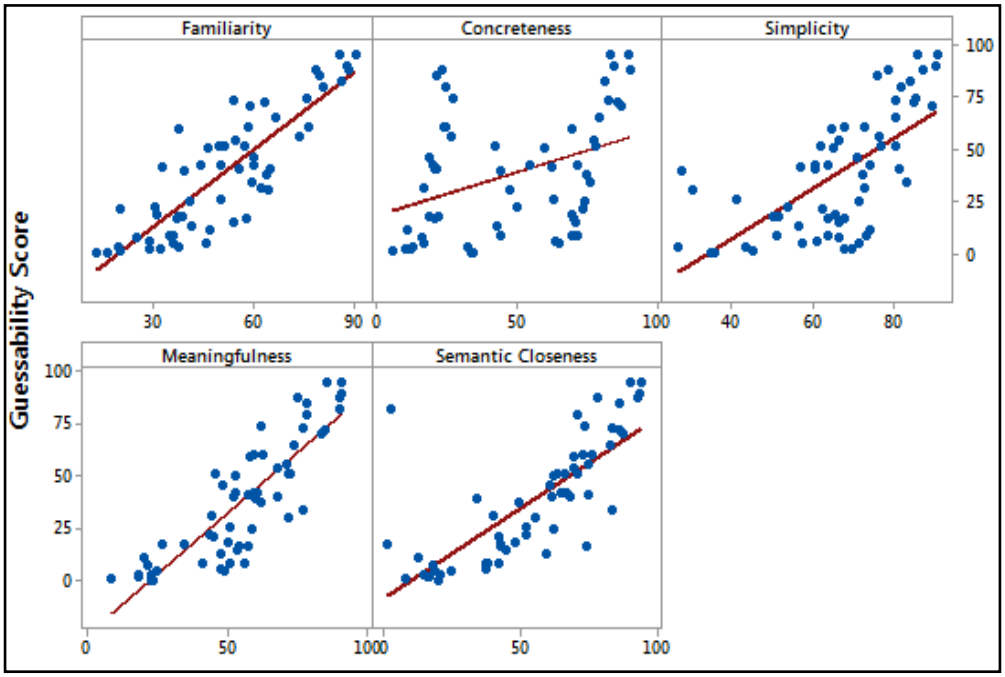


Figure 5.5. Scatter Graphs between Guessability Score and Sign Design Features

Table 5.12. Pearson correlations between Guessability Score and Sign Design Features

	Familiarity	Concreteness	Simplicity	Meaningfulness	Semantic Closeness
Guessability Score	0,855* <0,001**	0,385 <0,001	0,659 <0,001	0,856 <0,001	0,771 <0,001

*Pearson correlation coefficient; **p-value

5.2.4. Regression analysis of Guessability Score and Sign Design Features

To predict the guessability score, a multiple regression model was developed considering the sign design features as independent variables. As can be seen in Figure 5.6., regression assumptions were satisfied. Therefore, to develop the best equation Best Subsets Regression Analysis was used. The general form of the guessability score regression form is as mentioned below:

$$y = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \beta_3 x_3 + \beta_4 x_4 + \beta_5 x_5 + \varepsilon \quad (5.1)$$

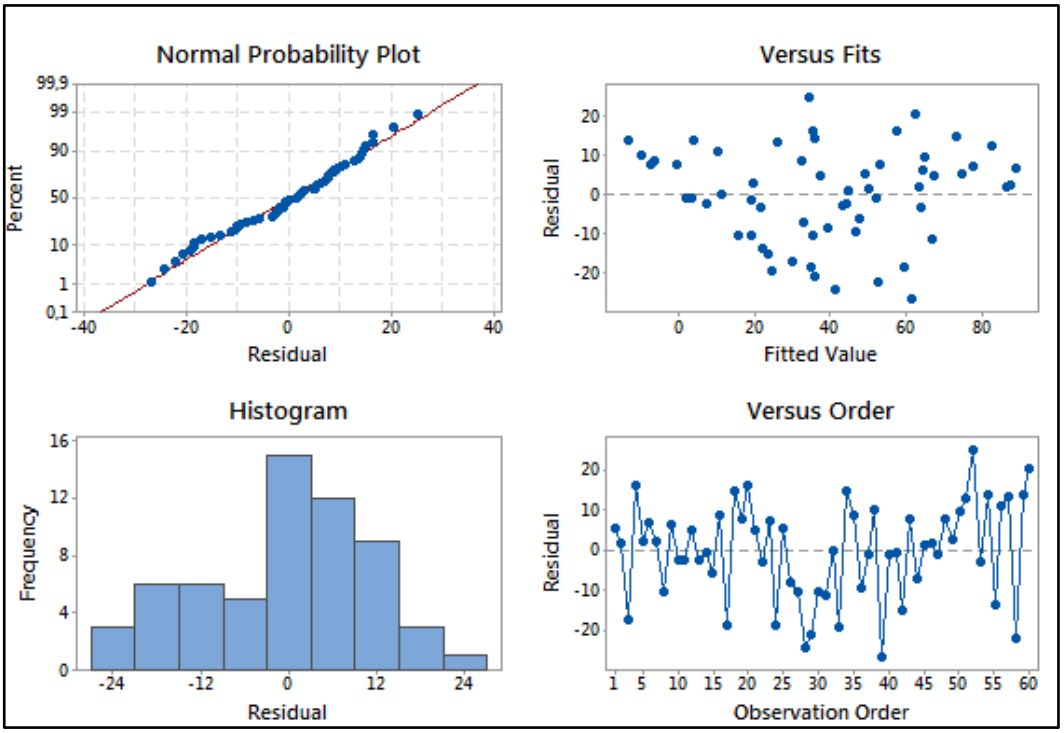


Figure 5.6. Residual Plots for Guessability Score

where,

β_0 : Constant

β_1 : Regression coefficient of familiarity

β_2 : Regression coefficient of concreteness

β_3 : Regression coefficient of simplicity

β_4 : Regression coefficient of meaningfulness

β_5 : Regression coefficient of semantic closeness

ε : Error term

x_1 : Regressor variable of familiarity

x_2 : Regressor variable of concreteness

x_3 : Regressor variable of simplicity

x_4 : Regressor variable of meaningfulness

x_5 : Regressor variable of semantic closeness

According to the Best Subsets technique, models and results are presented in Table 5.13. After all of the models are checked, Model 5 was selected as the most suitable model in order to predict guessability score. As a result, the regression equation with the maximum R^2 adj value 82.9 and Mallows' statistic C_p 2.5 was chosen as follows:

Guessability Score = -32.66 + 0.698 Familiarity + 0.369 Meaningfulness
+ 0.274 Semantic Closeness

Table 5.13. Different Regression Model Alternatives.

Regression Model	**S**	**R^2 (%)**	**R^2 adjusted (%)**	**Cp**
Model 1	15.123	73.3	72.9	29.2
Model 2	15.170	73.1	72.7	29.7
Model 3	12.711	81.5	80.8	5.2
Model 4	12.930	80.8	80.2	7.2
Model 5	12.306	82.9	82.0	2.5
Model 6	12.620	82.1	81.1	5.3
Model 7	12.389	83.0	81.8	4.2
Model 8	12.403	83.0	81.7	4.3
Model 9	12.477	83.1	81.5	6.0

Model 1: Meaningfulness

Model 2: Familiarity

Model 3: Familiarity, semantic closeness

Model 4: Familiarity, meaningfulness

Model 5: Familiarity, meaningfulness, semantic closeness

Model 6: Familiarity, concreteness, semantic closeness

Model 7: Familiarity, concreteness, meaningfulness, semantic closeness

Model 8: Familiarity, simplicity, meaningfulness, semantic closeness

Model 9: Familiarity, concreteness, simplicity, meaningfulness, semantic closeness

Analysis variance of regression model 5 was presented in Table 5.14. Since p value < 0.05, it can be concluded that at least one of the regressor variables contributes significiantly to the model.

Table 5.14 Analysis of Variance table of Regression Model

Source	**DF**	**Adj SS**	**Adj MS**	**F-Value**	**P-Value**
Regression	3	41.233,4	13.744,5	90.77	<0.001
Familiarity	1	3.867,5	3.867,5	25.54	<0.001
Meaningfulness	1	729	729	4.81	0.032
Semantic Closeness	1	1.050,2	1.050,2	6.94	0.011
Error	56	8.480	151.4		
Total	59	49.713,4			

To satisfy the autocorrelation assumption of regression, Durbin–Watson statistics was checked and the result was 2.0. Therefore it was concluded that there was no correlation between errors.

5.3. User Factors

5.3.1. Descriptive statistics for User factors and Guessability Score

In the Table 5.15, seven user factors, number and the mean guessability performance which was the average of the percentage of the fully correct answers were shown. Calculation details are shown in Appendix B. According to the results, the age of 18-29 yrs., marital status “married” and active car driving experience “more than 1 year and less than 5 years” had the highest mean guessability score (41.7%; 41.1%; 42.2%).

The average numbers of given fully correct, partially correct, no comment, wrong and opposite answers are calculated for each response group for each factor in Table 5.16.

5.3.2. Descriptive statistics for User Factors and Cognitive Sign Design Features

In the Table 5.16, seven user factors and the scores of each level in terms of sign design features are shown. Calculation details are shown in Appendix B.

Table 5.15. A summary of responses for the User Factors and the Guessability Score for different groups of factors

			Guessability Score (%)	
User Factor	**Groups**	**Number of participant**	**Mean**	**Standart Deviation**
Age	18-29 yr.	41 (20.4%)*	41.7	15.9
	30-39 yr.	43 (21.4%)	37.9	14.1
	40-49 yr.	40 (19.9%)	36.5	12.7
	50-59 yr.	41 (20.4%)	38.6	10.3
	60-69 yr.	36 (17.9%)	27.7	14.1
Gender	Male	100 (49.8%)	37.5	14.6
	Female	101 (50.2%)	36.0	13.7
Marital Status	Single	67 (33.3%)	35.4	14.8
	Married	39 (19.4%)	41.1	12.6
	Married w Children	95 (47.3%)	35.8	14.1
Educational Background	Primary / Secondary School	67 (33.3%)	35.4	15.1
	High School	66 (32.8%)	35.9	15.0
	University or higher degree	68 (33.8%)	38.8	12.2
Drivers Licence Holder	Yes	156 (77.6%)	37.9	12.7
	No	45 (22.4%)	33.1	18.0
Active Car Driving Experience	Not an active car driver	53 (26.4%)	31.5	16.3
	Less than or equal to 1 yr	53 (26.4%)	35.9	12.9
	More than 1 yr and less than 5 yrs	49 (24.4%)	42.2	12.5
	More than or equal to 5 yrs	46 (22.9%)	37.9	12.5
Smoking Habits	Yes	72 (35.8%)	37.8	14.2
	No	129 (64.2%)	36.1	14.2

*First number shows the number of participants in given group, second number (%) shows the percentage distribution of given group within the whole group of the given factor.

5.3.3. Effect of User Factors on Guessability

All factors that are age, gender, marital status, educational background, active car driving experience and smoking habits were categorical variables.

Table 5.16. Average number of answers given for each answer type in terms of each response group of factors.

User Factor	Group	Average Number of Given Answers*				
		Fully Correct	**Partially Correct**	**No Comment**	**Wrong**	**Opposite**
Age Group	18-29 yrs.	23.7*	14.0	13.6	6.5	2.2
	30-39 yrs.	21.7	12.2	21.1	3.7	1.3
	40-49 yrs.	21.0	15.6	18.1	3.3	2.1
	50-59 yrs.	22.1	12.8	17.3	5.1	2.6
	60-69 yrs.	15.6	10.0	30.4	2.9	1.1
Gender	Male	21.3	13.3	19.0	4.5	2.0
	Female	20.6	12.7	20.8	4.2	1.8
Marital Status	Single	20.2	13.7	19.9	4.3	1.9
	Married	23.5	12.0	18.1	4.6	1.8
	Married w Child	20.5	12.8	20.6	4.2	1.9
Educational Background	Primary / Secondary	20.2	12.8	21.1	4.1	1.8
	High School	20.5	13.0	19.6	5.0	1.9
	University / Higher	22.1	13.1	19.0	3.9	1.9
Active Car Usage	Not an active driver	18.0	12.6	23.2	4.5	1.6
	< or = to 1 yr	20.5	14.0	18.4	5.0	2.2
	>1 yr and <5 yrs	24.2	12.7	16.9	4.2	2.0
	> or = 5 yrs	21.5	12.5	20.9	3.5	1.7
Driver Licence Holder	Yes	20.7	12.1	21,0	4.1	2.1
	No	21.0	13.1	19.7	4.4	1.8
Smoking Habits	Yes	21.6	13.8	18.0	4.6	2.0
	No	20.6	12.5	20.9	4.2	1.8

* 23.7 indicates the average number of fully correct answers given by 41 participants in 18-29 yrs group for 60 traffic signs.

Table 5.17. A summary of responses for Cognitive Sign Design Features in terms of User Factors.

User Factor	Groups (Number of participants)	Familiarity		Concreteness		Simplicity		Meaningfulness		Semantic Closeness	
		$\bar{x}$	St. Dev.	$\bar{x}$	St. Dev.	$\bar{x}$	St. Dev.	$\bar{x}$	St. Dev.	$\bar{x}$	St. Dev.
Age	18-29 yrs. (41)	59	31	54	35	70	26	60	30	59	33
	30-39 yrs. (43)	54	32	52	35	68	30	57	32	56	34
	40-49 yrs. (40)	59	31	52	36	68	28	62	31	60	34
	50-59 yrs. (41)	54	33	50	37	67	32	61	33	57	35
	60-69 yrs. (36)	35	33	41	36	59	34	45	35	49	38
Gender	Male (100)	53	33	51	36	66	30	58	32	57	34
	Female (101)	52	34	49	36	67	30	56	33	55	35
Marital Status	Single (67)	52	32	50	36	67	29	57	32	56	34
	Married (39)	56	33	52	35	68	30	57	32	58	35
	Married with child (95)	51	34	49	36	66	31	57	34	55	36
Educational Background	Primary-Secon dary (67)	51	33	49	36	66	31	57	33	55	35
	High School (66)	52	33	51	36	67	30	58	33	57	34
	University or higher (68)	54	33	50	36	66	30	56	32	57	35
Active Car Driving Experience	Not active car driver (53)	46	34	46	36	64	32	54	34	55	36
	< = 1 yr (53)	55	32	53	35	68	29	58	32	56	34
	1- 5 yrs (49)	57	32	52	36	68	30	61	31	59	34
	>= 5 yrs (46)	51	33	49	36	66	30	56	33	55	35
Smoking Habits	Yes (72)	55	33	51	36	69	29	59	32	57	34
	No (129)	51	33	50	36	65	31	56	33	56	35
Sign Type	Warning (22)	59	33	55	36	71	28	64	33	69	33
	Regulatory (21)	50	32	44	34	67	29	52	32	46	34
	Information (13)	45	33	50	36	66	29	50	33	52	34
	Standing Parking (3)	47	30	43	36	28	35	62	24	39	25
	Roadwork(1)	74	28	88	18	82	21	86	22	91	13

In order to examine whether there were any statistically significant differences in guessability score among different levels of each factors, Analysis of Variance (ANOVA) was conducted.

5.3.3.1. Checking ANOVA assumptions. ANOVA assumptions were checked in below:

- Normality Test: In order to examine ANOVA, residuals of guessability score must fit to normal distribution. Therefore, normality of the residuals of the guessability score data were tested by using Anderson-Darling normality test ($\alpha = 0.05$) in Minitab 17. Since main data do not fit the normal, square-root transformation ($\sqrt{}$) is applied. As can be seen from the results, the p-value of residuals of transformed guessability score is > 0.05. Another indicator to prove normality is that the normal probability plot is resemble a straight line. The normal probability plot of residuals almost resembles a straight line which was shown in Figure 5.7. Therefore the error distribution of guessability score is normal.

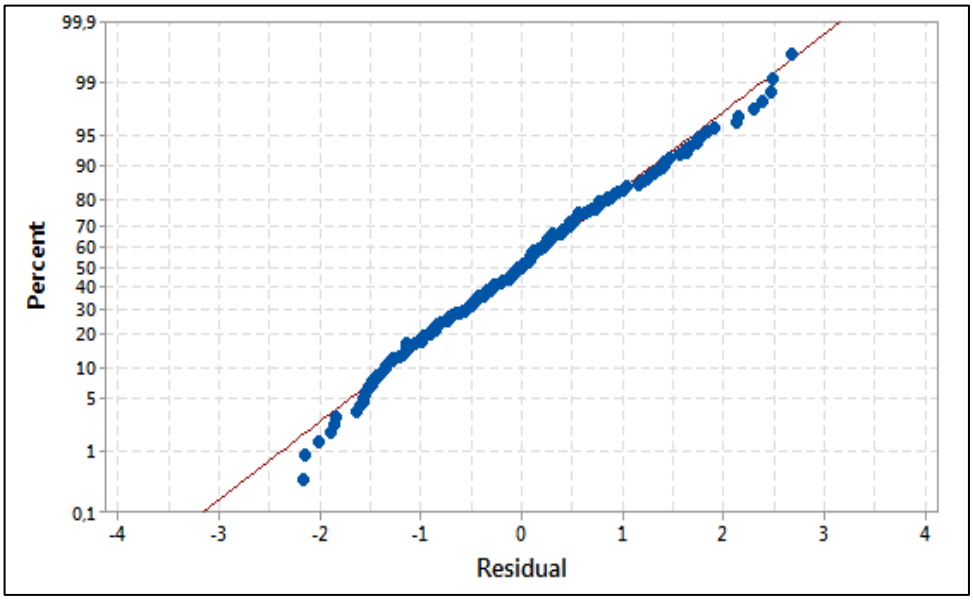


Figure 5.7. Normal Probability Plot of Residuals of Guessability Score

- Independence Test: The other ANOVA assumption is the independence assumption which means that there must not be any correlation between residuals and between independent variables and residuals. Plot of residuals versus observation is a way to

verify that there is no correlation between the residuals. The plot of the residuals versus observation order for guessability score is shown in Figure 5.8. As can be seen, independence assumption is approved since there is no reason to suspect from violation. In addition, for independence assumption there must not be correlation between independent variables (user factors) and residuals. According to the results using Pearson correlation, that shown in Table 5.18, there is not any significant correlation between independence variables and residuals. So, independence assumption is approved.

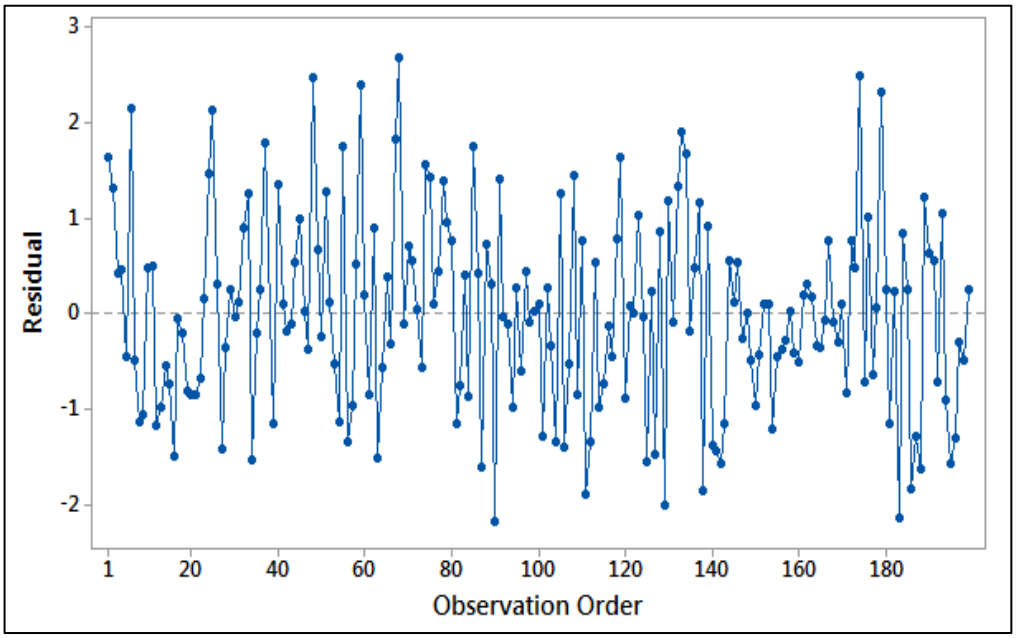


Figure 5.8. Plot of Residuals Versus Observation Order

Table 5.18. Correlation coefficients between independent variables and residuals

Residuals	Age Group	Educational Background	Active Car Usage	Marital Status	Smoking Habits	Gender
Residuals	0.025	0.022	0.024	0.024	-0.032	-0.019
P-value	0.722	0.754	0.741	0.742	0.656	0.792

- Variance Equality Test: Another assumption of ANOVA is the homogeneity assumption which means variance of response variables for each treatment must not be different from each other. Plot the residual versus fitted values is used to check

this assumption. If there is not any obvious pattern in graph, homogeneity assumption is proved. As can be seen in Figure 5.9, no unusual pattern was observed. As a result, variance equality or another meaning homogeneity of variance assumption was satisfied.

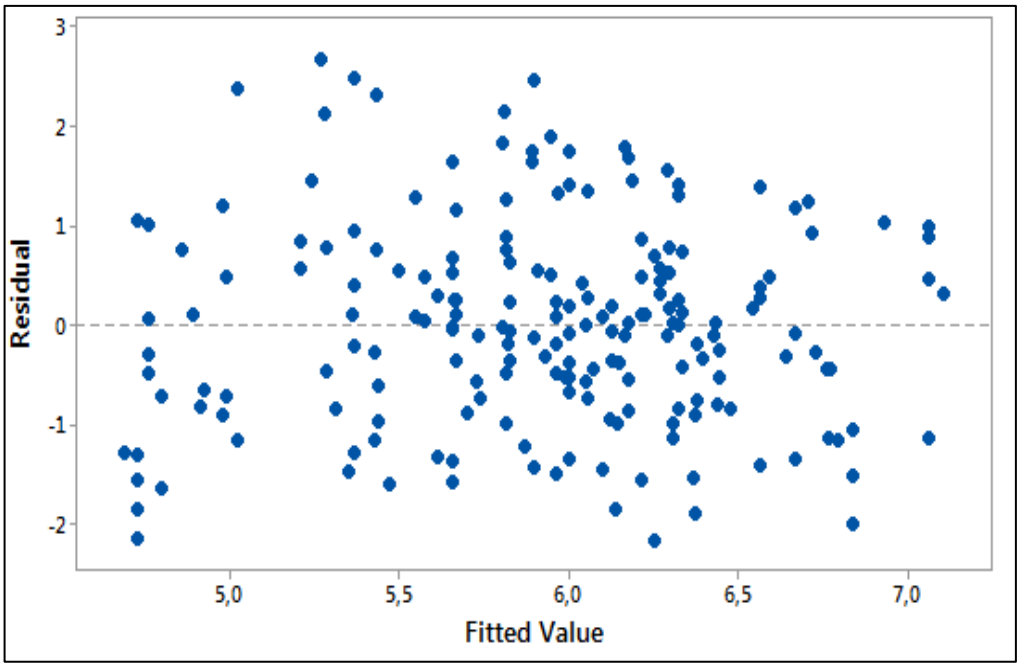


Figure 5.9. Plot of Residuals Versus Fitted Values

5.3.3.2. Analysis of Variance (ANOVA). Completely randomized design was performed for ANOVA analyses. Table 5.19 summaries the results.

The results of completely randomized design show that guessability score is significantly affected by age factor , active car usage and marital status ($p<0.05$)

5.3.3.3. Multiple Comparisons. The ANOVA results show that the age group, marital status group and active car usage had significant effect on guessability score. In order to understand which factors' level was significantly different, multiple comparisons was conducted for further analysis. Tukey's tests were performed for significant factors which have more than two levels: age factor, active car usage and marital status.

Table 5.19. Analysis of Variance table for Guessability Score

Source	**DF**	**Adjusted SS**	**Adjusted MS**	**F**	**P**
Age Group	4	24.821	6.20521	5.54	<0.001
Educational Background	2	2.846	1.42315	1.27	0.283
Active Car Usage	3	13.453	4.48418	4.01	0.009
Marital Status	2	10.023	5.01137	4.48	0.013
Gender	1	0.235	0.23495	0.21	0.647
Smoking Habits	1	0.064	0.06378	0.06	0.812
Error	184	205.913	1.11909		
Total	197	268.823			

- Age Factor: The box plot of the guessability score for different age groups was shown in Figure 5.10. The results of the Tukey's test for different age groups can be seen in Table 5.20. According to the results, fifth age group (60 – 69 yrs) had significantly low mean guessability performance. The performance of the fifth group significantly differs from the other groups.
- Active Car Driving Experience Effect: The box plot of guessability score for different active car driving experience groups is shown in Figure 5.11. Active car driving experience between 1-5 years had the highest mean guessability score whereas the not an active driver group had the lowest score. The results of Tukey's test are also shown in Table 5.21. Group 1 represents no active car driving experience, Group 2 represents the active car driving experience less than 1 year, Group 3 represents the active car driving experience between 1 to 5 years and Group 4 represents the active car driving experience more than 5 years. The results of Tukey's test shows that the mean guessability score of Group 1 significantly differs from Group 3 and 4 whereas other groups do not significantly differ from each other.

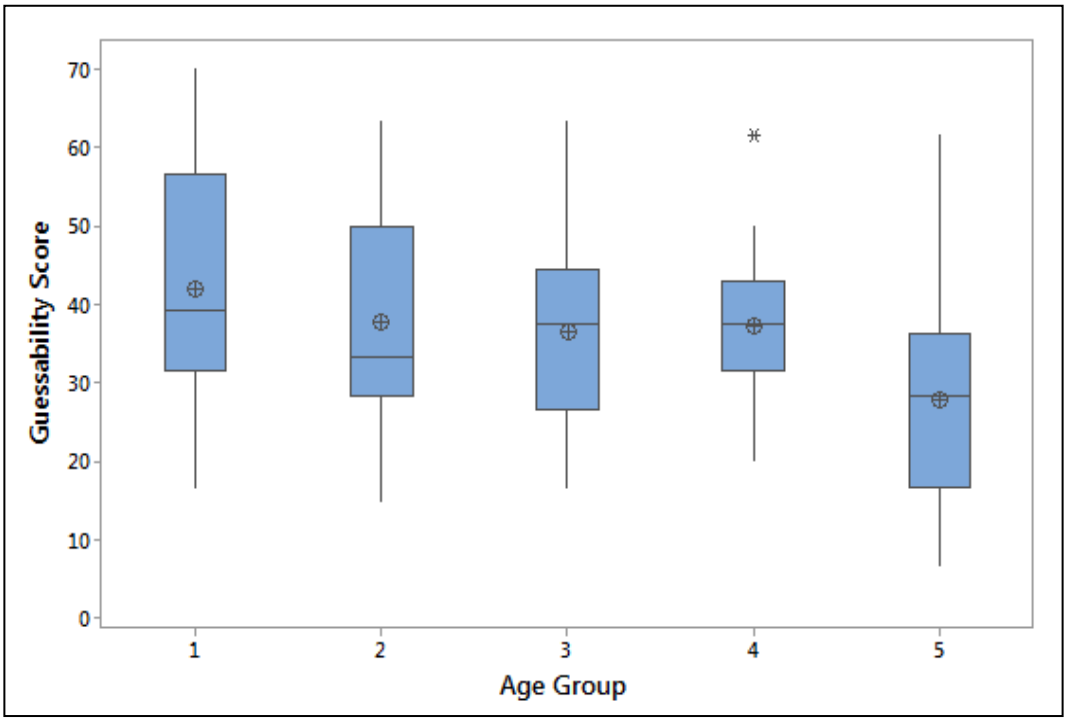


Figure 5.10. Box Plot of Guessability Score for Age Groups

Table 5.20. Results of Tukey's Test for Age Groups

Age Group Difference	Difference of Means	SE of Difference	T- value	Adj. P-value
Group 2 - Group 1	-0.223	0.243	-0.92	0.890
Group 3 - Group 1	-0.319	0.247	-1.29	0.697
Group 4 - Group 1	-0.148	0.247	-0.60	0.975
Group 5 - Group 1	-1.179	0.254	-4.64	<0.001
Group 3 - Group 2	-0.096	0.243	-0.40	0.995
Group 4 - Group 2	0.075	0.243	0.31	0.998
Group 5 - Group 2	-0.956	0.250	-3.83	0.002
Group 4 - Group 3	0.171	0.247	0.69	0.958
Group 5 - Group 3	-0.859	0.254	-3.38	0.008
Group 5 - Group 4	-1.031	0.254	-4.06	0.001

Group 1: 18-29 yrs; Group 2: 30-39 yrs, Group 3: 40-49 yrs; Group 4: 50-59 yrs; Group 5:60-69 yrs

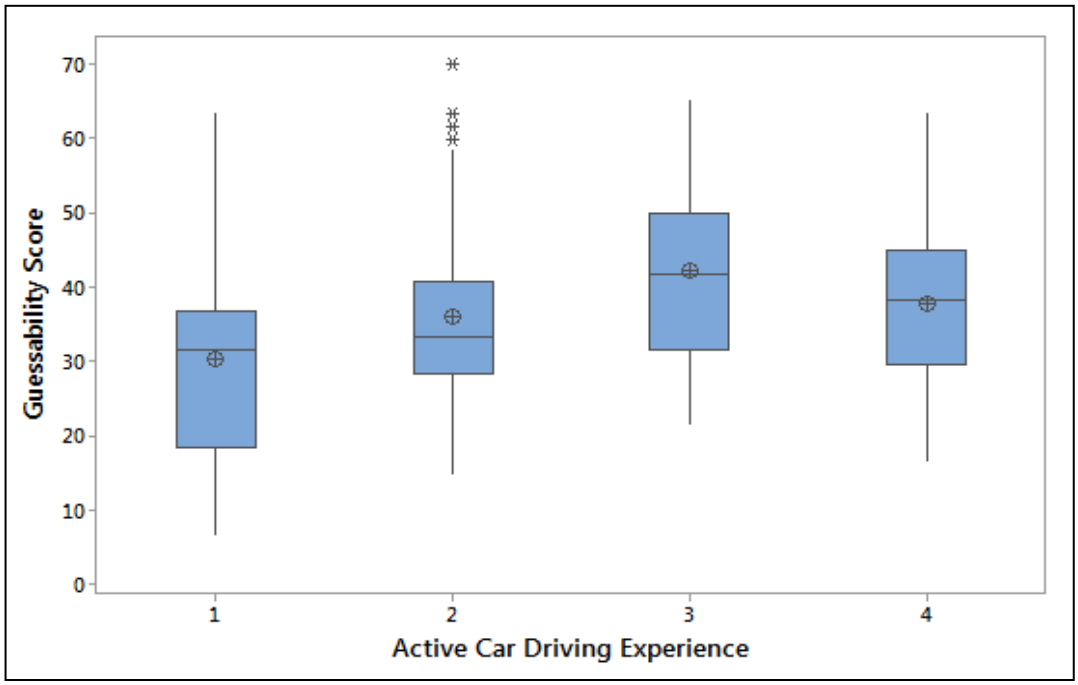


Figure 5.11. Box Plot of Guessability Score for Active Car Driving Experience

Table 5.21. Results of Tukey's test for Active Car Driving Experience Groups

Tool Group Difference	Difference of Means	SE of Difference	T- value	Adjusted p - value
Group 2 - Group 1	0.550	0.271	2.54	0.057
Group 3 - Group 1	1.049	0.222	4.72	<0.001
Group 4 - Group 1	0.713	0.225	3.71	0.009
Group 3 - Group 2	0.499	0.221	2.26	0.112
Group 4 - Group 2	0.163	0.224	0.73	0.886
Group 4 - Group 3	-0.337	0.229	-1.47	0.458

Group 1: Not an active driver; Group 2: Less than 1 yr.; Group 3: Between 1-5 yrs.; Group 4: More than 5 yrs.

- Marital Status Effect: Box plot of guessability score for different marital status was shown in Figure 5.12. The married group had the highest mean guessability score performance. Group 2 represents married, Group 3 represents married with children,

Group 1 represents single. The results of Tukey's test show that married group significantly differs from others, whereas the other groups do not differentiate.

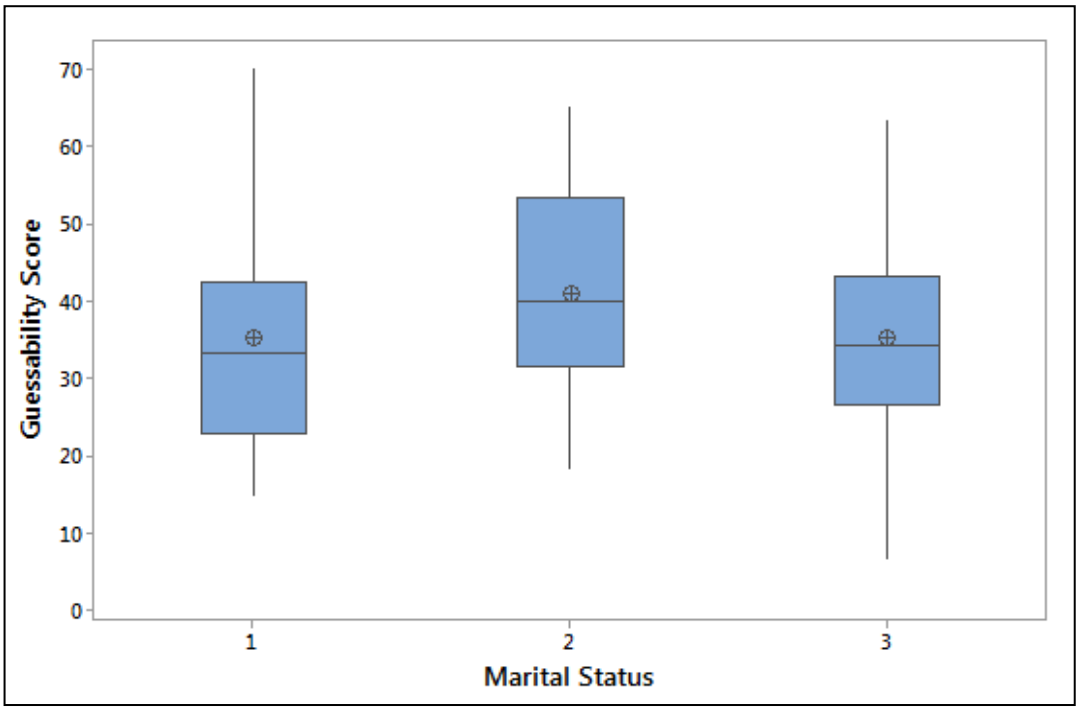


Figure 5.12. Box Plot of Guessability Score for Marital Status

Table 5.22. Results of Tukey's test for Marital Status Groups

Age Group Difference	Difference of Means	SE of Difference	T- value	Adj. P-value
Group 2 - Group 1	0.562	0.233	-2.42	0.044
Group 3 - Group 1	0.063	0.185	0.34	0.938
Group 3 - Group 2	-0.499	0.220	-2.27	0.062

Group 1: Single Group 2:Married; Group 3: Married with children,

5.3.4. Effect of User Factors on Cognitive Sign Design Features

Residuals of five cognitive sign design features; familiarity, concreteness, simplicity, meaningfulness and semantic closeness were not normally distributed. Therefore, ranking transformation for five sign design features was applied to conduct Analysis of Variance (ANOVA) and results were shared in below. Response was cognitive sign design features, whereas factors were user factors. In addition, sign type (warning, regulatory, information,

standing and parking, and roadworks) was included in factors to search the effect on each sign design features. Moreover, to understand which factor level was significantly different, multiple comparisons were conducted for further analysis. Fisher test was selected to be applied for factors which have more than two levels.

5.3.4.1. Analysis of Variance (ANOVA) for Familiarity and User Factors. Completely randomized design was performed for ANOVA analyses. Table 5.23 summaries the results.

Table 5.23. Analysis of Variance table for Familiarity

Source	DF	Adjusted SS	Adjusted MS	F	P
Age Group	4	8349068777	2087267194	197.10	<0.001
Educational Background	2	92861023	46430511	4,38	0.012
Active Car Usage	3	394769448	131589816	12.43	<0.001
Marital Status	2	701378769	350689384	33.12	<0.001
Gender	1	19957262	19957262	1.88	0.170
Smoking Habits	1	51864146	51864146	4.90	0.027
Sign Type	4	5969671900	1492417975	140.93	<0.001
Error	12.042	1.27524E+11	10589900		
Total	12059	1.44989E+11			

The results of completely randomized design show that familiarity is significantly affected by age factor, active car usage, marital status, sign type, educational background and smoking habits ($p<0.05$).

- Age Effect: The box plot of familiarity score for different age groups was shown in Figure 5.13. Group 1 represents 18-29 yrs, Group 2 represents 30-39 yrs, Group 3 represents 40-49 yrs, Group 4 represents 50-59 yrs, and Group 5 represents 60-69 yrs. According to the results of Fisher test, the familiarity score significantly differs from some groups. Group 1 and 3; Group 2 and 4 do not significantly differ from

each other whereas those couples differ from each other and also from Group 5. The Group 5 (60-69 years) had the lowest familiarity score performance.

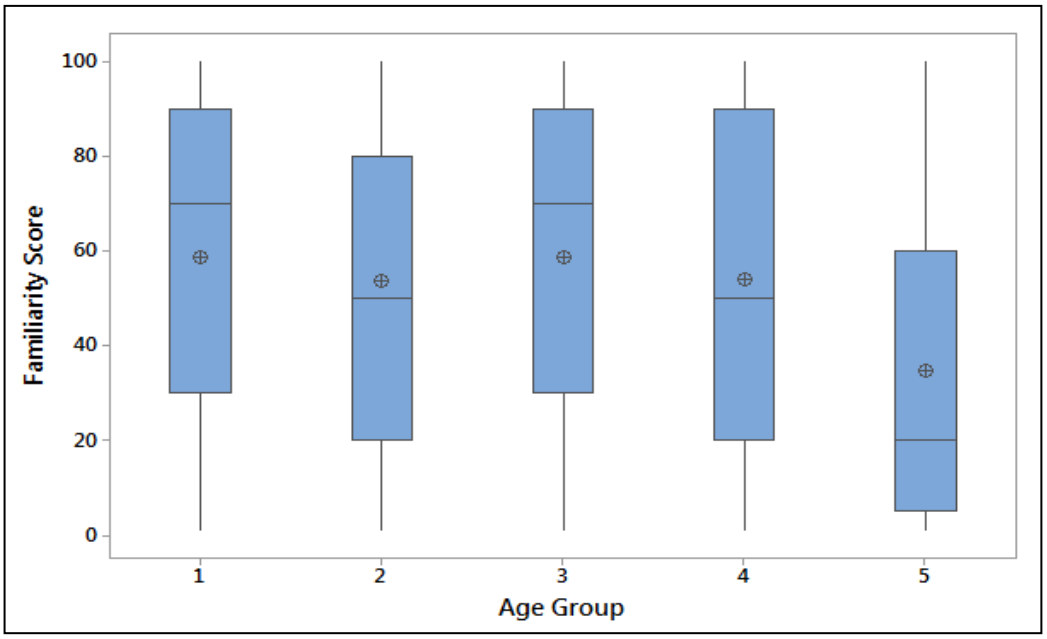


Figure 5.13. Boxplot of Age Groups

Table 5.24. Results of Fisher test for Age Groups

Difference of Levels	Difference of Means	SE of Difference	T- value	Adj. P- value
Group 2 - Group 1	-526.6	94.3	-5.59	<0.001
Group 3 - Group 1	-31.2	96.0	-0.32	0.745
Group 4 - Group 1	-478.6	95.4	-5.02	<0.001
Group 5 - Group 1	-2576.6	98.6	-26.12	<0.001
Group 3 - Group 2	495.5	94.9	5.22	<0.001
Group 4 - Group 2	48.1	94.3	0.51	0.610
Group 5 - Group 2	-2050.0	97.6	-21.01	<0.001
Group 4 - Group 3	-447.4	96.0	-4.66	<0.001
Group 5 - Group 3	-2545.5	99.2	-25.66	<0.001
Group 5 - Group 4	-2098.1	98.6	-21.27	<0.001

Group 1:18-29 yrs; Group 2:30-39 yrs; Group 3: 40-49 yrs; Group 4: 50-59 yrs; Group 5: 60-69 yrs.

- Active Car Usage Effect: The box plot of familiarity score for different active car driver groups was shown in Figure 5.14. Group 1 represents not active drivers, Group 2 represents <1 yr. experience, Group 3 represents 1-5 yrs. experience, Group 4 represents >5 yrs experience. Fisher test shows that Group 3 and 2 do not significantly differ from each other whereas differ from Group 1 and Group 4. Group 1 and Group 4 also differ from each other. Not an active car driver group had the lowest familiarity score performance.

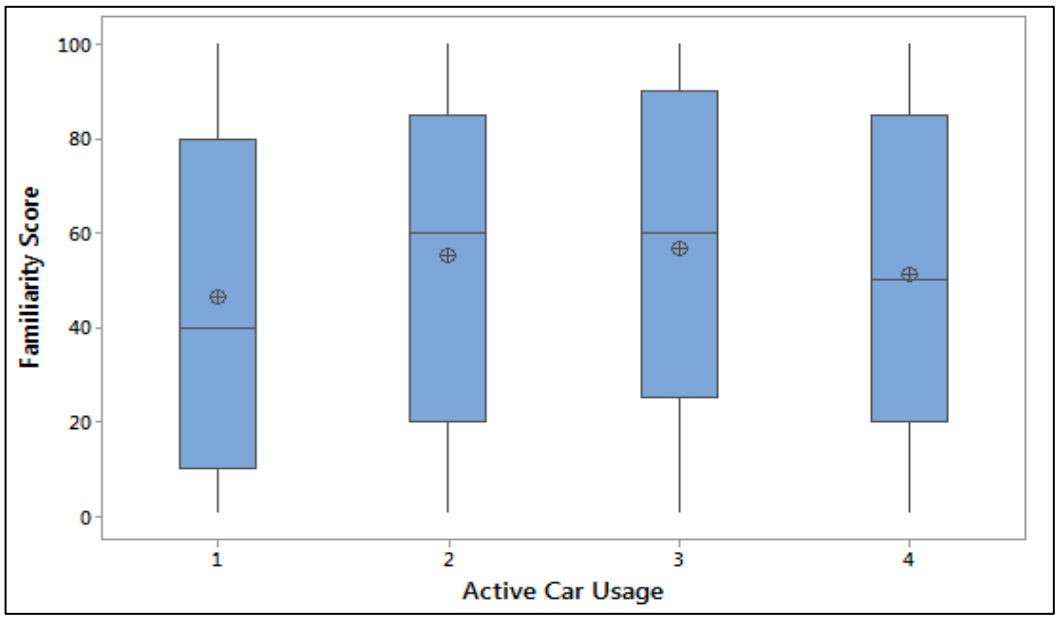


Figure 5.14. Boxplot of Active Car Driving Experience Groups

Table 5.25. Results of Fisher test for Active Car Driving Experience Groups

Difference of Levels	Difference of Means	SE of Difference	T- value	Adjusted p - value
Group 2 - Group 1	990.3	86.2	11.48	<0.001
Group 3 - Group 1	1148.7	88.0	13.06	<0.001
Group 4 - Group 1	554.2	89.4	6.20	<0.001
Group 3 - Group 2	158.5	88.0	1.80	0.072
Group 4 - Group 2	-436.1	89.4	-4.88	<0.001
Group 4 - Group 3	-594.5	91.1	-6.52	<0.001

Group 1: Not an active driver; Group 2: Less than 1 yr.; Group 3: Between 1-5 yrs; Group 4: More than 5 yrs

- Marital Status Effect: The box plot of familiarity score for different marital status groups was shown in Figure 5.15. Group 1 represents single, Group 2 represents married, and Group 3 represents married with children group. Results of Fisher test show that all groups significantly differ from each other. The Group 3 (married with children) had the lowest familiarity score performance.

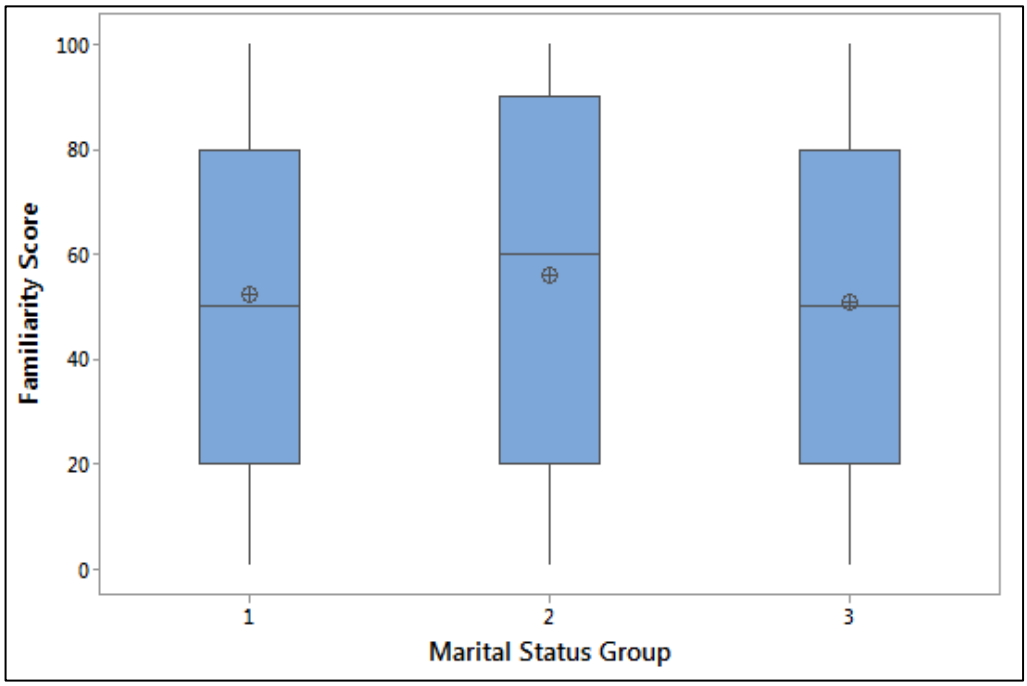


Figure 5.15. Boxplot of Marital Status Groups

Table 5.26. Results of Fisher test for Marital Status Groups

Difference of Levels	Difference of Means	SE of Difference	T- value	Adj. P-value
Group 2 - Group 1	403.0	90.0	4.48	<0.001
Group 3 - Group 1	-174.8	71.3	-2.45	0.014
Group 3 - Group 2	-577.8	85.0	-6.80	<0.001

Group 1:Single; Group 2: Married, Group 3: Married with children

- Sign Type Effect: The box plot of familiarity score for different sign types was shown in Figure 5.16. The results of Fisher test shows that all groups significantly differ from each other whereas standing and parking signs do not differ significantly

from regulatory and informational signs. Results show that roadwork sign had the highest familiarity score results

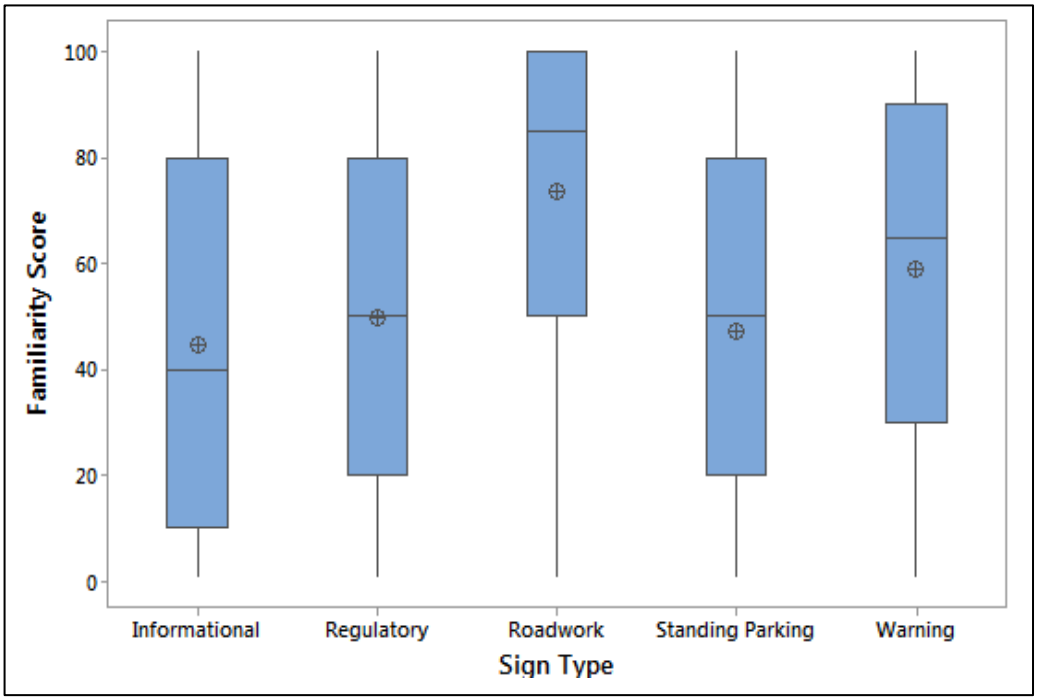


Figure 5.16. Boxplot of Sign Type Groups

Table 5.27. Results of Fisher test for Sign Type Groups

Difference of Levels	Difference of Means	SE of Difference	T- value	Adj. P- value
Regulatory-Information	483.7	84.5	5.72	<0.001
Roadwork-Information	3173	249	12.76	<0.001
Standing - Information	257	153	1.68	0.094
Warning - Information	1561.9	83.8	18.64	<0.001
Roadwork - Regulatory	2689	245	10.97	<0.001
Standing- Regulatory	-226	148	-1.53	0.126
Warning - Regulatory	1078.1	73.1	14.75	<0.001
Standing - Roadwork	-2915	277	-10.54	<0.001
Warning -Roadwork	-1611	245	-6.58	<0.001
Warning -Standing	1305	147	8.85	<0.001

- Educational Background Effect: The box plot of familiarity score for different educational background groups was shown in Figure 5.17. Group 1 represents primary/secondary school, Group 2 represents high school, and Group 3 represents university or higher degree background. The results of Fisher test show that Group 2 and 1 do not significantly differ from each other's whereas Group 3 significantly differs from all. The Group 3 (University or higher) had the highest familiarity score performance.

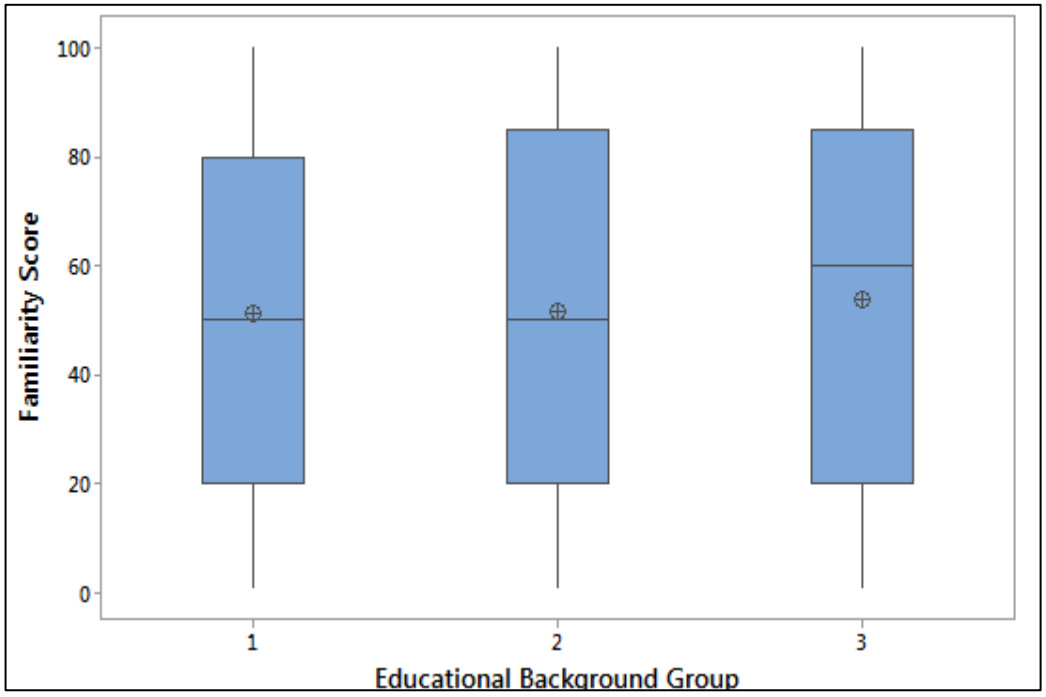


Figure 5.17. Boxplot of Educational Background Groups

Table 5.28. Results of Fisher test for Educational Background Groups

Difference of Levels	Difference of Means	SE of Difference	T- value	Adj. P- value
Group 2 - Group 1	17.8	77.9	0.23	0.819
Group 3 - Group 1	267.6	76.7	3.49	<0.001
Group 3 - Group 2	249.7	77.3	3.23	0.001

Group 1: Primary Secondary School; Group 2: High School; Group 3: University and / or higher degree

- Smoking Habits Effect: The box plot of familiarity score was shown in Figure 5.18. Smokers performed better than non-smokers.

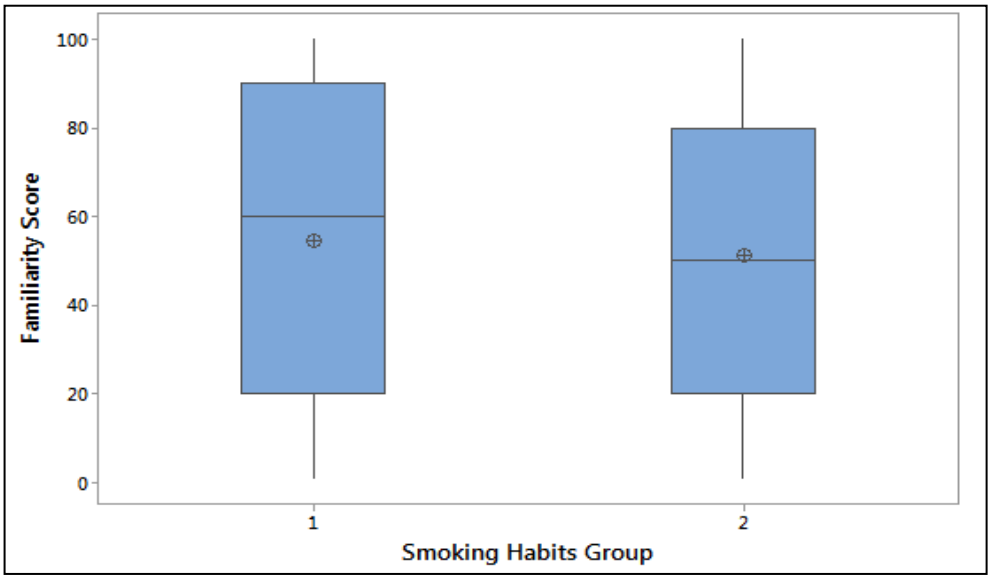


Figure 5.18. Boxplot of Smoking Habit Groups

5.3.4.2. Analysis of Variance (ANOVA) for Concreteness and User Factors. Results of ANOVA showed in Table 5.29. The results of completely randomized design show that concreteness is significantly affected by age factor, active car usage, sign type, marital status and gender ($p<0.05$).

Table 5.29. Analysis of Variance table for Concreteness

Source	DF	Adjusted SS	Adjusted MS	F	P
Age Group	4	2127524415	531881104	47.54	<0.001
Educational Background	2	20529376	10264688	0.92	0.400
Active Car Usage	3	168659401	56219800	5.01	0.001
Marital Status	2	133040264	66520132	5.95	0.003
Gender	1	76455881	76455881	6.83	0.009
Smoking Habits	1	642236	642236	0.06	0.811
Sign Type	4	6814649691	1703662423	152.26	<0.001
Error	12.042	1.34736E+11	11188865		
Total	12.059	1.44960E+11			

- Age Effect: The box plot of concreteness score for different age groups was shown in Figure 5.19. Group 1 represents 18-29 yrs, Group 2 represents 30-39 yrs, Group 3 represents 40-49 yrs, Group 4 represents 50-59 yrs, and Group 5 represents 60-69 yrs. The results of Fisher test show that Group 1,2 and 3 do not significantly differ from each other whereas differ from Group 4 and Group 5. The Group 5 (60-69 yrs) had the lowest concreteness score performance and differs from all groups.

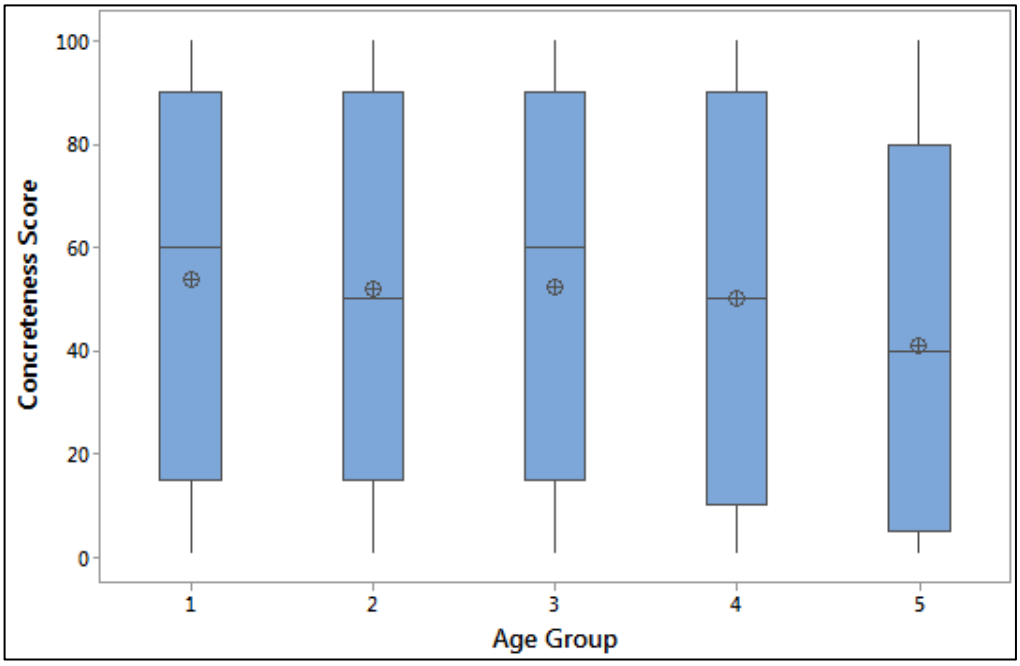


Figure 5.19. Boxplot of Age Groups

- Active Car Usage Effect: The box plot of familiarity score for different active car driver groups was shown in Figure 5.20. Group 1 represents not an active driver, Group 2 represents <1 yr. experience, Group 3 represents 1-5 yrs experience, Group 4 represents >5 yrs experience. The results of Fisher test show that Group 3 and 2 do not significantly differ from each other whereas differ from Group 1 and Group 4. Group 1 and Group 4 also differ from each other. The Group 1 (not an active car driver) had the lowest concreteness score performance.
- Sign Type Effect: The box plot of concreteness score for different sign types was shown in Figure 5.21. The results of Fisher test shows that all groups significantly differ from each other except standing and parking and regulatory signs. The

standing and parking and regulatory signs had the lowest concreteness score performance.

Table 5.30. Results of Fisher test for Age Groups

Difference of Levels	Difference of Means	SE of Difference	T- value	Adj. P- value
Group 2 - Group 1	-164.4	96.7	-1.70	0.089
Group 3 - Group 1	-134.5	98.5	-1.37	0.172
Group 4 - Group 1	-407.4	97.9	-4.16	<0.001
Group 5 - Group 1	-1423	101	-14,06	<0.001
Group 3 - Group 2	29.9	97.3	0.31	0.759
Group 4 - Group 2	-243.0	96.7	-2.51	0.012
Group 5 - Group 2	-1258	100	-12.57	<0.001
Group 4 - Group 3	-272.9	98.5	-2.77	0.006
Group 5 - Group 3	-1288	102	-12.66	<0.001
Group 5 - Group 4	-1015	101	-10.03	<0.001

Group 1: 18-29 yrs; Group 2: 30-39 yrs; Group 3: 40-49 yrs; Group 4: 50-59 yrs; Group 5: 60-69 yrs

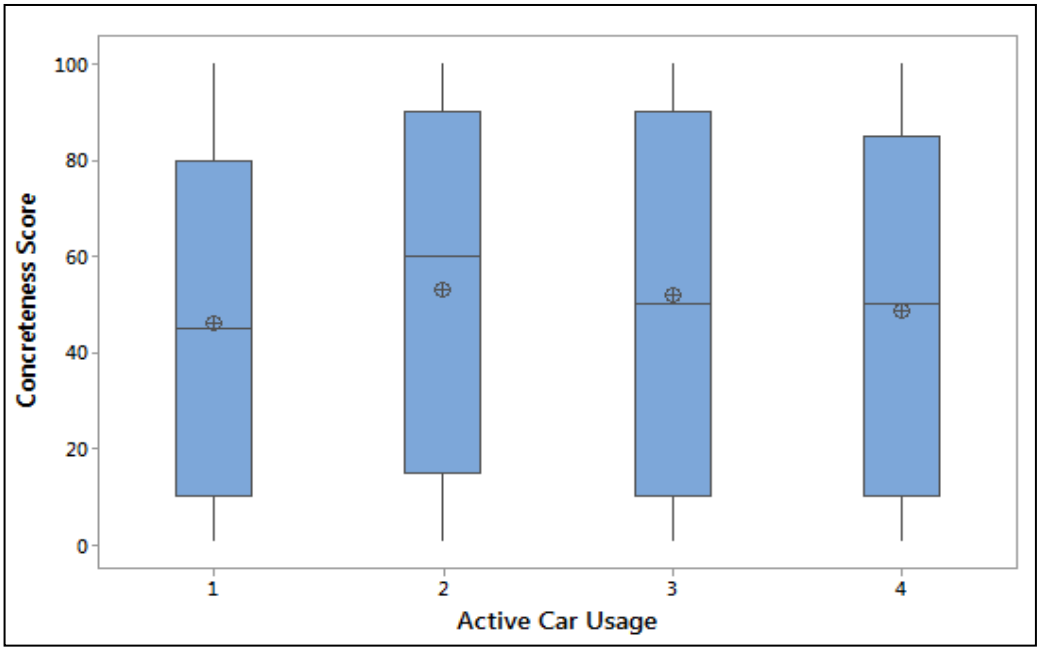


Figure 5.20. Boxplot of Active Car Driver Experience Groups

Table 5.31. Results of Fisher test for different Active Car Driving Experience Groups

Difference of Levels	Difference of Means	SE of Difference	T- value	Adjusted p - value
Group 2 - Group 1	735.9	86.6	8.49	<0.001
Group 3 - Group 1	639.0	88.4	7.23	<0.001
Group 4 - Group 1	260.6	89.9	2.90	0.004
Group 3 - Group 2	-97.0	88.4	-1.10	0.273
Group 4 - Group 2	-475.3	89.9	-5.29	<0.001
Group 4 - Group 3	-378.4	91.6	-4.13	<0.001

Group 1: Not an active driver; Group 2: Less than 1 yr.; Group 3: Between 1-5 yrs.; Group 4: More than 5 yrs.

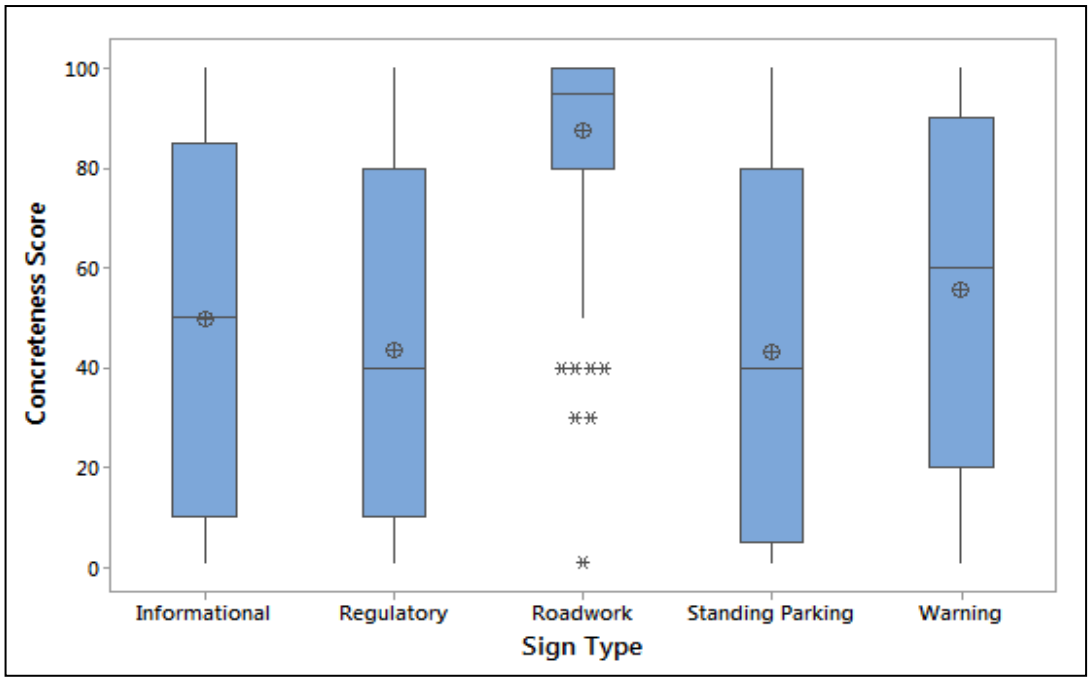


Figure 5.21. Boxplot of Sign Type Groups

- Marital Status Effect: The box plot of concreteness score for different marital status groups was shown in Figure 5.22. Group 1 represents single, Group 2 represents married, and Group 3 represents married with children. The results of Fisher test show that Group 2 significantly differs from Group 3.

Table 5.32. Results of Fisher Test for Sign Type Groups

Difference of Levels	Difference of Means	SE of Difference	T- value	Adj. P- value
Regulatory-Information	-491.0	84.3	-5.83	<0.001
Roadwork-Information	3932	248	15.87	<0.001
Standing - Information	-666	153	-4.35	<0.001
Warning - Information	778.2	83.5	9.32	<0.001
Roadwork - Regulatory	4423	244	18.10	<0.001
Standing- Regulatory	-175	147	-1.19	0.236
Warning - Regulatory	1269.2	72.8	17.42	<0.001
Standing - Roadwork	-4598	276	-16.68	<0.001
Warning -Roadwork	-3154	244	-12.92	<0.001
Warning -Standing	1444	147	9.83	<0.001

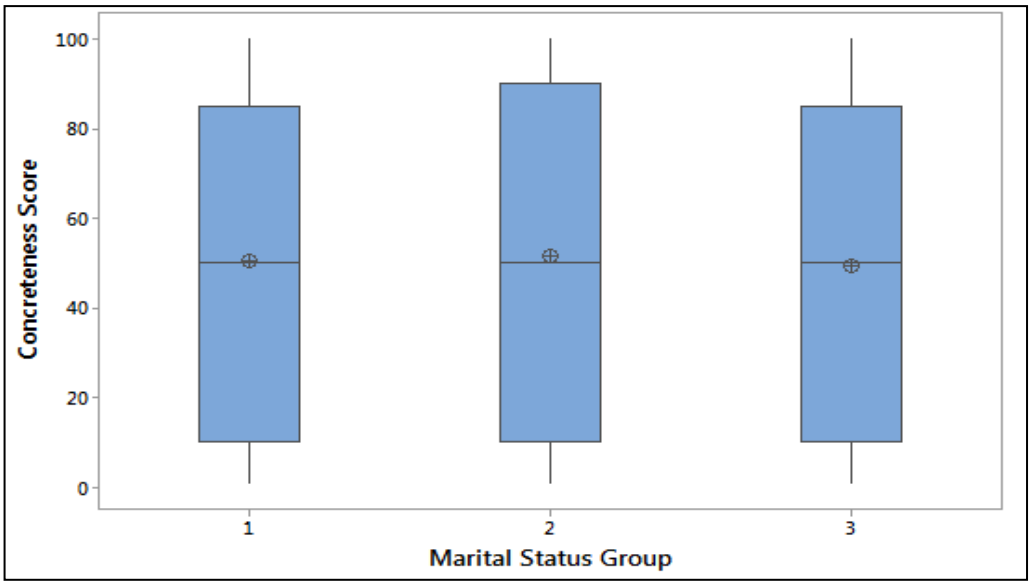


Figure 5.22. Boxplot of Marital Status Group

Table 5.33. Results of Fisher Test for different Marital Status Groups

Difference of Levels	Difference of Means	SE of Difference	T- value	Adj. P- value
Group 2 - Group 1	89.9	90.1	1.00	0.319
Group 3 - Group 1	-138.8	71.4	-1.94	0.052
Group 3 - Group 2	-228.7	85.1	-2.69	0.007

Group 1: Single; Group 2: Married; Group 3: Married with children

- Gender Effect: The box plot of concreteness score for gender group was shown in Figure 5.23. Group 1 represents men, Group 2 represents women. Results show that groups significantly differ from each other and men had the higher concreteness score performance than women.

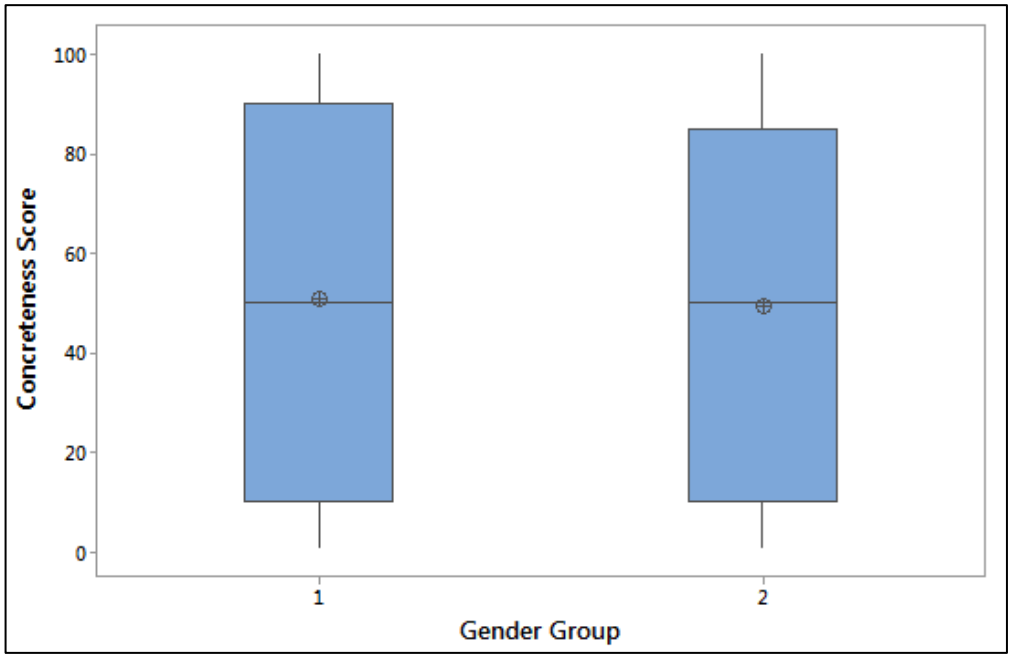


Figure 5.23. Boxplot of Gender Groups

5.3.4.3. Analysis of Variance (ANOVA) for Simplicity and User Factors. Completely randomized design was performed for ANOVA analyses. Table 5.34 summaries the results. The results of completely randomized design show that simplicity is significantly affected by age factor, sign type, educational background, active car usage, marital status and smoking habits ($p<0.05$).

- Age Effect: The box plot of simplicity score for different age groups was shown in Figure 5.24. Group 1 represents 18-29 yrs, Group 2 represents 30-39 yrs, Group 3 represents 40-49 yrs, Group 4 represents 50-59 yrs, and Group 5 represents 60-69 yrs. The results of Fisher test show that Group 1, 2 and Group 4 do not significantly differ from each other. Group 4 and Group 3 do not differ from each other. Group 5 differs significantly from all groups. The Group 5 (60-69 years) had the lowest concreteness score performance.

Table 5.34. Analysis of Variance table for Simplicity

Source	DF	Adjusted SS	Adjusted MS	F	P
Age Group	4	1169892249	292473062	26.62	<0.001
Educational Background	2	91069669	45534834	4.15	0.016
Active Car Usage	3	122125453	40708484	3.71	0.011
Marital Status	2	100390167	50195083	4.57	0.010
Gender	1	20258704	20258704	1.84	0.174
Smoking Habits	1	111596374	111596374	10.16	0.001
Sign Type	4	9817526196	2454381549	223.43	<0.001
Error	12.042	1.32283E+11	10985101		
Total	12.059	1.44298E+11			

- Active Car Usage Effect: The box plot of simplicity score for different active car driver groups was shown in Figure 5.25. Group 1 represents not an active driver, Group 2 represents <1 yr. experience, Group 3 represents 1-5 yrs experience, Group 4 represents >5 yrs experience. The results of Fisher test show that Group 3 and 2 do not significantly differ from each other whereas differ from Group 1 and Group 4. Group 1 and Group 4 also differ from each other. The Group 1 (not an active car driver) had the lowest simplicity score performance.
- Sign Type Effect: The box plot of simplicity score for different sign types was shown in Figure 5.26. The results of Fisher test shows that all groups significantly differ from each other .
- Smoking Habit Effect: The box plot of simplicity score for smoking habit was shown in Figure 5.27. Group 1 represents the smokers whereas Group 2 is the non-smokers. Results show that smokers performed better simplicity score.
- Marital Status Effect: The box plot of simplicity score for different marital status groups was shown in Figure 5.28. Group 1 represents single, Group 2 represents married, and Group 3 represents married with children group. The results of Fisher

test show that Groups 1 and 3 do not significantly differ from each other but differ significantly from Group 2. The Group 1 and 3 (single and married with children) had the lowest simplicity score performance.

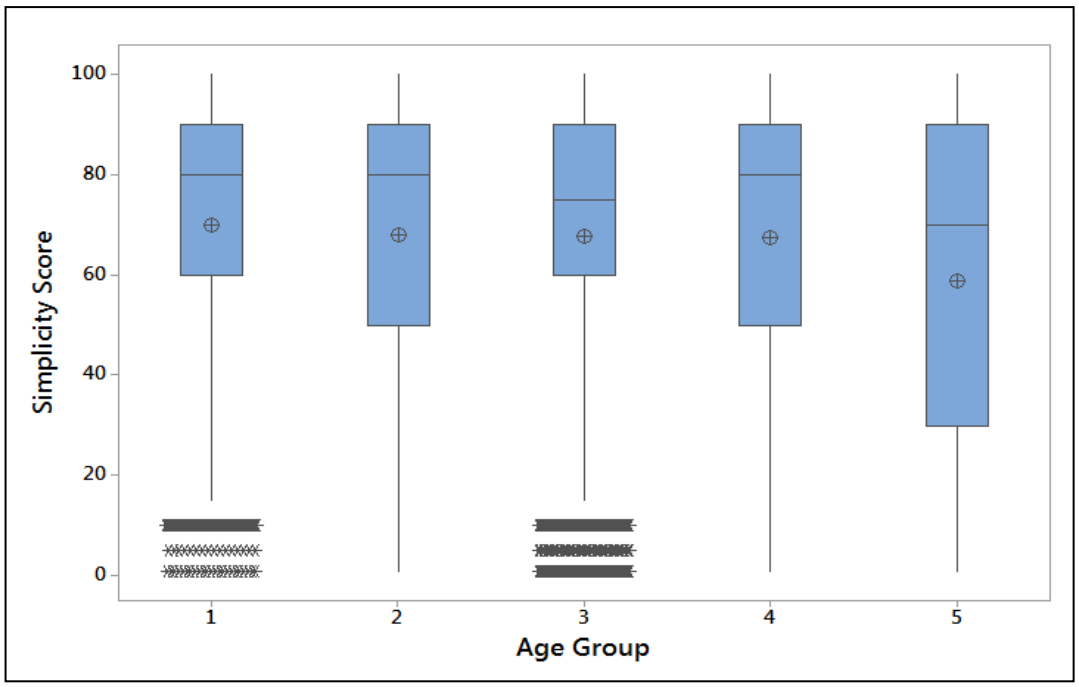


Figure 5.24. Boxplot of Age Groups

Table 5.35. Results of Fisher test for Age Groups

Difference of Levels	Difference of Means	SE of Difference	T- value	Adj. P- value
Group 2 - Group 1	-7.5	96.9	-0.08	0.939
Group 3 - Group 1	-225.4	98.7	-2.28	0.022
Group 4 - Group 1	-60.8	98.0	-0.62	0.535
Group 5 - Group 1	-1042	101	-10.27	<0.001
Group 3 - Group 2	-217.9	97.5	-2.23	0.025
Group 4 - Group 2	-53.4	96.9	-0.55	0.582
Group 5 - Group 2	-1034	100	-10.31	0.000
Group 4 - Group 3	164.5	98.7	1.67	0.095
Group 5 - Group 3	-816	102	-8.00	<0.001
Group 5 - Group 4	-981	101	-9.67	<0.001

Group 1:18-29 yrs; Group 2:30-39 yrs; Group 3: 40-49 yrs; Group 4: 50-59 yrs; Group 5: 60-69 yrs

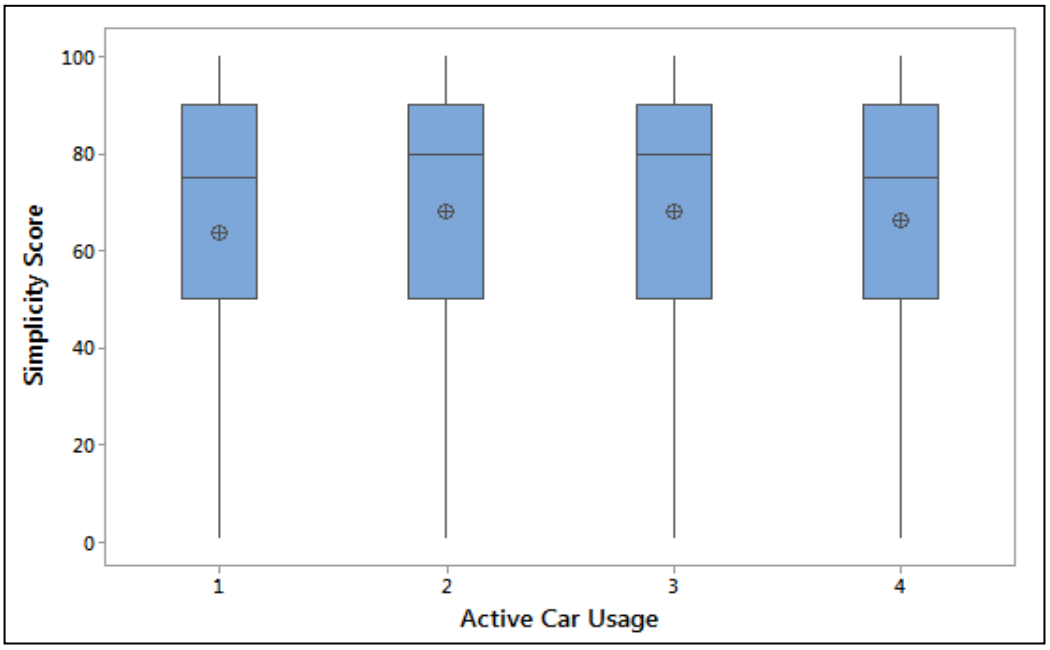


Figure 5.25. Boxplot of Active Car Driving Experience Group

Table 5.36. Results of Fisher test for Active Car Driving Experience Groups

Difference of Levels	Difference of Means	SE of Difference	T- value	Adjusted p - value
Group 2 - Group 1	462.1	86.6	5.34	<0.001
Group 3 - Group 1	541.0	88.3	6.12	<0.001
Group 4 - Group 1	242.4	89.8	2.70	0.007
Group 3 - Group 2	78.9	88.3	0.89	0.372
Group 4 - Group 2	-219.7	89.8	-2.45	0.014
Group 4 – Group 3	-298.7	91.5	-3.26	0.001

Group 1: Not an active driver; Group 2: Less than 1 yr.; Group 3: Between 1-5 yrs.; Group 4: More than 5 yrs.

- Educational Background Effect: The box plot of simplicity score for different educational background groups was shown in Figure 5.29. Group 1 represents primary/secondary school, Group 2 represents high school, and Group 3 represents university or higher degree background. The results of Fisher test shows that Group

2 and 3 do not significantly differ from Group 1, whereas significantly differ from each other.

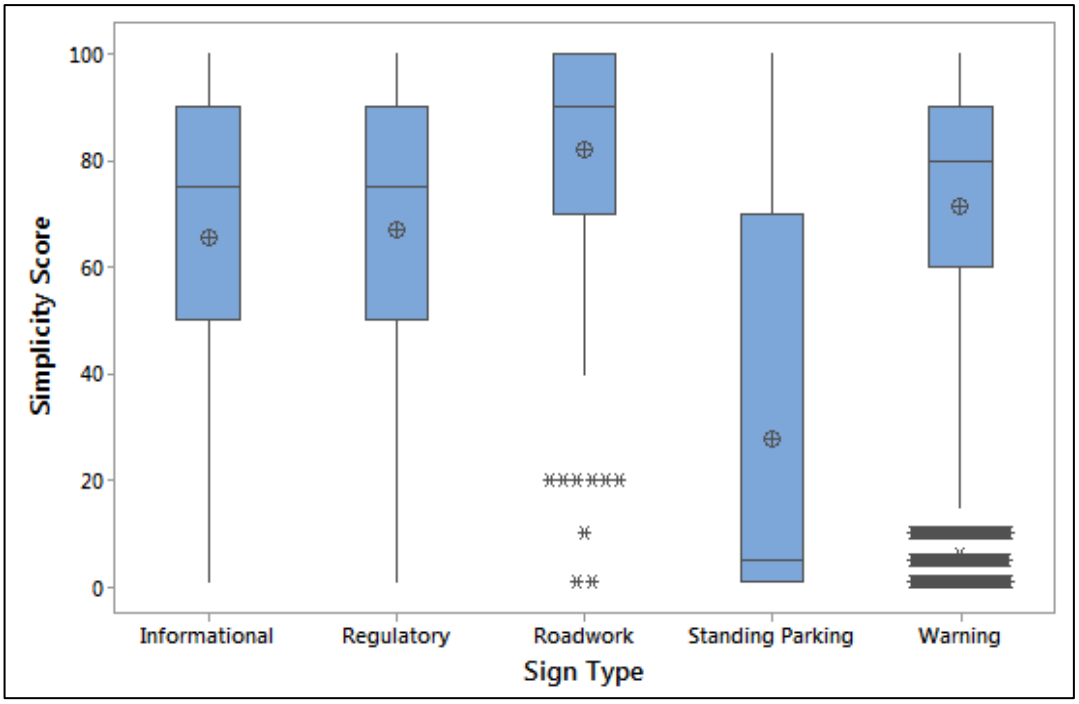


Figure 5.26. Boxplot of Sign Type Groups

Table 5.37. Results of Fisher test for Sign Type Groups

Difference of Levels	**Difference of Means**	**SE of Difference**	**T- value**	**Adj. P-value**
Regulatory-Information	192.7	83.1	2.32	0.020
Roadwork-Information	2077	244	8.49	<0.001
Standing - Information	-3344	151	-22.16	<0.001
Warning - Information	769.2	82.4	9.33	<0.001
Roadwork - Regulatory	1884	241	7.81	<0.001
Standing- Regulatory	-3537	145	-24.33	<0.001
Warning - Regulatory	576.5	71.9	8.02	<0.001
Standing - Roadwork	-5421	272	-19.93	<0.001
Warning -Roadwork	-1307	241	-5.43	<0.001
Warning -Standing	4114	145	28.87	<0.001

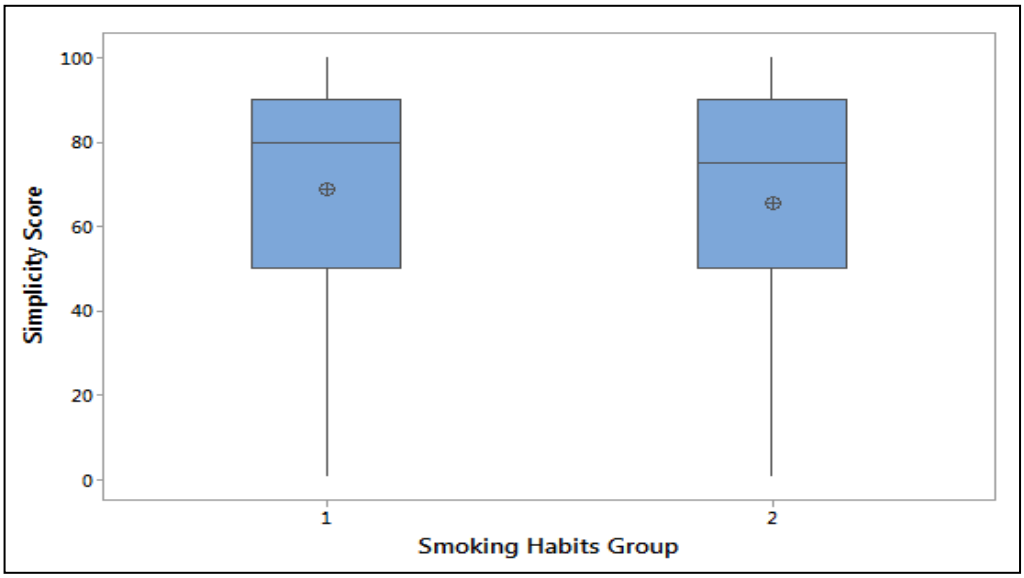


Figure 5.27. Boxplot of Smoking Habit Groups

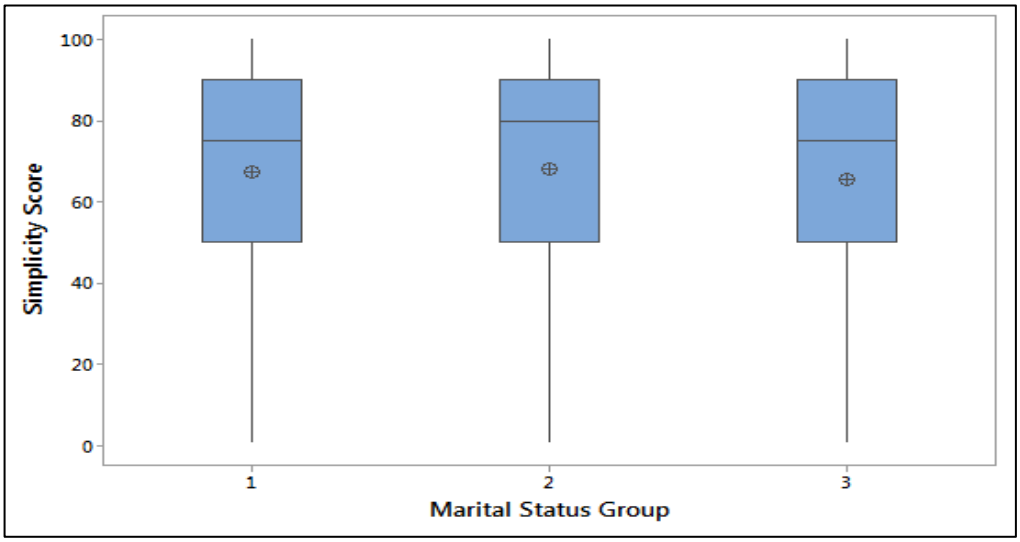


Figure 5.28. Boxplot of Marital Status Groups

Table 5.38. Results of fisher test for Marital Status Groups

Difference of Levels	Difference of Means	SE of Difference	T- value	Adj. P- value
Group 2 - Group 1	206.2	89.9	2.29	0.022
Group 3 - Group 1	-133.3	71.2	-1.87	0.061
Group 3 - Group 2	-339.5	84.9	-4.00	<0.001

Group 1:Single; Group 2: Married; Group 3: Married with children

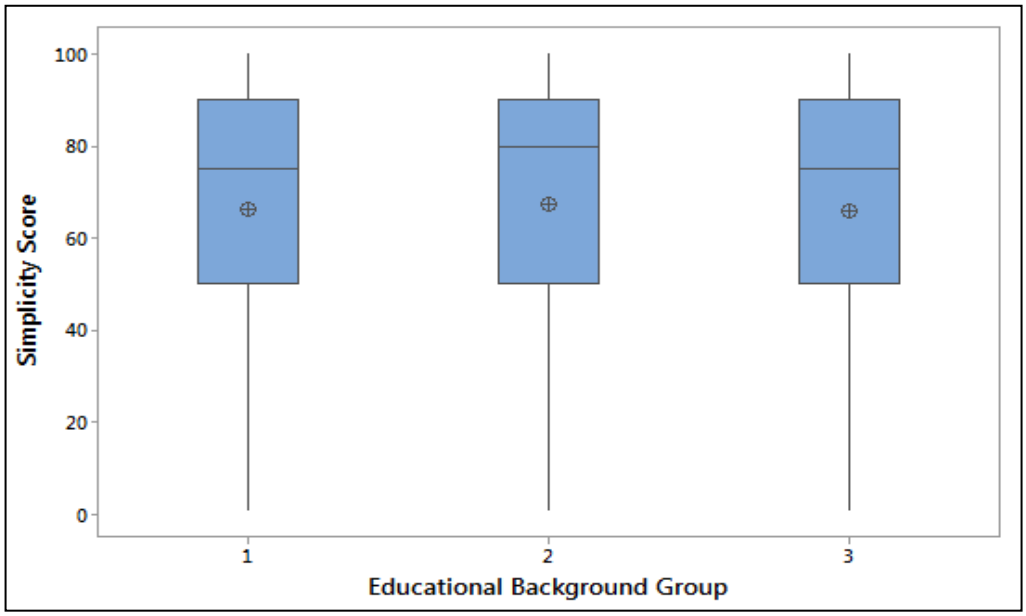


Figure 5.29. Boxplot of Educational Background Groups

Table 5.39. Results of Fisher test for different Educational Backgrounds

Age Group Difference	Difference of Means	SE of Difference	T- value	Adj. P- value
Group 2 - Group 1	84.9	77.7	1.09	0.275
Group 3 - Group 1	-71.9	76.6	-0.94	0.348
Group 3 - Group 2	-156.8	77.2	-2.03	0.042

Group 1: Primary/Secondary School; Group 2: High School; Group 3: University and/or higher degree

5.3.4.4 Analysis of Variance (ANOVA) for Meaningfulness and User Factors. Completely randomized design was performed for ANOVA analyses. Table 5.40 summaries the results. The results of completely randomized design show that meaningfulness is significantly affected by age factor, active car usage, sign type, gender and smoking habits ($p<0.05$).

- Age Effect: The box plot of meaningfulness score for different age groups was shown in Figure 5.30. Group 1 represents 18-29 yrs, Group 2 represents 30-39 yrs, Group 3 represents 40-49 yrs, Group 4 represents 50-59 yrs, and Group 5 represents 60-69 yrs. The results of Fisher test show that Group 3, 4 and 1 do not significantly

differ from each other, whereas differ from Group 5 and Group 2. The Group 5 (60-69 yrs) had the lowest meaningfulness score performance.

Table 5.40. Analysis of Variance table for Meaningfulness

Source	DF	Adjusted SS	Adjusted MS	F	P
Age Group	4	3714240516	928560129	84.68	<0.001
Educational Background	2	56262474	28131237	2.57	0.077
Active Car Usage	3	149948475	49982825	4.56	0.003
Marital Status	2	138055675	69027837	6.29	0.052
Gender	1	142980132	142980132	13.04	<0.001
Smoking Habits	1	111596374	111596374	10.16	0.020
Sign Type	4	7827517753	1956879438	178.45	<0.001
Error	12.042	1.32054E+11	10966094		
Total	12.059	1.45073E+11			

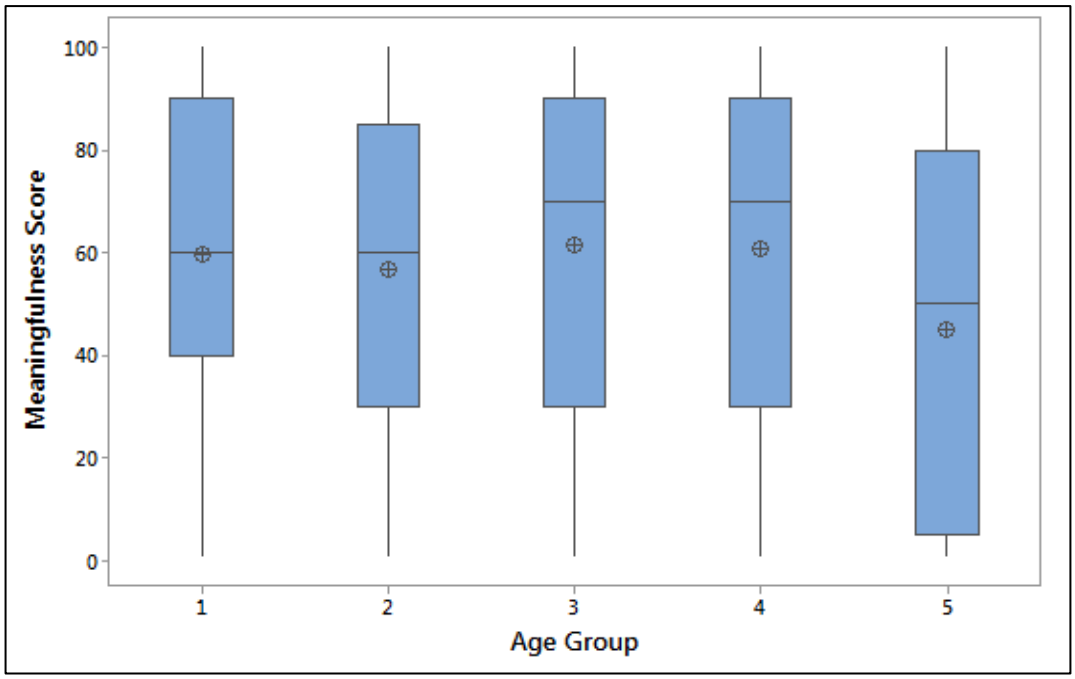


Figure 5.30. Boxplot of Age Groups

Table 5.41. Results of Fisher test for Age Groups

Difference of Levels	Difference of Means	SE of Difference	T- value	Adj. P- value
Group 2 - Group 1	-276.9	96.2	-2.88	0.004
Group 3 - Group 1	184.1	98.0	1.88	0.060
Group 4 - Group 1	156.8	97.4	1.61	0.107
Group 5 - Group 1	-1516	101	-15.06	<0.001
Group 3 - Group 2	461.0	96.9	4.76	<0.001
Group 4 - Group 2	433.8	96.2	4.51	<0.001
Group 5 - Group 2	-1239.5	99.6	-12.45	<0.001
Group 4 - Group 3	-27.3	98.0	-0.28	0.781
Group 5 - Group 3	-1701	101	-16.79	<0.001
Group 5 - Group 4	-1673	101	-16.62	<0.001

Group 1: 18-29 yrs; Group 2: 30-39 yrs; Group 3: 40-49 yrs; Group 4: 50-59 yrs, Group 5: 60-69 yrs

- Active Car Usage Effect: The box plot of meaningfulness score for different active car driver groups was shown in Figure 5.31. Group 1 represents not an active driver, Group 2 represents <1 yr. experience, Group 3 represents 1-5 yrs experience, Group 4 represents >5 yrs experience. The results of Fisher test show that all groups significantly differ from each other. The Group 1 (not an active car driver) had the lowest meaningfulness score performance.
- Sign Type Effect: The box plot of meaningfulness score for different sign types was shown in Figure 5.32. The results of Fisher test shows that all groups significantly differ from each other.
- Smoking Habit Effect: The box plot of meaningfulness score for smoking habit was shown in Figure 5.33. Results show that smokers performed better than non-smokers.
- Gender Effect: The box plot of meaningfulness score was shown in Figure 5.34. Results show that men performed better in meaningfulness score than women.

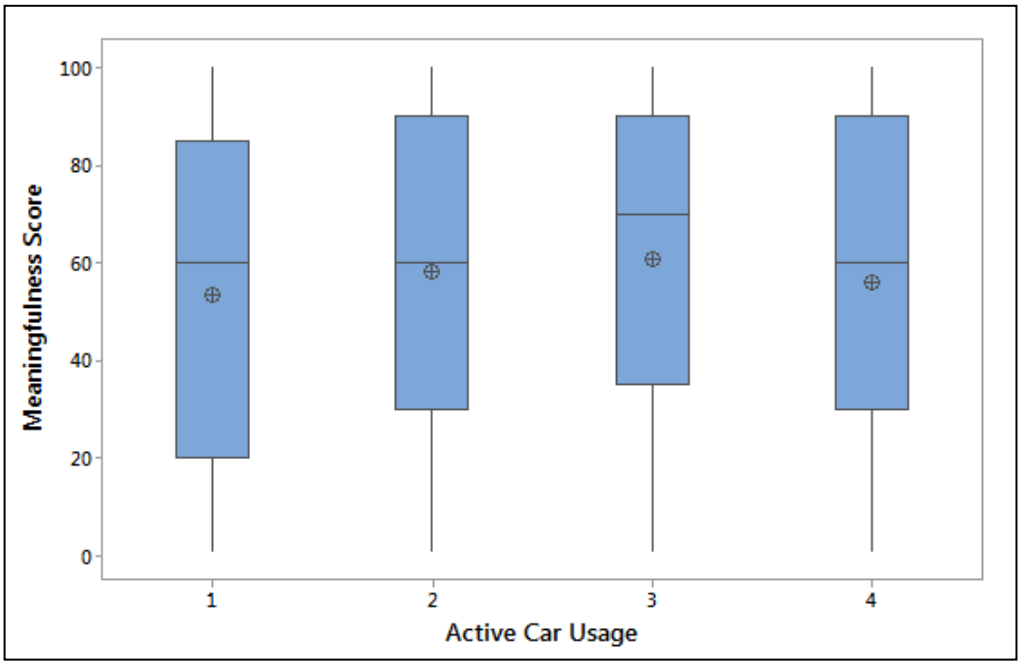


Figure 5.31. Boxplot of Active Car Driver Experience Groups

Table 5.42. Results of Fisher Test for Active Car Driving Experience Groups

Difference of Levels	Difference of Means	SE of Difference	T- value	Adjusted p-value
Group 2 - Group 1	512.9	86.7	5.92	<0.001
Group 3 - Group 1	748.7	88.5	8.46	<0.001
Group 4 - Group 1	247.0	89.9	2.75	0.006
Group 3 - Group 2	235.9	88.5	2.67	0.008
Group 4 - Group 2	-265.8	89.9	-2.96	0.003
Group 4 – Group 3	-501.7	91.6	-5.48	<0.001

Group 1: Not an active driver; Group 2: Less than 1 yr.; Group 3: Between 1-5 yrs.; Group 4: More than 5 yrs.

5.3.4.5. Analysis of Variance (ANOVA) for Semantic Closeness and User Factors. Completely randomized design was performed for ANOVA analyses. Table 5.44 summaries the results. The results of completely randomized design show that semantic closeness is significantly affected by age factor, active car usage, sign type, educational background, marital status and gender ($p<0.05$).

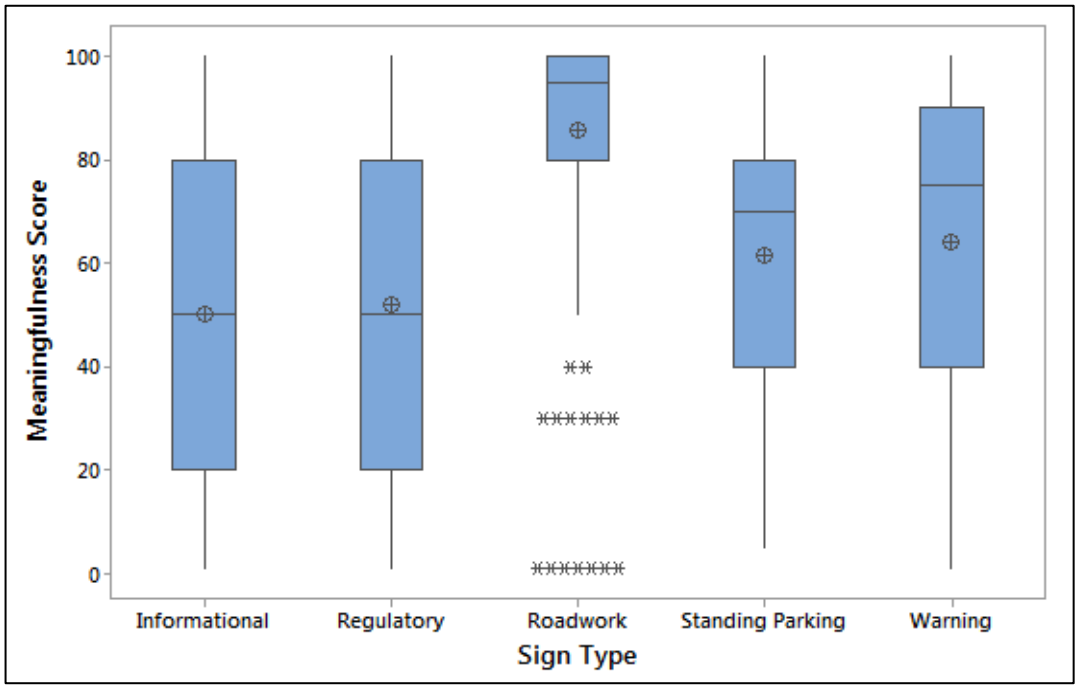


Figure 5.32. Boxplot of Sign Type Groups

Table 5.43. Results of Fisher test for Sign Type Groups

Difference of Levels	Difference of Means	SE of Difference	T- value	Adj. P- value
Regulatory-Information	199.7	84.0	2.38	0.017
Roadwork-Information	4059	247	16.43	<0.001
Standing - Information	1009	152	6.62	<0.001
Warning - Information	1551.3	83.3	18.63	<0.001
Roadwork - Regulatory	3859	244	15.84	<0.001
Standing- Regulatory	809	147	5.51	<0.001
Warning - Regulatory	1351.5	72.6	18.61	<0.001
Standing - Roadwork	-3050	275	-11.10	<0.001
Warning -Roadwork	-2507	234	-10.30	<0.001
Warning -Standing	542	146	3.70	<0.001

- Age Effect: The box plot of semantic closeness score for different age groups was shown in Figure 5.35. Group 1 represents 18-29 yrs, Group 2 represents 30-39 yrs, Group 3 represents 40-49 yrs, Group 4 represents 50-59 yrs, and Group 5 represents 60-69 yrs. The results of Fisher test show that Group 3, 4 and 1 do not significantly

differ from each other whereas all differ from Group 5. The Group 5 (60-69 yrs) had the lowest semantic closeness score performance.

- Active Car Usage Effect: The box plot of semantic closeness score for different active car driver groups was shown in Figure 5.36. Group 1 represents not an active driver, Group 2 represents <1 yr experience, Group 3 represents 1-5 yrs experience, Group 4 represents >5 yrs experience. The results of Fisher test shows that Group 3 significantly differs from the other groups whereas Group 1, 2, 4 do not differ from each other.

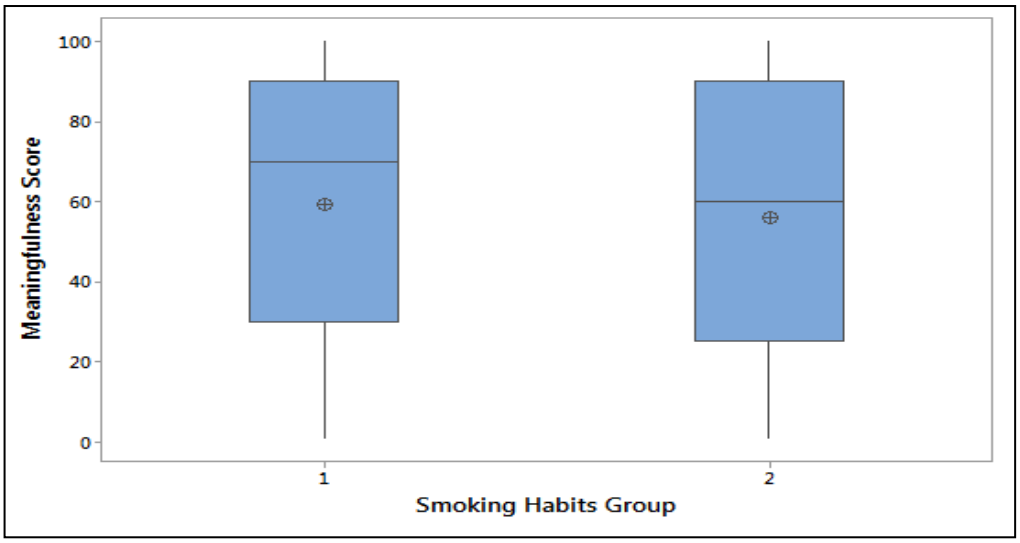


Figure 5.33. Boxplot of Smoking Habit Groups

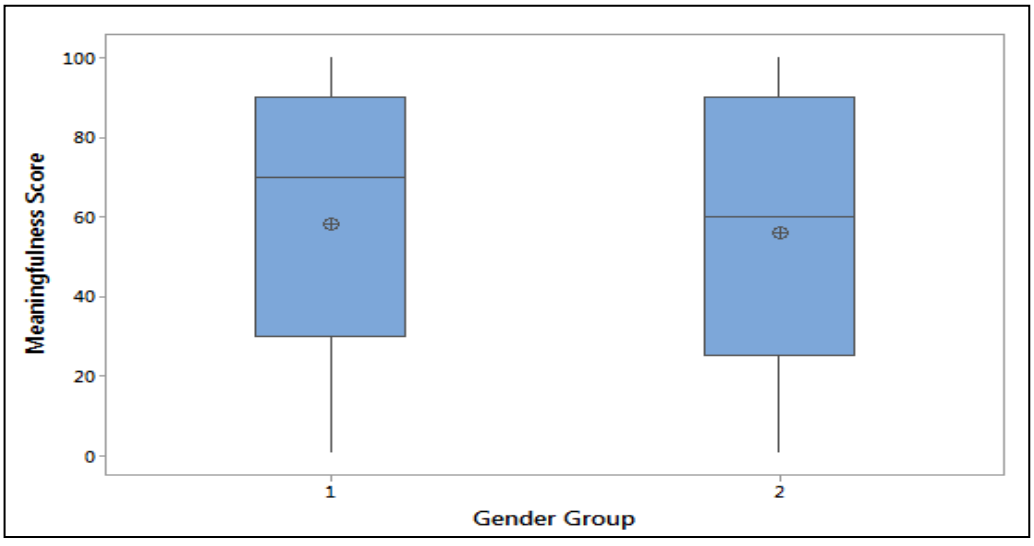


Figure 5.34. Boxplot of Gender Groups

Table 5.44. Analysis of Variance table for Semantic Closeness

Source	DF	Adjusted SS	Adjusted MS	F	P
Age Group	4	1480862443	370215611	35.47	<0.001
Educational Background	2	69805306	34902653	3.34	0.035
Active Car Usage	3	217283824	72427941	6.94	<0.001
Marital Status	2	90864584	45432292	4.35	0.013
Gender	1	115717301	115717301	11.09	0.001
Smoking Habits	1	5753896	5753896	0.55	0.458
Sign Type	4	17107168225	4276792056	409.8	<0.001
Error	12.042	1.25674E+11	10436325		
Total	12.059	1.44948E+11			

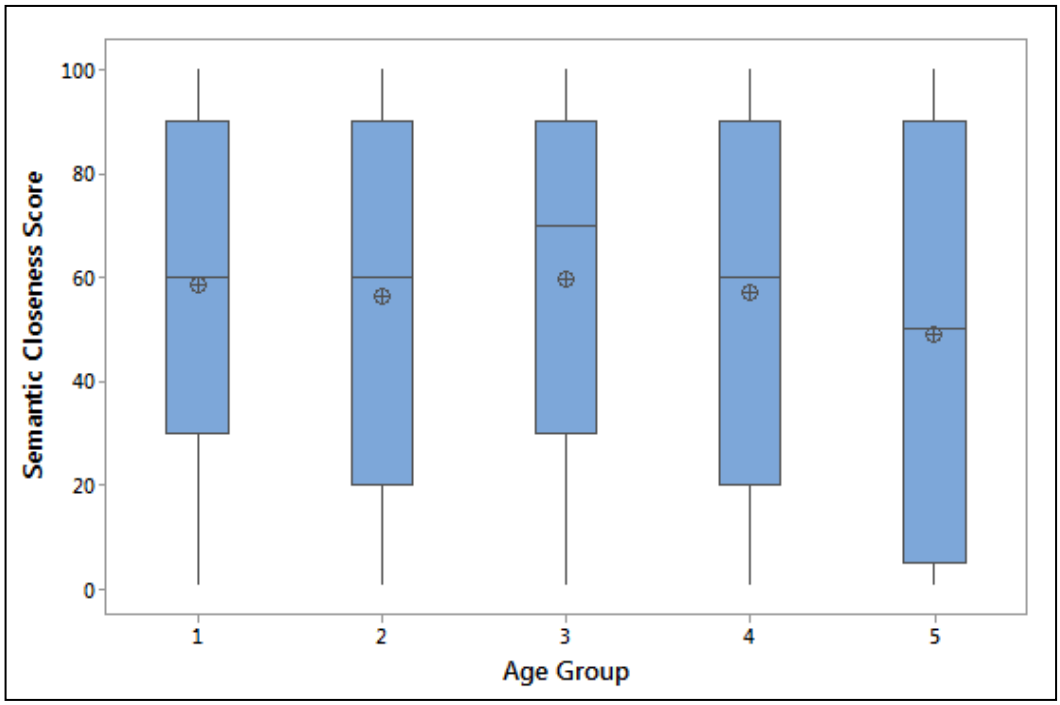


Figure 5.35. Boxplot of Age Groups

Table 5.45. Results of Fisher test for Age Groups

Difference of Levels	Difference of Means	SE of Difference	T- value	Adj. P- value
Group 2 - Group 1	-176.1	97.2	-1.81	0.070
Group 3 - Group 1	110.1	98.9	1.11	0.266
Group 4 - Group 1	-69.1	98.3	-0.70	0.482
Group 5 - Group 1	-966	102	-9.50	0.000
Group 3 - Group 2	286.2	97.8	2.93	0.003
Group 4 - Group 2	107.1	97.2	1.10	0.270
Group 5 - Group 2	-790	101	-7.85	<0.001
Group 4 - Group 3	-179.1	98.9	-1.81	0.070
Group 5 - Group 3	-1076	102	-10.52	<0.001
Group 5 - Group 4	-897	102	-8.82	<0.001

Group 1: 18-29 yrs; Group 2: 30-39 yrs; Group 3: 40-49 yrs; Group 4: 50-59 yrs; Group 5: 60-69 yrs

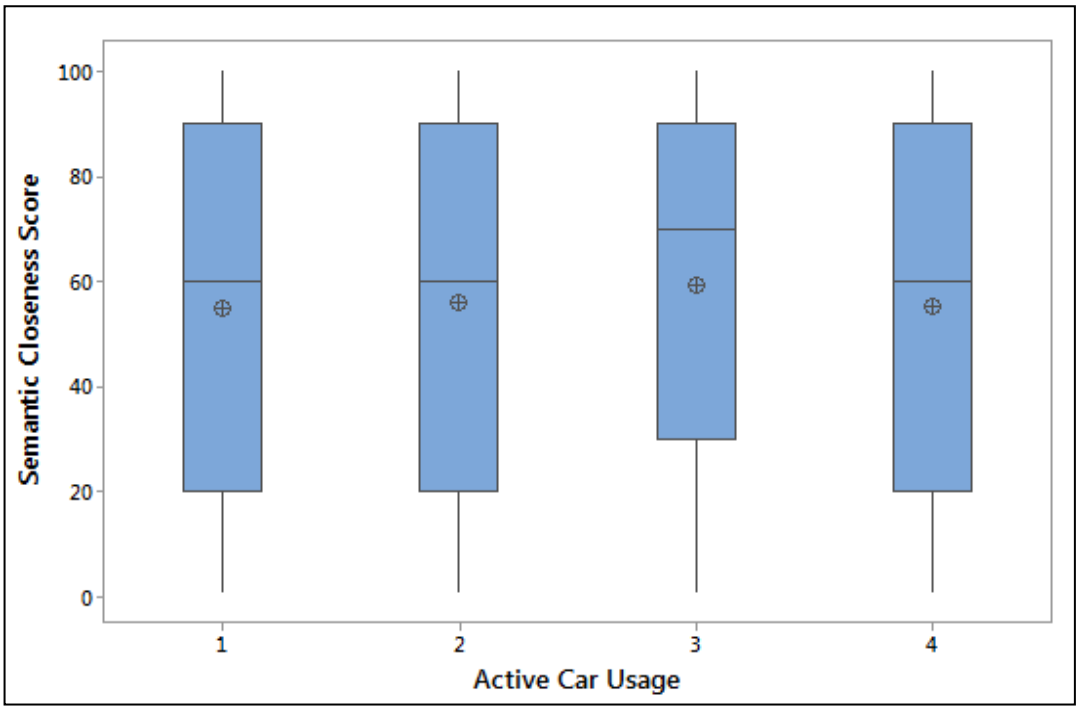


Figure 5.36. Boxplot of Active Car Driver Experience Groups

Table 5.46. Results of Fisher test for Active Car Driving Experience Groups

Difference of Levels	Difference of Means	SE of Difference	T- value	Adjusted p - value
Group 2 - Group 1	116.6	86.8	1.34	0.179
Group 3 - Group 1	511.8	88.6	5.78	<0.001
Group 4 - Group 1	49.5	90.1	0.55	0.582
Group 3 - Group 2	395.2	88.6	4.46	<0.001
Group 4 - Group 2	-67.1	90.1	-0.74	0.456
Group 4 – Group 3	-462.2	91.7	-5.04	<0.001

Group 1: Not active driver; Group 2: Less than 1 yr.; Group 3: Between 1-5 yrs.; Group 4: More than 5 yrs.

- Sign Type Effect: The box plot of semantic closeness score for different sign types was shown in Figure 5.37. The results of Fisher test show that all groups significantly differ from each other.

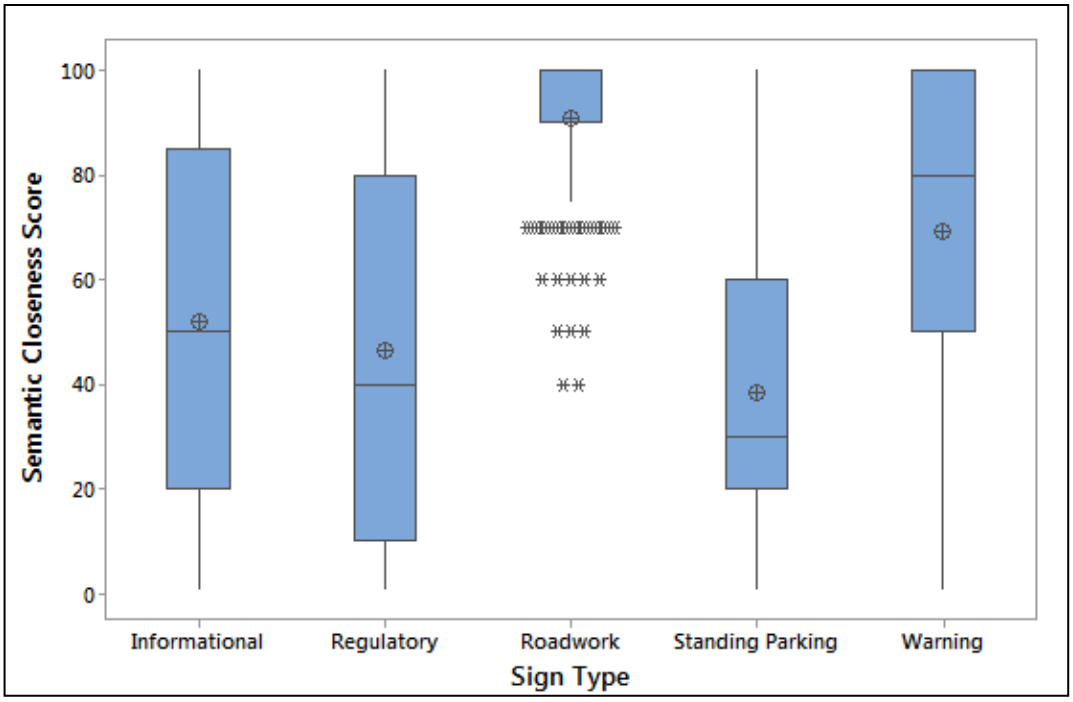


Figure 5.37. Boxplot of Sign Type Groups

Table 5.47. Results of Fisher test for Sign Type Groups

Difference of Levels	Difference of Means	SE of Difference	T- value	Adj. P- value
Regulatory-Information	-482.1	81.1	-5.95	<0.001
Roadwork-Information	4166	238	17.48	<0.001
Standing - Information	-1246	147	-8.47	<0.001
Warning - Information	1832	80.4	22.80	<0.001
Roadwork - Regulatory	4648	235	19.77	<0.001
Standing- Regulatory	-763	142	-5.38	<0.001
Warning - Regulatory	2314.1	70.1	33.02	<0.001
Standing - Roadwork	-5411	265	-20.40	<0.001
Warning -Roadwork	-2334	235	-9.94	<0.001
Warning -Standing	3077	141	21.77	<0.001

- Educational Background Effect: The box plot of semantic closeness score for different educational background groups was shown in Figure 5.38. Group 1 represents primary/secondary school, Group 2 represents high school, and Group 3 represents university or higher degree background. The results of Fisher test show that Group 1 and 3 significantly differ from each other.

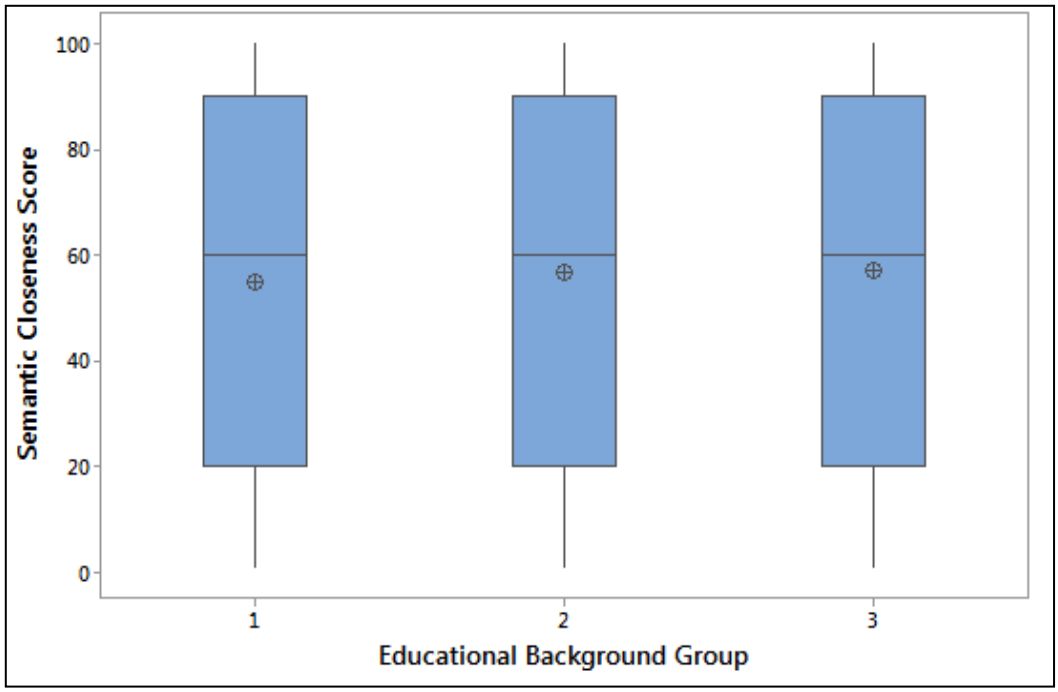


Figure 5.38. Boxplot of Educational Background Groups

Table 5.48. Results of Fisher test for Educational Background Groups

Difference of Levels	Difference of Means	SE of Difference	T- value	Adj. P- value
Group 2 - Group 1	146.3	77.9	1.88	0.060
Group 3 - Group 1	216.9	76.7	2.83	0.005
Group 3 - Group 2	70.6	77.3	0.91	0.361

Group 1: Primary/Secondary School; Group 2: High School; Group 3: University and/or higher degree

- Marital Status Effect: The box plot of semantic closeness score for different marital status groups was shown in Figure 5.39. Group 1 represents single, Group 2 represents married, and Group 3 represents married with children group. The results of Fisher test show that Group 2 differs from Group 1 and 3.
- Gender Effect: The box plot of semantic closeness score for gender was shown in Figure 5.40. Results show that men performed better than women in semantic closeness score.

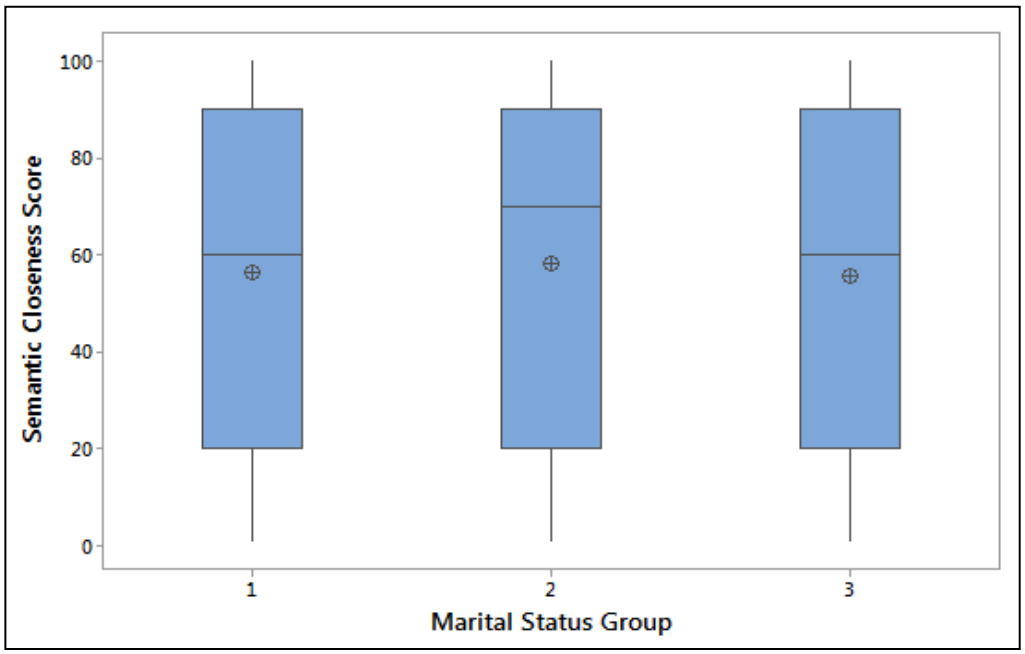


Figure 5.39. Boxplot of Marital Status Groups

Table 5.49. Results of Fisher test for Marital Status Groups

Difference of Levels	Difference of Means	SE of Difference	T- value	Adj. P-value
Group 2 - Group 1	181.2	90.1	2.01	0.044
Group 3 - Group 1	-62.7	71.4	-0.94	0.346
Group 3 - Group 2	-248.5	85.1	-2.92	0.004

Group 1: Single; Group 2: Married; Group 3: Married with children

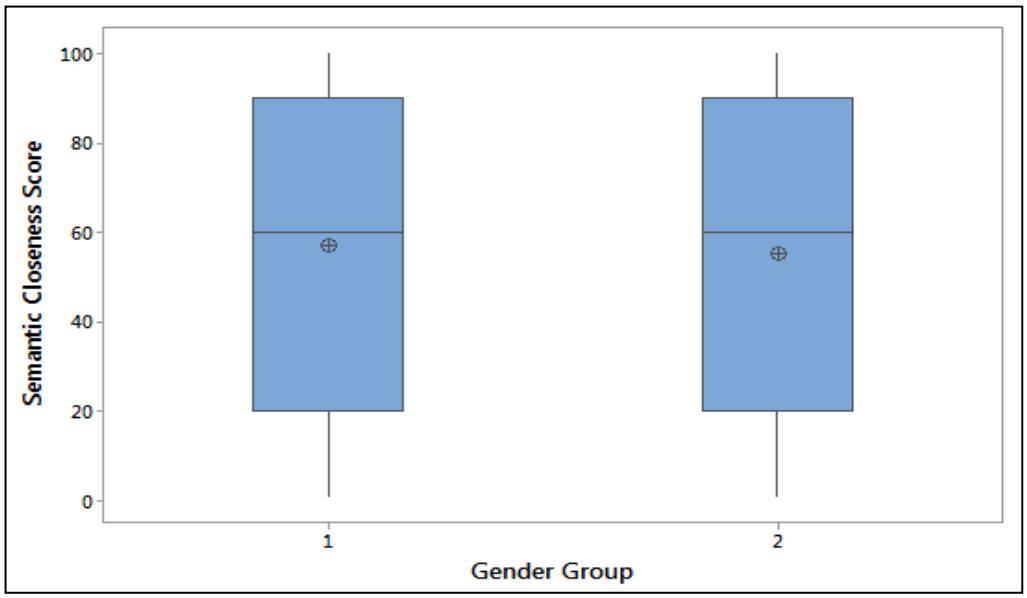


Figure 5.40. Boxplot of Gender Groups

6. DISCUSSION

6.1. Discussion of the Current Study Results

The purpose of the study was to investigate the comprehensibility of the current Turkish traffic signs and evaluate the effect of sign design features and user factors on comprehensibility.

6.1.1. Guessability Score of Current Turkish Traffic Signs

Traffic signs are evaluated in five different sign groups: warning, regulatory, informational, road work and parking and standing. In terms of guessability score results indicated below:

(i) Guessability score differed from sign to sign. Overall, the mean guessability score of current traffic signs was found lower than 50% with the score of 39.3%. Together with partially correct answers, the mean guessability score rose to 55.7%. It can be concluded that, current Turkish traffic signs can be guessed around 40% in average by the population.

(ii) In terms of sign type, only the mean guessability score of warning and road works sign were found as higher than 50% with the score of 50.5% and 82.6%. It can be said that, guessability of regulatory, informational, and standing and parking signs were found as difficult to guess for the population.

(iii) Signs were also analyzed according to the comprehension level of participants as well-known, partially known, left with no comment, misinterpreted and oppositely understood signs.

(iv) 12% of the traffic signs were comprehended as completely correct by 70% of the participants. Since the guessability score was associated with the percentage of completely correct answer, those well comprehended signs are the signs also with the highest guessability score.

(v) Results indicate that more than 50% of the participants could not guess the meaning of 21 traffic signs and left them with no comment. This traffic signs had also low guessability score.

(vi) The meanings of the 4 traffic signs were known as wrong by the 20% or more of the participants. For example, "No entry for vehicle carrying explosives" sign was known as wrong by 42,8 % of the participants. The most answers for this sign were "Car can explode" and "explosion". As a conclusion, these signs need to be re-evaluated again in order to prevent misunderstanding.

(vii) It is observed that 5 out of 60 traffic signs tested in this study were understood as the opposite of their true meanings. The finding is similar with the finding of the study by Kirmizioglu and Tuydes-Yaman (2011) where 5 traffic signs were found as the oppositely understood. 3 out of those 5 traffic signs were similar with the signs found in this study. Ben-Bassat and Shinar (2006) was also investigated that 10% or more of the participants misinterpreted some signs as the opposite of their true meanings. For example, "No entry for motor vehicle except motorcycles" was one of oppositely associated signs (46.77 %). This sign was understood by the participants as motor vehicle can use the road. Although the familiarity scores of those signs were not as much as high like the well-known signs, it is not also low either. The reason of this opposite understanding can be a result of EU Harmonization Process that was applied in Turkey in 2004 where the red crossing line was removed from the signs. As a result, this is very critical issue that understanding of the signs as the opposite of true meanings can increase accident rates on the road. In order to prevent any accident and keep road safety, those signs need to be discussed again considering the low comprehensibility of them.

6.1.2. Sign Design Features and Guessability Score

60 Turkish traffic signs were included in this study and five sign design features; familiarity, concreteness, simplicity, meaningfulness and semantic closeness were analyzed. Familiarity, meaningfulness and semantic closeness were found as significant predictors of guessability score, whereas concreteness had no effect on guessability score. Results indicated below:

(i) 60 Turkish traffic signs were perceived moderately familiar, meaningful, semantically close, simple and low in concreteness by the population. Simplicity score was higher than the other scores, whereas concreteness had the lowest score with 48.8%. Ratings for each sign design features varied differently from sign to sign and also category to category. Warning and Roadworks signs were the categories with the higher mean scores with more than 50%. Warning signs were evaluated as moderately concrete and familiar and highly simple, meaningful and semantically close. Road works signs were perceived highly familiar, concrete, simple, meaningful and semantically close. Regulatory signs were found simple, meaningful, moderately familiar but more abstract and semantically distant. Informational sings were evaluated as unfamiliar, moderately concrete, meaningful, and semantically close but highly simple. Lastly, Standing and parking signs were found as moderately meaningful, highly simple but unfamiliar, abstract and semantically distant.

(ii) An expected result is that guessability score was high for the signs with the higher familiarity score and low for the unfamiliar signs. Ng and Chan (2007c, 2008b) also found that guessability was higher for the familiar signs. Shinar *et al*. (2003) come up with the similar result and mentioned that "infrequent signs were more likely to be miscomprehended" by drivers. Ben-Bassat and Shinar (2006) showed a strong and positive relationship between comprehension and the frequency that a driver encounters traffic sign. Therefore, it is not surprising that familiarity was found significantly and positively correlated with the guessability score in our study and as the main predictor of the regression model.

(iii) Simplicity is the other factor that is found with moderately correlated with guessability score. According to the Bruyas *et al*. (1996), additional and non-important elements can disturb understanding. Simple signs can support higher comprehensibility. Chan and Chan (2011) reported that concrete signs have higher comprehensibility than abstract ones. However, in this study, simplicity was not found as the main predictor of the regression model. In the study of Ng and Chan (2008b) similarity was not also found related to the guessability score.

(iv) An expected result is that concreteness had no significant effect on guessability score and was not a main predictor. During the experiment, it was observed that Turkish population had a difficulty to separate the concrete and abstract concepts. That is

why the coefficient of variation was high for this feature. If a sign is meaningful enough, they tend to find it concrete even if the sign has an abstract symbol. Ng and Chan (2007) had noted that culture and thinking style of the population can lead to different results. It can be concluded that in the thinking style of Turkish population, meaningfulness can be interchangeable with concreteness which means abstract items can be found as concrete if it is meaningful enough.

(v) Meaningfulness is found as another main predictor of guessability score. The comprehensibility is high for the signs with high meaningful score and low for the ones with low meaningful score.

(vi) Semantic closeness is also found significant factor on guessability score. Semantic closeness defines the relationship between the image and the intended meaning that wanted to be given with that image. Therefore, it is not surprising that it is a strong predictor for guessability score. The Turkish traffic signs were evaluated as moderately semantically close.

(vii) According to the correlation analysis between sign design features, meaningfulness, familiarity and semantic closeness were found strongly and positively correlated with each other. It means that a sign with high familiarity has also found meaningful and semantically close by the population. Moreover, concreteness was found strongly correlated with meaningfulness, moderately correlated with semantic closeness and almost not correlated with familiarity. It can show that a sign with high familiarity and semantic closeness could not have to be concrete. That is why concreteness is not a main predictor of the guessability score.

6.1.3. User Factors and Guessability Score

Seven user factors, which are age, gender, marital status, educational background, active car driving experience, driver license ownership and smoking habits, are evaluated in this study. Age, active car driving experience and marital status factors are found as the significant predictors of the guessability score. The factors: educational background, gender and smoking habits had no significant effect on guessability score. Results are indicated below:

(i) An expected result is that age is the significant predictor of the guessability score. 5 different age groups were included in the study which are 18-29 yrs, 30-39 yrs, 40-49 yrs, 50-59 yrs, and 60-69 yrs. According to the results it is concluded that comprehension is decreasing with the age since the guessability score of 60-69 yrs were significantly differs from the other groups. It is also surprising that 60-69 yrs and 30-39 yrs gave more not commented answers and 18-29, 40-49 and 50-59 yrs gave more wrong and opposite answers. Lesch (2003) mentioned that "younger and older groups are different in terms of cognitive abilities. Older group has more limited working memory capacity than younger group". Al-Madani and Al-Janahi (2000), Liu and Ou (2012) and Shinar *et al*.(2003) came up with a similar result with this study and showed that age plays a major role in comprehension.

(ii) Results indicate that gender had no significant effect on guessability score. Female drivers gave more not commented and less wrong and opposite answers than male drivers. Al-Madani and Al-Janahi (2002) showed also the similar results with this study.

(iii) Marital status had a significant effect on guessability score. According to the result, married drivers performed well than the single and married with children drivers. They gave more fully correct, less no commented and opposite but more wrong answers. This can be explained by the increasing responsibility of the driver with the marriage. There is not much study that searched for the marital status. Al-Madani and Al-Janahi (2000) found also no significant effect of marital status on comprehension.

(iv) Active car usage was the significant predictor of guessability. Drivers with no active car driving experience gave less fully correct but more not commented and wrong answers than the experienced ones. Drivers more than and equal to 5 yrs experience gave less wrong and opposite answers than the other groups. It was surprising that drivers with less than 1 yr and between 1 to 5 yrs experience gave more opposite answers. Drivers with more than 1 yr and less than 5 yrs experience had the highest guessability score than the other groups. As an overall, experienced drivers performed well than non-experienced group. A similar result was found in the study of Al-Madani (1999). Ng and Chan (2008b) mentioned that "comprehension performance is better in the first year of holding a driving license and decreased with

years of being a licensed driver and comprehension performance is not related with the years of active driving.

(v) Another interesting result is that educational background does not have a significant effect on guessability score. University or higher degree group gave more fully correct and less no commented and wrong answers than the other groups. Primary/Secondary group gave more no commented answers. The average number of opposite answers was similar for each group. Chan and Chan (2011) found also no significant effect of educational background and Ng and Chan (2008b) mentioned that the performance of university or higher degree group is better in sign comprehension.

(vi) Results showed that smoking habits do not have a significant effect on guessability. Smoking group gave more opposite answers than the non-smoking group.

Briefly, guessability score was significantly affected by age, active car usage and marital status.

6.1.4. Sign Design Features and User Factors

The effect of six user factors, which are age, gender, marital status, educational background, active car driving experience, smoking habits, and sign type on each of five cognitive sign design features; familiarity, concreteness, simplicity, meaningfulness and semantic closeness were evaluated in this study.

Age factor, active car usage and sign type were found as a significant factors on the score of all five cognitive sign design features. Educational background was a significant factor for only familiarity, simplicity and semantic closeness. Gender has a significant effect on concreteness, meaningfulness and semantic closeness score. Table 6.1 shows the user factors that had a significant effect on cognitive sign design features score.

(i) Age group 60-69 yrs had the lowest score for all sign design features. Except meaningfulness and semantic closeness the score generally was higher in younger age (18-29 yrs), started to decrease in middle age (30-39 yrs or 40-49 yrs) and then increased in older age (40-49 yrs or 50-59 yrs) and was lowest in the oldest age (60-

69 yrs). For meaningfulness and semantic closeness, generally the scores were higher and similar in younger and older ages (18-29 yrs, 40-49 yrs and 50-59 yrs) and lower in middle age (30-39 yrs) but still the lowest in oldest age (60-69 yrs). It can be concluded that age is a significant factor on the score of cognitive sign design features.

Table 6.1. Significant User Factors for each Sign Design Features

User Factors	**Cognitive Sign Design Features**				
	Familiarity	Concreteness	Simplicity	Meaningfulness	Semantic Closeness
Age	+	+	+	+	+
Active car usage	+	+	+	+	+
Sign Type	+	+	+	+	+
Marital Status	+	+	+	-	+
Educational Background	+	-	+	-	+
Gender	-	+	-	+	+
Smoking Habits	+	-	+	+	-

+: significant effect observed; - : no significant effect observed

(ii) For the user factor, active car driving experience, the scores of not an active driver group was lowest for all cognitive sign design features as expected. Interesting result was that scores were higher in less experienced groups (less than 1 yr. and 1to5 yrs.), and lowest or same without experienced groups in more experienced group (more than 5 yrs). It can be concluded that the comprehension of signs in terms of cognitive sign design features increases with experience but after 5 years decreases dramatically.

(iii) Sign type (warning, regulatory, informational, standing and parking, roadworks) had a significant effect on all sign design features and significantly differ from each other's. For all sign design features, Roadworks sign type had the highest score. This type of sign was found more familiar, concrete, simple, meaningful and semantically close. Warning signs had the second high score for each user factors Informational sign type had the lowest score in terms of sign features in meaningfulness. Standing and parking signs had the lowest score in terms of simplicity, concreteness and

semantic closeness. It is concluded that informational, regulatory and standing and parking signs may be re-evaluated again considering sign design features.

(iv) Educational background was found as significant factor for familiarity, simplicity and semantic closeness. The group with university and/or higher degree educational background had the highest score for familiarity, and higher score and no significant difference with high school for semantic closeness. Interesting result was that for the design feature simplicity, university or higher degree group had the lower score and with no significant difference with primary/secondary school background. It can be concluded that with increase in level of education background, the comprehension of signs in terms of familiarity and semantic closeness increases.

(v) Marital status was a significant factor for familiarity, concreteness, simplicity and semantic closeness scores. Married group had the higher score of evaluation.

(vi) Gender was a significant factor for concreteness, meaningfulness and semantic closeness. For all design features, the score of men were higher than women.

(vii) Smoking habit had significant effect for familiarity, meaningfulness and simplicity. The score of smokers were higher than the non-smokers.

6.2. Comparison with Other Studies

In the literature, sign type, sample type, sample size, factors used, measure type, experimental procedure, age range, sign design features, user factors change study to study. Therefore, to make a one to one comparison with current study is not feasible since there is no study in the literature overlapping the current study to compare the results. Moreover, it is observed that studies in the literature were focused on to analyze the effect of sign design features or user factors only and generally chose one sample type. That is why making a direct comparison is difficult and not possible with the current study. Therefore, the most similar studies with current study in terms of methodology and objective were selected to evaluate.

For sign design features on guessability score, the most similar studies for comparison:

- Ng, Chan (2007a, 2007c, 2008b, 2011)
- Ou, Liu (2012)
- Ben-Bassat, Shinar (2006)

For user factors on guessability score, t1he most similar studies for comparison:

- Al-Madani, Al-Janahi (2000, 2002)
- Ng, Chan (2007c, 2008b)
- Shinar *et al.* (2003)
- Kirmizioglu, Tuydes-Yaman (2011)

For user factors on sign design features, the most similar studies for comparison:

There is no study in the literature that compares the effect of user factors on the score of cognitive sign design features.

Some specific information was given about the studies and comparison with current study was made in below:

- Current study vs. Ng, Chan (2007a, 2007c, 2008b, 2011): Ng and Chan (2007a, 2007c, 2008a, 2008b, 2011) had the most similar studies in terms of methodology to the current study. Ng and Chan (2007a) conducted a study to measure the effect of sign design features and Ng and Chan (2007c) enlarged the study by adding the effect of sign design features on guessability score. 120 Mainland Chinese traffic signs were used in the studies and Hong Kong Chinese engineering undergraduates 19 Male and 21 Female (18 to 27 yrs, median = 22,5 yrs) non-experienced subjects participated to the study. The number and type of participants are very limited when compare with the current study. The testing procedure and the experimental materials used in the study Ng and Chan (2007a) are similar to the current study. 120 traffic signs were shown to the participants and personal evaluation between 0-100 points were collected. Subjects were asked to rate for familiarity, concreteness, simplicity, meaningfulness before and then the real meaning of the traffic signs was shown and participants rated the semantic closeness. In Ng and Chan (2007c),

participants asked to guess the meaning of the sign before starting the ratings as similar to Ng and Chan (2007a) study. Significant relation among familiarity, meaningfulness and semantic closeness were observed in those studies similar to the current study. In the Ng and Chan (2007a) a strong relationship between concreteness and familiarity was observed as being different from current study. In the study Ng and Chan (2008b), in addition to the sign design features, the effect of user factors on comprehension score was considered. 21 traffic signs were used and 100 Male and 9 Female (age 18 to 57 yrs) experienced drivers were participated to the study. Convenience sampling is used. As a user factor: years with license, years of active driving, last time of driving, driving frequency, driving hours per trip, non-local driving experience, occupation, educational level, gender, age were discussed in the study. Although the factors are not totally similar with the current study, some of them offer to make comparison. As a comparison with the current study, there are some similar steps in the methodology but the participant size, distribution and type (especially in terms of gender) are limited for a general result. Moreover, convenience sampling can also restrict the generalization of the results. The age, education, active driving experience groups are different than the current study and also more restricted. In the study, familiarity was found as the best predictor of comprehension whereas others were found not related to the score. Different than the current study, no difference in comprehension performance regarding age was showed in the study. The group size of age is different in the study and last elderly group has a very large range which can effect the results.

- Current study vs. Ou, Liu (2012): The effect of sign design features and the training on sign comprehension using 65 traffic signs were analyzed in the study. Two separate groups Taiwanese and Vietnamese, 15 male and 15 female subjects (Taiwanese ages of 19 and 29 yrs, Vietnamese at the ages of 21 and 26 yrs) from each group were participated to the study. Almost same methodology was used to analyze the effect of sign design features on comprehension. In comparison with current study, the result of the study is very limited cause of the participant size and type, and restricted to the undergraduate education. Semantic closeness was found as the best predictor of comprehension. As different as the current study, the effect of training on comprehension was evaluated and positive effect of it was found.

Table 6.2. Summary of the results of User Factors for Each Sign Design Feature

User Factors	Cognitive Sign Design Features				
	Familiarity	**Concreteness**	**Simplicity**	**Meaningfulness**	**Semantic Closeness**
Age	60-69 yrs had lowest score and differentiate from others. No significant difference between 18-29 yrs and 40-49 yrs,30-39 yrs and 50-59 yrs.	60-69 yrs had lowest score and differentiate from others.No significant difference between others.	60-69 yrs had lowest score and differentiate from others. 18-29 yrs - 50-59 yrs groups do not differentiate..	60-69 yrs had lowest score and differentiate from others. No difference between 40-49 yrs and 50-59 yrs..	60-69 yrs had lowest score and differentiate from others. No significant difference between 18-29 yrs, 40-49 yrs and 50-59 yrs.
Active Car Usage	Not an active driver group had lowest score and differentiate from others..	Not an active driver group had lowest score and differentiate from others.	Not an active driver group had lowest score and differentiate from others	All groups differ Active driver 1to 5 yrs group had the highest score, not active driver group had the lowest score.	Active driver 1to 5 yrs group had the highest score and differentiate from others. Other groups do not differentiate.
Sign Type	All signs differ significantly, Informational signs had lowest score, Roadworks had highest score.	All signs differ significantly, except standing and parking and regulatory signs. Roadworks had highest score.	All signs differ significantly, Standing and parking signs had lowest score, Roadworks had highest score.	All signs differ significantly, Informational signs had lowest score, Roadworks had highest score.	All signs differ significantly, Standing and parking signs had lowest score, Roadworks had highest score
Marital Status	Married group had higher score than others.	Married group and single group had higher score than married with children group.	Married group had higher score than single and married with children group.	Not significant factor	Not significant factor
Educational Background	University & higher degree group had higher score and differentiate from others.	Not significant factor	University & higher degree group and primary/secondary group do not differentiate.	Not significant factor	. Primary/secondary school group had lowest score.
Gender	Not significant factor	Men had higher score than women	Not significant factor	Men had higher score than women	Men had higher score than women
Smoking Habits	Smokers had higher score than non-smokers	Not significant factor	Smokers had shigher score than non-smokers	Smokers had higher score than non-smokers	Not significant factor

- Current study vs. Ben-Bassat, Shinar (2006): Ben-Bassat and Shinar (2006) conducted a study to investigate the relationship of three ergonomic principles with the comprehension of the highway signs. 30 highway warning signs were used. 23 Male and 17 Female subjects (average age 28.8 yrs) with 3-15 yrs of driving experience were participated in the study. Ben-Bassat and Shinar (2006) defined ergonomic principles as compatibility, familiarity and standardization. Although those principles seem different from the sign design features used in the current study, the main purpose is similar except standardization. Standardization was about the color, shape and size of the signs which is not the objective of current study. In the study, the evaluation of the principles was done in 1-10 scale by the participants before and then it was also evaluated by the experts afterwards. As a comparison with current study, the number and type of the participants are very limited and driving experience gap of the group is very large. Moreover, moderate effect of active car usage on guessability had been found in the study so that large gap could have an effect on the results. Ng and Chan (2008a) mentioned that 0-100 points scale is preferred since it makes data more amenable. As a similar result with the current study, Ben-Bassat and Shinar (2006) showed that some of the signs guessed as opposite of their true meanings which is a threat for road safety. Studies also showed that ergonomically designed traffic signs are more understandable.
- Current study vs. Al- Madani and Al-Janahi (2000, 2002): Al-Madani and Al-Janahi (2000, 2002) assessed the effect of traffic, personal and social characteristics on the comprehension of drivers. A totally different methodology from the current study was used in this study. The population was sampled from five different Arabic countries and 28 posted signs were included in the study. Survey was sent to 9000 people and 56% of them responded back. Age, marital status, gender, nationality, educational background and monthly income were searched as a factor. In the study, Al-Madani and Al-Janahi (2000) selected groups: age (16-24 yrs., 25-34 yrs., 35-44 yrs., more than 45 yrs.), education (below secondary, secondary and diploma, B.Sc., higher degree) were different than the current study. According to the results, significant effects of age, gender, education and monthly income on comprehension were shown. As a similar result with the current study, it is shown that married drivers generally comprehended better than the single ones and the performance of male drivers were higher than the females. In the study Al-Madani and Al-Janahi

(2002), selected groups: age (16-24 yrs., 25-34 yrs., 35-44 yrs., 45-55 yrs., more than 55 yrs.), education (below secondary, secondary to under diploma, diploma holders, higher degree), driving experience (less than 5 yrs., 5-10 yrs., 10-15 yrs., more than 15 yrs.) were different than the current study. No significant effect of age, marital status, experience was found on drivers' comprehension in the study. For the driving experience group, experience range of the experienced group can be very large. Although the methodology is totally different from the current study, some of the results are similar. Moreover, that kind of non-face to face survey is a good way to reach much more participants but results could be misleading if the methodology is not fully understood by the participants.

- Current study vs. Shinar *et al.* (2003): Shinar *et al.* (2003) conducted a cross-cultural study to evaluate the comprehension level of 31 traffic signs used in different countries. Canada, Finland, Israel and Poland included in the study and five different driver groups: novice drivers, college students, tourists, problem drivers, and older drivers were chosen and 25 male and 25 female were collected from each group for each country. The general methodology was different than the current study but similar outcomes were also found. Shinar *et al.*(2003) mentioned some of the signs were comprehended as the opposite of their true meanings which is similar result with current study. Also in this study, the performance of elderly drivers were found lower than the other drivers.
- Current study vs. Kirmizioglu and Tuydes-Yaman (2011): Kirmizioglu, Tuydes-Yaman (2011) is one of the studies conducted to investigate the comprehensibility of Turkish traffic signs in the city of Ankara. 30 commonly used traffic signs were selected for the study. Some of the signs were similar with the current study however signs with number and words also included different than the current study. Since numbers can affect the comprehension, therefore it can also manipulate the results. 1134 male and 327 female were joined the study. Age (18-25 yrs., 26-35 yrs., 36-45 yrs., 46-60 yrs., more than 60 yrs.) and education (one-third of university graduates, other one-third high school graduates) groups and the methodology used were different than the current study. An open-ended questionnarie was used and paritipants were asked to guess the meaning of the traffic signs. Researcher evaluated the answers in scale of five as similar to the current study. It is shown that most of the signs were not known well by the drivers. Moreover, it was found that

five of the traffic signs were understood by the participants as the opposite of their true meanings. Three of this oppositely answered five signs were similar to the current study. This study is also an important indicator that traffic signs in Turkey are not well comprehendend. To make a comparison, this study is more restricted to a city and focused only the comprehensibility of traffic signs. In current study, only the mostly used signs that convey the message with symbols and not used in conjunction with other signs for giving a message were chosen. Also the objective of the current study was not only to investigate the comprehension but also reveal the effect of sign design features and user factors on comprehension of Turkish traffic signs.

7. CONCLUSIONS

7.1. Conclusions of the Study

The purpose of current study was to investigate the comprehension of current Turkish traffics signs, estimate the effect of five cognitive sign design features: familiarity, concreteness, simplicity, meaningfulness and semantic closeness on guessability score and reveal the effect of user factors: gender, age, marital status, educational background, active car driving experience, smoking habits on guessability score of the current Turkish traffic signs by the population of Turkey. The data were collected from 100 male and 101 female with the age of 18 to 70 yrs (average 43,6 yrs, median 43,9 yrs). Statistical analyses were performed to investigate both the effect of cognitive sign design features and user factors on guessability score. Prediction models for guessability score considering the sign design features were developed and comparisons were made with some studies in the literature. Based on analysis results, the following conclusions can be drawn:

- The comprehension and guessability score of current Turkish traffic signs were established.
- The mean guessability score of 60 Turkish traffic signs was found to be 39.3% which shows that current traffic signs are not guessed well by the population.
- In terms of sign type, only the mean guessability score of Warning (50.5%) and Road work sign (82.6%) were found higher than 50%. Regulatory, Informational and Standing and parking signs are difficult to guess for the population.
- Only 20% of the 60 traffic signs were comprehended fully correctly by the 70% of the population. More than 50% of the participants could not guess the meaning of 35% of traffic signs (21 signs) and left without comment. It explains the reason of low guessability score of the current Turkish traffic signs.
- 4 traffic signs were guessed as wrong by 20% of the population. 5 out of 60 traffic signs were comprehended as opposite of their true meanings. These are the signs that re-arranged with the EU Harmonization Process which was applied in Turkey in 2004 where the red crossing line was removed from the signs. Understanding of the

sign as the opposite of its true meaning can cause serious traffic accidents so these signs need to be controlled again.

- 60 traffic signs were perceived moderately familiar, meaningful, semantically close, simple and low in concreteness by the population. Road work sign were found more familiar; Warning and Regulatory signs were found moderately familiar, whereas Informational and Standing and parking signs were found unfamiliar by the population. Warning, Informational and Standing and parking signs were evaluated as highly simple. Regulatory and Standing and parking signs were taken as abstract and semantically distant. Only the Standing and parking signs were found as moderately meaningful whereas others considered as meaningful.
- Guessability score was found highly and positively correlated with familiarity, meaningfulness and semantic closeness whereas moderately correlated with simplicity and low correlated with the concreteness. It can be concluded that signs with high familiarity, meaningfulness and semantic closeness can be guessed well.
- In correlation analysis meaningfulness, familiarity and semantic closeness were found strongly and positively correlated with each other's. Moreover, concreteness was found strongly correlated with meaningfulness, moderately correlated with semantic closeness and almost not correlated with familiarity. It can show that a sign with high familiarity and semantic closeness could not be concrete.
- Turkish population had a difficulty to separate the abstract and concrete concepts and tend to find symbol concrete if they know the meaning of it enough. That is why concreteness is not a significant predictor of guessability whereas meaningfulness and familiarity are.
- Age is significant predictor of the guessability score. Comprehension decreases with the age. 60-69 yrs group had the lowest score. It is also observed that younger and older groups (18-29 yrs, 40-49 yrs and 50-59 yrs) gave more wrong and opposite answers than the oldest group whereas oldest group 60-69 yrs gave more not commented answers.
- No significant effect of gender was observed on guessability score but female drivers gave more not commented and less wrong and opposite answers than male drivers.
- No significant effect of marital status was observed on guessability score.
- Driving experience had a significant effect on guessability score. More experienced drivers performed better than the non-experienced ones. Moreover, drivers with no

active car driving experience gave less fully correct but more not commented and wrong answers than the experienced ones. Drivers with more than 1 yr. and less than 5 yrs. experience had the highest guessability score than the other groups. However, they gave more opposite answers.

- Guessability performance is not related with the educational background.
- No significant effect of smoking habits on guessability score was found. However, Smoking group gave more opposite answers than the non-smoking group.
- Age was found as a significant factor on evaluation of cognitive sign design features. 60-69 yrs group had the lowest score for all cognitive sign design features.
- Active car driving experience was a significant factor for all sign design features. The scores of not an active driver group were lowest for all cognitive sign design features. Comprehension of signs in terms of all cognitive sign design features increases with experience but after 5 years decreases dramatically.
- Sign types had a significant effect on all sign design features and significantly differ from each other's. For all sign design features Roadworks sign type had the highest score meaning that was found more familiar, concrete, simple, meaningful and semantically close followed by Warning type. It is concluded that Informational, Regulatory and Standing and parking signs may be re-evaluated again considering sign design features.

This study investigated the guessability score of current Turkish traffic signs and examined effect of both cognitive sign design features and user factors on guessability score, which these two different aspects were not investigated before in any study in Turkey. Moreover, the effect of user factors on evaluation of cognitive sign design features was investigated in this study which was not investigated before in worldwide. In the literature there are only three studies with similar design for guessability score analysis but with different user factors (Ng and Chan, 2007c, 2008b; Chan and Chan, 2011). Also, studying with that kind a large sample size with face to face surveys enables the current study to resemble Turkey and generalize the results.. Moreover, this study has focused on analyzing the relationship of cognitive sign design features and user factors, and the relationship of them with guessability score and each other's.

Overall it can be said that this study is a comprehensive study to investigate the comprehension of current Turkish traffic signs combining almost all aspects about the design of signs and the user factors in a single study. It is also the first study of Turkey that is realized in a broader aspect.

7.2. Recommendations for Future Research

Consideration of cognitive design features in the design of traffic signs is useful better comprehension of them. It is hoped that this study is an initial step in evaluation of current Turkish traffic signs and understanding of them by the population of Turkey. The findings of this study may provide useful information and recommendation for the designers for designing more user-friendly and effective traffic signs. Also, they may also help to decision makers of traffic education to organize driver license education program and better allocate their resources for better understanding of the traffic signs.

The traffic signs that were guessed by the population as opposite of true meaning were the ones that were rearranged with EU Harmonization Process that was applied in Turkey in 2004. Therefore, those signs need to be reconsidered again since the misunderstanding may lead undesirable results in traffic and threat the road safety.

Meaningful, semantically close and familiar signs were guessed well by the population. Therefore, in the design of traffic signs the message intended to be given and the icon should be compatible and more practice in driver license education should be planned. Extra attention should be given to the elderly drivers. It is recommended that as a further study behavioral and comprehensible aspects of elderly drivers can be investigated in more details.

Familiarity was found as a significant predictor of guessability in this study. Therefore, it is recommended that in the driving license education some programs that include more practice sessions to improve driving behavior and give experience on traffic sign comprehension may be planned. Moreover, as a further study, the effect of training on comprehension of traffic signs may be studied.

Lastly, some studies in literature showed that cultural issues are important in comprehension of traffic signs. Therefore it will be also useful to examine the cultural issues of Turkish population and the effect of them on comprehension and in the design of traffic signs.

It is hoped that further studies will be undertaken in order to clarify the problems which remain unsolved in the current study.

REFERENCES

Al-Madani, H., 2000, "Influence of Drivers' Comprehension of Posted Signs on Their Safety Related Characteristics", *Accident Analysis and Prevention,* Vol 32, pp. 575-581.

Al-Madani, H., Al-Janahi, A.R., 2002a, "Assessment of Drivers' Comprehension of Traffic Signs Based on Their Traffic, Personal and Social Characteristics", *Transportation Research Part F: Traffic Psychology and Behaviour,* Vol 5 (1), p. 63.

Al-Madani, H., Al-Janahi, A.R., 2002b, "Role of Drivers' Personal Characteristics in Understanding Traffic Sign Symbols", *Accident Analysis and Prevention,* Vol 34, pp. 185–196.

ANSI Z535.3, 2007, "American National Standard Criteria for Safety Symbols", *American National Standards Institute,* 2011.

Al-Yousifi, A.E., 1999, "Investigation of Traffic Signs to Improve Road Safety", *10th International Conference on Traffic Safety on Two Continents*, Malmo, Sweden, pp. 73–87.

Ben-Bassat, T., Shinar, D., 2006, "Ergonomic guidelines for traffic sign design increase sign comprehension", *The Journal of Human Factors and Ergonomics,* Vol 48(1), pp. 182‑195.

Bruyas, M.P., Le Breton, B., Pauzié, A., 1998, "Ergonomic guidelines for the design of pictorial information", *Int. J. Ind. Ergon,* Vol 21, pp 407-13.

Byrne, M. D., 1993, "Using icons to find documents: Simplicity is critical", *Interchi '93 Bridges between worlds: Proceedings of the Conference on Human Factors in Computing Systems,* In S. Ashlund, K. Mullet, A. Henderson, E. Hollnagel, & T.

While (Eds.), pp. 446–465, MA: Addison-Wesley.

Canfield, R.R., 1999, *Traffic signs and markings.* In: Pline, J. (Ed.), Traffic Engineering Handbook, 5th edn. Institute of Transportation Engineers, Washington DC, pp. 411–452.

Chan, A. H.S., Han S. H., Ng, A. W.Y., Park, W., 2009, "Hong Kong Chinese and Korean comprehension of American security safety symbols", *International Journal of Industrial Ergonomics*, Vol 39, pp. 835-850.

Chan, K.L. and Chan, A. H.S., 2011, "Understanding Industrial Safety Signs: Implications for Occupational Safety Management", *Industrial Management & Data Systems*, vol. 111, pp. 1481 – 1510.

Ells, J. G. and Dewar, R. E., 1979, "Rapid comprehension of verbal and symbolic traffic sign message", *The Journal of Human Factors and Ergonomic Society*, Vol 21, pp. 161–168.

García, M., Badre, A. N., Stasko, J. T., 1994, "Development and validation of icons varying in their abstractness," *Interacting with Computers*, Vol. 6, pp. 191-211.

General Directorate of Highway, 2004, "Trafik İşaretleri El Kitabı", Ankara: Bakim Dairesi Başkanliği Trafik Şubesi Müdürlüğü.

General Directorate of Security, 2014, *Traffic Statistical Bulletin*, http://www.trafik.gov.tr/SiteAssets/istatistik/2014_bulten_ozet.xls, accessed at June 2015.

Gupta, K., Baker-Prewitt, J., Miller, J., 2001, "Scaling: the Never-ending Debate - an Examination of the Factors Influencing Scale Construction and Selection", *CASRO Journal*, pp. 125-129.

Hancock, H. E., Rogers, W. A., Fisk, A. D., 1999, "Understanding Age-related Differences

in the Perception and Comprehension of Symbolic Warning Information", *Human Factors and Ergonomics Society Annual Meeting.*

Hawkins, H.G., Picha, D.L., Lopez, C.A., 1998, "Mexican Driver Comprehension of U.S. Traffic Control Devices", *Transportation Research Record,* Vol 1628, pp. 15–24.

Isherwood, S. J., McDougall, S. J. P., Curry, M. B., 2007, "Icon Identification in Context: The Changing Role of Icon Characteristics With User Experience", *The Journal of Human Factor and Ergonomics Society*, Vol 49, pp. 465-476.

ISO 9241-11, 1998 "Ergonomics Requirements for Office Work with Visual Display Terminals Guidance on Usability", https://www.iso.org/obp/ui/#iso:std:iso:9241:-11:ed-1:v1:en, accessed at December 2014.

ISO 15535, 2006, "General Requirements for Establishing Anthropometric Databases", *British Standards Institution*, 2007, London.

ISO 3864-3, 2012, "Graphical symbols -- Safety colours and safety signs - Part 3: Design principles for graphical symbols for use in safety signs".

Jordan, P.W., Draper, S.W., MacFarlane, K.K., McNulty, S. A., 1991, "Guessability, Learnability, and Experienced User Performance", In: Diaper, D., Hammond, N., *People and Computers VI. Cambridge, Cambridge University Press on behalf of the British Computer Society*, pp. 237–245.

Jordan, P.W., 1994, " What is Usability?" In: Robertson, S., *Contemporary Ergonomics 1994*, pp. 454–458. Taylor and Francis, London.

J. P. McDougall, M. B. Curry, and O. de Bruijn, 1999, "Measuring Symbol and Icon Characteristics: Norms for Concreteness, Complexity, Meaningfulness, Familiarity, and Semantic Distance for 239 Symbols", *Behavior Research Methods, Instruments, and Computers,* Vol. 31 No. 3, pp. 487-519.

Kirmizioglu, E., 2010, *Analysis of comprehension of traffic signs: a pilot study in Ankara, Turkey,* M.Sc. Thesis, METU Civil Engineering, Transportation Department, Ankara, Turkey.

Kirmizioğlu, E. and Tuydes-Yaman, 2012, "Comprehensibility of Traffic Signs Among Urban Drivers in Turkey", *Accident Analysis and Prevention,* Vol 45, pp. 131–141.

Koyuncu, M. and Amado, S., 2008, "Effects of Stimulus Type, Duration and Location on Priming of Road Signs: Implications for Driving", *Transportation Research Part F,* Vol 11, pp. 108-125.

Lesch, F. M., 2003, "Comprehension and Memory for Warning Symbols: Age-related Differences and Impact of Training", *Journal of Safety Research,* Vol 34, pp. 495-505.

Lin, R. and Kreifeldt, J. G., 1992, " Understanding the Image Functions for Icon Design", *Human Factors and Ergonomics 36th Annual Meeting,* 1992.

Liu, Y. C. And Ho, C. H., 2012, "The Effects of Age on Symbol Comprehension in Central Rail Hubs in Taiwan", *Applied Ergonomics,* vol 43, pp. 1016-1025.

Malaterre, G., 1990, "Error Analysis and In-depth Accident Studies", *Ergonomics,* Vol 33.

Ng, A.W.Y., Chan, A.H.S., 2007a., "Culture Issues in Traffic Sign Usability", *Usability and Internationalization, Part I, HCII 2007, LNCS 4559,* pp. 379–387.

Ng, A.W.Y., Chan, A.H.S., 2007b., "Cognitive Sign Design Features on Traffic Signs", *Enginnering Letters,* Vol 14:1, pp.13-18.

Ng, A. W. Y., Chan, A. H. S., 2007c., "The Guessability of Traffic Signs: Effects of Prospective-user Factors and Sign Design Features", *Accident Analysis and Prevention*, Vol 39, pp. 1245−1257.

Ng, A.W.Y., Chan, A.H.S., 2008a. "Visual and Cognitive Features on Icon Effectiveness" *Proceedings of the International MultiConference of Engineers and Computer Scientists Vol II IMECS* , 19-21 March 2008, Hong Kong.

Ng, A. W. Y., Chan A. H. S., 2008b, "The effects of driver factors and sign design features on the comprehensibility of traffic signs", *Journal of Safety Research,* Vol 39, pp. 321-328.

Ng, A. W. Y., Chan, A. H. S., 2011, "Investigation of the Effectiveness of Traffic Sign Training in Terms of Training Methods and Sign Characteristics", *Traffic Injury Prevention*, Vol 12, pp. 283−295.

Official Journal of the Republic of Turkey, 1983, "Karayolları Trafik Kanunu", Vol 22, no 5, p 687 (Resmi Gazete).

Ou Y., and Liu, Y., 2012, "Effects of Sign Design Features and Training on Comprehension of Traffic Signs in Taiwanese and Vietnamese User Groups", *International Journal of Industrial Ergonomics,* Vol 42, pp.1-7.

Peden, M., Scurfield R, Sleet D, *et al*, 2004, "World report on road traffic injury prevention", *World Health Organization*, Geneva.

Preston, C. C. and Colman, A. M., 2000, "Optimal Number of Response Categories in Rating Scales: Reliability, Validity, Discriminating Power, and Respondent Preferences", *Acta Psychologica,* Vol. 104, pp. 1-15.

Retting, R. A., Ferguson, S. A. McCartt, A. T., 2003, "A Review of Evidence-Based Traffic Engineering Measures Designed to Reduce Pedestrian–Motor Vehicle Crashes", *Am J Public Health,* Vol 93(9), pp. 1456–1463.

Shinar, D., Dewar, R.E., Summala, H., Zakowska, L., 2003, "Traffic Sign Symbol Comprehension - A Cross-Cultural Study", *Ergonomics,* Vol 46, pp. 1549–1565.

Shinar, D. and Vogelzang, M., 2013, "Comprehension of Traffic Signs with Symbolic Versus Text Displays", *Transportation Research Part F,* Vol 18, pp. 72-82.

Summala, H., 1998, "American Drivers in Europe - Different Signing Policy may cause Safety Problems at Uncontrolled Intersections", *Accident Analysis and Prevention*, Vol 30, pp. 285–289.

Turkish Statistical Institute, 2014, *Traffic Accidents Statistics,* http://www.tuik.gov.tr/PreTablo.do?alt_id=1051, accessed at April 2015

.

United Nations, 2008a, *Vienna Convention on Road Signs and Signals*, http://www.unece.org/trans/conventn/signalse.pdf, accessed at June 2015

United Nations, 2008b, *European Agreement Supplementing the 1968 Convention on Road Traffic*, http://www.unece.org/fileadmin/DAM/trans/conventn/Conv_road_traffic_ EN. pdf, accessed at June 2015.

Ward, S. J., Wogalter, M. S., Mercer, A. W., 2004, "Comprehension and Training of International Road Signs", *Human Factors and Ergonomics 48th Annual Meeting*.

Wolff, J.S. and Wogalter, M.S., 1998, "Comprehension of Pictorial Symbols: Effects of Context and Test Method", *Human Factors*, Vol. 40, pp. 173-86.

Yakut, S., 2006, *Comprehensibility of Traffic Signs*, M.Sc. Thesis, Gazi University, Ankara.

Yannis, G., Antoniou, C., Papadimitriou, E., 2007, "Road Casualties and Enforcement: Distributional Assumptions of Serially Correlated Count Data", *Traffic Injury Prevention,* Vol 8, pp. 300-308.

APPENDIX A: FORMS

The forms that were used during the experiment included in this part of the appendix. These forms are color deficiency declaration form, personal consent form, personal data collection forms and instructions. The experiment was conducted in Turkey, so all forms are also prepared in Turkish.

(i) Color Deficiency Declaration Form (see in Figure A.1): All participants had to be no color deficiency, so participants declared they were not suffering from color deficiency before starting the experiment. Only ones who are free from any color deficiency are accepted to participate in the study.

(ii) Personal Consent Form (see in Figure A.2): "Personal Consent Form" includes a detailed description of the objectives and procedures of the study. In this form, it is mentioned that all personal information will be kept in confidence. To ensure the voluntary participation of the subjects to the study, this form was being signed. Personal Consent

(iii) Personal Data Form: "User Factor Questionnaire" (see in Figure A.3), "Guessability Questionnaire" (see in Figure A.4), "Cognitive Sign Design Feature Evaluation Sheet" (see in Figure A.5) are used to collect personal data. The participants gave the information about age, occupation, family origin, mother's and father's birthplace, birthday, gender, marital status, educational background, active car driving experience, smoking habits. This information is recorded on form called "User Factor Questionnaire". The guessed meaning of signs was recorded on "Guessability Questionnaire". The ratings for cognitive sign design features were recorded on "Cognitive Sign Design Features Evaluation Sheet". Moreover, "Cognitive Sign Design Feature Explanation Sheet" (see in Figure A.6) was used to explain the meanings of features and rating instructions to the participants.

(iv) Instructions (see in Figure A.7): In order to prevent the confusion of the subjects, the experimenter guided the subjects using these instructions.

Color Deficiency Declaration Form

Please check your selected answer.

Do you have any problem to separate colors? Have you ever been diagnosed color deficiency?

☐ Yes

☐ No

Renk Körlüğü Deklarasyon Formu

Lütfen cevabınızı tik işareti ile işaretleyiniz.

Renkleri ayırt etmek ile ilgili bir probleminiz var mı? Daha önce renk körlüğü tanısı aldınız mı?

☐ Evet

☐ Hayır

Figure A.1. Color Deficiency Declaration Form in English and Turkish

Personal Consent Form

This thesis study aims to determine the comprehensibility of Turkish traffics signs and examine the effects of cognitive sign design features and user factors on comprehension. You do not have any color deficiency problem which affects your participation to the experiments adversely.

This comprehension study will help to understand the comprehensibility level of current Turkish traffic signs, help to evaluate them with ergonomic tools and be reference for the further development of them for effective traffic control and safe driving. All personal data will be confident only the general results will be used as a data.

If you decided to participate, please take into consideration the issues below:

i. Before the experiments your birthday, birth place, your family origin, occupation, gender, marital status, educational background, active car driving experience, age and smoking habits information will be asked. You will fill the "User Factors Questionnaire" that took about in 5 minutes to fill out first.

ii. Experiments will be performed in two phases in two different days. In the first day, you were asked to guess the meaning of selected 60 Turkish traffic signs and write down your answers to the given "Guessability Questionnaire". The second phase of experiment will be realized one day after the first experiment. You will be asked to rate each sign between 0-100 points for five cognitive sign design features; familiarity, concreteness, simplicity, meaningfulness and semantic closeness and write down to the "Sign Design Features Evaluation Sheet". There will be no time pressure to complete the survey. No feedback will be given during the experiment. To avoid fatigue, 1 minute break will be given after 30 signs answered.

Your participation is completely voluntary. You may choose to withdraw from participation at any time. All information obtained during this study will be held in strict confidence.

Figure A.2. Personal Consent Form in English and Turkish

If at any time you have questions regarding this research, you may contact either Gizem Zengin or Ph. D. Mahmut Ekşioğlu from Department of Industrial Engineering of Boğaziçi University.

By placing your signature below, you will accept that your participation to this study is voluntary. However, you can choose to withdraw from participation at any time at no cost or obligation to you.

Signature of Participant: Date:

Kişisel Kabul Formu

Bu tez çalışması, mevcut trafik işaretlerinin, yol kullanıcıları tarafından ne kadar anlaşılabilir olduğunu tespit etmeyi ve ergonomik kavramsal dizayn faktörleri ve kullanıcı karakterlerinin anlaşılabilirlik üzerindeki etkisini değerlendirmeyi hedeflemektedir. Bu çalışmaya engel teşkil edecek bir renk körlüğü probleminizin olmamasından dolayı, deneylere katılmak için uygun durumda bulunmaktasınız.

Bu çalışmadan elde edilecek sonuçlar, mevcut Türk trafik işaretlerinin ne kadar anlaşılabilir olduğunu gösterecek, ergomik araçların anlaşılabilirlik üzerindeki etkisini ortaya çıkaracak ve trafik işaretlerinin amacına hizmet edilecek şekilde geliştirilmesi ve düzenlenmesine yardımcı olacaktır. Böylece hem etkili trafik kontolü yapılacak hem de sürüç güvenliği sağlanacaktır. Çalışmada tüm kişisel datalar gizli tutalacak sadece genel sonuçlar data olarak kullanılacaktır.

Eğer katılmaya karar verdiyseniz. lütfen aşağıdaki hususlara dikkat ediniz:

Figure A.2. Personal Consent Form in English and Turkish (cont.)

(i) Deneye başlamadan önce doğum tarihiniz, doğum yeriniz, ailenizin doğum yeri, mesleğiniz, cinsiyetiniz, medeni durumunuz, eğitim geçmişiniz, aktif sürüş deneyiminiz, yaşınız ve sigara kullanma alışkanlığınız sorulacaktır. Bu bilgileri "Kullanıcı Faktörleri Anketi" ne girmeniz gerekmektedir.

(ii) Deney iki farklı günde gerçekleştirilecek iki aşamadan oluşmaktadır. Deneyin birinci aşamasında seçilen 60 trafik işaretinin anlamını tahmin ederek, cevabınızı "Tahmin Anketi"ne girmeniz istenecektir. Birinci aşamadan 1 gün sonra gerçekleştirilecek ikinci aşamada ise her bir işareti beş kavramsal işaret dizayn faktörüne; aşinalık, somutluk, basitlik, anlamlılık, verilmek istenen anlamı karşılama, 0 ile 100 arasında puanlamanız gerekmektedir. Bu bilgileri "İşaret Dizayn Değerlendirme Anketi"ne girmeniz gerekmektedir. Bu süreçte hiç bir zaman kısıt olmayacak, deney sırasında deneyi yapan kişi tarafından geri bildirim verilmeyecektir. Yorgunluğu engellemek için 30 trafik işareti puanlandıktan sonra 1 dakikalık ara verilecektir.

Katılımınız tamamen gönüllü olup, katılmanız için herhangi bir zorlamayla karşılaşmayacaksınız. Dilediğinizde, çalışmanın herhangi bir aşamasında çalışmayı terk edebilirsiniz. Elde edilecek kişisel bilgiler kimseyle paylaşılmayacak.

Bu çalışmayla ilgili sorularınız ve katkılarınız olması durumunda Boğaziçi Üniversitesi Endüstri Mühendisliği Bölümü'nde Gizem Zengin veya Doç. Dr. Mahmut Ekşioğlu ile temasa geçebilirsiniz.

Aşağıya atacağınız imza, bu çalışmaya gönüllü olarak katılmak istediğinizi belirtmektedir; ancak çalışmayı yarıda bırakmanız durumunda, size herhangi bir yükümlülük getirmemektedir.

Katılımcının İmzası: Tarih:

Figure A.2. Personal Consent Form in English and Turkish (cont.)

USER FACTORS QUESTIONNAIRE		
	Information will be used only as data for the experiment without reporting the name of the participant.	
Name and Surname	:	
Date of birth (day/month/year)	:	**Birth of place :** **Current city:**
Family origin city	:	**Mother's birthplace :** **Father's birthplace :**
Occupation	:	
Age	:	Please select the age range below and write your exact age in the given area(-----). ☐ 18-29 ------ ☐ 40-49 ------ ☐ 60-69 ----- ☐ 30-39 ------ ☐ 50-59 ------ ☐ 70-79 -----
Gender	:	☐ Female ☐ Male
Marital status	:	☐ Single ☐ Married ☐ Married with children
Do you smoke cigarettes?	:	☐ Yes ☐ No
Educational background (Last graduated or continuing degree)	:	☐ Primary or secondary school ☐ High School ☐ University or higher
Hold a driver's licence?	:	☐ Yes ☐ No **Type of driver's licence:** ----------
Active car driving experience	:	☐ Not an active driver ☐ Drive less than or equal to 1 year ☐ Drive more than 1 year and less than 5 years ☐ Drive more than or equal to 5 years

Figure A.3. User Factors Questionnaire in English and Turkish

KATILIMCI BİLGİ FORMU		
Burada yer alan bilgiler, isim belirtilmeden sadece araştırmada veri olarak kullanılacaktır.		
İsim ve Soyisim	:	
Doğum Tarihi (gün/ay/yıl)	:	**Doğum yeri :** **Yaşadığı yer :**
Aslen Nerelisiniz?	:	**Annenizin doğum yeri :** **Babanızın doğum yeri :**
Mesleginiz	:	
Yaş	:	Lütfen bulunduğunuz yaş aralığını seçerek, ----- ile belirtilen alana yaşınızı yazınız. ☐ 18-29 ------ ☐ 40-49 ------ ☐ 60-69 ------ ☐ 30-39 ------ ☐ 50-59 ------ ☐ 70-79 ------
Cinsiyeti	:	☐ Bayan ☐ Bay
Medeni Durum	:	☐ Bekar ☐ Evli ☐ Evli ve Çocuklu
Sigara içiyor musunuz?	:	☐ Evet ☐ Hayır
Eğitim Durumu	:	☐ İlkokul + Ortaokul ☐ Lise ☐ Üniversite ve sonrası
Aktif araç kullanımı	:	☐ Aktif kullanıcı değilim. ☐ 1 yıldan az ☐ 1 ile 5 yıl arasında ☐ 5 yıl ve üstü

Figure A.3. User Factors Questionnaire in English and Turkish (cont.)

GUESSABILITY QUESTIONNAIRE											
"What does this sign mean?"											
No	Sign	Meaning	No	Sign	Meaning	No	Sign	Meaning	No	Sign	Meaning
1			16			31			46		
2			17			32			47		
3			18			33			48		
4			19			34			49		
5			20			35			50		
6			21			36			51		
7			22			37			52		
8			23			38			53		
9			24			39			54		
10			25			40			55		
11			26			41			56		
12			27			42			57		
13			28			43			58		
14			29			44			59		
15			30			45			60		

Figure A.4. Guessability Questionnaire in English and Turkish

İŞARET TAHMİN FORMU											
"Bu işaret sizce ne anlama gelmektedir?"											
No	İşaret	Anlamı	No	İşaret	Anlamı	No	İşaret	Anlamı	No	İşaret	Anlamı
1			16			31			46		
2			17			32			47		
3			18			33			48		
4			19			34			49		
5			20			35			50		
6			21			36			51		
7			22			37			52		
8			23			38			53		
9			24			39			54		
10			25			40			55		
11			26			41			56		
12			27			42			57		
13			28			43			58		
14			29			44			59		
15			30			45			60		

Figure A.4. Guessability Questionnaire in English and Turkish (cont.)

	Sign Design Features	Familiarity	Concreteness	Simplicity	Meaningfulness	Semantic Closeness
		"How much is this signs familiar to you? How often do you encounter daily?"	"How much concrete is the symbol only (not sign)?"	"How much simple is the symbol only (not sign)?"	"Even you do not know the meaning, does the sign mean anything to you? How much?"	"How much the symbol used and the given meaning are related in your opinion?"
No	**Sign**					
1						
2						
3						
4						
5						
6						
7						
8						
9						
10						

Figure A.5. Cognitive Sign Design Feature Evaluation Sheet in English and Turkish

	İşaret Dizayn Faktörleri	Aşinalık	Somutluk	Basitlik	Anlamlılık	Verilmek İstenen Anlamı Karşılama
		"Bu işarete ne kadar aşinasınız? Sıklıkla karşılaştığınız bir işaret mi?"	"İşaret üzerindeki sembol ne kadar somut?Gerçek dünyada karşılığı var mı?	"İşaret üzerindeki sembol ne kadar basit? Ne kadar az parça ile ifade edilmiş?"	"İşaretin anlamını bilmeseniz bile ilk bakışta size ne kadar anlam ifade ediyor?"	"İşaretteki sembol ile verilmek istenen gerçek anlamı ne kadar uyumlu"
No	**İşaret**					
1						
2						
3						
4						
5						
6						
7						
8						
9						
10						

Figure A.5. Cognitive Sign Design Feature Evaluation Sheet in English and Turkish (cont.)

Sign Design Features with Explanations	
Familiarity	
Definition	The signs that frequently encountered in daily life. Highly encountered signs are more familiar.
Evaluation Criteria	Very unfamiliar → Very familiar 0 50 100
Example	 
Concreteness	
Definition	The things that have a physical referrant in real world and detectable with sense organs are concrete. The things that are not concrete are called as abstract.
Evaluation Criteria	Very abstract → Very concrete 0 50 100
Example	 
Simplicity	
Definition	Symbol is complex if it contains a lot of detail, simple if it has a few elements and little detail. When the number of shapes(arrow, arch,line etc.) are increased, symbol gets more complex.
Evaluation Criteria	Very Complex → Very Simple 0 50 100
Example	 
Meaningfulness	
Definition	Even the exact meaning of the symbol was not known, how meaningful the symbol is perceived.
Evaluation Criteria	Very meaningless → Very meaningful 0 50 100
Example	 
Semantic Closeness	
Definition	The closeness of the relationship between what is depicted in sign and the the meaning intented to give.
Evaluation Criteria	Very weakly related → Very strongly related 0 50 100
Example	 Peace  Elevator  Typewriter

Figure A.6. Cognitive Sign Design Feature Explanation Sheet in English and Turkish

Tasarım Faktörü	Açıklama	
Tanıdık Gelmek, Aşina olmak (Familiarity)	Tanım	Günlük yaşamda sık karşılaşılan, alışkın olduğumuz işaretlerdir.
	Değerlendirme Skalası	Hiç Tanıdık Değil → Çok Tanıdık 0 50 100
	Örnek	
Somutluk, Ne kadar somut ifadeler içeriyor? (Concreteness)	Tanım	Elle tutulabilen ya da gözle görülebilen gerçek objelerin, nesnelerin ya da insan ve hayvanların resmedildiği işaretler somuttur. Somut olmayan işaretler ise soyuttur.
	Değerlendirme Skalası	Tamamen Soyut → Tamamen Somut 0 50 100
	Örnek	
Basitlik, İşaret ne kadar basit? (Simplicity)	Tanım	İşaret çok fazla detay içeriyorsa karmaşık, az detay ve parça içeriyor ise basit olarak tanımlanır.
	Değerlendirme Skalası	Çok Karmaşık → Çok Basit 0 50 100
	Örnek	
İşaretin, anlamını bilmeseniz bile bir anlam ifade etmesi (Meaningfulness)	Tanım	İşaretin ne kadar anlamlı bulunduğudur. İşarete bakıldığında bir anlam çıkarılabilmesidir.
	Değerlendirme Skalası	Hiç anlam ifade etmeyen → Tam anlam ifade eden 0 50 100
	Örnek	
Resmedilen görsel ile anlatılmak istenen arasındaki ilişki (Semantic Closeness)	Tanım	İşaret üzerinde resmedilen görsel ile anlatılmak istenen arasındaki ilişkidir. İşaret, anlatılmak istenen ile ne kadar iyi ifade ediyor?
	Değerlendirme Skalası	Hiç ilişkili değil → Çok İlişkili 0 50 100
	Örnek	 Barış Asansör Daktilo

Figure A.6. Cognitive Sign Design Feature Explanation Sheet in English and Turkish (cont.)

Instructions

(i) You have read the personal consent form and received information about the experiment. Now, please fill "User Factors Questionnaire".

(ii) 60 traffic signs will be presented in random order. Please write down your opinion about the meaning of sign on the blank placed on Guessability Questionnaire".

(iii) The meaning of each cognitive sign design features and 0-100 points rating information mentioned in "Cognitive Sign Design Feature Evaluation Sheet". When you are ready for the experiments, inform the experimenter.

(iv) 60 traffic signs will be presented in random order. For each sign please rate between 0-100 points for familiarity, concreteness, simplicity and meaningfulness. After rating for four sign design feature please press "Next". After pressing, the real meaning of the sign will be presented on the screen. After that please rate for the semantic closeness feature. There is no time pressure on experiment, please take your time when rating. You can get 1 minute rest after 30 signs are completed.

Talimatlar

(i) Kişisel kabul formunu okudunuz ve deney hakkında bilgi edindiniz. Simdi lütfen "Katılımcı Bilgi Formunu" doldurun.

(ii) 60 trafik işareti rastgele bir düzende gösterilecektir. Lütfen her işaret için tahmin ettiğiniz anlamı, size verilen "İşaret Tahmin Formuna" yazınız.

(iii) Kavramsal tasarım faktörlerinin açıklaması ve 0 ile 100 arasında yapacağınız puanlamanın skalası "Kavramsal İşaret Tasarım Faktörleri Açıklaması" sayfasında belirtilmiştir. Lütfen hazır olunca deneyi gerçekleştirecek kişiye haber veriniz.

(iv) 60 trafik işareti ratgele bir şekilde gösterilecektir. Her işaret için, işaret gösterildikten sonra öncelikle ilk dört kavramsal tasarım faktörü; aşinalık, somutluk, basitlik ve anlamlılık için 0 ile 100 arasındaki puanlamanızı yapınız. Sonrasında ileri tuşuna basınız. İşaretin gerçek anlamı işaretin altında gösterilecektir. Sonra lütfen verilen anlamı karşılama faktörü için puanlamanızı yapınız. Puanlamanızı "Kavramsal İşaret Tasarım Faktörleri Değerlendirme Formuna" giriniz. Deneyde bir süre kısıtı yoktur. 30 trafik işareti tamamlandıktan sonra 1 dakikalık dinlenme verilecektir.

Figure A.6. Instructions in English and Turkish

APPENDIX B: GUESSABILITY SCORE AND SIGN DESIGN FEATURES SCORE CALCULATIONS

B.1. Calculation Formulas for Guessability Score

Guessability Score of any sign $= \left(\frac{x}{n}\right) 100$ where,

x : Number of completely correct answer given by participants for any sign

n : Total number of participants (201 participants)

Guessability Score of sign group $= \left(\frac{x}{nm}\right) 100$ where,

x : Number of completely correct answer given by participants for the signs in any group

n : Total number of participants (201 participants)

m : Number of signs in any group (for example, 22 signs are included in warning sign group)

Guessability Score of overall signs $= \left(\frac{x}{nk}\right) 100$ where,

x : Number of completely correct answer given by participants for all signs

n : Total number of participants (201 participants)

k : Total number of signs (60 traffic signs)

Guessability Score of any participant $= \left(\frac{x}{m}\right) 100$ where,

x : Number of completely correct answer given by any participant for the all signs

k : Total number of signs (60 traffic signs)

B.2. Calculation Formulas for Sign Design Feature Score

Overall Mean score of any sign design feature$= \left(\frac{p}{nk}\right)$ where,

p : Total of given scores for all signs for any sign design features by all participants

n : Total number of participants (201 participants)

k : Total number of signs (60 traffic signs)

Mean score of any sign design feature for any sign group $= \left(\frac{r}{nm}\right)$ where,

p : Total of given scores for the signs included in any sign group for any sign design features by all participants

n : Total number of participants (201 participants)

m : Number of signs in any group (for example, 22 signs are included in warning sign group)

Mean score of any sign design feature for any user factor group $= \left(\frac{s}{xk}\right)$ where,

p : Total of given scores for 60 signs by the participants included in any group of user factors

x : Total number of participants included in any group of user factors (for example 41 participants are included in 18-29 yrs. group of age factor)

k : Total number of signs (60 traffic signs)

APPENDIX C: REGRESSION EQUATIONS

C.1. Equation Models for Full Data

Model 1. The regression equation is

Guessability Score = -26,73 + 1,1726 Meaningfulness

Model 2. The regression equation is

Guessability Score = -26,73 + 1,1726 Meaningfulness

Model 3. The regression equation is

Guessability Score = -30,12 + 0,894 Familiarity + 0,4200 Semantic Closeness

Model 4. The regression equation is

Guessability Score = -32,84 + 0,685 Familiarity + 0,650 Meaningfulness

Model 5. The regression equation is

Guessability Score = -32,66 + 0,698 Familiarity + 0,369 Meaningfulness + 0,274 Semantic Closeness

Model 6. The regression equation is

Guessability Score = -32,65 + 0,910 Familiarity + 0,0940 Concreteness + 0,3656 Semantic Closeness

Model 7. The regression equation is

Guessability Score = -31,92 + 0,633 Familiarity - 0,054 Concreteness + 0,472 Meaningfulness + 0,265 Semantic Closeness

Model 8. The regression equation is

Guessability Score = -34,49 + 0,679 Familiarity + 0,052 Simplicity + 0,362 Meaningfulness + 0,270 Semantic Closeness

Model 9. The regression equation is

Guessability Score = -34,27 + 0,593 Familiarity - 0,066 Concreteness + 0,071 Simplicity + 0,485 Meaningfulness + 0,257 Semantic Closeness