2023

INCREASING THE SAFETY OF BICYCLISTS USING A CYCLIST BEHAVIOR QUESTIONNAIRE AND A SMARTPHONE BASED APPLICATION

Anika Jannat Rimu

INCREASING THE SAFETY OF BICYCLISTS USING A CYCLIST BEHAVIOR QUESTIONNAIRE AND A SMARTPHONE BASED APPLICATION

by

ANIKA JANNAT RIMU

A Dissertation

Presented to the Faculty of the Graduate School of

The University of Texas at Arlington in Partial Fulfillment of the Requirements

for the Degree of Doctor of Philosophy at

The University of Texas at Arlington

August 2023

Arlington, Texas

Supervising Committee:

Dr. Shuchisnigdha Deb, Supervising Professor

Dr. Victoria Chen

Dr. Yuan Zhou

Dr. Ming Li

ACKNOWLEDGEMENTS

I am very grateful to my supervisor and committee chair, Dr. Shuchisnigdha Deb, for her support, mentorship, and guidance. I am also grateful to Dr. Victoria Chen, Dr. Yuan Zhou, and Dr. Ming Li for joining my committee and providing me with helpful advice, continuous support, and guidance during my doctoral program.

I would also like to express my sincere gratitude to the faculty and staff of the Industrial Engineering Department at UT Arlington for their continuous support. Finally, I would like to thank my parents who have sacrificed a lot to support me since I moved to this country. Without my family's unconditional support, I wouldn't be here now. I would also like to thank all my friends, both here and back home, who always supported me and believed in me.

Above all, I thank God, the Almighty for blessing me with this opportunity.

Thank you all.

ABSTRACT

Bicycling is beneficial for health, the environment, road users' flexibility, and personal expenses. Compared to motor vehicles, they are an active mode of transport, cause minimum pollution, are affordable, and can easily navigate through the increasing traffic all over the world. This increase in traffic, however, also increases the possibility of crashes with motor vehicles. Bicyclists, being more exposed to traffic than drivers, suffer fatal consequences from a crash. Therefore, a standard tool is required to understand bicyclist behavior on the road. This tool can provide insights into bicyclists' behavior so that appropriate infrastructure or policy changes can be implemented. Furthermore, affordable technology can be utilized to assist bicyclists by alerting them of imminent danger ahead of time. The objectives of this research are to 1) develop and validate a Cyclist Behavior Questionnaire (CBQ) for the US population and 2) identify an effective warning system for a smartphone-based app to alert bicyclists. To accomplish the first objective, a CBQ was developed and administered online. A Principal Component Analysis (PCA) determined the 11-item 4 factorial structure of CBQ, which was later verified using a Confirmatory Factor Analysis (CFA). An innovative methodology was developed and implemented to validate self-reported responses of CBQ with bicyclists' actual responses from a bike-simulator study. For the second objective, a focus group study with experts was conducted. Experts identified a list of potential warning signals including red/yellow flashing signals, and tone/speech audible signals. A bike-simulator experiment further investigated the efficacy of these signals under different environmental factors. The results were analyzed using cyclists' response to the warnings, as well as their physiological and emotional reaction. Results identified a multimodal combination of red visual and tone audible warning to be the most efficient at alerting cyclists. The findings of these studies will improve the understanding of bicyclists' behavior and their interaction with

technologies while riding. Future research should focus on how the adoption of these technologies would affect bicycling skills and behavior (for example, situation awareness).

LIST OF FIGURES

LIST OF TABLES

TABLE OF CONTENTS

CHAPTER I: THE OVERVIEW OF THE PROPOSED STUDY

Bicycles have been a convenient and ecofriendly transportation mode all over the world for decades. According to BBC, there has been a dramatic increase in the usage of bicycles during this COVID-19 era (Bernhard, 2020). The report also mentions that the reason behind this increase could be due to the anxiety over using public transportation where maintaining social distancing is difficult, or it could also be due to the surge in exercise over nationwide lockdown. According to the United States Census Bureau, in 2019, roughly 870,000 people reported commuting by a bicycle; this is most common inside cities of metropolitan areas (Jin et al., 2021; Burrows, 2019). Cycling is not only an ecofriendly mode of transport but also possesses remarkable health benefits, which are essential now with the increasing obesity rates in developed countries. This is the reason why policymakers in countries like Austria, Finland and Norway have been promoting cycling (Utriainen, 2020). In the United States, The Fixing America's Surface Transportation (FAST) Act was passed by Congress and passed into law by President Obama in 2015 (Law, 2015). FAST Act included a budget of $4 billion in total to make roads safer for bicyclists by including more bike lanes, paths, trails, and other improvements of infrastructure. Despite having so many health benefits, the one drawback of promoting this mode of transportation is the increase in crashes, which can diminish the quality of life temporarily or permanently, as research has shown (IIHS, 2021; NSC, 2021; NHTSA, 2020).

Road users such as pedestrians, bicyclists, and powered two-wheeler riders are considered vulnerable road users (VRU) as they are at a higher risk of injury in the event of a collision with motor vehicles (Ngwu et al., 2022a; Yannis et al., 2020). For instance, according to the National Center for Health Statistics mortality data, there was a total injury of 41,615 bicyclists in 2021 (T. Stewart, 2023). The National Highway Traffic Safety Administration (NHTSA) reported bicyclist deaths account for 2% of all motor vehicle traffic fatalities (NHTSA, 2022; T. Stewart, 2023). In the "Overview of Motor Vehicle Traffic Crashes in 2021" report published by NHTSA, 966 pedal cyclists were reported to die in a motor vehicle traffic crash in the United States (T. Stewart, 2023). The report further mentions that this was an increase in crashes compared to 2020, which was 948 crashes. Additionally, bicyclist deaths have increased by 55% since 2010, when the number reached its lowest point. The NHTSA report includes further classification of bicyclist fatality by other categories such as gender, age, road condition, road type, and time. The fatality rate of males has been higher than that of females, with males accounting for 84% of bicyclist deaths. In terms of age, 88% of the bicyclists who died were aged 20 or older, which is triple the number of deaths in the same age group snice 1975. Around 22% of the bicyclist deaths occurred between 6 pm and 9 pm and 35% of the deaths occurred at intersections (IIHS, 2023; T. Stewart, 2023). Furthermore, bicycle-related deaths seemed to have peaked in the summer months from June through October, where the highest number of deaths occurred in September (119) and lowest in February (45) in 2021 (IIHS, 2023; T. Stewart, 2023).

These crash statistics indicate the need to conduct research on finding ways to assist bicyclists that can prevent crashes with motor vehicles. The objective of this paper is to aid bicyclists on the road by doing extensive research on their behavior and assistive technology. Bicyclists often exhibit unpredictable behavior on the road and since the consequences of such behavior are severe, it is necessary to understand bicyclist behavior and develop solutions based on those behaviors. The literature review showed several research gaps including not having a validated questionnaire to study the US bicyclists' behavior on roads and inadequate research on cost-effective technology to assist bicyclists for safe travel. This research aims to bridge these two gaps by first, understanding bicyclist's behavior on the road with the development and validation of the Cyclist Behavior Questionnaire (CBQ) for the U.S cyclist population and then, developing warning signals to use in a smartphone-based application by using findings from a focus-group study followed by a bike-simulator study. Cyclists' behavior questionnaire, developed in the first study, will be used to find the effect of their behavior on their responses to the assistive smartphone application, developed in the second study.

Two studies were designed for this research to accomplish the abovementioned goals. Chapter II describes the development of CBQ from an online survey and the validation of the questionnaire using a bike-simulator study. Chapter III describes the development of a warning system based on cyclist behavior and cognitive principles. These studies have the following objectives:

Study I: Evaluating cyclist behavior on the road: An innovative approach to validate cyclists' behavior questionnaire for the US population.

1. Developing a CBQ based on the existing behavior questionnaires from extensive literature review and by including modifications for the U.S cyclist population.
2. Identifying the factors of the questionnaire using an online survey study with the U.S cyclist population.
3. Validating the CBQ using a bike-simulator study with factors-specific scenarios.
4. Analyzing the influence of demographics on the CBQ.

Study II: Identifying a warning system for a smartphone-based assistive application for bicyclists:

1. Identifying warning signals from a focus-group study with experts and stakeholders in the field of transportation.

2. Testing potential warning signals and their efficacy under different environmental conditions using a bike-simulator study.
3. Testing cyclists’ comfort using the warning system from their physiological responses (heart rate) and experience surveys.

CHAPTER II: STUDY I

EVALUATING CYCLIST BEAVIOR ON THE ROAD: AN INNOVATIVE APPROACH TO VALIDATE CYCLIST BEHAVIOR QUESTIONNAIRE

2.1 Introduction

The crash statistics in Chapter I indicate that it is of utmost importance to understand bicyclists' behaviors so that the deaths and injuries of cyclists can be mitigated. In previous studies conducted on factors that contributes to cyclists' crashes, researchers have found that cyclists' age, gender, and exposure to cycling are significantly important (Boele-Vos et al., 2017; Martínez-Ruiz et al., 2014). Cyclists who are older in age, male cyclists regardless of their age, as well as cyclists who travel more frequently had a higher risk of getting into a crash (Chen & Shen, 2016). NHTSA has reported some common causes of crashes, such as lack of proper infrastructures designed to assist cyclists; losing control of the bicycle; riding bicycles in an unfamiliar area; not complying with traffic laws, or inappropriate usage of facilities dedicated for cycling or crossing roads (NHTSA, 2019). Additionally, aberrant cyclists' behaviors, such as cycling at high speed or participating in competitive riding; bicyclists' negligence by not wearing helmets or not wearing proper reflecting clothing while cycling in the dark; riding under the influence of alcohol or illicit substances; riding while using cellphones; and lastly, riding despite having environmental hazards such as adverse weather or road anomalies are also reported by NHTSA (NHTSA, 2019).

In most situations, it is up to the cyclists to be aware of their surroundings and take proper precautions to avoid an accident. In an accident between a cyclist and a motor vehicle, it is most likely that the cyclist, being the vulnerable road user (VRU), will suffer severe consequences. However, from the NHTSA report it is evident that cyclist behavior is a major factor that contributed to many crashes which could have been avoided if the cyclist had been more careful

(NHTSA, 2019). Therefore, having a valid survey tool to investigate cyclist's generic behavior on the road will help government officials and policymakers understand the causes of such accidents and will provide guidance and direction to improve bicyclist safety.

There is extensive research conducted on the behavior and the interaction of cyclists with other road users all around the world; however, surveys and questionnaires describing the behavior of the U.S population are still not well developed and validated. Many past studies have identified the factors that can be used to understand aberrant bicyclist behaviors (O'Hern et al., 2020; Useche et al., 2018; Hezaveh et al., 2018; Cristea & Gheorghiu, 2016; Fernández-Heredia et al., 2014). These factors specify each aspect of bicyclists' cycling behavior, compliance with traffic rules, and interactions with infrastructures and road users. On an average, bicyclists' perception of safety varies based on gender, age, and infrastructure (Wang et al., 2019, 2020b; Useche et al., 2018; Carter et al., 2007). Male bicyclists, as well as younger bicyclists, exhibit more aberrant behaviors compared to female and older bicyclists (Wang et al., 2019). The factors that they used to associate both gender and age were aggressive violations, ordinary violations, personal control errors, and distractions. In another research, age was associated with error and violation, where older bicyclists reported fewer errors and violations (O'Hern et al., 2020). Gender, however, was only associated with errors, where male bicyclists reported fewer errors than their female counterpart. For male cyclists, age, and risk perceptions had been significant behavioral predictors (Useche, Montoro, Alonso, et al., 2018). When comparing demographics among different countries, Li et al. (2022) found male cyclists to have higher frequencies of traffic violation, and lower positive behaviors in Australia, China, and Colombia. However, they found male cyclists in Colombia to have higher number of crashes than female cyclists. Similar to the O'Hern et al. (2020) study mentioned above, Li et al. (2022) also found older cyclists in Australia and Colombia to report fewer errors and more positive behaviors. Useche et al. (2021) found male cyclists in Belgium to report higher rates of traffic violation and female cyclist to report higher rates of positive behavior. Therefore, demographics is an important factor to understand bicyclists' behavior. This study achieves this goal by developing a behavioral questionnaire for the US population. In the next chapter, an extensive literature review will be conducted to summarize existing questionnaires and the factors identified from them, as well as their association with the demographics of bicyclists.

The overall objective of this study is to develop and validate a Cyclists Behavior Questionnaire for the U.S. population. The factorial structure and model of the questionnaire were also validated using specific factor-based scenarios using a bike simulator.

2.1.1 Research Objective and Research Questions

***Objective I: Developing Cyclist Behavior Questionnaire for the U.S population*:** A self-report survey was created based on literature review. Questions regarding demographics and behavioral factors were included in the survey. The numerical survey responses were collected on a 5-point Likert scale from never (1) to almost always (5).

***Objective II: Identifying specific factors from the survey results*:** The factorial structure that explained the majority of the cyclists' behavior was identified using the Principal Component Analysis (PCA). The internal reliability of the subscales obtained from the PCA was tested using Cronbach's alpha (α).

***Objective III: Validating the factorial structure of the questionnaire*:** The factorial structure and associated model of the CBQ was validated using a Confirmatory Factor Analysis (CFA).

***Objective IV: Validating responses from CBQ using a scenario-based bike-simulator study*:** The factors identified from PCA, and confirmed from CFA, were used to develop scenarios on the bike simulator. An ordinal logistic regression confirmed the association between participants' scenario-based responses with survey-based scores for the specific behavior.

This study investigates the following research questions based on the objectives:

> **Question 1A**: How effectively did the PCA explore that the US cyclist's behavior can be explained by the factors- violation, error, positive behavior, and distraction and forgetfulness?
>
> **Question 1B**: How did the chi square test and goodness of fit indices of the CFA show that the model is consistent with the pattern of covariation among the observed variables from the PCA?
>
> **Question 1C**: How accurately did the observable behavioral responses for specific factor-based scenarios for the simulator study validate the cyclists' behavior questionnaire?

2.2 Literature Review

2.2.1 Existing Research on Developing Behavior Questionnaire

There are many different methods of surveying and collecting perceived data on how cyclists behave on the road and interact with other road users. In the study by Carter et al. (2007), a bicycle intersection safety index (Bike ISI) was developed, and a survey was created that utilized 67 video clips of various sites, and participants were asked to rate how safe they perceived the sites to be. The respondents for this survey were experts in bicyclist safety research and were asked to rate the behaviors of cyclists in the videos on a scale from 1 to 6, with 1 being the safest and 6 being the most dangerous (Carter et al., 2007). The same survey was administered online to the general population in the USA and the clips shown in the survey were recorded from a camera positioned on the leg of interest (of the cyclist) that was facing toward the intersections. The results of this survey showed that the two safety measures, crash avoidance maneuvers and subjective rating on intersection safety, provided a multi-faceted perception about the safety of a particular intersection. However, the model developed in this study is not comprehensive, does not include and classify all the potential risky or safe behaviors of cyclists, and should be considered for future validation. In another study by Bernhoft & Carstensen (2008), the preferences and behavior of both pedestrians and cyclists were investigated for specific populations in Denmark. There were two populations that were considered: older pedestrians and cyclists over 70 years, and younger pedestrians and cyclists aged 40-49 (Bernhoft & Carstensen, 2008). The questionnaire developed for this study was very short with only three questions about preferences, and four questions about their behaviors on the road (not including demographic questions). The results from this study showed that although older participants preferred safer infrastructure and following traffic laws, they might take contradictory actions to avoid a detour due to their poor health. Researchers in this study mentioned that older pedestrians and cyclists did not follow as many safety margins as the younger participants. One limitation that comes with comparing younger and older age groups is that whether the difference is due to the age or changing society is often indistinguishable. Therefore, the changes in policy and laws to accommodate the present-day older population might not be applicable in the future.

Looking at the impact of age on cyclist behavior, Feenstra et al. (2010) developed an adolescent cycling behavior questionnaire (ACBQ) to determine the relevance of certain key misbehaviors and risk factors. The survey tool was created based on the Driver Behavior Questionnaire (DBQ) and included various risky cycling behaviors. Factor analysis revealed a three-factor structure which includes errors, common violations, and exceptional violations. The confirmatory factor analyses suggested a good model fit and the reliability of this 3-factor scale was confirmed. ACBQ was found to be a useful instrument in determining the risky behaviors among young cyclists in the Netherlands. The validation of ACBQ was supported in a follow-up study that observed participants in a smaller age group among adolescents and found similar results in terms of risky behaviors in young cyclists (Feenstra et al., 2011). The shortcomings of this study included underreported accidents due to the self-report questionnaire and using cross sectional data. As a result, the accident experience had an influence on the behavior rather than behavior having an influence on the accident experience. ACBQ only consisted of negative or risky behaviors; having safe and positive behaviors can improve the questionnaire.

In a field study, a site survey was used to look at the crossing behavior of cyclists at urban intersections and data was gathered in real-time to minimize bias or skewed results that typically come with self-reporting surveys. The survey was conducted at three selected signalized intersections near Jiaotong University in China by placing video cameras at each location during peak and off-peak hours (Yang et al., 2012).The results of the survey showed that the longer the waiting time elapsed, the more likely cyclists would end the wait and run through the red light. By obtaining results this way, the study was able to explain the behavior of cyclists and their intentions more accurately, instead of self-reporting measures. However, for a standardized tool, it is important to include all the potential interactions and behaviors of the general cyclist population. A field study located in a few signalized intersections of a university area cannot be a comprehensive representation of all possible interactions and of general cyclist behavior. Similar to the Bernhoft & Carstensen (2008) study, a survey was used in Mexico to determine the relationship between the behaviors of cyclists and their perceptions toward cycling and risks (Fernández-Heredia et al., 2014). The main conclusion from this study was that convenience and exogenous restrictions (danger, vandalism, facilities) to be the most important factors to understand cyclists' perceptions toward cycling. Danger is described as the perception of risk in relation to accidents or falls, vandalism is described as fear of the bicycle being stolen, and facilities would be the need for complimentary bicycle facilities. In another study, researchers tested the factors of the theory of planned behavior (TPB) to explain risk taking behaviors among cyclists. A scenario-based questionnaire was used in this study. In the questionnaire, participants were expected to read each of the two traffic situations and the two corresponding scenarios (running a red light and turning left at an intersection). They answered the items that assessed three TRB factors: attitude, subjective norm, and perceived behavioral control. The general objective of the study was to understand what factors of TPB could possibly predict cyclist's behavior and intentions. The results showed that the final model developed in this study indicated that perceived behavioral control and attitude are the best predictors of behavioral intention. This same survey except for a few modifications was later used in Italy, yielding similar results (Marín Puchades et al., 2018).

Another group of researchers considered factors from DBQ to develop the Cycling Behavior Questionnaire (CBQ) for the general cyclist population (Useche et al., 2018a). This questionnaire had multiple parts looking at demographics and risky behavior patterns. In this CBQ, participants are expected to report their perceptions of infrastructural conditions and other roads users, such as complexity of urban roads and interactions with other cyclists and civilians. The results from this questionnaire proved that cyclists' safety must be approached from different ways and that multiple factors should be assessed and included in future behavioral questionnaires for cyclists. The results also supported the existence of a relationship between infrastructure characteristics, human factors, and negative road safety outcomes reported by the international sample of cyclists from 20 countries. The validation of this survey was published in another article and explained that CBQ's factorial structure included three factors: errors, violations, and positive behaviors (Useche et al., 2018b). Both the Useche et al. studies used participants from North America (*N*=72); however, this small sample size cannot adequately represent the general population of the USA. This survey was later used in conjunction with the Big Five Inventory (BFI: agreeableness, extroversion, conscientiousness, neuroticism, and openness to experience) to investigate the relationship between self-reported crashes and behaviors and personality traits of cyclists to develop further understanding of risk factors associated with cycling (O'Hern et al., 2020). A positive association was found between extroversion and both errors and violations; however, agreeableness and conscientiousness were negatively associated with both factors. In a later study O'Hern et al. (2021), looked at the positive behaviors and attitude towards the knowledge of traffic rules and risk perception of Australian bicyclists. They used the error and violation to represent the risky behavior factors of the CBQ and added positive factors to the questionnaire. They found a negative relationship between positive behaviors and total crashes. This finding indicates that positive behavior might act as a protective factor for the risk of a crash (O'hern et al., 2021).

Taking the aberrant behavior of bicyclists into consideration, the Bicycle Rider Behavior Questionnaire (BRBQ) was developed using the factors from the motorcycle rider behavior questionnaire (MRBQ) (Hezaveh, Zavareh, et al., 2018). The factors- perceptual errors and violations were taken from DBQ. The remaining behaviors questions were drawn from available literature. The results showed that the distinction between intentional violations and unintentional errors is consistent with the taxonomy of human errors in aberrant behaviors and can be classified into errors and violations. A five-factor structure illustrated the following scales: stunts and distractions, traffic violations, notice failures, control errors, and signaling violations. Another survey study assessed the safety of shared space between cyclists and pedestrians; the authors developed their own survey and collected responses from college students (Beitel et al., 2018). The goal of the survey was to generate a heat-map of locations where cyclists and pedestrians had an accident, or where they had a near-miss incident. This was carried out by sending out the online survey and asking participants if and where either of these two events happened. The heat-map later helped identify three points of interest on campus and helped the university in making traffic decisions to help with the safety of pedestrians and cyclists.

In a study that specifically investigated cyclist behaviors related to violations and errors, observed, and interviewed cyclists in order to design the Chinese Cycling Behavior Questionnaire (CCBQ) (Wang et al., 2019). CCBQ is divided into three parts: demographics and cycling habits, self-reported behavior patterns, and subjective rule-knowledge over rules pertaining to cycling. A four-factor model was found in the final questionnaire: rule and aggressive violations, ordinary violations, personal control errors, and distractions. The results of this study showed that gender was significantly associated with rule and aggressive violations, ordinary violations and distractions, and age was associated with rule and aggressive violations, personal control errors and distractions. Male bicyclists were reportedly engaged in aberrant behaviors more frequently than females, and young cyclists were more engaged in aberrant behaviors than older cyclists. It was later administered again in China and yielded similar results in the categories of sex and age (Wang et al., 2020b). In addition to those factors, cyclists with lower scores for perceived cycling skills were more likely to be engaged in risky behaviors. Around the same time that the CCBQ was developed, so was a shortened version of the CBQ, otherwise known as the simplified cycling behavior questionnaire (SCBQ). This questionnaire was validated for Chinese roads and assesses the same risky cycling behaviors of the CBQ, however, with only 6 items each for violations, errors, and positive behaviors (Qi et al., 2019). This shortened version still holds the same factors that are addressed in the CBQ (errors, violations, and positive behaviors) and guarantees psychometric value for studying the behaviors of cyclists. BRBQ was used in another study determining overall behavior of cyclists on the road; how they interact with vehicles and how they are willing to share the road (Kaplan et al., 2019). This study concluded that self-perceptions are associated with cycling habits, other travel habits, socio-demographics, and how cyclists perceive drivers and their beliefs about how drivers see them.

Another video-based questionnaire was designed to study the anticipation of cyclists during safety-critical situations (Kovácsová et al., 2019). In this questionnaire, video clips from the point of view of a cyclist were collected from public online websites where safety-critical situations were selected. These were included to assess whether participants could discriminate between safe and unsafe traffic situations at intersections. From these videos and existing literature on other questionnaires, an eight-item survey was developed where participants could indicate responses to questions regarding perceived risks. The findings from this study indicated that cycling experience was not significantly associated with predicting the cyclist's behavior when watching videos.

Table 1 summarizes survey-based studies investigating cyclist behaviors while cycling and their perception towards cycling.

Table 1: Summary of survey-based cyclist safety research

Study	Country N	Age Group	Survey Used	Factors Investigated
Carter et al., 2007	US 97	18 and over	67 video clips taken at traffic sites	Perceived safety, prediction of behavior
Bernhoft and Carstensen, 2016	Denmark 1905	40-49, >70	11-item preference and behavior questionnaire	Traffic preferences, cycling against red lights, cycling on sidewalks
Feenstra et al., 2010	Netherlands 1446	16-24	22-item ACBQ	Risky intention, risky behavior, attitude toward violations, safety for self and others, near accidents
Feenstra et al., 2011	Netherlands 1749	13-18	22-item ACBQ	Occurrence of risky cycling behavior, near accidents
Yang et al., 2012	China 459	Unknown	Site survey of 459 cyclist crossing videos	Crossing violations, normal crossings, waiting duration before crossing
Fernández-Heredia et al., 2014	Mexico 3048	Students and workers	Unknown item cyclist perception questionnaire	Mobility, cyclist choice factors, trip purpose, topography, network, risk perception, exercise chance
Cristea and Gheorghiu, 2016	France 224	19-27	Scenario based questionnaire	Intention, attitude, risk judgements, perceived behavioral control, overconfidence
Puchades et al., 2018	Italy 455	19-72	Scenario based questionnaire used in reference above	Perceived control, overconfidence, perceived risk, avoidance to cycle, near miss.
Useche et al., 2018a, 2018b	20 countries 1064	17-80	44-item CBQ	Violations, errors, positive behaviors, perceptions about infrastructural conditions, road crash rates
Hezaveh et al., 2018	Iran 306	Mean age 33.1	34-item BRBQ	Distractions, traffic violations, notice failures, control errors, signaling violations
Beitel et al., 2018	Canada 872	Students, faculty, and staff	Self-report survey of collisions with cyclists	Real and perceived risk of collisions, near misses, heatmap for risky areas on campus
Wang et al., 2019	China 547	13-64	36-item CCBQ	Rule and aggressive violations, personal control errors, number of distractions
Qi et al., 2019	China 338	College students	18-item SCBQ	Violations, errors, and positive behaviors
Kaplan et al., 2019	Israel 474	20-60	Unknown item BRBQ with additional questions	Willingness to share road, perceived risk,
Kovácsová et al., 2019	65 countries 1384	18-70	8-item video-based questionnaire	Perceived risk, prediction of driver behavior, priority, number of observations
Wang et al., 2020	China 448	15-24	9-item behavior questionnaire	Traffic violations, distractions, impulsive behavior, errors, safety skills, perceptual skills
O’Hern et al., 2020	Australia 625	>18	16-item CBQ	Personality factors and behavior (violations, errors, positive behaviors)

2.2.2 Validation Methods Used in Current Research

The existing literature on cyclist behavior questionnaire shows that most of the questionnaires use perceived responses for validation and/or used specific incidents with a smaller scope in real-world scenarios. There are many other important views and relationships that should be considered when validating the self-reported survey data. Current research over cyclist behavior is more of a one-sided approach, with participants self-reporting behaviors from either scaled responses or just scenario-based questionnaires. Although several of these questionnaires have been validated (Table 1), the relationship between participants' scaled responses and how they actually act in these scenarios has not been tested. The surveys that are currently available, however, have proven to be an effective method of understanding cyclists' behavior further. The ACBQ, CBQ, and BRBQ all have been validated with similar constructs by looking at the models' goodness of fit (Feenstra et al., 2011; Hezaveh et al., 2018; O'Hern et al., 2020; Useche et al., 2018). CBQ has been validated for different countries as well; for instance, in the study to compare the behavior of Australian, Chinese, and Colombian cyclists, researchers found all three countries to have a stable and uniform factorial structure (Li et al., 2022). Australian cyclists, however, indicated more positive behaviors and fewer violations compared to other countries which is an indicator of differences in road environment, policies, and enforcements among the three countries. CBQ was also validated in Belgium in both two languages- Dutch and French (Useche et al., 2021). Results showed strong factor structure, fair psychometrical properties, and good convergent validity which makes CBQ a tool that can be used to study cyclist's behavior in French and Dutch speaking countries from the perspective of Human Factors. The surveys, however, are not validated with participants' actual responses. These types of surveys are still very important to use as a means of gathering self-reported responses.

In addition to the research gap mentioned above, most of the current studies that are scenario based are in written format, where the participant reads upon different scenarios and answers questions. Although this method is popular, it has its drawbacks, such as the lack of interaction the participant will get. In a study that looked at the differences between video-based and written situational judgement tests (SJTs), the following findings were concluded. The interpersonally oriented video based SJT used in this study had significantly higher predictive validity and incremental validity for predicting interpersonally oriented criteria than the written SJT (Lievens & Sackett, 2006). Another study also assessed whether repeated video-based testing is more effective than repeated text-based testing in training medical students to choose appropriate diagnostic tests (Ludwig et al., 2018). Students in this study were randomly allocated to read text cases or watch videos and answer questions on the treatment of the presented patient. The results from the exit exam and retention test found that the repeated video-based testing method produced superior short-term learning compared to the text-based method of testing. Other research has shown that the use of immersive simulation in studies can be very promising and more useful than video or traditional methods (Rahman et al., 2022; Deb et al., 2018, 2020; Deb, Carruth, et al., 2017). Although these studies do not directly pertain to cyclist behavior, it can be inferred that video-based or more immersive methods of scenario-based surveying can be more effective across all subjects. When participants watch videos or act out the scenario themselves, there is a higher level of involvement and realistic experience than with just written surveys.

2.2.3 Lack of an Effective CBQ for the U.S. Population

While there has been research on the behavior of cyclists which has helped form questionnaires (Appendix A), they are very seldom compared to the amount of conducted research with other road users, especially those developed in the US. In the transportation system, the role of the road users tends to get over exaggerated and their behavior gets identified as the cause of crashes and injuries (Hauer, 2020). However, poor road design can also contribute to the risky behavior adaptation of the road user (Li et al., 2022). This interactive relationship between the transportation system and the road user behavior is the reason why it is very important to study road user behavior from a comprehensive point of view (Li et al., 2022). Analyzing this behavior gives the opportunity to understand whether the road users are following the rules set by the system designers as it is a shared responsibility among road user, road designers, road authorities, vehicle manufacturers, and regulators (Welle et al., 2018; Larsson et al., 2010; Tingvall & Haworth, 1999). Furthermore, this also allows to trace back some of the protective or risky behaviors to the transport system and identify the factors behind them (Shinar & Gurion, 2019).

Li et al. (2022) found Australian cyclists to have more positive behaviors and fewer violations compared to cyclists from China and Colombia. Hastuti et al. (2022) also found positive behavior factor to be the highest among other dimensions for Malaysian cyclists. Chinese and Colombian cyclists on the other hand, had a higher rating of traffic violations and errors (Li et al., 2022). The subscales under this factor mostly related to the mixed traffic situations; for instance, cycling against traffic, failing to notice a crossing pedestrian, travelling at a higher speed, or braking abruptly to avoid collision with a turning or a parked vehicle. These factors indicate the problem of traffic management that results in the risky behavior adaptation of the cyclist. The common road characteristics that can be observed between Colombia and China is the lack of separate bike lanes, or wherever they are available, they are occupied by other vehicles. This leads to the cyclist sharing the road with vehicles resulting in an increased likelihood of dangerous swerves and cyclists' overtaking behaviors (Hastuti et al., 2022; Rubie et al., 2020; Bujang et al., 2018). This research illustrates the importance of having a behavioral questionnaire for cyclists so the factors behind the aberrant behaviors can be accurately identified.

Although CBQ has several benefits, it also has some limitations. One major limitation that comes with self-reporting surveys is that participants tend to under or over-report their behaviors. This was mentioned in Feenstra's studies on adolescent cycling behavior questionnaire (ACBQ) where underreporting was one of the shortcomings of both studies (Feenstra et al., 2010, 2011). Previous research has also found that motor-vehicle drivers tend to under-report undesirable behaviors which can most likely be applied to cyclists and e-bike riders (Useche et al., 2018). To eliminate this bias, it is necessary to observe the cyclists' actual behavior in practice under a field study or simulated scenario. As it is difficult to control all the factors in a field study and participants are exposed to real crash scenarios, a simulator study can be very useful to address these issues. The goal of this study was to review current literature and questionnaires available for the field of cyclist behavior research and validate the questionnaire for the U.S. cyclist population by confirming the association between cyclists' actual responses with their self-reported survey responses.

In summary, there are no current validated questionnaires for the US population, and most of the questionnaires that are available have not been tested outside of the country it was developed in. However, most of the questionnaires have similar factorial structures that look at violations, errors, risky behaviors, and safe behaviors (ACBQ, CBQ, and BRBQ) (Table 1). These four factors provide a base for future cyclist behavior studies to be conducted in the US and tested against their road laws and population. These factors were used to develop CBQ for the US population and later, factor specific scenarios were developed in a bike simulator to validate the CBQ.

2.3 Methodology

An online survey study was performed with a newly developed cyclist behavior questionnaire for the U.S. cyclist population. The survey was carried out on an online survey platform- QuestionPro. The approval for this study was collected from the Institutional Review Board at The University of Texas at Arlington.

2.3.1 Survey Instrument

A new questionnaire was designed using CBQ, CCBQ, and individual knowledge of road rules. The survey was conducted using QuestionPro software provided by Microsoft Office. The complete survey is attached in appendix A. The literature review in section 2.2 shows studies that were reviewed to get an idea of important factors, and it was found that many factors overlapped between questionnaires. The most prominent factors were taken and further reviewed to get rid of any redundant questions. The final questionnaire included questions regarding demographics, frequency of cycling, history of crashes, cycling knowledge, and behaviors.

In the first part of the questionnaire, 10 questions were asked regarding demographics (gender, age, background) and general cycling information, such as the duration of the trip, how long they have been cycling, frequency of crashes with other road users, and their knowledge of traffic rules for bicycling. The next part of the questionnaire comprised of five different sections with each section relating to a behavioral factor. The final factors included in the newly developed version were violations, errors, positive behaviors, aggressive violations, and distractions and forgetfulness. The questions were placed in random order in the final survey. Table 2 shows the definition of each of these factors. Participants were asked to respond to each behavior using a numerical 5-point Likert scale ranging from 'never' which was recorded as a '1' to 'almost always' which was recorded as a '5'. The last section consisted of a scenario-based questionnaire that had ordinal responses with increasing severity (Appendix A). Each participant received one scenario-based question in random. One participant received 43 questions in total.

Table 2: Definition of CBQ factors

Subscale	Definition	Example
Violation	Risky cycling behaviors that deviate from traffic rules, but they do not have the intent to cause injury. (Deb, Strawderman, Carruth, et al., 2017; S. A. Useche et al., 2022)	Biking under the influence of drugs, alcohol
Error	Risky behaviors that are indeliberate and are due to the lack of knowledge of traffic rules. (Deb, Strawderman, Carruth, et al., 2017; S. A. Useche et al., 2022)	Not using proper hand signals while turning
Aggressive Violation	Risky cycling behaviors that deviate from social rules to express annoyance or anger. (Deb, Strawderman, Carruth, et al., 2017; Hezaveh, Nordfjærn, et al., 2018)	Yelling at a pedestrian for crossing the street in front of the bicycle
Positive Behavior	Behaviors or habits that are protective in nature to avoid violation or error. (Deb, Strawderman, Carruth, et al., 2017; S. A. Useche et al., 2022)	Maintaining proper distance from other vehicles and pedestrians
Distraction and Forgetfulness	Losing focus from cycling due to lack of concentration or forgetting the task. (Deb, Strawderman, Carruth, et al., 2017; S. A. Useche et al., 2022)	Not checking the traffic before changing lanes

2.3.2 Study Protocol

2.3.2.1 Survey Protocol

The survey was distributed to participants in different cycling groups using the social media-Facebook. In the beginning of each survey, there was a brief description of the surveys' intentions and contents. They were first asked if they agreed to participate prior to moving to the survey items. If the participant agreed to participate, they selected yes after reading the consent and moved to the next section. Participants needed to answer each question from the entire survey. However, they were free to leave the survey anytime they wanted. Along with the demographics, and behavioral survey items, there were two check questions. The check questions ensured respondents' attention to the survey questions. Responses from participants who answered at least one of the check questions incorrectly were not included in the data analysis.

2.3.2.2 Simulator Study Protocol

The factors that explained U.S. cyclists' behaviors on the road that were revealed from the PCA and CFA analysis, were used to create factors-based scenarios on a bike simulator. The bike simulator is developed by RTI Technology that consists of a mounted bicycle and three projectors projecting on three screens creating an immersive environment (Figure 1). The scenario-based questions of the CBQ were developed on the simulator. Each scenario was repeated twice to ensure similar behavior from the participants.

Figure 1: Bike simulator setup at UTA Human Factors lab

Participants were greeted by the research team, and they were acquainted with the general structure of the experiment. They were then provided with an infectious disease screening form followed by a consent form explaining the process of the experiment and the benefits and risks of their participation. Covid-19 risk information was read out to each participant to ensure that they do not have any signs of Covid-19. Participants were then asked to fill out the CBQ on QuestionPro. Once they completed the survey, they were then provided with Simulation Sickness Questionnaire (SSQ) to ensure their fitness for the participation (Kennedy et al., 1993). The questionnaire is attached in Appendix B. The questions fall under three categories- Oculomotor, Disorientation, and Nausea. Participants scored SSQ items on a 4-point scale with none '0', slight

‘1’, moderate ‘2’, and severe ‘3’ scores. A 5-point score was set as a threshold for withdrawing participants from the study, with zero being the highest score for fitness. Based on this, the participant’s health was kept in check along with the participant’s ability to give accurate feedback. After filling up and qualifying with the first SSQ (<5), participants were taken to the bike simulator room. They rode the bicycle for about 5 minutes to get acclimated to the simulated traffic environment and the bicycle. They were then asked to fill out the SSQ again. Once they were qualified to begin the study, they were instructed to find the construction barrels placed in one of the intersections (Figure 2), Their responses to the subscales of the CBQ were recorded by the researchers. Once the trials were over, they were asked to first complete the final SSQ and if they qualified, they were asked to complete the scenario-based questionnaires on QestionPro (Abstract A). At the end of the experiment, they were compensated for their time.

Figure 2: Construction barrels that participants had to find

2.3.3 Participants

2.3.3.1 Survey

To partake in the study, participants had to meet the following qualifications: must live in the U.S., know English, and be over 18 years of age. For the purposes of the study, a cyclist is defined as a person who uses a bicycle at least once a week and has at least one year of cycling experience prior to the beginning of the study. These criteria were included at the beginning of the survey. 249 people completed the survey and after eliminating responses based on the criteria and wrong answers for check question, 224 participants data were used to develop the CBQ scale. Table 3 summarizes the characteristics of the participants who completed the survey. The mean age of the participants was 48 years (*SD*= 15.5). The range of age was between 18-80 years old. The number of male participants was 1.6 times higher than female or other gender. The majority of the participants use cycling for exercise (91%) and use urban roads (71%). Most participants had no crashes with non-motor vehicles (87%) and motor vehicles (88%). The participants also had a very well knowledge of traffic rules (61%).

Table 3: Characteristics of the survey participants (N= 224)

Characteristics		***n* (%)**
Gender	Male	138 (62)
	Female	83 (37)
	Other	2 (1)
Age	18-30	39 (18)
	31-50	75 (34)
	50+	107 (48)
Cycling Duration	<15 min	13 (6)
	15-30 min	22 (10)
	>30 min	177 (79)
Cycling Experience	<1 year	1 (0.4)
	2-5 years	31 (14)
	>5 years	190 (85)
3 years crash history with non-motor vehicles	0 times	195 (87)
	1-5 times	26 (12)
	>5 times	0 (0)
3 years crash history with motor vehicles	0 times	197 (88)
	1-5 times	23 (10)
	>5 times	1 (0.4)
Traffic rules knowledge	Not at all	2 (1)
	Well	86 (39)
	Very well	135 (61)

2.3.3.2 Bike Simulator Study

The requirements to participate in this study were similar to the survey study: must live in the U.S., know English, be over 18 years of age, and uses a bicycle at least once a week and has at least one year of cycling experience. Participants were recruited via word of mouth. 15 participants participated in the study. Table 4 summarizes the characteristics of the participants who completed the survey. The mean age of the participants was 24 years old (*SD*= 3.13). The range of age was between 19-29 years old. The number of male participants was 4 times higher than female. The majority of the participants use cycling for exercise (87%) and use residential roads (73%). Most participants had no crashes with non-motor vehicles (87%) and motor vehicles (93%). The participants also had well knowledge of traffic rules (53%).

Table 4: Characteristics of the simulator study participants (N= 15)

Characteristics		***n* (%)**
Gender	Male	12 (80)
	Female	3 (20)
	Other	0 (0)
Age	18-30	15 (100)
	31-50	0 (0)
	50+	0 (0)
Cycling Duration	<15 min	9 (60)
	15-30 min	2 (13)
	>30 min	2 (13)
Cycling Experience	<1 year	0 (0)
	2-5 years	0 (0)
	>5 years	15 (100)
3 years crash history with non-motor vehicles	0 times	13 (87)
	1-5 times	1 (7)
	>5 times	1 (7%)
3 years crash history with motor vehicles	0 times	14 (93)
	1-5 times	0 (0)
	>5 times	1 (7)
Traffic rules knowledge	Not at all	1 (7)
	Well	8 (53)
	Very well	6 (40)

2.3.5 Data analysis

2.3.5.1 Scale development

The scale was developed using Principal Component Analysis (PCA) on SPSS Statistics 29. PCA is a widely used method to reduce dimensionality of data while preserving the variability (Jollife & Cadima, 2016). It uses algorithms to reduce variables into correlated factors that can be used to understand the construct of interest. PCA requires rotation of the factors which helps to achieve a simpler structure. The factors can be rotated either following oblique rotation where factors are correlated; or orthogonal rotation where the factors are uncorrelated. For this study, first the correlation analysis was assessed to determine which rotation to apply. The correlation coefficient among each item of the questionnaire was low (<0.5). Therefore, orthogonal rotation was used (Tabachnick & Fidell, 2007). Varimax rotation, a popular orthogonal rotation, was used for factor extraction.

After performing PCA, the correlation matrix of the items was first analyzed for the possibility of serious multicollinearity. Kaiser- Meyer- Olkin (KMO) measure of sample adequacy (>.5) and Bartlett's Test ($p < 0.05$) were used to determine whether the sample data met the requirements for the analysis (Napitupulu et al., 2017). The eigenvalues, i.e., the sum of the squared factor loadings across all items, were then analyzed. The eigenvalues explain the amount of variance explained by the data (DiStefano et al., 2009). The number of components that explain the majority of the variance were identified from this table. Scree plot is a visual presentation of eigenvalues against the number of components. It further showed how many components should be used in the model. The rotated factor matrix was analyzed to determine how the items measure each component. For factor loading, a cut-off point of 0.4 was used.

After the factors were determined, mean scores of the items under each factor were calculated and used as composite scores for each subscale of the cyclist's behavior. Cronbach's alpha (α) was used to test the internal reliability. Cronbach's alpha rages from 0 to 1, where a high alpha value (>0.9) would indicate some items are redundant, while a low alpha (<0.7) value would indicate inadequate items to measure the construct, heterogeneous constructs, or poor interrelatedness among the items of the survey (Bujang et al., 2018; Tavakol & Dennick, 2011).

2.3.5.2 Scale validation

The factorial structure obtained from PCA results was then validated using the Confirmatory Factor Analysis (CFA) (Deb, Strawderman, DuBien, et al., 2017). CFA is used to determine the relationship between the observed measures or *indicators* and latent variables or *factors* (Brown & Moore, 2012). AMOS 29 was used for this analysis. Maximum likelihood estimation (MLE) procedure was applied for CFA. The factor loadings for the model were assessed for each item and any item with a low factor loading was removed. The model's overall goodness of fit was assessed using model fit measures: Root Mean Square Error of Approximation (RSMEA), the chi-square test statistics, comparative fit using Comparative Fit Index (CFI), Tucker-Lewis index (TLI), Goodness of fit index (GFI), and Standardized Root Mean Square Residual (SRMR).

Construct Reliability to assess the consistency of the variables in what they measure was analyzed using Composite Reliability (CR) (Straub & Gefen, 2004). CR was calculated from the factor loadings. Construct validity was then analyzed to determine how well the items selected measures the construct. The two ways it was analyzed were- convergent validity, and discriminant validity. Convergent validity determines the degree to which the measures of the construct that should be theoretically related are, in fact, related; whereas Discriminant validity determines how the measures that are unrelated to each other are indeed unrelated (Gefen et al., 2000; Anderson & Gerbing, 1988). Convergent Validity was assessed using Average Variance Extracted (AVE). AVE determines how much of the variance is explained by the latent unobserved variable (Fornell & Larcker, 1981). Discriminant validity was assessed using Fornell and Larcker Criterion and Heterotrait-Monotrait (HTMT) Ratio. Fornell And Lacrcker Criterion compares the square root of AVE of each latent variable to that of the correlations with other latent variables. HTMT, on the other hand, compares the ratio of between-trait correlations to that of within-trait correlations of constructs. The demographics and the cycling related factors were analyzed using one-way ANOVA for mean comparisons. Table 5 summarizes the acceptable measures that were used to assess model fit, and construct reliability and validity.

Table 5: Accepted values for Confirmatory Factor Analysis (CFA)

Sources	**Parameter estimates**	**Measure**	**Recommended value**
Schumacker & Lomax, 2004	χ^2/ df	Assess the fit between the hypothesized model and the observed variables	2-5
Hair et al., 2010	Goodness of fit index	To what degree sample data fits what is expected of a population	>.90
Kenny, 2012; Bentler, 1990	Comparative Fit Index	Measure of the relative improvement in fit of the model under evaluation compared to the baseline model	>.90
Kenny, 2012; Bentler, 1990	Tucker-Lewis index	Measure of the misfit per degree of freedom for the model under evaluation compared to the baseline model	>.90
Alavi et al., 2020; Hu & Bentler, 1998	Standardized Root Mean Square Residual	An index to compare the average of the standardized residuals between the observed and the hypothesized covariance	<.08
Hu & Bentler, 1998	Root Mean Square Error of Approximation	Measure of the estimated discrepancy between the population and model-implied population	<.08
Hair et al., 2010	Composite Reliability	Measure of internal consistency	>0.7
Fornell & Larcker, 1981	Average Variance Extracted	Degree of variance explained by latent variables	>0.5
Henseler et al., 2015	Heterotrait-Monotrait ratio	Degree of similarity between latent variables	<.85

In order to confirm whether each sub-scale scores and/or the composite CBQ score can predict bicyclists' actual responses, the scenario-based ordinal responses and actual behavior (categorical variable) during the simulator study were compared with the scale scores (continuous variable) using ordinal logistic regression. The association between the scenario-based behavior and subscale scores shed light on the successful development of the questionnaire or necessity for modifications.

2.4 Results

2.4.1 Survey Item Descriptive Statistics

Table 6 summarizes the descriptive statistics of each survey item responses by percentage. For the convenience of this descriptive analysis only, quite infrequently and infrequently were merged into "Infrequently". Similarly, quite frequently and frequently were merged into "Frequently".

Most participants (88.8%) indicated that they never or infrequently violate traffic laws. 82% participants selected that they never or infrequently make errors on the road; however, 67.3% participants indicated that they always or frequently ride their bicycles across the crosswalk instead of getting off and walking. In terms of positive behavior, 89.2% participants exhibit positive behavior on the road. 92.6% participants do not exhibit aggressive violation on the road. Lastly, 93.6% participants indicated they do not get distracted or forgetful on the road.

Table 6: Descriptive statistics of the CBQ scale

Survey Item		Responses by Percentages			
		Never	Infrequently	Frequently	Always
V1	I cycle against traffic (the wrong way)	85.7	11.7	0.9	1.8
V2	Even though there is an exclusive bicycle lane nearby, I cycle on the vehicular lane or on the sidewalk	45.3	33.1	17.9	3.6
V3	I cycle under the influence of alcohol and / or other drugs or hallucinogens	85.2	12.6	1.3	0.9
V4	I zigzag between vehicles when I am using a mixed lane to go faster	75.3	18.4	5.3	0.9
V5	I cross the road when it appears to be a clear crossing, even if the traffic light is red	35.4	35.4	22.0	7.2
V6	I carry potentially obstructive objects while riding the bicycle (food, packs, cigarettes, etc.)	74.0	14.8	9.0	2.2
E1	I do not brake on a “stop” or “yield” sign and come close to colliding with other vehicles or pedestrians	86.1	11.6	1.8	0.4
E2	I misjudge a turn and hit the curb on the road	85.2	13.5	0.9	0.4
E3	I try to brake, but I am not able to do so because of poor hand or foot positioning (for brakes) or a slippery surface	83.0	15.7	1.3	0
E4	When using crosswalks, I stay on my bicycle and ride across, instead of getting off my bicycle and walking	17.9	14.8	33.2	34.1
E5	I try to overtake vehicles that had previously used indicators to signal that they were about to turn	87.9	10.7	1.3	0
E6	I sometimes mistake a traffic signal for another one, and maneuver according to the latter	89.7	9.4	0.9	0
PB1	I stop and look both sides before crossing a corner or intersection	2.2	9.0	25.5	63.2
PB2	I try to move at an appropriate speed to avoid sudden collision or braking	4.0	3.6	26.9	65.5
PB3	I usually keep a safe, recommended distance from vehicles and other road users	1.3	3.1	41.2	54.3
PB4	I avoid cycling under poor weather conditions (heavy rain, sleet, hail, high winds, etc.)	7.6	16.1	33.2	43.0
PB5	I use the helmet for cycling	3.6	3.1	7.1	86.1
PB6	When I travel at night, I use the necessary safety equipment (lights, vest, and reflectors)	7.2	3.6	17.5	71.7
AV1	I change course (such as turning, avoiding obstacles, passing pedestrians) without giving any signal to other road users, making a sudden sharp turn	58.3	35	6.7	0
AV2	I yell at other road users if they do not follow the rules	51.6	35.8	10.3	2.2
AV3	I make rude gestures (hand, face gestures) to other road users if they do not follow the rules	52.9	33.6	11.6	1.8
AV4	I cycle around other vehicles, cyclists, or pedestrians and “cut them off”, forcing them to brake or stop	90.1	9.4	0.4	0
AV5	I have races with other cyclists or drivers	65.5	23.8	10.3	0.4
AV6	I hit other road users if they slow down and block my way	97.8	1.7	0	0.4
DF1	I get distracted and unintentionally hit a parked vehicle	97.8	1.7	0	0.4
DF2	I cycle with one hand and execute other actions with the other hand (holding up umbrellas, eating, using phones, etc.)	58.7	34.9	5.8	0.4
DF3	I listen to audio (news or music) while cycling and do not hear audible cues	72.6	19.8	5.8	1.8
DF4	I get forgetful and think about other things while cycling	39.5	44.4	14.8	1.3
DF5	I get distracted and do not see that there is an object or parked vehicle on the road and get close to hitting them	87.0	13	0	0
DF6	I get distracted and forget to gesture that I am turning left or right	60.5	31.8	6.7	0.9

Note: The first column indicates the subscale of the CBQ scale where V is denoted for Violation, E is denoted for Error, PB is denoted for Positive Behavior, AV is denoted for Aggressive Violation, and DF is denoted for Distraction and Forgetfulness

2.4.2 Scale Development

The factor structure of the CBQ scale was developed using Principal Component Analysis (PCA) (Table 8). The first output that was analyzed was the correlation matrix of the original variables. If the correlation is too high (>0.7), it would indicate the possibility of multicollinearity. In this case, the correlation among the variables was less than 0.7. In the next step, the rotated factor matrix was analyzed and items failing to load above 0.4 were incrementally deleted and the PCA procedure was repeated (Rattray & Jones, 2007; Wang et al., 2019). This resulted in the elimination of item V2, V3, V6, E1, E2, E3. E4, E5, AV4, AV5, AV6, PB1, PB2, PB3, PB4, DF1, DF2, DF3, and DF4. The resulting CBQ had a Kaiser-Meyer-Olkin (KMO) i.e., measure of sample adequacy value of 0.708 and Bartlett's Test of Sphericity to determine whether the sample data met the requirements for the analysis was significant. According to Napitupulu et al., (2017) both Kaiser- Meyer- Olkin (KMO) measure of sample adequacy (>.5) and Bartlett's Test ($p < 0.05$) met the requirements. Four factors with eigenvalue greater than 1 were identified from the Scree plot (Figure 3). The resulting structure was 11-item questionnaires with 4 factors that cumulatively explained 66.9% of the total variance. The first component, violation explained 28.4% of the variation and consisted of the items V4, AV1, V5, V1, and DF6. The second component, aggressive violation explained 16.7% of the variation and consisted of items AV2, and AV3. The third component, positive behavior explained 11.6% of the variation and consisted of items PB5 and PB6. Lastly, the fourth component, distraction and forgetfulness explained 10.2% of the variation and consisted of items DF5 and E6 (Table 7).

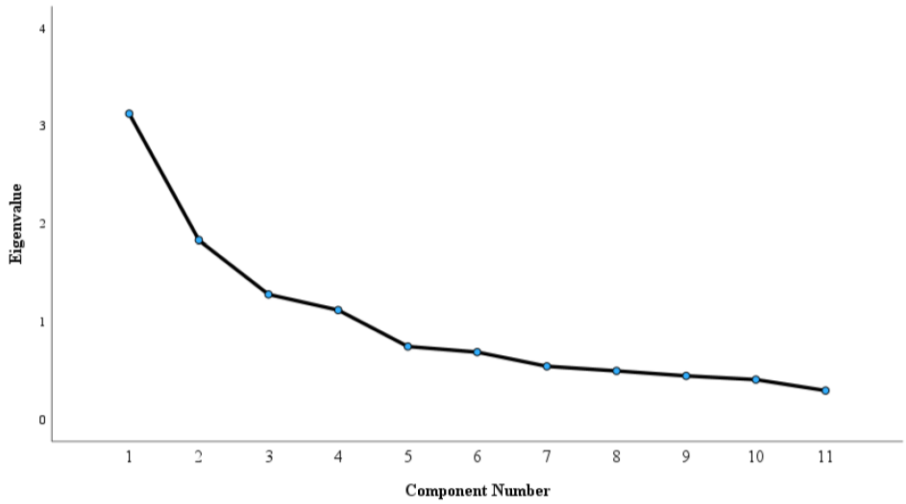


Figure 3: Scree plot for the eigenvalues for each component

Table 7: The principal component analysis of the CBQ scale. The numeric values indicate the factor loading where higher factor loading indicates stronger association.

Survey Items	**Components**			
	Violation	**Aggressive Violation**	**Positive Behavior**	**Distraction and Forgetfulness**
V4: I zigzag between vehicles when I am using a mixed lane to go faster	.755			
AV1: I change course (such as turning, avoiding obstacles, passing pedestrians) without giving any signal to other road users, making a sudden sharp turn	.717			
V5: I cross the road when it appears to be a clear crossing, even if the traffic light is red	.702			
V1: I cycle against traffic (the wrong way)	.689			
DF6: I get distracted and forget to gesture that I am turning left or right	.595			
AV2: I yell at other road users if they do not follow the rules		.906		
AV3: I make rude gestures (hand, face gestures) to other road users if they do not follow the rules		.891		
PB5: I use the helmet for cycling			.838	
PB6: When I travel at night, I use the necessary safety equipment (lights, vest, and reflectors)			.800	
DF5: I get distracted and do not see that there is an object or parked vehicle on the road and get close to hitting them				.847
E6: I sometimes mistake a traffic signal for another one, and maneuver according to the latter				.839

The reliability of the scale was determined by calculating mean scores of items under each subscale. Cronbach's alpha tested the internal reliability. Cronbach's Alpha values for all the subscales were greater than 0.7. Violation had an alpha value of 0.721, aggressive violation had an alpha value of 0.806, positive behavior had a value of 0.710, and lastly, distraction and forgetfulness had a value of 0.741.

2.4.3 Scale Validation

The structure of the survey was validated using the Maximum Likelihood Estimate (MLE) procedure of Confirmatory Factor Analysis (CFA). The normality of the model was assessed using kurtosis and skewness values. According to Cohen et al. (2013), the skewness should be between -2 and +2, and the kurtosis should be between -7 and +7. However, according to Brown (2015), for Structural Equation Modeling, the acceptable values can be between -3 and +3 for skewness, and -10 and +10 for kurtosis. All the skewness and kurtosis values in this study were within the range mentioned by Brown (2015). The critical ratio for the multivariate Kurtosis was also <5 as required by Bentler (2005). Therefore, the normality assumption for MLE was satisfied.

The 11-item questionnaire confirmed the factor structure obtained using PCA in section 2.4.2. Figure 4 shows the standardized solution for the CBQ scale. Alternative structures were also explored to find the ideal factorial structure. The goodness of fit measures was assessed, and first order 4-factor structure fits the data best (Table 8). The goodness of fitness indicated the model fit is adequate (χ^2/df = 2.072, RMSEA= 0.069, CFI= 0.927, GFI= 0.942, TLI= 0.9, and SRMR= 0.069).

Table 8: Comparison of alternative structures of CBQ

	χ^2/ df	**GFI**	**CFI**	**TLI**	**SRMR**	**RMSEA**
Recommended Value	2-5	>.90	>.90	>.90	<.08	<.08
Model 1: First order model 1-factor (Behavior) model	6.711	0.809	0.551	0.439	0.171	0.160
Model 2: First order 4-factor model (Violation, Aggressive Violation, Positive Behavior, Distraction and Forgetfulness)	**2.072**	**0.942**	**0.927**	**0.900**	**0.069**	**0.069**
Model 3: Second order (Behavior) 4-factor model (Violation, Aggressive Violation, Positive Behavior, Distraction and Forgetfulness)	2.172	0.935	0.916	0.885	0.088	0.073

The Composite Reliability (CR) of the construct for all the factors was greater than 0.7. The Average Variance Extracted (AVE) value for all the factors was greater than 0.5, except violation. Violation had an AVE of 0.4. The Heterotrait-Monotrait Ratio for all the factors was less than 0.85.

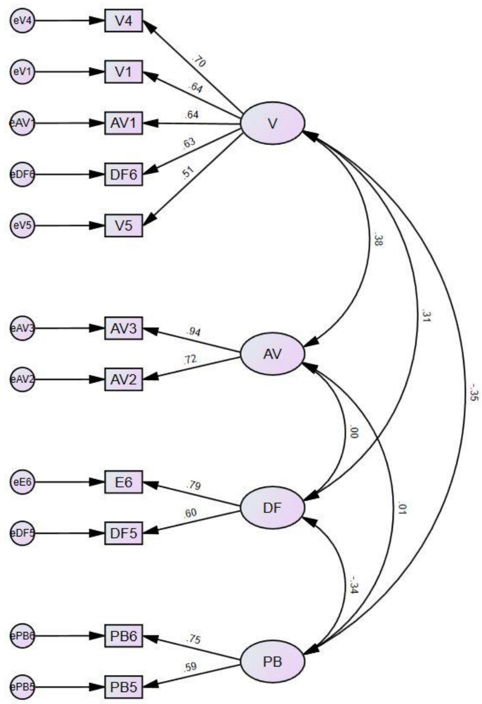


Figure 4: Standardized solution for the CBQ scale

2.4.4 Scenario-based Validation

To further validate the survey, scenario-based questions for both the survey and the bike simulator were analyzed using ordinal logistic regression for each scale. The scale for positive behavior questions was reversed before the analysis to match it with the CBQ scale. The assumption that the dependent variable must be ordinal was satisfied since the simulator behavior was in an ordinal scale from less risky to high risky behavior. The proportional odds assumption i.e., the relationship between each pair of outcome groups is the same, was tested using the test of

Parallel Lines. The hypothesis whether the coefficients in the model are the same across response categories was tested. Chi-square statistic for all the models in Table 9 had a *p* value >0.05, which indicated the coefficients were the same across response categories.

From the ordinal regression outputs, it was observed that for the survey-based scenario question, none of the questions predicted the scale it was intended for. However, for the simulation-based questions, all the factors were significantly predicted by the subscale. Table 9 summarizes the findings for the scenario-based questionnaire.

Table 9: Association between scenario-based questions and the factors of CBQ

	Survey Scenarios	**Simulation Scenarios**
Violation	*b*= 0.256 χ^2= 0.385	*b*= 1.632 χ^2= 5.695*
Aggressive Violation	*b*= 0.173 χ^2= 0.331	*b*= 2.653 χ^2= 5.101*
Positive Behavior	*b*= -0.121 χ^2= 0.058	*b*= 1.497 χ^2= 7.024*
Distraction and Forgetfulness	*b*= -0.087 χ^2= 0.006	*b*= 2.405 χ^2= 7.022*

Note: b denotes parameter estimate
* $p < 0.05$

2.4.5 Effects of Demographic Variables

Analysis of Variance (ANOVA) was used to determine the influence of gender (3 levels), age (4 levels), frequency of crashes with motor vehicles (3 levels), frequency of crashes with non-motor vehicles (4 levels), and cyclists' knowledge of traffic laws of the CBQ survey participants (6 levels). The mean score for each subscale was used to conduct ANOVA. Gender had a significant effect on aggressive violation; age and frequency of crashes with non-motor vehicles had significant effects on and distraction and forgetfulness. Table 10 summarizes the ANOVA results.

Table 10: Demographic influences on CBQ subscale

Demographics	Statistics			
	Violation	Aggressive Violation	Distraction and Forgetfulness	Positive Behavior
Gender (*df*: 2, 222)	F= 1.777	**F= 5.051***	F= 0.198	F= 0.506
Age (*df*: 3, 219)	F= 1.607	F= 2.088	**F= 4.690***	F= 2.047
Crashes with non-motor vehicles (*df*: 2, 222)	F= 0.879	F= 0.820	**F= 2.780***	F= 0.370

Note: df denotes degrees of freedom for one-way ANOVA
* $p < 0.05$

Table 11: Multiple comparison (Bonferroni) of gender for aggressive violation

Gender	Score LSMEAN	Bonferroni Grouping
Other	4.50	A
Male	1.96	B
		B
Female	1.93	B

Note: Means with same letter are not statistically significant

Table 12: Multiple comparison (Bonferroni) of age for distraction and forgetfulness

Age	Score LSMEAN	Bonferroni Grouping
31-50	1.24	A
		A
18-30	1.22	A
		A
51-70	1.04	A
		A
70+	1.04	A

Note: Means with same letter are not statistically significant
The letters do not reflect all significant comparisons. (31-50) and (51-70) are significantly different.

Table 13: Multiple comparison (Bonferroni) of collision with non-motor vehicles for distraction and forgetfulness

Number of crashes with non-motor vehicles	Score LSMEAN	Bonferroni Grouping
1-5 times	1.31	A
		A
0 times	1.12	A
		A
>5 times	1.00	A

Note: Means with same letter are not statistically significant
The letters do not reflect all significant comparisons. 1-5 times and 0 times are significantly different.

Table 11-13 shows the post hoc analysis using Bonferroni test at α=0.05 significance level. It revealed that male and 'other' cyclists exhibited significantly higher aggressive violation.

Middle aged cyclists aged 31-50 exhibited significantly higher distraction and forgetfulness compared to cyclists aged 51-70. Lastly, cyclists who had frequency of crashes with non-motor vehicles 1-5 times had a significantly higher mean score for distraction and forgetfulness compared to cyclists who had no crashes.

2.5 Discussion

2. 5.1 Development and validation of CBQ

This research aimed to develop a Cyclist Behavior Questionnaire (CBQ) for the US population. The factors-violation, error, aggressive violation, positive behavior, and distraction and forgetfulness were taken from previous literature (Useche et al., 2022; Wang et al., 2019). These factors have been tested and validated in several countries in the world (Li et al., 2022; Useche et al., 2021; Useche, Montoro, Tomas, et al., 2018; Wang et al., 2020). In this study, Principal Component Analysis (PCA) with varimax rotation indicated violation, aggressive violation, positive behavior, and distraction and forgetfulness explained 66.9% of the variance of the model. The Cronbach's alpha indicated good internal reliability of the subscales ($\alpha > 0.7$). The Confirmatory Factor Analysis (CFA) further validated the factorial structure of the questionnaire with goodness of fit values within the acceptable range (Table 8). The Composite Reliability was also >0.7. The Average Variance Extracted (AVE) value to determine the construct validity, however, was low for violation. According to Fornell & Larcker (1981), if the CR of the construct is >0.7, but the AVE of the construct is <0.5, the construct can still be considered adequate.

CBQ for the US population was adapted from the Chinese cyclist Behavior Questionnaire developed by Wang et al. (2019) and the Cyclist Behavior Questionnaire developed by Useche et al. (2018). CCBQ consisted of the factors aggressive violations, violations, personal control errors, and distraction, whereas CBQ consisted of violations, errors and positive behaviors. Both scales had error as a factor that explained the behavior of their respective population; however, it did not explain US cyclists' behavior. As mentioned in Table 2, error is considered as risky behaviors that are indeliberate and are due to the lack of knowledge of traffic rules (Deb, Strawderman, Carruth, et al., 2017; S. A. Useche et al., 2022). In the case of US cyclists, the majority of cyclists had well or very well knowledge of traffic rules (Table 3, Table 4). Therefore, lack of knowledge leading to an error did not contribute to their behavior. Violation, aggressive violation, and lapses (distraction and forgetfulness) were also observed in the Pedestrian Behavior Questionnaire developed by Deb et al. (2017). This indicates that the vulnerable road users in the US exhibit similar behavior on the road.

The 11-item CBQ was then validated using a scenario-based questionnaire using survey method and the bike simulator in the Human Factors Lab at UTA (Attached in Appendix D). Ordinal Logistic regression analysis of the survey results showed the CBQ subscales did not predict the risky/conservative cyclist behavior significantly; however, when the simulator-based scenarios were used, CBQ subscales predicted the risky/conservative behavior significantly (Table 9). This difference in prediction indicates the discrepancy in using scenario-based questionnaires to understand realistic behavior on the road. This discrepancy can also be observed in the current study, especially for the subscale aggressive violation. When participants were asked to complete the same scenario-based questions after they finished the bike simulator trial, 80% of them mentioned that they will patiently wait for the slow-moving vehicle. However, during the trial, only 20% of the participants followed the vehicle. Most participants overtook the vehicle either using the other lane or the right side of the lane closer to the curb. When their CBQ scores were compared, they indeed had a higher score for the aggressive violation subscale. This indicates that the self-reported survey responses of CBQ collected the realistic responses of cyclists while riding their bikes. Although, many researchers suggest using scenario-based survey responses to collect road-user behavior, they do not necessarily produce road-users' realistic behavior on the road (Deb et al., 2017; Li et al., 2023).

2.5.2 Demographic influence on cyclist behavior

Gender

While exploring the influence of demographic on cyclist behavior, the male population is 1.6 times the population of female and other gender, which is a good representation of the US population. According to the US census Bureau, in 2019, the number of male cyclists who commuted to work using a bicycle was 2.4 times than that of female cyclists (US Census Bureau, 2022). In terms of behavioral difference based on gender, cyclists who selected other for gender exhibited significantly higher aggressive violation than male or female gender. This response was not conclusive due to the extremely small sample size (N=2). The mean score for aggressive violation for male participants was higher than that of female participants. This is consistent with previous studies where researchers found male cyclists to have aggressive expressions of anger compared to their female counterpart (Møller & Haustein, 2017; Stephens et al., 2020). Wang et al. (2019) also found male cyclists to report more aggressive behavior than female cyclists.

Age

Middle aged cyclists aged 31-50 had a significantly higher mean score for distraction and forgetfulness than the other age groups. This is consistent with Useche et al. (2019) study, where researchers found cyclists aged 46-55 had the highest mean score for distraction, whereas cyclists under 26 had the lowest mean score. Another explanation for this difference is the cognitive ability of more experienced cyclists. As cyclists become more experienced, they may require paying less attention for the mental processes related to cycling (Wierda & Brookhuis, 1991). As a result, they can get distracted and forget their tasks related to cycling.

Frequency of crashes with non-motor vehicles

Cyclists who had a crash frequency of 1-5 times with non-motor vehicles had a significantly higher mean score for distraction and forgetfulness than the other groups. Cyclists' distraction and forgetfulness has been identified as a major factor leading up to crashes in several previous studies (Useche, Alonso, et al., 2018; Wang et al., 2019; Wolfe et al., 2016). In one study, researchers found in traffic crashes where cyclists were at fault, 21% of them were distracted prior to the crash occurring (Useche, Alonso, et al., 2018). Some examples of distractions for cyclists on the road include technology (cell phones, headphones, and navigators), billboards, weather conditions and road features (Dukic et al., 2013; Jägerbrand & Sjöbergh, 2016; Oikawa et al., 2016; Wolfe et al., 2016). These distractions can cause cyclists to be unfocused on the road, leading them to crash into non-motor vehicles.

2.5.3 Limitation

This study developed and validated a questionnaire for the US cyclist population; however, there are several limitations of this study. The first one being the US having different cycling laws in different states. For instance, in Alabama, according to the Ala. Code § 32-5A-52 (1975), bicycles are considered vehicles. As a result, cyclists are prohibited from riding on the sidewalk; however, in Florida according to the Florida Statute Fla. Stat. §§316.2065(9)-(10) (2012), bicyclists are allowed to operate their bike on the sidewalk (The League of American Bicyclists, 1880). This discrepancy in laws was also evident from the data in this study. In Arkansas, the mean score for aggressive violation for male (*M*= 2.2, *SD*= 1.01) was significantly higher than that of female cyclists (*M*= 1.0, *SD*= 0). In New York, the mean score for violation was significantly higher for male (*M*=2.08, *SD*=1.18) cyclists than their female (*M*=1.67, *SD*=0.42) counterpart. In Texas, however, the violation mean score was significantly higher for female cyclists (*M*=2.21, *SD*= 0.78) than male cyclists (*M*=1.64, *SD*=0.67). In Florida, cyclists who had adequate knowledge of traffic rules, had significantly higher mean scores for positive behavior (*M*= 5.69, *SD*= 0.53). Similarly, cyclists who were aged 18-31 had a significantly higher mean score for violation (*M*=1.28, *SD*= 0.90), and participants aged 31-50 had significantly higher score for distraction and forgetfulness (*M*= 1.38, *SD*= 0.25). In Georgia, cyclists aged 31-50 had significantly higher score for positive behavior (*M*=5.88, *SD*= 0.25). Further feedback received from participants shed light on some other important factors. For example, cyclists in New York city mentioned the bike lanes are often obstructed by delivery trucks or pedestrians which forces them to share the road with motor vehicles. This is evident from the high score for violation in New York City. Most participants mentioned riding the bike on the road makes them feel more visible and safer, especially in busy urban areas where the speed limit is low. Additionally, some participants mentioned they often do not feel the same responsibility to wait in the traffic when they can maneuver between the traffic to reach their destination. These differences in laws indicate the need to do state-wise validation of the CBQ. This will clarify the differences observed in the subscales.

Another limitation is the use of a bike simulator as a validation method. While the results showed it was able to predict all the subscales of the CBQ, cyclists still wouldn't behave in a similar way on the road. It is important to do further validation using a naturalistic study on the road.

2.6 Conclusion

This study explored an 11-item survey questionnaire to understand the behaviors of US cyclists on the road. The factorial structure was developed after conducting PCA. The subscales included violation, aggressive violation, positive behavior, and distraction and forgetfulness. Each subscale had a reliable Cronbach alpha. The factorial structure was then validated using CFA. Construct reliability and validity was found to be adequate as well. The questionnaire was further validated using scenario-based questionnaire on the survey and on the bike simulator. Results showed the behavior on the bike simulator to predict all four subscales of the CBQ. This indicates the importance of having observational validation method investigating realistic cyclist behaviors. Overall, the study answered all three research questions. The PCA structure identified the top four factors from the questionnaire- violation, aggressive violation, positive behavior, and distraction and forgetfulness (Question 1A). The CFA results validated the factors and structure of the questionnaire by providing an acceptable range for chi square test results and goodness of fit indices (Question 1B). Lastly, the scenario-based behavior questionnaires further validated the questionnaire by predicting the scales accurately (Question 1C).

There were differences observed in terms of gender, age, and history of crashes with non-motor vehicles. Male participants had a higher mean score for aggressive violation, cyclists aged 31-50 and cyclists who had 1-5 times crashes with non-motor vehicles had a higher mean score for distraction and forgetfulness. There were several differences observed in different states which indicates this CBQ could be a very powerful tool in finding the issues that lead cyclists to take risky behavior on the road. Further research is needed to find the underlying issues in terms of infrastructure or traffic laws in different states using the CBQ.

CHAPTER III: STUDY II

DESIGNING A WARNING SYSTEM FOR A SMARTPHONE-BASED ASSISTIVE APPLICATION FOR BICYCLISTS

3.1 Introduction

Developing a questionnaire can aid in understanding the behavior of bicyclists which can guide researchers and policymakers to develop and implement safer and improved infrastructures, and assistive technologies. There has been a lot of research on bicyclist-centered traffic infrastructure (McNeil et al., 2022a, 2022b; Ngwu et al., 2022b). However, there is still the need to design an affordable and effective assistive tool for bicyclists to reduce the crash risks with motor vehicles. Bicycle-related crash prevention involves developing educational programs, changing policies, and infrastructure designs (Pucher et al., 1999). The infrastructure design plays a key role in making bicyclists feel safe on the roads. In a survey study, participants were asked to provide their perceptive response on the type of infrastructure that they would find the safest (Stülpnagel & Binnig, 2022). Participants rated bicycling tracks (infrastructure exclusively built for bicyclists and physically separated from the motor vehicle lanes and sidewalks) to be the safest followed by bicycling lanes (a part of the roadway exclusively built for bicyclists, not always physically separated). Bicycling lanes with physical separation from the car lane, greater lane width, and the colored surface had a higher perceived safety from bicyclists' ratings. During the COVID-19 pandemic, several European cities added bicycle lanes to ensure safe cycling and maintain appropriate social distancing (European Cyclists' Federation, 2020). Research has shown that this new change had led to a significant increase in the bicycling rate (Kraus & Koch, 2021).

Clearly, the impact of infrastructure change on bicycling is strong; however, developing or modifying infrastructures can be time-consuming and expensive, and thereby hard to implement.

With the current advances in technology, assistive systems can be developed in a cost-effective way and can be more efficient and easier to implement than policy or infrastructure changes. Several studies have been developing such systems to alert motor vehicle drivers of possible obstacles in their way. The Advanced Driver Assistance Systems (ADAS) is such a computer-controlled technology that can help drivers by controlling vehicle acceleration, deceleration, lane position, and by providing alerts for impending collisions from any side (Zahabi et al., 2020). The Intersection Collision Warning System (ICWS) is another example of a safety feature that can detect any approaching car in real-time via the sensors in the car and/or base station at intersections (Zhang et al., 2020). Similarly, the Motorcyclist Safety Assistant Application (MSAA) can detect a motorcycle's speed in real real-time and alert the driver if their speed has exceeded the speed limit (Fernando et al., 2020). It can also detect collision threats and can alert the driver via visual and audio cues (Fernando et al., 2020). Bicyclists, being one of the vulnerable road users and being exposed to the severe consequences of a traffic crash, need to have similar opportunities to use these types of assistive technologies for their safety. Nevertheless, when there is much research on assistive technologies for motor vehicle users, there have not been many investigations on such systems for bicycle riders. This can be due to the need for expensive sensors which will increase the price of this affordable and low-cost means of transport. In addition, bicyclists need to be always aware of their surroundings for making immediate responses to traffic threats and an assistive system can have two-fold outcomes: supporting or distracting.

This research gap regarding the successful implementation of bicyclist assistive technology needs to be addressed. Bicyclists should be able get timely and useful head starts with effective warnings for an approaching collision event while not causing any distraction to their usual cycling behavior. The current study aims to investigate potential warning signals for a bicyclist assistance system, CycleGuard (Jin et al., 2021). CycleGuard is a prototype smartphone-based collision detection app, developed out of the scope of this research, that can assist bicyclists to detect hazardous traffic conditions. This app can also provide them with warnings to avoid those hazards. The objective of this research is to identify effective warning signals which will successfully convey collision alerts to the user while riding and will not be distracting and/or startling to avert them from their attention to the surroundings. This study was completed in two phases: the first phase administered a focus-group discussion with stakeholders and researchers in transportation. The objective of this study was to find a list of potential warning signals for CycleGuard. A final list of warning signals was developed based on the focus-group study. The signals were implemented in CycleGuard to assess their efficiency and feasibility with respect to users' reaction time, their physiological and emotional statuses. In order to do these, the second phase used a bike-simulator study where bicyclists were exposed to various traffic environments with different types of collision threats. Under these scenarios, bicyclists experienced each warning signal from the final list along with a control condition with no warning. From bicyclists' responses toward the collision event, emotional statuses, physiological responses (heart rate), and cycling behavior, the efficacy of these warning signals were evaluated for their inclusion in CycleGuard. With the advancements of simulators and virtual reality, it has become easier to assess such assistance systems in a safer environment without causing real threats to human subjects. The overall objective of this research is to alert the bicyclist to prevent fatalities and enjoy the benefits of cycling.

3.1.1 Research Objective and Research Questions

Objective 1: Identifying potential alerts to develop a warning system for a BAS. A focus group study with experts including researchers (transportation safety, cyclist safety), policymakers (City and DOT), and stakeholders (bike manufacturers, People for Bike, GoHopr Bikeshare, Strava, Verizon) was performed to identify a potential list of designs that can be incorporated as warnings to cyclists for different road hazards. Different types of multimodal warning systems such as visual, auditory, and haptic alerts were considered for experts to discuss in this focus group study. Data regarding pros and cons for each design along with experts' ratings and rankings for each individual design and different combinations were collected.

Objective 2: Evaluating alerts to develop an effective warning system for cyclists. A bike simulator study was conducted to assess the multimodal alerts and their effectiveness in warning cyclists about impending hazards while riding their bikes. Recommended visual and auditory alerts from the focus group study were added under different environmental conditions and collision scenarios to analyze the usefulness of each alert. The outcomes of this study can be used in developing an effective warning system for the bicyclist assistance app.

This study investigates the following research questions based on the objectives:

Question 2A: What type of designs were more preferred; visual, audible, or haptic?

Question 2B: Was there a preference for multimodal designs over unimodal designs?

Question 2C: How did different alerts affect users' heart rate and emotional statuses?

Question 2D: How effectively did the use of warnings improve overall cycling experience in terms of reaction time, stable heart rate, and emotional statuses?

3.2 Literature Review

3.2.1 Effect of warnings on user performance

Human Factors Engineering (HFE) is the discipline and process that discovers and applies the knowledge of human limitations and capabilities to design systems and equipment (Luquetti dos Santos et al., 2009). It also ensures that the design of a system, as well as human tasks and the work environments are compatible with sensory, perceptual, and cognitive capabilities of the operator. Man-Machine Interface or Human-Computer Interface (HCI) of a system incorporates Human Factors Principles and Cognitive Engineering principles to design a user-friendly system. In these interfaces, designers often use an alarm/warning system that compensates for human limitations, such as limited attention span or inability to track several parameters simultaneously (Xiao & Seagull, 1999). Researchers have shown that alarm systems can improve human performance while performing complex tasks (Sun et al., 2010; Bustamante et al., 2007; Gupta et al., 2002; Sorkin et al., 1988). Bustamante et al. (2007) investigated the effect of changing threshold of alarms on human performance. They found that human performance increases in the presence of a warning system regardless of the threshold. Highest performance was observed when the threshold was lower despite having a higher number of false alarms. Gupta et al. (2002) similarly found that human performance can improve when drivers had a warning of adverse road and weather conditions ahead of time. They observed drivers to have fewer skids when they had a warning system with lower threshold and a graded auditory warning. Sorkin et al. (1988) investigated operator performance in the presence of an alarm. Similar to the previous studies, they also showed that warning can improve operator performance on primary and secondary tasks. Operator’s attention allocation can be improved as well, which was integrated into their decision. Sun et al. (2010) asked driver participants to perform some steering tasks, in the presence of error feedback. They concluded that participants' accuracy increased significantly in the presence of error feedback as a guidance.

All these studies indicate that an alarm system can significantly enhance human performance; however, a poorly functioning alarm system can be a contributing factor for major accidents (Hollifield & Habibi, 2007). A poorly designed alarm system is considered a usability issue. In 1994, there was an explosion and fires at Texaco Milford Haven refinery that caused several injuries, as well as £48 million in loss (Smith et al., 2003). The poor warning system was the main reason behind the accident as the operators had to recognize, acknowledge, and act on 275 alarms in just 11 minutes prior to the accident (Smith et al., 2003). With proper warning, this accident could have been prevented (Newsholme, 2007). The principles behind an alarm are- an operator response is required because of an alarm; the same thing should not be signified by multiple alarms; and alarms should be based on the true condition it is designed for, not for expected cases (Hollifield & Habibi, 2007). Having too many false alarms can reduce the compliance to the alarm system(Dixon et al., 2007). This is known as the cry wolf effect (Naujoks et al., 2016). Mistrusting the system can lead to the operator ignore or shut down the warning system (Bliss, 1993).

The perception-reaction time of bicyclists is 2.5 s (Taylor & Davis, 1999). It is essential that the warning signal can convey the message and assist the cyclist within this short period of time. Therefore, the principles for an effective alarm system should be used in transportation to alert the road users. Having a warning that takes long to process is a poor warning design and can cause fatal accident (Hollifield & Habibi, 2007).

From the research mentioned above, it is evident that warning systems have contributed significantly to improving human performance. Therefore, it is worth investigating how this form of HCI can be utilized to increase the safety of vulnerable road users.

3.2.2 Importance of effective warnings in traffic safety

Traffic safety is a major research area for researchers and engineers as it has a direct impact on human lives (Oh et al., 2005). There are several electronic driving aids available to eliminate human errors occurred during the information processing stage of navigating through complex traffic (Brookhuis et al., 2009). However, having these aids can increase the mental workload of drivers. Therefore, several research has been conducted on developing a warning system that can reduce the mental workload on the road (Biondi et al., 2017; De Angelis et al., 2018; Murata et al., 2013; Whitmire et al., 2011; Zuki & Sulaiman, 2016). Multimodal warning system has been a popular alerting system due to its redundant alerting features. In several studies, multimodal warning system has proven to be more effective than unimodal warning system (Erdei et al., 2020; Matviienko et al., 2018; Waard et al., 2016; Yun & Yang, 2020).

Warning System for Drivers

In transportation, several research focused on reducing the reaction time to prevent fatal consequences (Druta & Alden, 2020; Mohebbi et al., 2009). Druta & Alden (2020) investigated drivers' alertness when animal crossing was detected. They found 80% of participants either braked or reduced their speed when the warning sign was activated. One year later, the warning sign was deployed in practice, and researchers found that deer-vehicle crashes reduced by 75%. Mohebbi et al. (2009) explored this type of collision further by including secondary task of having a phone conversation. They found tactile warnings to have a faster brake reaction time than both auditory and no warning. This indicates that the inattention that is caused by doing secondary tasks while driving can be mitigated by using a warning system. There has been several research conducted on developing warning systems for motor vehicle users (Whitmire et al., 2011; Geitner et al., 2019a). Geitner et al. (2019) conducted a car simulator experiment to evaluate the efficiency of auditory, tactile, and multimodal warnings while driving an automated car. They found participants to have shorter reaction time for multimodal (auditory, tactile) and auditory only warning systems. One limitation was the inaccurate *Reaction Time;* as participants already had their foot on the brake pedal, they had a quicker reaction as opposed to a real-life scenario where they would take some time to move their foot on the brake. Whitmire et al. (2011) investigated driver's behavior in work zones when presented with multimodal signals. They found in-vehicle warning has a significant effect on the driver's response; researchers also revealed that auditory warning creates a quicker response from the drivers. They concluded their study by suggesting a combined warning system consisting of both visual and auditory warnings. Several researchers have expressed concern in using visual warning system on the road as it can cause mental workload for road users (Ho et al., 2007; McKeown & Isherwood, 2007; Senders et al., 1967). This is due to the simultaneous processing of both traffic environment and warning signals using visual stimuli. Being consistent with this concern, Ho et al. (2007) investigated the effect of utilizing auditory warning consisting of a car honk, vibrotactile warning consisting of a Velcro belt, and audio-tactile warning signal to alert the drivers about a front-to-rear-end collision. Results indicated that multimodal warnings have a significant advantage in alerting drivers over unimodal warnings. They also found vibrotactile warning captured drivers' attention while they were distracted. One limitation of incorporating vibrotactile warning is that thickness of clothing might contribute to how effectively this mode of warning is perceived.

Warning System for Vulnerable Road Users (Cyclists and Pedestrians)

While the research on developing warning system for the vulnerable road users is scarce, the preference of multimodal warning signals and the concern for visual signals is consistent for bicyclists and pedestrians (von Sawitzky et al., 2022; Strohaeker et al., 2022; Erdei et al., 2020; de Angelis et al., 2018; Matviienko et al., 2018). Matviienko et al. (2018), conducted a simulator-based study on implementing multimodal warning system for children, whereas Erdei et al. (2020), ran similar field experiment for adults. Mtvieenko et al. (2018) found implementing visual, auditory and vibrotactile cues enabled children bicyclists to perceive the information faster and have shorter reaction time. Similarly, Erdei et al. (2020) found vibro-tactile signals to be the most efficient warning and auditory signals to be easily perceivable. Nevertheless, the researchers stated that visual signals were frequently missed by bicyclists. Additionally, Erdei et al. (2020) found route types; for example, high traffic density, bumpy road surface, and loud ambient noise had effects on the perception of the signals, especially the vibro-tactile signals. It is important to note that both studies did not include all age groups and they did not measure the application of a warning system to avoid a crash.

Strohaeker et al. (2022) conducted a test track experiment where the warning system had acoustic signal, vibro-tactile signal, or no signal. The results showed that the participants who received acoustic signals had the shortest reaction time followed by the participants who received vibrotactile signal. The participants of this study, however, were the employees of the research institution due to COVID-19 restrictions and therefore, lacks the representation of general population. Another web-based study was performed by Angelis et al. (2018) on the preference of passive vs active warning systems. In this study, passive warning system warns the driver about a bicyclist and active warning system warns the bicyclist about potentially dangerous situations.

Researchers found that the majority of the participants (67.8%) preferred a passive warning system followed by an active audio-visual warning system for the bicyclists as a complementary precautionary method. This study, however, is inclined to bias and reporting errors as the participants self-reported the survey questionnaire, and they also completed the survey without having exposure to the scenario. Sawitzky et al. (2022) developed a warning system to alert bicyclists about a car door opening. They incorporated three different warnings in a smart helmet: visual messages, visual messages and auditory tone, and visual and voice messages. They found the warning systems significantly increased the lateral distance between the door and the participant. Participants showed an inclination towards the bimodal warning signals. This study, however, had a very small sample size (*N*= 24) and didn't include a varying age group. While research has shown auditory and vibrotactile to be more effective than visual, Erdei et al. (2020), pointed out the environmental factor of using vibrotactile warnings. These warnings were often missed by their participants in the presence of haptic interference, such as bumpy and uneven roads. This indicates having a vibrotactile warning may not be efficient while cycling.

In a more recent study KeriBig et al. (2022) evaluated a cyclist warning system consisting of the trimodal warnings- visual, audible, and tactile. The visual signal consisted of red hexagon with an exclamation mark, audible consisted of alerting sounds, and tactile consisted of vibration of the handlebar in the direction of the critical event. Researchers evaluated the warning system in terms of users' trust, acceptance, feeling of safety, and workload using a bicycle simulator and questionnaires. They also investigated the differences in assessing the warning system by the older and younger cyclists. Results revealed that participants had increased perceived safety, as well as trust in the warning system in the presence of the critical event. Researchers found differences in age groups in terms of workload and acceptance. While the older population had higher ratings for workload, the younger population had higher ratings for acceptance. This study identified the differences in how different age population perceives new technology; however, this analysis was based on subjective assessment and lacked any objective measure to understand the efficacy of the warning system.

It can be observed in this section that, although there is a significant amount of research conducted on alerting motor vehicle drivers, there is very limited research on alerting bicyclists. Some researchers also opt for evaluating the safety of cyclists from the perspective of drivers, where the drivers are alerted about the bicyclist on the road. Additionally, there is no research that currently evaluates the efficacy of the warning system based on cyclists' physiological, and emotional variables. Cyclists being very exposed on the road need to be very alerted, therefore slight distraction can cause severe consequences. To prevent a warning signal from being distracting for the cyclist, it is important to monitor the physiological parameters, such as heart rate. Their situational awareness, which can be analyzed from their emotional data, also needs to be assessed in the presence of the warning system. There is a huge research gap in finding an effective warning system for vulnerable road users which needs to be addressed. This study aims to bridge this gap by assessing the efficacy of the warning system from both subjective and objective perspectives.

The following table summarizes the related work reviewed in this section.

Table 14: Summary of safety research in transportation

Study reference	Study population	Platform	Factors considered	Warning considered	Effectiveness tested?
KeriBig et al. (2022)	Elderly: 32 Younger: 33 Age: 18-40 years	Buke Simulator	- Rating Scale of Mental Effort (Questionnaire) - Subjective safety assessment (Questionnaire) - Trust in automated systems (Questionnaire)	Trimodal	Yes
Swaitzky et al. (2022) Germany	Male: 17 Female: 5 Age: 18- 40	Bike simulator	- Cycling behavior (speed and lateral position) - Perceived safety (Questionnaire) - Ease of use (Questionnaire)	Visual Visual-tone Visual-voice message	No Yes Yes
Strohaeker et al. (2022) Germany	Male: 14 Female: 7 Age: 20- 44	Test track	- Reaction time and time to collision - Speed difference - Minimum speed - Average acceleration - Maximum front wheel brake pressure - Brake duration - Adopted absolute yaw rate	Auditory Vibrotactile No Warning	Yes Yes No
Matviienko et al. (2018) Germany	Male: 8 Female: 7 Age: 6- 13	Bicycle simulator	- Reaction time - Duration and frequency of glances - Number of accidents - Understandability (Likert scale) - Distraction (Likert scale)	Visual Auditory Vibrotactile Trimodal	No No No Yes
Erdei et al. (2020) Germany	Male: 37 Female: 15 Age:18-65	Field study	- Reaction time - Duration - Mean speed - Maximum speed	Visual Auditory Vibrotactile	No Yes Yes
De Angelis et al. (2018) EU	Male: 1171 Female: 1210 Transgender: 8 Age:18-86	Survey	- Warning mode preference - Cycling frequency and country of residence - Demographics	Trimodal Active haptic Active audio-visual Passive	No No Yes Yes
Ho et al. (2007) UK	Male: 15 Age: 17- 41 years	Driving Simulator	- Response time - Shortest distance headway - Percentage of collisions	Auditory Vibrotactile Audio tactile	No No Yes
Geitner et al. (2019) UK	Male: 19 Female: 26 Age:20-39	Driving Simulator	- Reaction time - Subjective variable: ratings to the warnings	Tactile Auditory Auditory-tactile	No Yes Yes
Whitmire et al. (2011) USA	Male: 27 Female: 33 Age:20-63	Driving Simulator	- Lane position, steering wheel positions - acceleration, - braking inputs - Time and speed - Mental workload using NASA-TLX	Traditional signage Visual Auditory	No No Yes

3.2.3 Research methods to design warning systems

Qualitative approach

Developing an efficient warning system requires both a qualitative and a quantitative approach. A very popular and widely used qualitative method is conducting focus group studies (Dimitrakopoulou, 2021). Focus group study enables participants to have an interaction among themselves as a group based on their experiences. Compared to self- reporting qualitative approaches, that has the limitation of bias, focus group studies act as a tool to understand peoples' perception and mental models relying on prior experiences, opinions, and decisions (Dimitrakopoulou, 2021). The sample size and composition of the focus doesn't have any set method of calculation as it depends on characteristics, and age of the participants, as well as the questions that are being investigated (Then et al., 2014). In literature, many researchers have recommended to have between four to twelve participants in the focus group so that the group is neither too large nor too small (Bloor et al., 2012; Dilorio et al., 1994; Morgan et al., 1998). The total time used for focus group studies should not exceed two hours in total (Doody et al., 2013; Morgan & Krueger, 2014; Plummer-D'Amato, 2008). Focus group study has been used as a primary source of data collection method across several basic and applied disciplines, such as education, sociology, communications, health sciences, organization behavior, psychotherapy, political science, social psychology, gerontology, policy research, sociology, anthropology, information systems, marketing, and management (D. W. Stewart & Shamdasani, 2014).

In transportation research, qualitative method, such as focus group studies are recommended for areas that are seldom studied as this allows to have an in-depth investigation of the issue (Thomas et al., 2022; Clifton & Handy, 2003). This can also act as a powerful tool to understand complex travel behaviors; in fact, focus group is considered an effective way to understand the behavior of young children as the interactive group discussion can reveal different points of view among them via comparing and sharing (Thomas et al., 2022b; Simons et al., 2014; Clifton & Handy, 2003). Focus group has been used to study the transportation of disadvantaged groups of people on the road as well; some examples include, the transportation preferences of older population (Coughlin & Aarp, 2001); transportation needs and obstacles of adults in New Jersey who fall under the autism spectrum (Lubin & Feeley, 2016); and the limitation in physical activities of older population with intellectual disabilities (van Schijndel-Speet et al., 2014). With the advent of autonomous vehicles and assistive technology, cognitive researchers have shifted their attention to investigate the perception of the technology by different age group population (Ngwu et al., 2022b; Dichabeng et al., 2021; Etminani-Ghasrodashti et al., 2021) and their trust in the system (Zoellick et al., 2021; Faber & van Lierop, 2020) using focus group studies.

The insights obtained from these studies are commonly used to develop heuristics for user-centric designs (Delgado et al., 2020; Machín et al., 2020; Lee & Marlowe, 2003). Heuristic analysis is a popular method in interface design as it gives the usability experts "rules of thumb" to evaluate the design (Cassano-Piché, 2015). Heuristic analysis is typically performed on the HCI of software designs to evaluate its usability (Combs et al., 2020; Markov et al., 2016; Atkinson et al., 2007). Researchers have developed several well-established guidance to design a good user interface (Shneiderman et al., 2016; Khaun, 2013; J. Zhang et al., 2003; Nielsen, 1995). Some leading experts in developing these heuristics include Nielsen (1995) with his *10 Usability Heuristics for User Interface Design,* Schneiderman (2016) with his *Eight Golden Rules of Interface Design,* and Zhang (2003) with the *14 Usability Heuristics* for evaluating medical devices (Stanton et al., 2013). The heuristic techniques of Nielsen and Schneiderman are very general and can be used to analyze any type of user interfaces (Stanton et al., 2013). Nielsen's heuristics are very commonly utilized in literature as they are reviewed, modified, and validated several times to ensure the heuristics can capture the changes and advances in technology (de Almeida Pacheco et al., 2019; Sivaji et al., 2011). Nielsen's 10 heuristic lists contain- visibility of system status; match between system and the real world; user control and freedom; consistency and standards; error prevention; recognition rather than recall; flexibility and efficiency of use; aesthetic and minimalist design; help users recognize, diagnose, and recover from errors, and help documentation (Nielsen, 1994). Finding such heuristics requires experts and stakeholders in the area to come together and have a discussion, which can be achieved via a focus group study.

Quantitative approach

The quantitative aspect of the HCI to evaluate its efficacy can be conducted using a field test, experimental platform test, and simulation (Zhao et al., 2022). While field tests provide the most realistic data, it has high requirements for the driver, scene, and the equipment (Li et al., 2019). Moreover, any task that might compromise the driver's safety, cannot be performed in a field study (Bham & Leu, 2018). Experiment platform tests, in comparison are safer than field tests; however, they are very expensive to implement, and they lack versatility (Kim et al., 2020). Virtual reality and simulated environments allow testing users' interaction with new and under-developed technologies in a safe setting. They also have a low cost, as well as highly reproducible tests which can be repeated multiple times without resetting entire scenario, as required by the experimental platform tests (Hang et al., 2022; Watanabe & Sakai, 2021). Virtual Reality (VR) headsets, Cave Automatic Virtual Environment (CAVE), or traffic simulations are a few examples of these types of settings. These simulated environments allow users to repeat trials multiple times, save time and money in designing and retrieving different experimental settings, and test products and services that are currently unavailable (Baek et al., 2020; Bella & Silvestri, 2021; Chang & Chang, 2011).

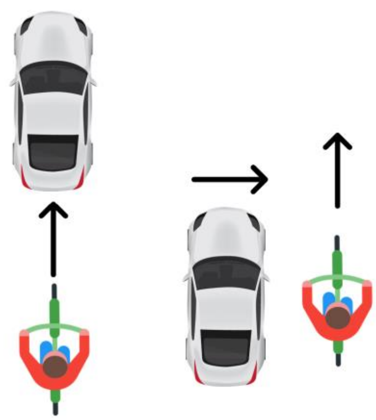
Baek et al., (2020) tested their multi-senior Collision Warning System using a driving simulator study. They used visual and audible signals to test vehicle-to-vehicle collision and vehicle-to-pedestrian collision scenarios. From their time-to-collision (TTC) measures, researchers have concluded that the system could increase the safety of drivers and pedestrians by improving their perception about their surroundings. In another study, Chang & Chang (2011) used a bus simulator and investigated the efficacy of visual and auditory signals for bus-to-pedestrian collisions at an intersection. They measured accident occurrences, glance frequency, glance duration, perception-reaction time, and emergency deceleration rate. Researchers stated that their warning system was effective; however, it required better integration of detection and telecommunication techniques. Bella & Silvestri, (2021) conducted a driving simulator study to look at drivers' behavior while interacting with pedestrians at and outside of crosswalks. Researchers provided the driver participants with auditory and visual alerts when a pedestrian was detected. They collected variables such as, driver's initial speed, distance from the conflict point, minimum speed at the end of the deceleration phase, and the distance from the conflict point to the point where the minimum speed value is located. The outcomes of the study revealed positive effect of warnings on most drivers' interaction with pedestrians.

One of the major concerns about conducting simulation studies is that they are not completely realistic (Zhao et al., 2022). Despite having this limitation, simulation studies provide a platform to understand how the design can be perceived by the users. The efficiency of the HCI can also be analyzed from the perspective of Human Factors Engineering. Moreover, this research method helps understand risky driving behavior under different conditions without actually putting the driver at risk.

In summary, continuous rising number of bicyclist fatalities and injuries and inconsequential focus on bicyclist assistance system research, there is a need for effective warning system for bicyclists. Bicycling can promote healthy living, facilitate affordable mobility for low-income and disadvantaged populations, and become a useful transport mode for inexperienced population like teenagers. Therefore, an assistive system to improve bicyclists' situational awareness about their surroundings can significantly enhance road-users' mode choice. The goal of this research is to design an effective and affordable smartphone-based warning system for bicyclists to improve traffic safety and encourage more people to bicycle. To accomplish this goal, a unique approach of evaluating HCI in designing a warning system was utilized. It is of utmost importance to have an HCI that can assist bicyclists instead of distracting them. Therefore, interface analysis of the warning system for CycleGuard was conducted in two phases. In the first phase, an online focus group study was organized with stakeholders and transportation researchers to identify a list of potential warning signals. Heuristics, relevant to this interface design, was used for experts' discussion data analysis. In the second phase, an experimental study was conducted by using a bike-simulator. The simulator study tested the feasibility and usefulness of the selected signals for their inclusion in the application using a safe and controlled simulated traffic environment. These studies were approved by the Institutional Review Board (IRB) at the University of Texas at Arlington. The following chapters are divided according to the two study phases.

3.3 Phase One: Focus Group Study

This research was conducted to determine a warning system for the smartphone-based application-CycleGuard, developed and tested for validity by Jin et al., (2021). CycleGuard is an acoustic-based collision detection app that is developed using a low-cost and commercially available off-the-shelf portable speaker and a smartphone. The goal of this app is to alert bicyclists of potential *Right Hook Collisions* and *Frontal Collisions* by detecting the surrounding traffic in real time. Figure 5 illustrates and defines these two collisions as considered in this research.



(a) (b)

Figure 5: (a) Frontal collision: a car abruptly applies brake in front of the bicyclist. (b) Right Hook Collision: Driver turning right fails to notice the bicyclist.

3.3.1 Methodology

3.3.1.1 Warning Signals

A list of signals that have been found effective in previous road-user safety studies was created to present to the experts during the focus group study. The visual warnings included - a static image containing very descriptive illustration of the imminent danger (i.e., frontal collision and right hook collision) (Kingsley et al., 2020); static image of black arrow pointing in the direction of the collision (Neurauter, 2005); flashing image containing an easy to perceive icon, such as a STOP sign or a red square (Erdei et al., 2020; Matviienko et al., 2018); flashing directional cues with a car icon flashing either in the left or in the middle of the screen; and flashing red or yellow circles, and flashing red or yellow boxes with a message against a black background (Politis et al., 2013; Whitmire et al., 2011). The auditory warnings included- 4 repeated base frequency of 2573 Hz and 60 dB loud tones (Waard et al., 2016); a 2.0 kHz and 3.0 kHz tone of three seconds duration and 0.2 seconds intermittent cycle (Whitmire et al., 2011); tones differentiating high (1000 Hz) vs. low (400 Hz) urgencies; car honk (Geitner et al., 2019; Graham, 1999); and speech saying, “Slow Down!”, and “Please Stop!” (Whitmire et al., 2011). The haptic warnings included- phone vibration, wearables, seat vibration, and handlebar vibration (Erdei et al., 2020; Matviienko et al., 2018). The list of signals is summarized in Figure 6. The unimodal signals, as well as combination of signals (multimodal) were used in the focus group study to be included in the assistive application.

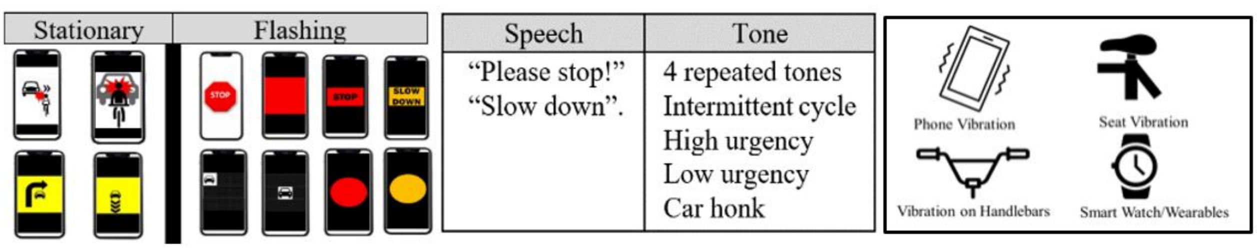


Figure 6. Warning signals included in the study

3.3.1.2 Participants

Researchers, stakeholders, and policymakers were invited in the focus group study. Some of the organization where participants were contacted included Atlanta Regional Commission, People for Bikes, Verizon, City of Portland, Georgia Department of Transportation (GDOT), Bike Dallas-Fort Worth, and a transportation researcher in the field of smart-phone app technology. The stakeholders were invited to participate in the research via email. Upon receiving the confirmation of their participation, they were provided with the consent form and an online meeting invitation for the focus-group discussion. There was a total of 7 experts who participated in the focus group discussion.

3.3.1.3 Study Protocol

In the beginning of the study, participants were introduced to the researchers and to each other. They were given a description of the study procedure followed by a description of the CycleGuard. Participants watched a demonstration explaining the usage and functionalities of the app while being used by a researcher in a real-world scenario. After this demo, they were presented with each class of signal, individually, i.e., first visual, then audible, and finally haptic. When the visual signals were presented, the sub-classifications, such as static and stationary signals were also presented. Followed by the presentation of each class of signals, participants were asked to answer questions on how they liked the mode of the signals, which sub-classification they would prefer, what message each signal conveyed to them, etc. The process was repeated for audible and haptic signals. Once all the signals were presented and discussions were completed for each class, the participants were asked to fill out a survey containing each signal and to rate them on a numerical 7-point Likert scale for ‘1’ being ‘Strongly Disagree’ to ‘7’ being ‘Strongly Agree’. The survey also included questions about what modality and combination of signals they preferred the most.

3.3.1.4 Data Analysis

The transcript was produced from the Microsoft Teams Transcription service and was edited for a clean verbatim by removing filler words, repeated words, and stutters. The analysis of the transcript was completed using Python’s Natural Language Processing (NLP) package. Once the process of Data Scrubbing was completed, the keywords were identified to determine which warnings they preferred the most. The keywords for the heuristic analysis were analyzed by coding the words into the appropriate heuristic defined by Nielsen (Nielsen, 1995). Nielsen's heuristics included visibility of system status; match between system and the real world; user control and freedom; consistency and standards; error prevention; recognition rather than recall; flexibility and efficiency of use; aesthetic and minimalist design; help users recognize, diagnose, and recover from errors, and help documentation. The NLP analysis was used to find the frequency of those heuristics.

3.3.2 Results

3.3.2.1 Warning system features

The transcript was cleaned by removing all capitalization, punctuation, number and alphanumeric strings, and all new lines and carriage returns. The top words and their frequency distribution related to the feature of the warning was then identified. Words such as “bright”, “animated”, “flashing”, “high pitch”, etc. were considered as features of the warning system. Figure 7 shows the frequency distribution for the features of the warning. The top preference for the warning was flashing/animated visual signal (86%). Experts also wanted the color to be consistent with traffic colors for stopping or yielding, such as a bright red or a yellow color. They wanted the color to be bright and eye catching as well.

For audible signals, most experts preferred an abstract tone (71%) with high pitch to be ideal warning. Some experts preferred speech or to provide the user with the option to choose tone or speech (29%). The experts did not prefer having haptic signals as a warning signal for CycleGuard, as having a haptic signal might create confusion for the cyclist while distinguishing between other phone notifications and the warning.

3.3.2.2 Heuristic factors influencing the selection of the warning system

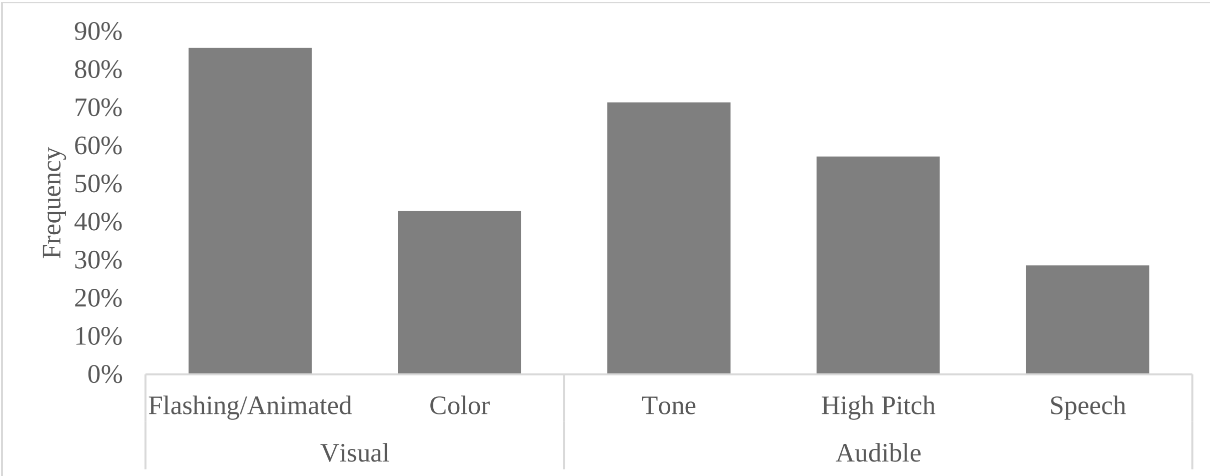


Figure 7: Visual and audible warning signal features

Table 15: Factors influencing experts' perception of the warning system design for CycleGuard

Heuristics (Percentage)	**Definition**	**Example of Verbatim**
Compatibility (34%)	The appearance of the system and how it works should be compatible with user's expectation and conventions	Not too much information, not distract bicyclist, not over-alert
Visibility (29%)	While carrying out the task, the system should be able to meet the need and requirements of the user	Alert, catch attention
Explicitness (18%)	The user should be clear about how the system works and how it is structured	Clear, understand, interpret
Consistency (8%)	The appearance of the system and how it works should be consistent all the time	Consistent, standard
User guidance and support (8%)	The system should be easy to use, informative and have relevant guidance and help to understand and operate	Simple
Error prevention (3%)	The system should be able to minimize user error	Error

As described in section 3.3.1.4, the heuristics for the usability of the warning system were identified from the transcript. Each participant's comments were recorded onto one Excel file separating each class of designs. The words with high frequency were identified from NLP analysis and were matched with the comment to understand the context. All comments regarding potential designs were coded based on the heuristics developed by (Nielsen, 1995). The percentage of factor usage and examples of verbatim are presented in Table 15.

3.3.2.3 Survey Results Analysis

Due to the limitation of time, the survey could not be completed at the end of the focus group session. Experts were reminded of the survey in an email afterwards; however, only two responses out of the seven participants were received. Therefore, the survey results were not analyzed.

3.3.3 Discussion

This preliminary study was conducted to identify warning signals that can be incorporated for further testing in the second phase of the study using a bike simulator. Experts were invited to participate in the focus group discussion to identify the ideal warning signals from a list containing visual, audible, and haptic signals. Their discussion was transcribed from Microsoft Teams and was analyzed using Python's Natural Language Processing (NLP) and Microsoft Excel.

3.3.3.1 Preference for unimodal vs multimodal warnings

During the focus group discussion, participants had a strong inclination towards having a combination of visual and audible signals. Some participants compared it to having Google maps where both visual and audible cues are used to direct the users. The two survey results that were received, both participants strongly preferred the use of multimodal warnings. In previous studies, multimodal warnings have been mostly preferred over unimodal warnings (De Angelis et al., 2018; Erdei et al., 2020; Matviienko et al., 2018; von Sawitzky et al., 2022). Experts were strongly against using haptic in any combination with the warning signals as they can create confusion regarding whether it is a phone notification or a critical warning signal. In terms of having a seat or handlebar vibration, they expressed their concerns about the vibration interfering with the user's cycling experience.

3.3.3.2 Preference for the features of the warning signal

As shown in Figure 6, experts preferred flashing/animated visual signals over static signals. They used sentences like "I like flashing something that catches your attention", "Having flashing color, instead of the word stop" to emphasize their preference. They argued that having flashing stop sign might prompt the user to come to a hard stop instead of being aware of their surroundings. Experts mentioned color to be an important feature for the warning. While some experts supported using a red color to warn the cyclists, others preferred a yielding color, such as yellow to be ideal.

For audible signals, most participants preferred tone and speech (Figure 6). One expert expressed his concern regarding using "Please Stop" as a speech warning, as that can startle the cyclist. The experts were strongly against using a car horn to alert the cyclists as that might interfere with the surrounding traffic and create confusion. While the preference for an abstract high pitch tone signals was higher, one expert mentioned giving the user an option to choose between tone and speech would be a better option than selecting either one of them.

Lastly, in terms of using different signals for different warning, experts were also strongly against that idea. They said, "the simpler the better", therefore, having one signal to identify both crashes is better than remembering which signal would indicate front collision vs. which one would indicate right hook collision.

3.3.3.2 Heuristics influencing experts' perception of the warning system

Heuristics are used in qualitative research to assess the usability of the product especially, software (Combs et al., 2020; Markov et al., 2016; Atkinson et al., 2007). Table 14 summarizes the heuristics that were obtained after coding the word with Nielsen's heuristics (Nielsen, 1995). For example, words like "not too much information", "not distract bicyclist", "not over-alert", were coded for compatibility. These words were used in the context of user expectations and conventions, so the app can assist cyclists instead of causing distraction leading to a fatal accident. According to the experts, compatibility was the most important feature of the CycleGuard app followed by functionality, explicitness, consistency and standards, user guidance and support, and lastly error prevention. These heuristics are essential to evaluate the usability of the app before releasing it to the public.

3.3.3.3 Limitations

This limitation of this study was that it was conducted only in one focus group session. Multiple sessions with different experts in transportation might have shed light on different perspectives.

3.3.4 Conclusion

This study was conducted as phase one for the Bicycle Assistant System (BAS) app-CycleGuard. The objective was to identify the ideal warning signals for the next phase of the study including a bike simulator, as well as identify heuristics from experts' opinion regarding the usability of the app. The session was conducted on Microsoft Teams and was recorded and transcribed to analyze using Python's Natural Language Processing (NLP) package. The results provided a compelling answer to the first research question (Question 2A) by revealing that experts preferred visual, and audible warnings the most. They were strongly against using haptic warning as haptic warning in any form could cause confusion. In response to the second question (Question 2B), experts preferred a multimodal warnings system consisting of visual and audible signals over unimodal warning. Limitations for this study included a single focus group session. Based on the preference for the features of the warning system, red and yellow flashing visual sign, and high pitch tone and speech saying “slow down” were used in bike simulator study to identify the ideal warning that can be incorporated in the CycleGuard app. The heuristics identified will be used in assessing the usability of the final app before making it public.

3.4 Phase Two: Bicycle Simulator Study

The second phase of the study involving the bike simulator at the Human Factors lab at UTA was conducted using the warning signals obtained from the first phase of the study. The following sections will go over the methodology, results, discussion, and conclusion of the study.

3.4.1 Methodology

3.4.1.1 Scenario Design

The environment for the study was designed using the Internet Scene Assembler package provided by Realtime Technologies (RTI). Historic crash data and literature review was used to determine the road types. Two road types were created: a two-lane road with midblock crossing for the frontal collision scenario, and a two-lane road with a separate bike lane at a signalized intersection for the right hook collision scenario (Figure 8). SimCreator DX, another package provided by RTI, was used to incorporate three factors: collision type (frontal vs. right hook), speed limit (30 mph vs. 45 mph), and time of the day (day vs. night). A simpler smartphone application, similar to CycleGuard, was developed in order to prompt a warning on the phone screen using Android Studio and JavaScript.

The signals obtained from the focus group discussion were incorporated in this study to identify the most efficient warning system. These signals were compared with the control scenario i.e., in the absence of a warning.

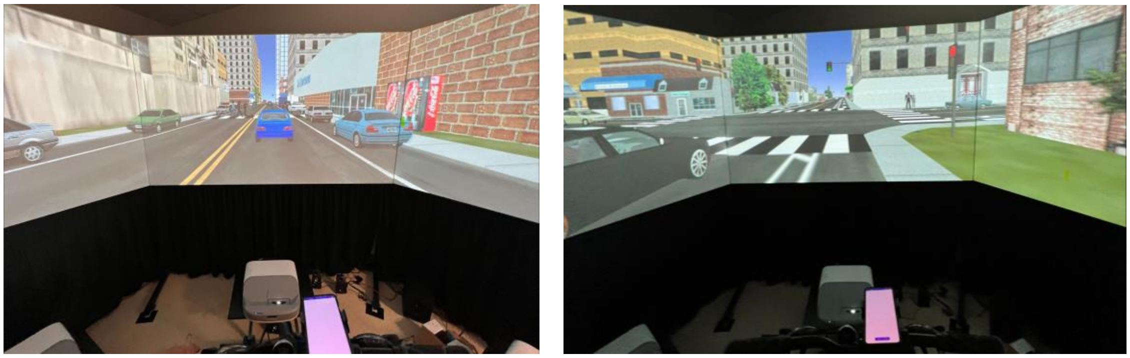


Figure 8: (a) Front collision scenario (b) Right hook collision scenario

3.4.1.2 Participants

A total of 32 participants were recruited for the study. Participants were informed about the study in multiple engineering courses at UTA as a means of earning extra credit. To be eligible to participate in the study, the participants had to be 18 years old; cycle at least once a week or for 20 minutes a week; have no visual disabilities or have corrected visual abilities with contact lenses, have no hearing disabilities or have corrected hearing abilities with hearing aids; fluent English speakers, and can follow instructions during the study. Participants who had motion or simulation sickness, are fitted with a heart pacemaker or automatic defibrillator, have pre-existing irritation or trauma around the chest, cannot bicycle for up to 20 minutes with a small break, were not included in the study. Participants were informed about the BioHarness, physiological data (heart rate) collection sensor, that was strapped around their chest.

3.4.1.3 Design of experiment

The experiment was designed following the two-level half-fractional factorial design. Since there were five factors of study, each having two levels, number of trials was calculated using., $2^{5-1} = 16$. The five factors included time of the day (day/night), speed limit (low/high), collision type (right hook/frontal), visual cues (red/yellow), and audible cues (speech/tone). The model had 16 combinations and each combination was replicated twice, hence 32 participants were required. Each participant was randomly assigned to one of the 32 scenarios. They experienced five combinations of visual and audible signals and five no-signal conditions with the assigned scenario. Each combination was run five times to get an average value, resulting in participants getting ten trials in total. The assignment of warnings and controls was randomized. The model had Resolution V confounding structure. In this design, no effect is aliased with the main effect or the two-factor interaction, however, the two-factor interaction is aliased with at least one three factor interaction. As a result, all two factor interactions can be observed here. The ANOVA model tested was:

$$Y_{ijklmt} = \mu_{\ldots..} + \alpha_i + \beta_j + \gamma_k + \delta_l + \lambda_m + (\alpha\beta)_{ij} + (\alpha\gamma)_{ik} + (\alpha\delta)_{il} + (\alpha\lambda)_{jm} + (\beta\gamma)_{jk} + (\beta\delta)_{jl} + (\beta\lambda)_{jm} + (\gamma\delta)_{kl} + (\gamma\lambda)_{km} + (\delta\lambda)_{lm} + \varepsilon_{ijklmt}$$

where $\mu_{\ldots..}$ is the overall mean of the population, α is the parameter for visual, β is the parameter for audible, γ is the parameter for time of day, δ is the parameter for speed, λ is the parameter for collision type, $\alpha\beta$ is the parameter for the interaction of visual and audible, $\alpha\gamma$ is the parameter for the interaction of visual and time of the day, $\alpha\delta$ is the parameter for the interaction of visual and speed, $\alpha\lambda$ is the parameter for the interaction of visual and collision type, $\beta\gamma$ is the parameter for the interaction of audible and time of day, $\beta\delta$ is the parameter for the interaction of audible and speed, $\beta\lambda$ is the parameter for the interaction of audible and collision type, $\gamma\delta$ is the parameter for the interaction of time of day and speed, $\gamma\lambda$ is the parameter for the interaction of time of day and collision type, $\delta\lambda$ is the parameter for the interaction of speed and collision type, and ε_{ijklmt} is the random error term.

3.4.1.4 Study Protocol

Participants upon their arrival were first acquainted with the general structure of the experiment. They were then provided with an infectious disease screening form followed by a consent form explaining the process of the experiment and the benefits and risks of their participation. Covid-19 risk information was read out to each participant to ensure that they do not have any signs of Covid-19. Participants were then provided with the demonstration of strapping the BioHarness and if they agreed to continue, they were asked to put them on. Similar to the simulator study in Chapter 2, the participants were provided with Simulation Sickness Questionnaire (SSQ) to ensure their fitness for the participation (Kennedy et al., 1993). The questionnaire is attached in Appendix B. A 5-point score was set as a threshold for withdrawing participants from the study, with zero being the highest score for fitness. Based on this, the participant's health was kept in check along with the participant's ability to give accurate feedback. After filling up and qualifying with the first SSQ, participants were taken to the bike simulator room. Participants were instructed about operating the bicycle using the smartphone app. They were riding the bicycle for about 5 minutes to get acclimated to the simulated traffic environment and the bicycle. They were then asked to fill out the SSQ again to assess their fitness to continue. The participants were asked to take the cyclist behavior questionnaire, developed in the first study described in chapter II and the personal innovativeness questionnaire (Agarwal & Prasad, 1998) to show their willingness towards accepting this assistive technology while riding bicycles. The questionnaire is attached in Appendix C. The participant then completed all 10 trials. Upon the completion of these trials, participants took the final SSQ and a feedback survey with demographics and their likeliness of using the app rated on a 7-point Likert scale. Participants were asked to provide any additional feedback that they felt like the surveys did not cover before they left. The experimental process is summarized in Figure 9.

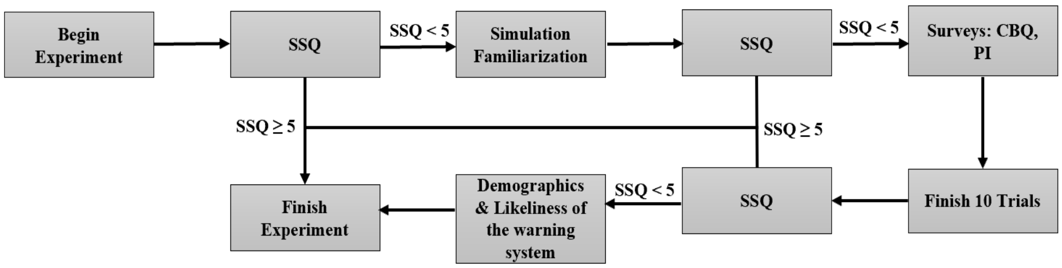


Figure 9: Experimental process flow

Note: SSQ= Simulation Sickness Questionnaire; CBQ= Cyclist Behavior Questionnaire; PI= Personal Innovativeness Questionnaire

3.4.1.5 Data Analysis

The data was analyzed using Python, SPSS Statistics 29, and SAS Studio software. Python was used to compile the data and plot preliminary graphs. SPSS and SAS were used to run all assumption tests, Analysis of Variance (ANOVA), Pearson correlation test, and t-tests. The independent variables were categorical: type of collision, time of the day, speed limit, audio, and visual warning type. The following dependent variables were collected:

Reaction Time: Duration between when the alarm went off and when the participant started applying brakes. This time was obtained from the video data. An audible cue was placed when the warning would go off and brake lights were installed on the bike. Reaction time was calculated by subtracting the time when the cyclists started to press the brakes from the time the audible cue went off. This is a continuous variable.

Heart Rate: This data was obtained from the BioHarness manufactured by BIOPAC Systems. The BioHarness is equipped with a sensor that can detect, monitor, and record physiological parameters, such as the heart rate. This is also a continuous variable.

Additionally, participant's emotional variables- arousal, surprised, scared, angry, disgusted, valence, sad, happy, and neutral were also collected by running the videos through the FaceReader software developed by Nodulus. FaceReader has been tested for its accuracy by using Amsterdam Dynamic Facial Expression Set (ADEFS), where FaceReader resulted in a 100% accuracy in detecting emotions (van der Schalk et al., 2011). FaceReader works by first detecting the face using a deep learning-based face-finding algorithm; then it does facial modeling technique based on deep neural network that can find 468 key points on the face (Bulat et al., 2017; Zafeiriou et al., 2015). The key points are then compressed using Principal Component Analysis. Lastly, the facial expression classification is carried out using deep artificial neural network for recognizing patterns in the face (Gudi et al., 2015). Each expression is calculated in terms of intensity where "0" would mean the emotion is absent and "1" would mean the emotion is present (Loijens & Krips, 2021). The resulting variable is a continuous variable that can measure emotion. While all the outputs are based on intensity, valence is the calculation of positive or negative emotions, therefore it is calculated by subtracting the intensity of negative emotions from the intensity of positive emotions.

For this study, the data was collected in a paired approach where each participant did both control and trials. Therefore, to assess the improvement in cyclists' performance, the difference in means for the control and trial was used as response variables. The data was first divided into two sections: data collected in the presence of visual-audible warning system, and data collected in the absence of the warning system (control). The mean value for each participant for each dependent variable was calculated. The mean value in the absence of warning was then subtracted from the mean value in the presence of the warning for all the dependent variables. This difference in mean value resulted in observing the improvement in the presence of the warning signals under different factors. A more negative number for the difference in mean reaction time would indicate an improved cyclist's reaction time in the presence of the warning signals. Additionally, since the heart rate was expected to increase in the presence of the warning signals, a positive difference is expected. A lower number for the difference in mean heart rate would indicate less increase in heart rate in the presence of the warning system. The ANOVA assumptions of no outliers, normally distributed error, and constant variance were satisfied before conducting the ANOVA. The first assumption was tested using Bonferroni outlier test. The decision rule used here was to identify any absolute value of the studentized deleted residual that exceeds the Bonferroni cutoff point that is calculated using a t-test. For this study the t_{ijklmt} value was 3.854. The second assumption for normally distributed errors was tested using the normal probability plot of the residuals and normality test. The normality test for whether the data confirms normality was conducted. The decision criteria for the normality test was if $\hat{\rho} < c(\alpha,n)$, the null hypothesis H_0 can be rejected, and normally was violated at 0.05 significance level. From the table for critical values for "coefficients of correlation between ordered residuals and expected values under normality when distribution of error terms is normal", c(0.05,32) ~ 0.964. The third assumption for constant error variance was tested using the plots for residual vs. each factor and residual vs. fitted values, as well as modified Levene's Test. To conduct the modified Levene's Test, first the observations were divided into two groups. An F test was carried out to determine whether the variances of the two groups were equal, where the null hypothesis H_0 was that the variances of the two groups are equal. Next, a *t*-test was conducted with the hypothesis that, means of the two groups were equal. A fail to reject H_0 would indicate constant variance at 0.05 significance level. Once the assumptions were satisfied, each model was developed using F-tests. Post-hoc analyses were performed for significant effects. The effect of the questionnaires- Cyclist Behavior Questionnaire

(CBQ) and Personal Innovativeness (PI) questionnaire on the difference in mean reaction time, and difference in mean heart rate was explored using Pearson correlation analysis. Lastly, the effect of the warnings on the emotional variables was assessed using t-test.

3.4.2 Results

3.4.2.1 Effect of the environmental factors on reaction time, and heart rate

The statistical analysis of the factors was conducted using Analysis of Variance (ANOVA). As mentioned in section 3.4.1.5, five factors were used in this study to design a two-level half fractional factorial design. The dependent variables were calculated by subtracting the mean value of the five trials from the five control runs. This would indicate the improvement after implementing the warning system.

Reaction Time

The difference in mean reaction time was calculated by subtracting the mean reaction time (RT_0) without any warning from the mean reaction time (RT) in the presence of the warning ($RT-RT_0$). The assumptions for ANOVA were then tested. For the outlier test, none of the absolute value of the studentized deleted residual was greater than 3.854. Therefore, there were no outliers detected in this model. For the normality test, the correlation coefficient, $\hat{\rho}$ obtained was 0.984. Since, $\hat{\rho}$ (0.984) > c(0.964), the conclusion was fail to reject H_0; i.e., the test did not detect any non-normal distribution of errors at 0.05 significance level. This is however a weak conclusion. The normal probability plot in Figure 10, shows shorter tails than normal distribution. From the plots it can be concluded that normality was violated. therefore, the constant variance assumption was tested using the residual plots.

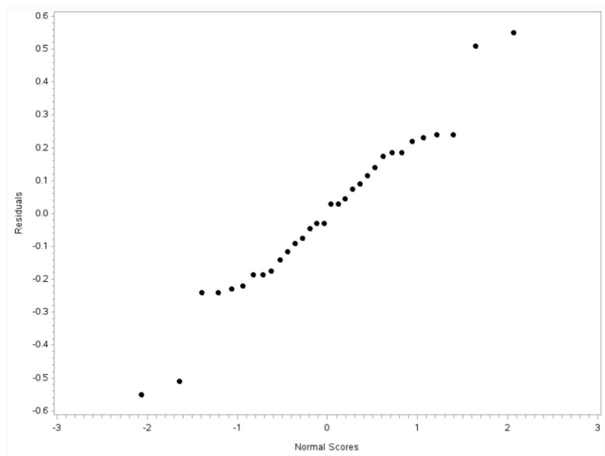


Figure 10: Normal probability plot of the residuals of the difference in mean reaction time

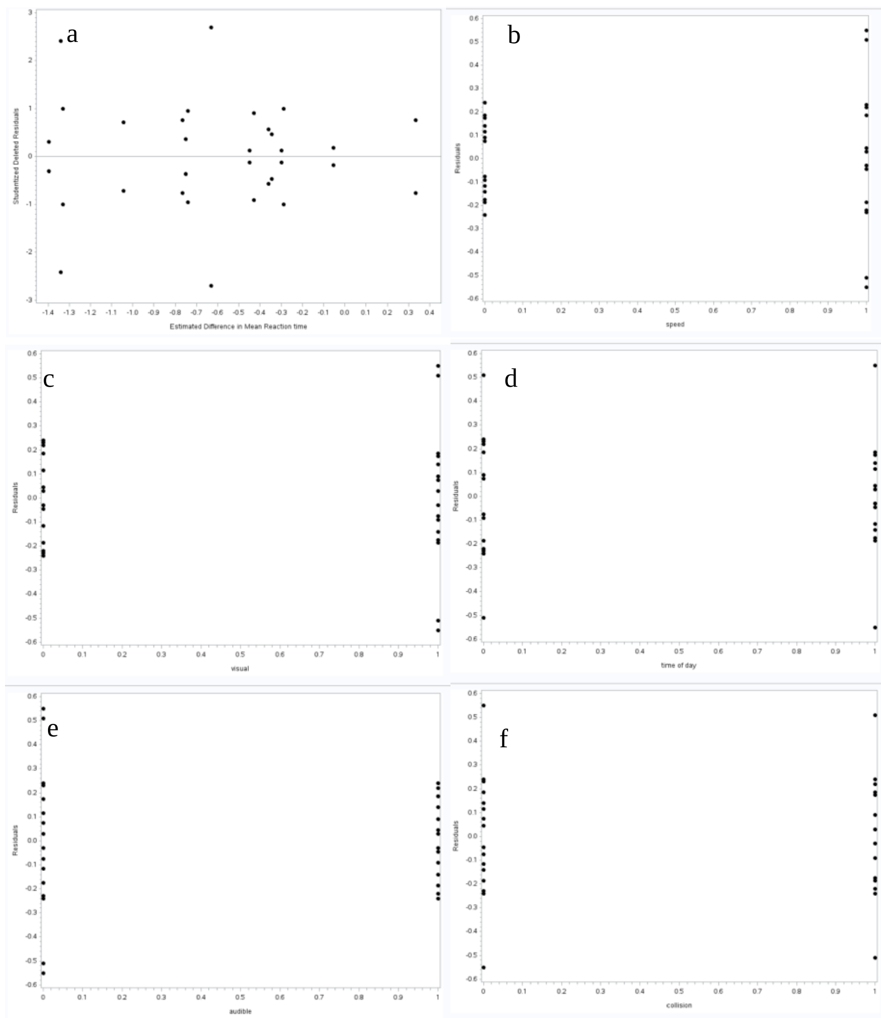


Figure 11: Residual plots for *difference in mean* rection time

(a) residual vs. fitted (b) residual vs. speed (c) residual vs. visual (d) residual vs. time of the day (e) residual vs audible (f) residual vs. collision type.

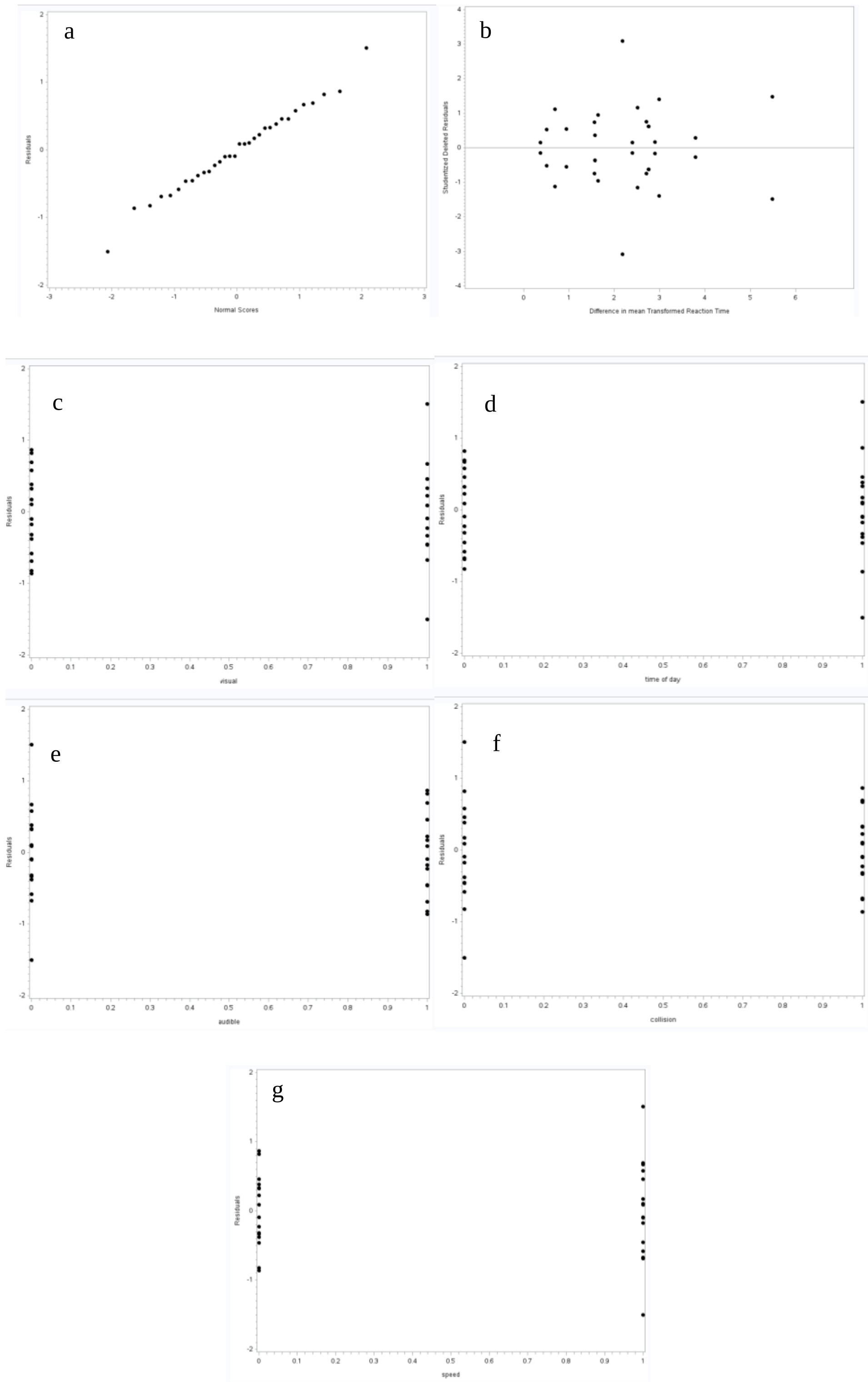


Figure 12: Residual plots for difference in the transformed mean rection time

(a) Normal probability plot (b) residual vs. fitted (c) residual vs. visual (d) residual vs. time of the day (e) residual vs. audible (f) residual vs. collision type (g) residual vs. speed.

Table 16: Effect of the factors on $(RT\text{-}RT_0+2)^2$

Source	*df*	Type III SS	Mean Square	F statistics	p value
Visual	**1**	**12.813**	**12.813**	**17.54**	**0.0007**
Audible	**1**	**15.207**	**15.207**	**20.81**	**0.0003**
Time of the Day	**1**	**15.989**	**15.989**	**21.88**	**0.0003**
Collision Type	1	0.118	0.118	0.16	0.6937
Speed	1	0.015	0.015	0.02	0.8878
Time of the Day*Speed	1	0.324	0.324	0.44	0.5149
Collision Type*Speed	1	0.016	0.016	0.02	0.8859
Visual*Speed	1	0.573	0.573	0.78	0.3890
Audible*Speed	**1**	**3.538**	**3.538**	**4.84**	**0.0428**
Time of the Day*Collision Type	1	0.306	0.306	0.42	0.5267
Visual* Time of the Day	1	1.308	1.308	1.79	0.1996
Audible* Time of the Day	1	0.005	0.005	0.01	0.9324
Visual*Collision Type	1	0.261	0.261	0.36	0.5584
Audible*Collision Type	1	1.013	1.013	1.39	0.2563
Visual*Audible	1	1.068	1.068	1.46	0.2442
Note: Significant effects are bolded					

The residual plot of residual vs. fitted value in figure 11 does show some variability which was confirmed from the residual vs. factors plot. The residual plots show variability which indicated constant variance assumption was violated. Therefore, variance stabilizing transformation was used on the response variable. The transformations that were tested are summarized in Appendix G. The final transformation used was: $(RT\text{-}RT_0+2)^2$.

The normal probability plot and the residual plots for the transformed value $(RT\text{-}RT_0+2)^2$ are presented in Figure 12. The normal probability plot showed slightly longer tails than normal distribution; however, it improved from Figure 10. For the normality test, the correlation coefficient, $\hat{\rho}$ obtained was 0.99224. Since, $\hat{\rho}$ (0.99224) > c(0.964), the conclusion was fail to reject the null hypothesis, H_0. The test did not detect any non-normal distribution of errors at 0.05 significance level. Therefore, it can be concluded that the assumption that errors are normally distributed was not violated.

The residual plots had also improved compared to Figure 11. The residual vs. fitted values showed more scattered pattern, and the residual vs. factors show less variability compared to the original plot. The results of Modified Levene's test showed p-value= 0.2108, therefore fail to reject H_0, and equal variances can be used for the t-test. The t-test results for equal variances also had a p-value of 0.5681, hence, fail to reject H_0. Therefore, the modified Levene's test did not detect any non-constant variance in the model at 0.05 significance. The SAS output is attached in Appendix E.

After all the assumptions were satisfied, the model was selected using a series of F-tests. First the two-way interactions were tested where the null hypothesis H_0 was that interaction is negligible at 0.05 significance level. The ANOVA results are summarized in Table 16. From the results presented in Table 16 the following conclusions were made:

a) For the interaction effect of visual and audible, $p>0.05$, therefore, the interaction effect is not statistically significant.

b) For the interaction effect of audible and collision type, $p>0.05$, therefore, the interaction effect is not statistically significant.

c) For the interaction effect of visual and collision type, $p>0.05$, therefore, the interaction effect is not statistically significant.

d) For the interaction effect of audible and time of the day, $p>0.05$, therefore, the interaction effect is not statistically significant.

e) For the interaction effect of visual and time of the day, $p>0.05$, therefore, the interaction effect is not statistically significant.

f) For the interaction effect of time of the day and collision type, $p>0.05$, therefore, the interaction effect is not statistically significant.

g) For the interaction effect of audible and speed, $p<0.05$, therefore, the interaction effect is statistically significant.

h) For the interaction effect of visual and speed, $p>0.05$, therefore, the interaction effect is not statistically significant.

i) For the interaction effect of collision type and speed, $p>0.05$, therefore, the interaction effect is not statistically significant.

j) For the interaction effect of time of the day and speed, $p>0.05$, therefore, the interaction effect is not statistically significant.

Next, the main effects of visual, time of day, and collision type were tested. The null hypothesis H_0 was that the main effect is negligible at 0.05 significance level. The results are listed below:

a) For the main effect of visual, $p<0.05$, therefore, visual is highly statistically significant at $\alpha = 0.001$.

b) For the main effect of time of the day, $p<0.05$, therefore, time of the day is highly statistically significant at $\alpha = 0.001$.

c) For the main effect of collision type, $p>0.05$, therefore, collision type is not statistically significant.

The resulting model for future work is:

$$Y1_{ijklm} = \mu_{....} + \alpha_i + \beta_j + \gamma_k + \delta_l + (\beta\delta)_{jl}$$

Table 17: Least square means and least square means numbers for different combinations of the Audible*Speed interaction

Audible	Speed	$(RT\text{-}RT_0+2)^2$ LSMEAN	LSMEAN Number
0	0	1.14326250	1
0	1	1.85162500	2
1	0	3.18698750	3
1	1	2.56530000	4

Note: Audible level 0: tone, 1: speech
Speed level 0: 30 mph, 1: 45 mph

Table 18: Simultaneous 95% Confidence intervals for the interaction effect of Audible*Speed using Tukey test

Least Squares Means for Effect Audible * Speed				
i	j	Difference Between Means	Simultaneous 95% Confidence Limits for LSMean(i)-LSMean(j)	
1	2	-0.708363	-1.931186	0.514461
1	3	-2.043725	-3.266548	-0.820902
1	4	-1.422038	-2.644861	-0.199214
2	3	-1.335362	-2.558186	-0.112539
2	4	-0.713675	-1.936498	0.509148
3	4	0.621688	-0.601136	1.844511

Post-hoc test using Tukey pairwise comparison was then conducted on the interaction effect of audible and speed, and the main effect of visual, and time of the day. The following hypotheses was tested for the interaction effect:

H_0: *Difference between treatment means is zero*

H_1: *Treatment between treatment means is not zero*

Using Tukey pairwise comparison, the decision rule was to fail to the reject the H_0 if 0 was in the 95% confidence interval. The indices i, j represent the LSMEAN number in Table 17. The two LSMEAN numbers compare two combinations of audible*speed interaction. The complete SAS output for Tukey test is attached in Appendix E. The following conclusions were made based on Table 17-18;

1. For i=1, j=2, 0 lied in the confidence interval, therefore, fail to reject H_0, $\mu_{11} \sim \mu_{12}$

2. For i=1, j=3, 0 did not lie in the confidence interval, therefore, reject H_0, $\mu_{11} \neq \mu_{21}$

3. For i=1, j=4, 0 did not lie in the confidence interval, therefore, reject H_0, $\mu_{11} \neq \mu_{22}$

4. For i=2, j=3, 0 did not lie in the confidence interval, therefore, reject H_0, $\mu_{12} \neq \mu_{21}$

5. For i=2, j=4, 0 lied in the confidence interval, therefore, fail to reject H_0, $\mu_{12} \sim \mu_{22}$

6. For i=3, j=4, 0 lied in the confidence interval, therefore, fail to reject H_0, $\mu_{21} \sim \mu_{22}$

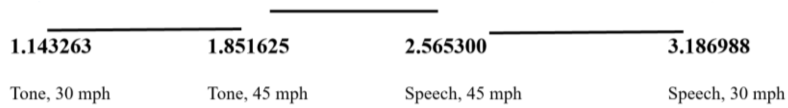


Figure 13: Line plot for Audible*Speed interaction effects

Figure 13 shows the line plot that can help visualize the interaction effect. There was no significant difference between the difference in mean reaction time for tone at 30 mph and at 45 mph. It was also not significant for speech at 30 mph and at 45 mph. This indicates that the difference in mean reaction time did not change at different speed levels for either tone or speech. It was significant for tone with speech at both 30 mph and 45 mph. Tone warning at 30 mph has the lowest difference in mean reaction time. In this case, as mentioned earlier, a lower value would indicate improved reaction time. Therefore, tone warning significantly improved reaction time compared to speech warning at both 30 mph and 45 mph.

Table 19: Simultaneous 95% Confidence intervals for the main effect of visual using Tukey test

Least Squares Means for Effect Visual				
i	j	Difference Between Means	Simultaneous 95% Confidence Limits for LSMean(i)-LSMean(j)	
1	2	1.265575	0.624915	1.906235

Table 20: Simultaneous 95% Confidence intervals for the main effect of time of day using Tukey test

Least Squares Means for Effect Time of Day				
i	j	Difference Between Means	Simultaneous 95% Confidence Limits for LSMean(i)-LSMean(j)	
1	2	-1.413738	-2.054397	-0.773078

Further post-hoc tests on the main effect of visual and time of the day were conducted. The i, j indices indicate the different levels of the factor. For visual, it is yellow and red; for time of day, it is night and day. The following conclusions on the main effects were made based on Table 19-20:

Visual: For i=1, j=2, 0 did not lie in the confidence interval, therefore, reject H_0, $\mu_1 \neq \mu_2$

Time of the day: For i=1, j=2, 0 did not lie in the confidence interval, therefore, reject H_0, $\mu_1 \neq \mu_2$

Red visual had a lower mean (*M*=1.55) compared to yellow warning (*M*=2.82) for the difference in mean reaction time. Similarly, night had a lower mean (*M*=1.48) compared to day (*M*=2.89). A lower mean would indicate the reaction time in the presence of the warning was lower than the reaction time in the absence of the warning, resulting in a more negative value (higher magnitude). This result indicates that the reaction time for the participants improved in the presence of red warning. Additionally, during night, participants also had a shorter reaction time in the presence of the warning signals indicating improved reaction time.

Heart Rate

The difference in mean heart rate was calculated by subtracting the mean heart rate (HR_0) without any warning signals from the mean heart rate (HR) in the presence of the warning signals ($HR\text{-}HR_0$). The assumptions for ANOVA were then tested. For the outlier test, none of the absolute value of the studentized deleted residual was greater than 3.854. Therefore, there were no outliers detected in this model. For the normality test, $\hat{\rho}$ (0. 985) > c(0.964), fail to reject H_0. The test did not detect any non-normal distribution of errors at 0.05 significance level; however, this is a weak conclusion. The normal probability plot in Figure 14 shows shorter tails than normal distribution. Therefore, from the normal probability plot, it can be concluded that the assumption that errors are normally distributed is violated.

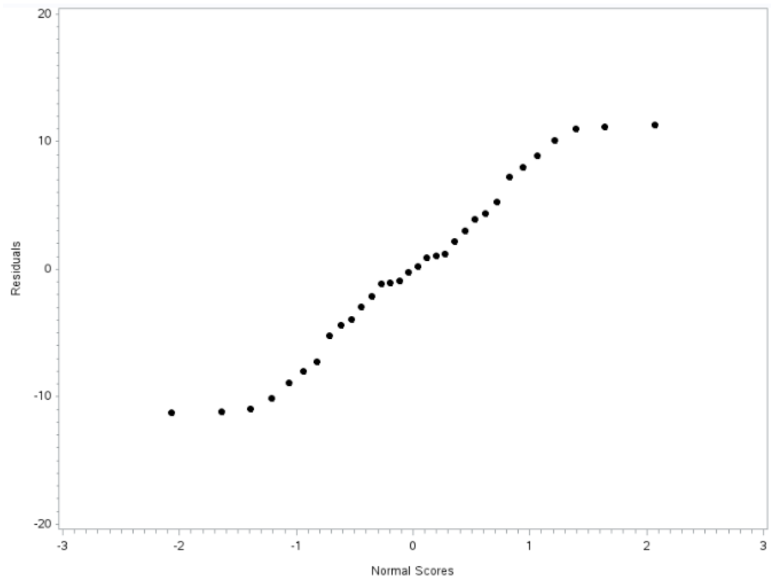


Figure 14: Normal probability plot of the residuals for the difference in mean heart rate

For the third assumption following the modified Levene's test, F-statistics was first assessed. Results showed the *p*-value= 0. 678, therefore fail to reject H_0, and equal variances can be used for the t-test. The t-test results for equal variances also had a *p*-value of 0. 2137, hence, fail to reject H_0. Therefore, the modified Levene's test did not detect any non-constant variance in the model at 0.05 significance level. This is further detected in the residual vs. fit and residual vs each factor graphs where the points did not show any wide variability (Figure 15). The SAS output is attached in Appendix E.

Aside from the normality assumption, other assumptions were satisfied. However, according to Lakens & Caldwell, (2021) ANOVA is robust against the violation of the assumption of normality, therefore, it can be used as long as the other assumptions are met. The model was selected using a series of F-tests. First the two-way interactions were tested where the null hypothesis H_0 is interaction is negligible at 0.05 significance level. The ANOVA results are summarized in Table 21.

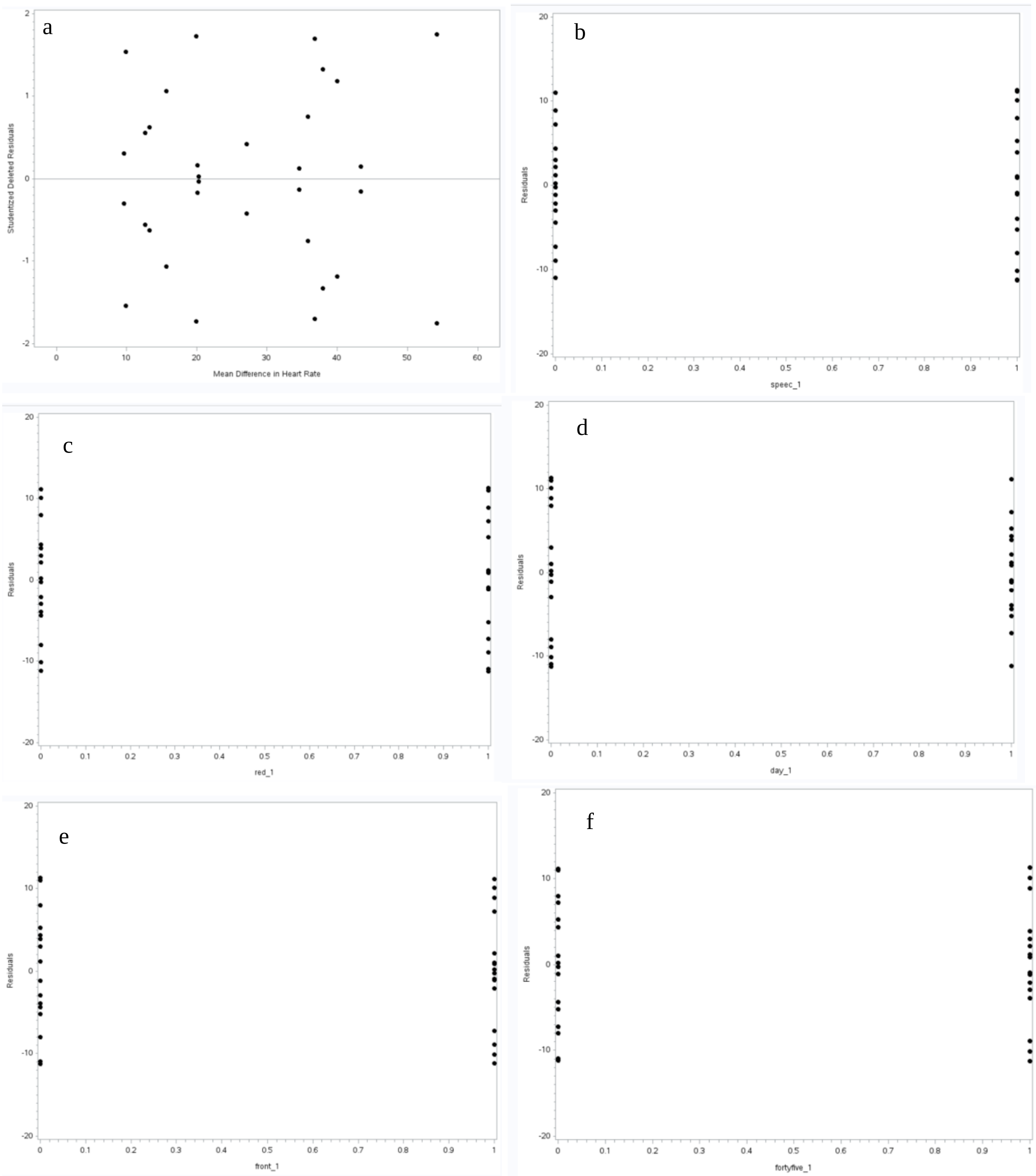


Figure 15: Residual plots for the difference in mean heart rate

(a) residual vs. fitted (b) residual vs. audible (c) residual vs. visual (d) residual vs. time of the day (e) residual vs collision type (f) residual vs. speed.

Table 21: Effect of the factors on difference in mean heart rate

Source	Sum of Squares	*df*	Mean Square	F statistics	p value
Visual	**1973.325**	**1**	**1973.325**	**21.042**	**<0.001**
Audible	**606.447**	**1**	**606.447**	**6.467**	**0.022**
Time of the Day	**1457.666**	**1**	**1457.666**	**15.544**	**0.001**
Collision Type	**294.930**	**1**	**294.930**	**3.145**	**0.095**
Speed	45.310	1	45.310	0.483	0.497
Time of the Day*Speed	1.772	1	1.772	0.019	0.892
Collision Type*Speed	2.736	1	2.736	0.029	0.867
Visual*Speed	**304.173**	**1**	**304.173**	**3.244**	**0.091**
Audible*Speed	167.672	1	167.672	1.788	0.200
Time of the Day*Collision Type	248.305	1	248.305	2.648	0.123
Visual* Time of the Day	74.641	1	74.641	0.796	0.386
Audible* Time of the Day	45.252	1	45.252	0.483	0.497
Visual*Collision Type	40.094	1	40.094	0.428	0.522
Audible*Collision Type	55.790	1	55.790	0.595	0.452
Visual*Audible	255.800	1	255.800	2.728	0.118

Note: Significant effects are bolded

The results are listed below:

a) For the interaction effect of visual and audible, p>0.05, therefore, therefore, the interaction effect is not statistically significant.

b) For the interaction effect of audible and collision type, p>0.05, therefore, the interaction effect is not statistically significant.

c) For the interaction effect of visual and collision type, p>0.05, therefore, the interaction effect is not statistically significant.

d) For the interaction effect of audible and time of the day, p>0.05, therefore, the interaction effect is not statistically significant.

e) For the interaction effect of visual and time of the day, p>0.05, therefore, the interaction effect is not statistically significant.

f) For the interaction effect of time of the day and collision type, p>0.05, therefore, the interaction effect is not statistically significant.

g) For the interaction effect of audible and speed, $p>0.05$, therefore, the interaction effect is not statistically significant.

h) For the interaction effect of visual and speed, $p>0.05$, however, it is significant at $\alpha <0.1$, therefore, the interaction effect is statistically significant.

i) For the interaction effect of collision type and speed, $p>0.05$, therefore, the interaction effect is not statistically significant.

j) For the interaction effect of time of the day and speed, $p>0.05$, therefore, the interaction effect is not statistically significant.

Next, the main effects of audible, time of the day, and collision type were tested. The null hypothesis H_0 was that the main effect is negligible at 0.05 significance level. The results are listed below:

a) For the main effect of audible, $p<0.05$, therefore, audible is statistically significant.

b) For the main effect of time of day, $p<0.05$, therefore time of day is statistically significant.

c) For the main effect of collision type, $p>0.05$, however, it is significant at $\alpha <0.1$, therefore, collision type is statistically significant.

The resulting model for future work is:

$$Y2_{ijklm} = \mu\ldots.. + \alpha_i + \beta_j + \gamma_k + \delta_l + \lambda_m + (\alpha\delta)_{il}$$

Table 22: Least square means and least square means numbers for different combinations of the Visual*Speed interaction

Visual	Speed	(HR-HR_0) LSMEAN	LSMEAN Number
0	0	23.3587500	1
0	1	14.8112500	2
1	0	32.8962500	3
1	1	36.6837500	4

Note: Visual level 0: yellow, 1: red
Speed level 0: 30 mph, 1: 45 mph

Table 23: Simultaneous 90% Confidence intervals for the interaction effect of Visual *Speed using Tukey test

Least Squares Means for Effect Visual * Speed				
i	j	Difference Between Means	Simultaneous 90% Confidence Limits for LSMean(i)-LSMean(j)	
1	2	8.547500	-3.249983	20.344983
1	3	-9.537500	-21.334983	2.259983
1	4	-13.325000	-25.122483	-1.527517
2	3	-18.085000	-29.882483	-6.287517
2	4	-21.872500	-33.669983	-10.075017
3	4	-3.787500	-15.584983	8.009983

Post-hoc test using Tukey pairwise comparison was then conducted on the interaction effect of visual and speed, and the main effect of audible, time of the day, and collision. The following hypotheses was tested for the interaction effect:

H_0: Difference between treatment means is zero

H_1: Treatment between treatment means is not zero

Using Tukey pairwise comparison, the decision rule was to fail to the reject the H_0 if 0 was in the 90% confidence interval. The indices i, j represent the LSMEAN number in Table 20. The two LSMEAN numbers compare two combinations of visual*speed interaction. The complete SAS output for Tukey test is attached in Appendix E. The following conclusions were made based on Table 22-23;

1. For i=1, j=2, 0 lied in the confidence interval, therefore, fail to reject H_0, $\mu_{11} \sim \mu_{12}$

2. For i=1, j=3, 0 lied in the confidence interval, therefore, fail to reject H_0, $\mu_{11} \sim \mu_{21}$

3. For i=1, j=4, 0 did not lie in the confidence interval, therefore, reject H_0, $\mu_{11} \neq \mu_{22}$

4. For i=2, j=3, 0 did not lie in the confidence interval, therefore, reject H_0, $\mu_{12} \neq \mu_{21}$

5. For i=2, j=4, 0 did not lie in the confidence interval, therefore, reject H_0, $\mu_{12} \neq \mu_{22}$

6. For i=3, j=4, 0 lied in the confidence interval, therefore, fail to reject H_0, $\mu_{21} \sim \mu_{22}$

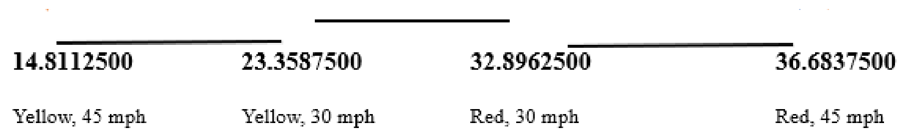


Figure 16: Line plot for Visual*Speed interaction effects

Figure 16 shows the line plot that can help visualize the interaction effect. There was no significant difference between the difference in mean heart rate between red visual at 30 mph and at 45 mph. It was also not significant between yellow at 30 mph and speech at 45 mph. This indicates that the difference in mean heart did not change at different speed levels. It was significant for red with yellow at both 30 mph and 45 mph. Red warning at 45 mph has the highest difference in mean heart rate. In this case, a higher value would indicate a more increased heart rate. Therefore, red warning caused more increase in heart rate than yellow visual warning did.

Table 24: Simultaneous 90% confidence intervals for the main effect of audible using Tukey test

Least Squares Means for Audible				
i	j	Difference Between Means	Simultaneous 90% Confidence Limits for LSMean(i)-LSMean(j)	
1	2	-8.706250	-14.578692	-2.833808

Table 25: Simultaneous 90% confidence intervals for the main effect of time of day using Tukey test

Least Squares Means for Effect Time of Day				
i	j	Difference Between Means	Simultaneous 90% Confidence Limits for LSMean(i)-LSMean(j)	
1	2	13.498750	7.626308	19.371192

Table 26: Simultaneous 90% confidence intervals for the main effect of collision type using Tukey test

Least Squares Means for Effect Collision Type				
i	j	Difference Between Means	Simultaneous 90% Confidence Limits for LSMean(i)-LSMean(j)	
1	2	6.070000	0.197558	11.942442

Further post-hoc tests on the main effect of audible, time of the day, and collision type was conducted. The i, j indices indicate the different levels of the factor. For audible, it is tone and speech; for time of day, it is night and day; for collision type it is right hook and frontal collision. The following conclusions on the main effects were made based on Table 24-26:

Audible: For i=1, j=2, 0 did not lie in the confidence interval, therefore, reject H_0, $\mu_1 \neq \mu_2$

Time of the day: For j=1, l=2, 0 did not lie in the confidence interval, therefore, reject H_0, $\mu_1 \neq \mu_2$

Collision type: For j=1, l=2, 0 did not lie in the confidence interval, therefore, reject H_0, $\mu_1 \neq \mu_2$

For audible warnings, speech warning had a higher difference in mean heart rate (*M*= 31.291) compared to tone warnings (*M*= 22.585). For collision type, right hook collision had a higher difference in mean heart rate (*M*= 29.973) compared to front collision (*M*= 23.903). Lastly, the difference in mean heart rate at night was also higher (*M*= 33.687) compared to day (*M*= 20.189). In this case, a higher heart rate indicates the heart rate in the presence of the warning was higher than the heart rate in the absence of the warning. A higher positive value would indicate an increased heart rate. Results suggest that speech warnings, right hook collision, and night resulted in a higher heart rate.

3.4.2.2 Influence of cyclists' behavior questionnaire and personal innovativeness on the warning system

Correlation analysis was used to find the linear correlation of the difference in mean reaction time and the difference in mean heart rate with the Cyclists' Behavior Questionnaire (CBQ) obtained in study I. Pearson correlation coefficient (*r*) was used to assess the level of the linear association among the variables. The subscales of CBQ included- violation, aggressive violation, positive behavior, and distraction and forgetfulness. The average value of each subscale was calculated.

Similarly, correlation analysis was conducted to find the linear correlation of the difference in mean reaction time and the difference in mean heart rate with the personal innovativeness questionnaire (PI) developed by Aggarwal & Prasad (1998). Responses for the third question stating, "In general, I am hesitant to try out new technologies", were first reversed before finding the mean score of the questionnaire. The average score of PI scale was then calculated to compare it with the difference in mean reaction time $(RT\text{-}RT_0)$ and the difference in mean heart rate $(HR\text{-}HR_0)$.

The plot for each analysis is summarized in Appendix G.

Reaction Time

The subscales of the CBQ (violation, aggressive violation, positive behavior, and distraction and forgetfulness) did not have significant linear correlation with the difference in mean reaction time (Table 27).

For the personal innovativeness (PI) questionnaire, the difference in mean reaction time had a significant negative correlation with the PI scale. The Pearson correlation coefficient, $r = -0.573$, indicates a moderate correlation among the variables. The negative sign indicates that participants who had higher PI score (more positive), had a more negative difference in mean reaction time ($R\text{-}R_0$). A more negative number (higher magnitude) indicates shorter reaction time in the presence of warning signals. Therefore, participants with higher PI scores had an improved reaction time in the presence of the warning signals.

Table 27: Effect of Cyclists Behavior Questionnaire and Personal Innovativeness Questionnaire on difference in mean reaction time

Questionnaire scale	Difference in mean reaction time
Violation	r = -0.107 *p* = 0.559
Aggressive Violation	r = 0.026 *p* = 0.887
Positive Behavior	r = -0.048 *p* = 0.792
Distraction and Forgetfulness	r = 0.173 *p* = 0.344
Personal Innovativeness	**r = -0.573** ***p* = <0.001**

r denotes Pearson correlation coefficient.

Heart Rate

The mean score of positive behavior of the CBQ scale had a significant negative correlation with the difference in mean heart rate ($HR\text{-}HR_0$) (Table 28). The Pearson correlation coefficient (r = -0.383) indicates a moderate correlation among the variables. None of the other subscales of CBQ scale or the PI scale had any significant correlation with the difference in mean heart rate.

Table 28: Effect of Cyclists Behavior Questionnaire and Personal Innovativeness Questionnaire on difference in mean heart rate

Questionnaire scale	Difference in mean heart rate
Violation	r = 0.014 *p* = 0.937
Aggressive Violation	r = 0.064 *p* = 0.729
Positive Behavior	**r = -0.383** ***p* = 0.031**
Distraction and Forgetfulness	r = 0.086 *p* = 0.641
Personal Innovativeness	r = -0.032 *p* = 0.860

r denotes Pearson correlation coefficient.

The negative correlation coefficient indicates a negative correlation of positive behavior with the difference in mean heart rate. Therefore, the higher the score participants had for positive behavior, the lower value they had for their difference in mean heart rate. Therefore, the lower increase they had in their heart rate in the presence of the warning signals. This indicates that participants who exhibit more positive behavior on the road had a low change in heart rate in the presence of the warning signals.

3.4.2.3 Influence of emotional variables on the warning types

The emotional variables- arousal, surprised, scared, angry, disgusted, valence, sad, happy, and neutral were analyzed using the FaceReader software developed by Nodulus as mentioned in section 3.3.2.5. The changes in each of these emotional variables in the presence of yellow vs red warning and tone vs speech warning was analyzed. Since each participant received either yellow or red for visual warnings, and either tone or speech for audible warnings, independent-sample t-test was conducted to find the significant differences in emotional variables based on the type of warning for each group of participants. The normality of the emotional variables was assessed using the normality plots attached in Appendix G. For arousal and scared, normality was satisfied. For surprise, normality was not satisfied; however, the distribution was symmetric with slightly shorter tails than normal distribution. Therefore, central limit theorem can be applied here (N=32), and t-test results can be considered approximate. For valence, normality was not satisfied; however, the distribution was symmetric with slightly longer tails than the normal distribution. Therefore, central limit theorem can be applied here (N=32), and t-test results can be considered approximate. Similarly for valence normality was not satisfied. The distribution was slightly right-skewed; however, central limit theorem can be applied here (N=32), and the t-test result can still be considered approximate due to the slight skewness. Angry, disgusted, happy, and sad had severe violation of normality. A log transformation was conducted as a remedy for angry, disgusted, and sad. A square root transformation was used for happy. The transformed distributions for log(angry) and log(sad) were close to the normal distribution, therefore, t test could be used. For log(disgusted) and $\sqrt{\text{happy}}$ there was still some skewness; however, it was improved and closer to the normal distribution, therefore, central limit theorem can be applied here (*N*=32), and t test can be used as an approximate.

Table 29: Effect of the visual signals on the emotional variables

Emotion Variables	Visual Warnings	Mean	SD	t	p value
Arousal	Red	0.320	0.073	0.595	0.557
	Yellow	0.310	0.066		
Surprised	Red	0.090	0.048	-0.869	0.392
	Yellow	0.110	0.067		
Scared	Red	0.040	0.028	0.430	0.670
	Yellow	0.040	0.027		
Valence	Red	-0.047	0.112	0.758	0.454
	Yellow	-0.076	0.106		
Neutral	Red	0.720	0.120	-1.055	0.300
	Yellow	0.765	0.121		
log(angry)	Red	-1.823	0.455	-0.940	0.354
	Yellow	-1.667	0.483		
log(disgusted)	Red	-1.663	0.600	-0.827	0.415
	Yellow	-1.506	0.464		
log(sad)	Red	-1.137	0.281	1.000	0.325
	Yellow	-1.274	0.469		
√happy	Red	0.257	0.127	1.089	0.285
	Yellow	0.209	0.124		

Table 30: Effect of the audible signals on the emotional variables

Emotion Variables	Audible Warnings	Mean	SD	t	p value
Arousal	**Tone**	**0.340**	**0.068**	**-2.331**	**0.027**
	Speech	**0.290**	**0.060**		
Surprised	**Tone**	**0.140**	**0.040**	**-7.068**	**<0.001**
	Speech	**0.050**	**0.032**		
Scared	**Tone**	**0.050**	**0.028**	**-2.699**	**0.011**
	Speech	**0.030**	**0.021**		
Valence	Tone	-0.071	0.116	0.479	0.635
	Speech	-0.052	0.103		
Neutral	Tone	0.741	0.136	-0.042	0.967
	Speech	0.743	0.109		
log(angry)	Tone	-1.755	0.533	-0.116	0.909
	Speech	-1.736	0.412		
log(disgusted)	Tone	-1.508	0.603	-0.810	0.424
	Speech	-1.662	0.462		
log(sad)	Tone	-1.269	0.407	-0.916	0.367
	Speech	-1.143	0.368		
√happy	Tone	0.243	0.131	0.424	0.675
	Speech	0.224	0.123		

Results revealed that visual signals did not have any significant effect on the emotional variables as shown in Table 29.

Audible signals had significant differences with arousal, surprised, and scared (Table 30). Tone warnings had higher mean score for arousal (*M*= 0.340, *SD*= 0.068), surprised (*M*= 0.140, *SD*= 0.040), and scared (*M*= 0.050, *SD*= 0.028) compared to the speech warnings (Figure 17).

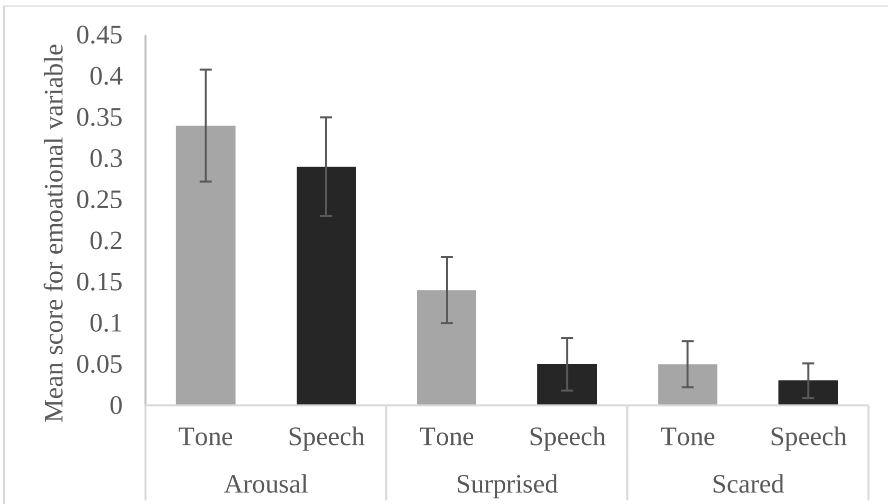


Figure 17: Mean score for significant emotional variables in the presence of audible signals

3.4.3 Discussion

The objective of this study was to determine the ideal warning system for the bicyclists' assistance app- CycleGuard. The efficacy of the warning system was explored considering factors such as collision type, time of the day, speed limit, visual warning signals, and audible warning signals. Several subjective and objective measures were analyzed to find the ideal warning signal.

3.4.3.1 Impact of the environmental factors on cyclists' performance

The Analysis of Variance (ANOVA) was conducted to find the effect of the factors on the dependent variables. The dependent variables were calculated by finding the difference in mean in the presence and absence of warning signals. The variables considered were reaction time, and heart rate. The difference in mean was calculated to identify whether implementing the warning system indeed improved the cyclists' performance in terms of shorter reaction time, and stable heart rate.

The ANOVA model for difference in mean reaction time revealed that visual, time of the day, and an interaction effect of speed and audible had a significant effect on the response variable (Table 16). Red visual warnings, tone audible warnings regardless of the high (45 mph) or low speed (30 mph), as well as night had a more negative difference in reaction time. This indicates the mean reaction time was shorter for red (M = 1.32 s, SD = 0.34), and night (M = 1.63 s, SD = 0.46) compared to where there was no warning signal present (Figure 18). The ANOVA model for heart rate revealed audible, time of the day, collision, and an interaction of visual and speed to be significant (Table 21). The heart rate was higher for speech audible signals (M = 107.04, SD = 20.94), and night (M = 100.31, SD = 18.53) (Figure 18). Heart rate was also significantly higher for right hook collision. Right hook collision is more challenging than front collision as the collision occurs from the cyclist's blind spot (Figure 5). This might have caused cyclists to have an increased heart rate compared to front collision.

The reason behind red light having a shorter reaction time and higher heart rate could be the societal conditioning of associating red with danger or a problematic situation (Pravossoudovitch et al., 2014). This finding in the current study is consistent with several other ergonomics research on warning system colors where the researchers found red to be the most efficient in conveying the danger (Chapanis, 1994; David Leonard, 1999; Pravossoudovitch et al., 2014; Wogalter et al., 1998). Tone warnings had shorter reaction time at both 30 mph and 45 mph. Previous research has shown speech warnings to not perform as efficiently as other audible cues such as, abstract tone, audio icons or spearcons (Šabić et al., 2021; von Sawitzky et al., 2022).

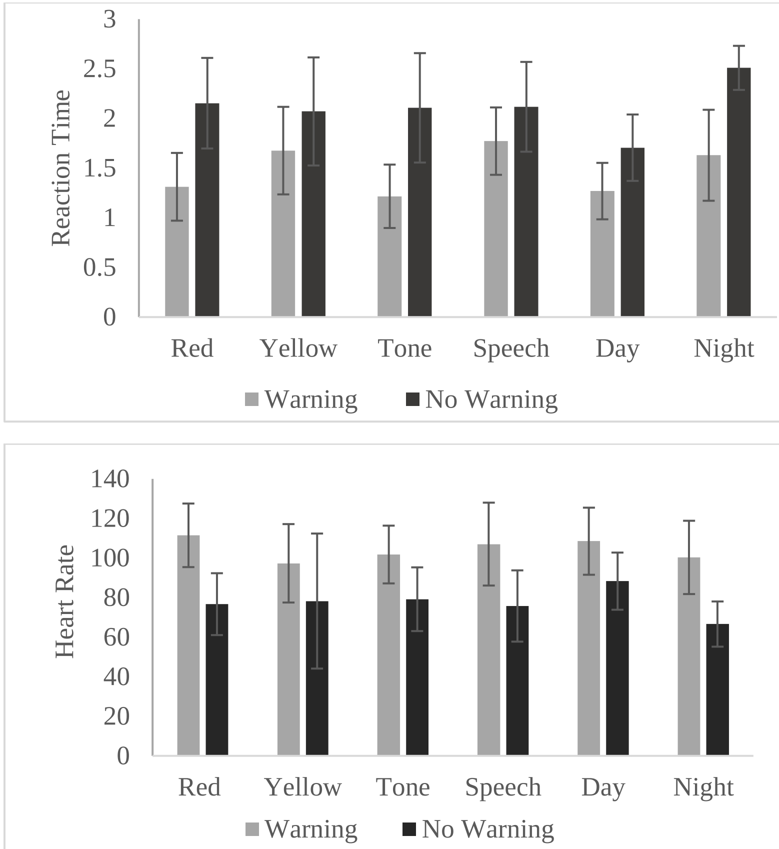


Figure 18: Reaction Time (Top), Heart Rate (Bottom) for visual, audible signals and time of the day with or without any warning

While speech warnings can be easily decoded, it usually takes longer to convey the message to the user and therefore, makes it difficult to comprehend the message in a critical situation (Freeman et al., 2017). This might be the reason why speech warnings also resulted in a higher heart rate. Some participants mentioned they didn't comprehend the speech warnings during their first trial. According to Simpson & Marchionda-Frost (1984), while speech warnings can be decoded though automatic cognitive process, it might not be fully comprehended until the message is complete which can delay the reaction time in critical situation. The reason why night had such an improvement in the presence of warning system could be due to increased confidence of cyclists. In previous studies researchers found poor visibility at night to be the main contributor for the crash between cyclists and motor vehicles (Lacherez et al., 2013; Wood et al., 2022). According to Johansson et al. (2009), cyclists at night are 55% more likely to get involved in a traffic crash in both urban and rural areas combined. In the human factors lab at UTA, the bike simulator is in a room where minimum light is allowed to go in so it can mimic the poor visibility at night. This might have been the reason why heart rate was observed to be high as well. After implementing the red visual warnings, despite the heart rate being higher, the reaction time reduced by 0.84 seconds.

3.4.3.2 Impact of cyclists' behavior and personal innovativeness

The effect of cyclists' personal behavior on their reaction time, and heart rate was assessed using the Cyclist Behavior Questionnaire (CBQ) obtained in study I. The subscales determined were- violation, aggressive violation, positive behavior and distraction and forgetfulness. Among the subscales, only positive behavior had a significant correlation with the difference in mean heart rate (Table 28). This indicates that participants who tend to exhibit more positive behavior on the road had a less increase in their heart rate compared to when they didn't have any warning signals present. This display of safe behavior and compliance with traffic laws could lead to cyclists become more cautious and less stressed. Increased heart rate has been associated in previous research with high level of stress (Lim et al., 2022, 2023). Since the participants were already cautious about their surroundings, having the warning system assisted them avoid the collision without causing any further stress.

The personal innovativeness (PI) questionnaire was developed by Agarwal & Prasad (1998) and is used to measure the willingness of people to try some new technology. PI only had a significant correlation with the mean difference in reaction time (Table 27). A negative correlation can be observed between the two variables indicating that participants who exhibited a more positive attitude towards trying new technology, had more improved reaction time. In previous studies, several research has found participants with higher PI score to have more openness towards accepting the new technology. Aharony (2013) found participants with higher PI score and extroversion could use Facebook better; in fact, Nov & Ye (2008), found a positive correlation between individuals' openness towards new technology and PI scores. In this study, the improved reaction time with a higher PI score is an indicator of participant's willingness to try the CycleGuard app.

3.4.3.3 Effect of the warnings on cyclists' emotional variables

Cyclists' emotional variables were assessed using the FaceReader software developed by Nodulus. The emotional variables that the software package can detect include neutral, happy, sad, arousal, surprised, scared, angry, disgusted, and valence. t-test results revealed that while the visual signals did not have any significant effect on the emotional variables, the audible signals had significant effect on arousal, surprised, and scared (Table 30). The mean arousal score for tone was higher than speech audible warning for all three variables. Arousal can be defined as the energy or activation that was required to begin the activity (Guasch, 2022). In a previous research, arousal has been considered as attention awareness that can lead to a timely response by increasing the situational awareness (J. Lee et al., 2019). Similarly, in this study, arousal is used to assess increased situational awareness. From the results it can be concluded that tone warning led to a higher arousal. This increase in arousal might also be the reason why heart rate was observed to be higher in the presence of the warning signals. Heart rate variability has been used to assess arousal in several research (Breuninger et al., 2017; Marín-Morales et al., 2021). When arousal level increases, the sympathetic activity in the autonomous nervous system (ANS) also increases which can lead to increased heart rate (Schaaff & Adam, 2013).

For the other two variables, surprised and scared, tone auditory warnings had a higher mean than the speech warning. According to Niepel et al. (1994), the purpose of surprise is to identify a discrepancy that is not related to their current input. This allows to have an accurate prediction of the event as well take control of the environment. Since tone had a higher surprise score, it can be deduced that this warning signal helped participants focus on the urgency of the warning better. Therefore, by surprising the participant, it might have helped them to be aware of the danger and take appropriate action. Lastly, tone warnings also had higher score for scared which could be due to the sudden and high pitch nature of the tone. Nevertheless, the mean intensity score for arousal was higher than surprised and scared combined indicating that a tone audible warning alerted the cyclist about the collision better than speech warning.

3.4.3.4 Limitation

There were several limitations in this study. The first one being the half fractional factorial design of the study. Due to the large number of participants required for full factorial design, fractional factorial study was selected which allowed to analyze up to 2-way interaction. Furthermore, in Human Factors research, participants are generally considered blocks; however, due to lack of resources a fractional factorial design was conducted. Moreover, the participants in this study were all UTA students who belong to the same age group. People from older/younger age population might not have the same effects. Lastly, although the simulator provides a safe environment to conduct research, it can never mimic how participants would behavior on an actual road. Nevertheless, the findings of this study can be used to proceed to the next step involving naturalistic study.

3.4.4 Conclusion

This study was conducted to find the warning system for the Bicyclist Assistance System-CycleGuard. The warning signals obtained from the first phase of the study- red or yellow visual signal, tone or speech audible signal were incorporated in this phase. Participants (*N*=32) were recruited for this study by offering participation as an extra credit for two engineering courses. The design for this experiment was a half fractional factorial model with a Resolution V confounding structure. This allowed the analysis of the 2-way interaction effects. Each participant received five trials and five controls for the study. Results from the ANOVA models revealed that red visual warning and tone audible warnings improved cyclists' performance by reducing the reaction time. The red visual signals also improved the performance at night and during a possibility of right hook collision. Additionally, heart rate was higher in the presence of the warning signals and at night, which could be due to the increased arousal, surprise, and scared intensity score. The findings for the emotional variable analysis were obtained from the FaceReader software. Lastly, the CBQ developed in the first study and the PI questionnaire revealed that participants who had higher PI score exhibited more willingness to try CycleGuard, thus had a reduced reaction time. The results revealed an interesting answer to the first research question (Question 2C). Red visuals had higher heart rate regardless of the speed, while tone audible warnings had a higher intensity for the emotional variables. Speech warnings on the other hand resulted in an increased heart rate instead. It also answered the second research question by improving cyclists' experience in terms of reduced reaction time, stable heart rate, and higher arousal leading to more situational awareness (Question 2D). Red visual warning and tone audible warning resulted in the highest improvement in cyclists, even at night. This study had limitations in terms of small sample size, same age group participants, and the inability of the simulator to mimic real world behavior. However, the effectiveness of red-tone multimodal warning can be concluded from the findings. Future studies include having naturistic study on the road. Different age group participants will need to be recruited in the study to observe different behavior.

CHAPTER IV: CONCLUSION OF THE DISSERTATION

Cycling has numerous benefits including physical, environmental, financial, etc. It causes no pollution while keeping the cyclist healthy. However, in recent years, cyclists have seen an increase in fatality despite having new infrastructures that are being added to aid the cyclists. This proves that there needs to be research on how to make cycling safer based on the behavior of the cyclists and then using technology to assist them while they are on the road.

To achieve this goal, this research was divided into two studies. In the first study, a Cyclsits' Behavior Questionnaire (CBQ) was developed for the US population. It was developed from several existing questionnaires that are validated all over the world. The survey was developed on QuestionPro. Participants were recruited from various cycling groups on Facebook. After collecting survey responses, Principal Component Analysis was conducted on the data to reduce it to an 11-item 4 factors questionnaire. It was then validated using Confirmatory Factor Analysis. To further validate it, a scenario-based bike simulation study was conducted in the Human Factors Lab at UTA. The simulation study validated the survey, while the survey-based behavior questions failed to validate it. This highlights one big limitation of using survey studies as validation method as participants often don't behave the same as they say they do on the survey. The results further showed some differences in terms of gender, age, and which state the participants were located. Future studies include validating the US CBQ in different states.

The second part of this research was to develop an ideal warning system for cyclists based on their physical and physiological responses to the warning, as well as their behavior and emotions. This study was conducted in two phases. In the first phase a group of experts in the area of transportation were invited. They were provided with a shortlist of warning signals that can be incorporated in the mobile app. The entire session was recorded, transcribed, and analyzed with Python's Natural Language Processing (NLP), and Microsoft Excel. The analysis showed that the experts preferred a combination of red or yellow visual signals, and high pitch tone or speech warnings. The experts were strongly against using haptic signals to alert cyclists. The limitation of this study included conducting a single session.

The second phase of the study included a bike simulator study by incorporating the warning signals identified in the first phase. A half fractional factorial design was designed for the factors- audio, visual, time of day, collision type, and speed. The response variables collected were reaction time, and heart rate. A total of 32 participants were recruited so each treatment combination can be repeated twice. ANOVA models were developed for each variable once the assumptions were satisfied. Results showed that tone audible warnings resulted in shorter reaction time, stable heart rate, and higher situational awareness. Red visual warnings also had a reduced reaction time. When the improvement was considered based on cyclists' behavior using the questionnaire developed in study I, it was observed that cyclists with highly positive behavior also had lower change in heart rate in the presence of the warning signals. The Personal Innovativeness scale indicated willingness to try CycleGuard by improving the reaction time. Lastly, the emotional variables indicated that the tone warnings had high arousal, surprised and scared intensity scores. Overall, based on the results, it can be concluded that a multimodal combination of red and tone warning would be ideal to incorporate into CycleGuard. The limitation of this study included small sample size, same age group, and the inability of the simulator to mimic real world experience. Future studies include a naturalistic study on the road.

REFERENCES

Abe, G., & Richardson, J. (2005). The influence of alarm timing on braking response and driver trust in low speed driving. *Safety Science*, *43*(9). https://doi.org/10.1016/j.ssci.2005.04.006

Agarwal, R., & Prasad, J. (1998). A Conceptual and Operational Definition of Personal Innovativeness in the Domain of Information Technology. *Information Systems Research*, *9*(2). https://doi.org/10.1287/isre.9.2.204

Aharony, N. (2013). Factors affecting the adoption of facebook by information professionals. *Proceedings of the ASIST Annual Meeting*, *50*(1). https://doi.org/10.1002/meet.14505001030

Alavi, M., Visentin, D. C., Thapa, D. K., Hunt, G. E., Watson, R., & Cleary, M. (2020). Chi-square for model fit in confirmatory factor analysis. In *Journal of Advanced Nursing* (Vol. 76, Issue 9). https://doi.org/10.1111/jan.14399

Anderson, J. C., & Gerbing, D. W. (1988). Structural Equation Modeling in Practice: A Review and Recommended Two-Step Approach. *Psychological Bulletin*, *103*(3). https://doi.org/10.1037/0033-2909.103.3.411

Atkinson, B. F. W., Bennett, T. O., Bahr, G. S., & Nelson, M. M. W. (2007). Development of a Multiple Heuristics Evaluation Table (MHET) to support software development and usability analysis. *Lecture Notes in Computer Science (Including Subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*, *4554 LNCS*(PART 1). https://doi.org/10.1007/978-3-540-73279-2_63

Baek, M., Jeong, D., Choi, D., & Lee, S. (2020). Vehicle trajectory prediction and collision warning via fusion of multisensors and wireless vehicular communications. *Sensors (Switzerland)*, *20*(1). https://doi.org/10.3390/s20010288

Beitel, D., Stipancic, J., Manaugh, K., & Miranda-Moreno, L. (2018). Assessing safety of shared space using cyclist-pedestrian interactions and automated video conflict analysis. *Transportation Research Part D: Transport and Environment*, *65*. https://doi.org/10.1016/j.trd.2018.10.001

Bella, F., & Silvestri, M. (2021). Vehicle–pedestrian interactions into and outside of crosswalks: Effects of driver assistance systems. *Transport*, *36*(2). https://doi.org/10.3846/transport.2021.14739

Bentler, P. M., & Wu, E. J. (2005). EQS 6.1 for Windows. Structural equations program manual. Encino, CA: Multivariate Software.

Bentler, P. M. (1990). Comparative fit indexes in structural models. *Psychological Bulletin*, *107*(2). https://doi.org/10.1037/0033-2909.107.2.238

Bernhard, A. (2020). The great bicycle boom of 2020. *BBC*. https://www.bbc.com/future/bespoke/made-on-earth/the-great-bicycle-boom-of-2020.html

Bernhoft, I. M., & Carstensen, G. (2008). Preferences and behaviour of pedestrians and cyclists by age and gender. *Transportation Research Part F: Traffic Psychology and Behaviour*, *11*(2). https://doi.org/10.1016/j.trf.2007.08.004

Bham, G. H., & Leu, M. C. (2018). A driving simulator study to analyze the effects of portable changeable message signs on mean speeds of drivers. *Journal of Transportation Safety and Security*, *10*(1–2). https://doi.org/10.1080/19439962.2017.1314398

Bhatia, P. (2003). Vehicle technologies to improve performance and safety. *University of California Transportation Center*.

Biondi, F., Strayer, D. L., Rossi, R., Gastaldi, M., & Mulatti, C. (2017). Advanced driver assistance systems: Using multimodal redundant warnings to enhance road safety. *Applied Ergonomics*. https://doi.org/10.1016/j.apergo.2016.06.016

Bliss, J. P. (1993). The Cry-Wolf Phenomenon and its Effect on Alarm Responses. In *Retrospective Theses and Dissertations*.

Bloor, M., Frankland, J., Thomas, M., & Robson, K. (2012). Focus Groups in Social Research. In *Focus Groups in Social Research*. https://doi.org/10.4135/9781849209175

Boele-Vos, M. J., van Duijvenvoorde, K., Doumen, M. J. A., Duivenvoorden, C. W. A. E., Louwerse, W. J. R., & Davidse, R. J. (2017). Crashes involving cyclists aged 50 and over in the Netherlands: An in-depth study. *Accident Analysis and Prevention, 105*. https://doi.org/10.1016/j.aap.2016.07.016

Breuninger, C., Sláma, D. M., Krämer, M., Schmitz, J., & Tuschen-Caffier, B. (2017). Psychophysiological Reactivity, Interoception and Emotion Regulation in Patients with Agoraphobia During Virtual Reality Anxiety Induction. *Cognitive Therapy and Research*, *41*(2). https://doi.org/10.1007/s10608-016-9814-9

Brookhuis, K. A., van Driel, C. J. G., Hof, T., van Arem, B., & Hoedemaeker, M. (2009). Driving with a congestion assistant; mental workload and acceptance. *Applied Ergonomics*, *40*(6). https://doi.org/10.1016/j.apergo.2008.06.010

Brown, T. (2015). Confirmatory Factor Analysis for Applied Research, Second Edition. In *Guilford Publications*.

Brown, T. A., & Moore, M. T. (2012). Confirmatory factor analysis. *Handbook of Structural Equation Modeling*, *361*, 379.

Bujang, M. A., Omar, E. D., & Baharum, N. A. (2018). A review on sample size determination for cronbach's alpha test: A simple guide for researchers. *Malaysian Journal of Medical Sciences*, *25*(6). https://doi.org/10.21315/mjms2018.25.6.9

Bulat, A., Tzimiropoulos, G., & Kingdom, U. (2017). How far are we from solving the 2D & 3D Face Alignment problem? *Iccv*. https://doi.org/10.1109/ICCV.2017.116

Burrows, M. (2019). *May 17 is National Bike to Work Day*. MAY 14, 2019. https://www.census.gov/library/stories/2019/05/younger-workers-in-cities-more-likely-to-bike-to-work.html

Bustamante, E. A., Bliss, J. P., & Anderson, B. L. (2007). Effects of varying the threshold of alarm systems and workload on human performance. *Ergonomics*, *50*(7). https://doi.org/10.1080/00140130701237345

Carter, D. L., Hunter, W. W., Zegeer, C. v., Stewart, J. R., & Huang, H. (2007). Bicyclist intersection safety index. *Transportation Research Record*, *2031*. https://doi.org/10.3141/2031-03

Cassano-Piché, A. (2015). *Heuristic analysis*. HumanEra@ UHN & IFMBE Clinical Engineering Division.

Chang, C. Y., & Chang, T. W. (2011). The Development of Parameters and Warning Algorithms for an Intersection Bus-Pedestrian Collision Warning System. *International Journal of Information Systems in the Service Sector*, *3*(4). https://doi.org/10.4018/jisss.2011100105

Chapanis, A. (1994). Hazards associated with three signal words and four colours on warning signs. *Ergonomics*, *37*(2). https://doi.org/10.1080/00140139408963644

Char, F., Serre, T., Compigne, S., & Puente Guillen, P. (2022). Car-to-cyclist forward collision warning effectiveness evaluation: a parametric analysis on reconstructed real accident cases. *International Journal of Crashworthiness*, *27*(1). https://doi.org/10.1080/13588265.2020.1773740

Chen, P., & Shen, Q. (2016). Built environment effects on cyclist injury severity in automobile-involved bicycle crashes. *Accident Analysis and Prevention*. https://doi.org/10.1016/j.aap.2015.11.002

Clifton, K. J., & Handy, S. L. (2003). Qualitative methods in travel behaviour research. In *Transport survey quality and innovation*. Emerald Group Publishing Limited.

Cohen, J., Cohen, P., West, S. G., & Aiken, L. S. (2013). *Applied multiple regression/correlation analysis for the behavioral sciences*. Routledge.

Combs, K., Fendley, M., & Bihl, T. (2020). A Preliminary Look at Heuristic Analysis for Assessing Artificial Intelligence Explainability. *WSEAS TRANSACTIONS ON COMPUTER RESEARCH*, *8*. https://doi.org/10.37394/232018.2020.8.9

Coughlin, J., & Aarp, P. P. I. (2001). Transportation and older persons: perceptions and preferences: a report on focus groups. *AARP, Public Policy Institute, Washington, DC*.

Cristea, M., & Gheorghiu, A. (2016). Attitude, perceived behavioral control, and intention to adopt risky behaviors. *Transportation Research Part F: Traffic Psychology and Behaviour*, *43*. https://doi.org/10.1016/j.trf.2016.10.004

David Leonard, S. (1999). Does color of warnings affect risk perception? *International Journal of Industrial Ergonomics*, *23*(5–6). https://doi.org/10.1016/S0169-8141(98)00015-8

de Almeida Pacheco, B., Guimarães, M., Correa, A. G., & Farinazzo Martins, V. (2019). Usability evaluation of learning objects with augmented reality for smartphones: A reinterpretation of nielsen heuristics. *Communications in Computer and Information Science*, *847*. https://doi.org/10.1007/978-3-030-05270-6_16

De Angelis, M., Fraboni, F., Prati, G., Giusino, D., Depolo, M., Zani, B., Pietrantoni, L., Shires, J., & Johnson, D. (2018). Preferences of European cyclists towards passive and active systems with audio-visual and with handlebar vibration warnings. *Human Factors and Ergonomics Society Europe Chapter 2018 Annual Conference. Berlin, Germany*.

Deb, S., Carruth, D. W., & Hudson, C. R. (2020). How communicating features can help pedestrian safety in the presence of self-driving vehicles: Virtual reality experiment. *IEEE Transactions on Human-Machine Systems*. https://doi.org/10.1109/THMS.2019.2960517

Deb, S., Carruth, D. W., Sween, R., Strawderman, L., & Garrison, T. M. (2017). Efficacy of virtual reality in pedestrian safety research. *Applied Ergonomics*. https://doi.org/10.1016/j.apergo.2017.03.007

Deb, S., Strawderman, L., Carruth, D. W., DuBien, J., Smith, B., & Garrison, T. M. (2017). Development and validation of a questionnaire to assess pedestrian receptivity toward fully autonomous vehicles. *Transportation Research Part C: Emerging Technologies*, *84*. https://doi.org/10.1016/j.trc.2017.08.029

Deb, S., Strawderman, L., DuBien, J., Smith, B., Carruth, D. W., & Garrison, T. M. (2017). Evaluating pedestrian behavior at crosswalks: Validation of a pedestrian behavior questionnaire for the U.S. population. *Accident Analysis & Prevention*, *106*, 191–201. https://doi.org/10.1016/J.AAP.2017.05.020

Deb, S., Strawderman, L. J., & Carruth, D. W. (2018). Investigating pedestrian suggestions for external features on fully autonomous vehicles: A virtual reality experiment. *Transportation Research Part F: Traffic Psychology and Behaviour*. https://doi.org/10.1016/j.trf.2018.08.016

Delgado, D., Zamora, D., Quiñones, D., Rusu, C., Roncagliolo, S., & Rusu, V. (2020). User experience heuristics for national park websites. *Lecture Notes in Computer Science (Including Subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*, *12195 LNCS*. https://doi.org/10.1007/978-3-030-49576-3_14

Dey, D., Habibovic, A., Pfleging, B., Martens, M., & Terken, J. (2020). Color and Animation Preferences for a Light Band eHMI in Interactions between Automated Vehicles and

Pedestrians. *Conference on Human Factors in Computing Systems - Proceedings*. https://doi.org/10.1145/3313831.3376325

Dichabeng, P., Merat, N., & Markkula, G. (2021). Factors that influence the acceptance of future shared automated vehicles – A focus group study with United Kingdom drivers. *Transportation Research Part F: Traffic Psychology and Behaviour*, *82*. https://doi.org/10.1016/j.trf.2021.08.009

Dilorio, C., Hockenberry-Eaton, M., Maibach, E., & Rivero, T. (1994). Focus groups: an interview method for nursing research. In *The Journal of neuroscience nursing : journal of the American Association of Neuroscience Nurses* (Vol. 26, Issue 3). https://doi.org/10.1097/01376517-199406000-00013

Dimitrakopoulou, D. (2021). Designing Generative Dialogue Spaces to Enhance Focus Group Research: A Case Study in the Context of COVID-19 Vaccination. *International Journal of Qualitative Methods*, *20*. https://doi.org/10.1177/16094069211066704

DiStefano, C., Zhu, M., & Mîndrilă, D. (2009). Understanding and using factor scores: Considerations for the applied researcher. *Practical Assessment, Research and Evaluation*, *14*(20).

Dixon, S. R., Wickens, C. D., & McCarley, J. S. (2007). On the independence of compliance and reliance: Are automation false alarms worse than misses? *Human Factors*, *49*(4). https://doi.org/10.1518/001872007X215656

Doody, O., Slevin, E., & Taggart, L. (2013). Preparing for and conducting focus groups in nursing research: Part 2. *British Journal of Nursing*, *22*(3). https://doi.org/10.12968/bjon.2013.22.3.170

Druta, C., & Alden, A. S. (2020). Preventing Animal-Vehicle Crashes using a Smart Detection Technology and Warning System. *Transportation Research Record*, *2674*(10). https://doi.org/10.1177/0361198120936651

Dukic, T., Ahlstrom, C., Patten, C., Kettwich, C., & Kircher, K. (2013). Effects of Electronic Billboards on Driver Distraction. *Traffic Injury Prevention*, *14*(5). https://doi.org/10.1080/15389588.2012.731546

Erdei, E. H., Steinmann, J., & Hagemeister, C. (2020). Comparing perception of signals in different modalities during the cycling task: A field study. *Transportation Research Part F: Traffic Psychology and Behaviour*. https://doi.org/10.1016/j.trf.2020.06.011

Etminani-Ghasrodashti, R., Ketankumar Patel, R., Kermanshachi, S., Michael Rosenberger, J., Weinreich, D., & Foss, A. (2021). Integration of shared autonomous vehicles (SAVs) into existing transportation services: A focus group study. *Transportation Research Interdisciplinary Perspectives*, *12*. https://doi.org/10.1016/j.trip.2021.100481

European Cyclists' Federation. (2020). *COVID-19 Cycling Measures Tracker* . European Cyclists' Federation.

Faber, K., & van Lierop, D. (2020). How will older adults use automated vehicles? Assessing the role of AVs in overcoming perceived mobility barriers. *Transportation Research Part A: Policy and Practice*, *133*, 353–363. https://doi.org/10.1016/J.TRA.2020.01.022

Feenstra, H., Ruiter, R. A. C., Schepers, J., Peters, G. J., & Kok, G. (2011). Measuring risky adolescent cycling behaviour. *International Journal of Injury Control and Safety Promotion*, *18*(3). https://doi.org/10.1080/17457300.2010.540334

Feenstra, H., Ruiter, R. A., & Kok, G. (2010). Social-cognitive correlates of risky adolescent cycling behavior. *BMC Public Health*, *10*. https://doi.org/10.1186/1471-2458-10-408

Fernández-Heredia, Á., Monzón, A., & Jara-Díaz, S. (2014). Understanding cyclists' perceptions, keys for a successful bicycle promotion. *Transportation Research Part A: Policy and Practice*, *63*. https://doi.org/10.1016/j.tra.2014.02.013

Fernando, A. H. V., Muthuarachchi, M. D. C., Anandakumar, D. R., Chamalka, W. N. R. B., Gamage, M. P., & Amarasena, N. C. (2020). Motorcyclists Safety Assistant App. *11th Annual IEEE Information Technology, Electronics and Mobile Communication Conference, IEMCON 2020*. https://doi.org/10.1109/IEMCON51383.2020.9284940

Fiske, S. T., & Taylor, S. E. (1991). *Social cognition*. Mcgraw-Hill Book Company.

Fornell, C., & Larcker, D. F. (1981). Evaluating Structural Equation Models with Unobservable Variables and Measurement Error. Journal of Marketing Research, 18(1), 39. doi:10.2307/3151312. *Journal of Marketing Research*, *18*(1).

Freeman, E., Wilson, G., Vo, D.-B., Ng, A., Politis, I., & Brewster, S. (2017). Multimodal feedback in HCI: haptics, non-speech audio, and their applications. In *The Handbook of Multimodal-Multisensor Interfaces: Foundations, User Modeling, and Common Modality Combinations - Volume 1*. https://doi.org/10.1145/3015783.3015792

Gefen, D., Straub, D., & Boudreau, M.-C. (2000). Structural Equation Modeling and Regression: Guidelines for Research Practice. *Communications of the Association for Information Systems*, *4*. https://doi.org/10.17705/1cais.00407

Geitner, C., Biondi, F., Skrypchuk, L., Jennings, P., & Birrell, S. (2019). The comparison of auditory, tactile, and multimodal warnings for the effective communication of unexpected events during an automated driving scenario. *Transportation Research Part F: Traffic Psychology and Behaviour*. https://doi.org/10.1016/j.trf.2019.06.011

Graham, R. (1999). Use of auditory icons as emergency warnings: Evaluation within a vehicle collision avoidance application. *Ergonomics*, *42*(9). https://doi.org/10.1080/001401399185108

Guasch, M. (2022). *Analysing facial expressions of basic emotions with FaceReader: An experimental study*.

Gudi, A., Tasli, H. E., Den Uyl, T. M., & Maroulis, A. (2015). Deep learning based facs action unit occurrence and intensity estimation. *2015 11th IEEE International Conference and*

Workshops on Automatic Face and Gesture Recognition, FG 2015, *2015-January*. https://doi.org/10.1109/FG.2015.7284873

Gupta, N., Bisantz, A. M., & Singh, T. (2002). The effects of adverse condition warning system characteristics on driver performance: An investigation of alarm signal type and threshold level. *Behaviour and Information Technology*, *21*(4). https://doi.org/10.1080/0144929021000013473

Hang, J., Yan, X., Li, X., Duan, K., Yang, J., & Xue, Q. (2022). An improved automated braking system for rear-end collisions: A study based on a driving simulator experiment. *Journal of Safety Research*, *80*. https://doi.org/10.1016/j.jsr.2021.12.023

Hastuti, R., Heng, P. H., & Zein, A. N. (n.d.). *Safety Cycling During Pandemic: A Descriptive Study of Cycling Behavior*.

Hauer, E. (2020). Crash causation and prevention. *Accident Analysis and Prevention*, *143*. https://doi.org/10.1016/j.aap.2020.105528

Henseler, J., Ringle, C. M., & Sarstedt, M. (2015). A new criterion for assessing discriminant validity in variance-based structural equation modeling. *Journal of the Academy of Marketing Science, 43*(1). https://doi.org/10.1007/s11747-014-0403-8

Hezaveh, A. M., Nordfjærn, T., Mamdoohi, A. R., & Şimşekoğlu, Ö. (2018). Predictors of self-reported crashes among Iranian drivers: Exploratory analysis of an extended driver behaviour questionnaire. *Promet - Traffic - Traffico*, *30*(1). https://doi.org/10.7307/ptt.v30i1.2412

Hezaveh, A. M., Zavareh, M. F., Cherry, C. R., & Nordfjærn, T. (2018). Errors and violations in relation to bicyclists' crash risks: Development of the Bicycle Rider Behavior Questionnaire (BRBQ). *Journal of Transport and Health*, *8*. https://doi.org/10.1016/j.jth.2017.11.003

Ho, C., Reed, N., & Spence, C. (2007). Multisensory in-car warning signals for collision avoidance. *Human Factors, 49*(6). https://doi.org/10.1518/001872007X249965

Hollifield, B. R., & Habibi, E. (2007). *Alarm management: seven effective methods for optimum performance*. Isa.

Hu, L. T., & Bentler, P. M. (1998). Fit Indices in Covariance Structure Modeling: Sensitivity to Underparameterized Model Misspecification. *Psychological Methods, 3*(4). https://doi.org/10.1037/1082-989X.3.4.424

IIHS. (2021). *Fatality Facts 2018 Bicyclists*. 501(c)(3) Organizations. https://www.iihs.org/topics/fatality-statistics/detail/bicyclists#when-they-died

IIHS. (2023). *Fatality Facts 2021 Bicyclists*. https://www.iihs.org/topics/fatality-statistics/detail/bicyclists#where-they-died

Jägerbrand, A. K., & Sjöbergh, J. (2016). Effects of weather conditions, light conditions, and road lighting on vehicle speed. *SpringerPlus*, *5*(1). https://doi.org/10.1186/s40064-016-2124-6

Jin, W., Murali, S., Cho, Y., Zhu, H., Li, T., Panik, R. T., Rimu, A., Deb, S., Watkins, K., Yuan, X., & Li, M. (2021). CycleGuard: A Smartphone-based Assistive Tool for Cyclist Safety Using Acoustic Ranging. *Proceedings of the ACM on Interactive, Mobile, Wearable and Ubiquitous Technologies*, *5*(4). https://doi.org/10.1145/3494992

Johansson, Ö., Wanvik, P. O., & Elvik, R. (2009). A new method for assessing the risk of accident associated with darkness. *Accident Analysis & Prevention*, *41*(4), 809–815.

John, O. P., & Robins, R. W. (1994). Accuracy and Bias in Self-Perception: Individual Differences in Self-Enhancement and the Role of Narcissism. *Journal of Personality and Social Psychology*, *66*(1). https://doi.org/10.1037/0022-3514.66.1.206

Jollife, I. T., & Cadima, J. (2016). Principal component analysis: A review and recent developments. In *Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences* (Vol. 374, Issue 2065). https://doi.org/10.1098/rsta.2015.0202

Kaplan, S., Luria, R., & Prato, C. G. (2019). The relation between cyclists' perceptions of drivers, self-concepts and their willingness to cycle in mixed traffic. *Transportation Research Part F: Traffic Psychology and Behaviour*, *62*. https://doi.org/10.1016/j.trf.2018.12.011

Kennedy, R. S., Lane, N. E., Berbaum, K. S., & Lilienthal, M. G. (1993). Simulator Sickness Questionnaire: An Enhanced Method for Quantifying Simulator Sickness. *The International Journal of Aviation Psychology*, *3*(3). https://doi.org/10.1207/s15327108ijap0303_3

Kenny, D. a. (2012). Measuring Model Fit. *Davidakenny.Net*, *1985*.

KHUAN, O. B. (2013). STUDY OF FSKKP PORTAL BASED ON SHNEIDERMAN"S EIGHT GOLDEN RULES. *Jurnal Teknologi*, *1*(1).

Kim, W. S., Kim, Y. J., Kim, Y. S., Baek, S. Y., Baek, S. M., Lee, D. H., Nam, K. C., Kim, T. B., & Lee, H. J. (2020). Development of control system for automated manual transmission of 45-kW agricultural tractor. *Applied Sciences (Switzerland)*, *10*(8). https://doi.org/10.3390/APP10082930

Kingsley, C., Thiry, E., Flowers, A., & Jenkins, M. (2020). Augmented Riding: Multimodal Applications of AR, VR, and MR to Enhance Safety for Motorcyclists and Bicyclists. *Lecture Notes in Computer Science (Including Subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*. https://doi.org/10.1007/978-3-030-59990-4_27

Kovácsová, N., de Winter, J. C. F., & Hagenzieker, M. P. (2019). What will the car driver do? A video-based questionnaire study on cyclists' anticipation during safety-critical situations. *Journal of Safety Research*, *69*. https://doi.org/10.1016/j.jsr.2019.01.002

Kraus, S., & Koch, N. (2021). Provisional COVID-19 infrastructure induces large, rapid increases in cycling. *Proceedings of the National Academy of Sciences of the United States of America*, *118*(15). https://doi.org/10.1073/pnas.2024399118

Kreißig, I., Springer, S., & Krems, J. F. (2022, June). Cycling Safety at the Fore–Evaluation of a Cyclist Warning System from a User Perspective. In International Conference on Human-Computer Interaction (pp. 223-230). Cham: Springer International Publishing.

Lacherez, P., Wood, J. M., Marszalek, R. P., & King, M. J. (2013). Visibility-related characteristics of crashes involving bicyclists and motor vehicles - Responses from an online questionnaire study. *Transportation Research Part F: Traffic Psychology and Behaviour*, *20*. https://doi.org/10.1016/j.trf.2013.04.003

Lakens, D., & Caldwell, A. R. (2021). Simulation-Based Power Analysis for Factorial Analysis of Variance Designs. *Advances in Methods and Practices in Psychological Science*, *4*(1). https://doi.org/10.1177/2515245920951503

Larsson, P., Dekker, S. W. A., & Tingvall, C. (2010). The need for a systems theory approach to road safety. *Safety Science*, *48*(9). https://doi.org/10.1016/j.ssci.2009.10.006

Law, P. (2015). *FIXING AMERICA ' S SURFACE An Act*. 1–491.

Lee, J. D., McGehee, D. v., Brown, T. L., & Reyes, M. L. (2002). Collision warning timing, driver distraction, and driver response to imminent rear-end collisions in a high-fidelity driving simulator. *Human Factors*, *44*(2). https://doi.org/10.1518/0018720024497844

Lee, J., Hirano, T., Hano, T., & Itoh, M. (2019). Conversation during Partially Automated Driving: How Attention Arousal is Effective on Maintaining Situation Awareness 2019. *2019 IEEE International Conference on Systems, Man and Cybernetics (SMC), Bari, Italy*, 3718–3723. https://doi.org/0.1109/SMC.2019.8914632

Lee, J., & Marlowe, J. (2003). How consumers choose a financial institution: Decision-making criteria and heuristics. *International Journal of Bank Marketing*, *21*(2). https://doi.org/10.1108/02652320310461447

Li, X., Oviedo-Trespalacios, O., Afghari, A. P., Kaye, S. A., & Yan, X. (2023). Yield or not to yield? An inquiry into drivers' behaviour when a fully automated vehicle indicates a lane-changing intention. Transportation research part F: traffic psychology and behaviour, 95, 405-417

Li, X., Oviedo-Trespalacios, O., Rakotonirainy, A., & Yan, X. (2019). Collision risk management of cognitively distracted drivers in a car-following situation. *Transportation Research Part F: Traffic Psychology and Behaviour*, *60*. https://doi.org/10.1016/j.trf.2018.10.011

Li, X., Useche, S. A., Zhang, Y., Wang, Y., Oviedo-Trespalacios, O., & Haworth, N. (2022). Comparing the cycling behaviours of Australian, Chinese and Colombian cyclists using a behavioural questionnaire paradigm. *Accident Analysis and Prevention*, *164*. https://doi.org/10.1016/j.aap.2021.106471

Lievens, F., & Sackett, P. R. (2006). Video-based versus written situational judgment tests: A comparison in terms of predictive validity. *Journal of Applied Psychology*, *91*(5). https://doi.org/10.1037/0021-9010.91.5.1181

Lim, T., Kalra, A., Thompson, J., Caldwell Odgers, J., & Beck, B. (2022). Physiological measures of bicyclists' subjective experiences: A scoping review. *Transportation Research Part F: Traffic Psychology and Behaviour*, *90*. https://doi.org/10.1016/j.trf.2022.09.007

Lim, T., Thompson, J., Tian, L., & Beck, B. (2023). A transactional model of stress and coping applied to cyclist subjective experiences. *Transportation Research Part F: Traffic Psychology and Behaviour*, *96*, 155–170. https://doi.org/10.1016/J.TRF.2023.05.013

Loijens, L., & Krips, O. (2021). *FaceReader Methodology Note*.

Lubin, A., & Feeley, C. (2016). Transportation Issues of Adults on the Autism Spectrum. *Transportation Research Record: Journal of the Transportation Research Board*, *2542*(1). https://doi.org/10.3141/2542-01

Ludwig, S., Schuelper, N., Brown, J., Anders, S., & Raupach, T. (2018). How can we teach medical students to choose wisely? A randomised controlled cross-over study of video-versus text-based case scenarios. *BMC Medicine*, *16*(1). https://doi.org/10.1186/s12916-018-1090-y

Luquetti dos Santos, I. J. A., dos Santos Grecco, C. H., Abreu Mol, A. C., & Rodrigues Carvalho, P. V. (2009). The use of questionnaire and virtual reality in the verification of the human factors issues in the design of nuclear control desk. *International Journal of Industrial Ergonomics*. https://doi.org/10.1016/j.ergon.2008.08.005

Machín, L., Antúnez, L., Curutchet, M. R., & Ares, G. (2020). The heuristics that guide healthiness perception of ultra-processed foods: A qualitative exploration. *Public Health Nutrition*, *23*(16). https://doi.org/10.1017/S1368980020003158

Marín Puchades, V., Fassina, F., Fraboni, F., de Angelis, M., Prati, G., de Waard, D., & Pietrantoni, L. (2018). The role of perceived competence and risk perception in cycling near misses. *Safety Science*, *105*. https://doi.org/10.1016/j.ssci.2018.02.013

Marín-Morales, J., Higuera-Trujillo, J. L., Guixeres, J., Llinares, C., Alcañiz, M., & Valenza, G. (2021). Heart rate variability analysis for the assessment of immersive emotional arousal using virtual reality: Comparing real and virtual scenarios. *PLoS ONE*, *16*(7 July). https://doi.org/10.1371/journal.pone.0254098

Markov, A. S., Fadin, A. A., & Tsirlov, V. L. (2016). Multilevel metamodel for heuristic search of vulnerabilities in the software source code. *International Journal of Control Theory and Applications*, *9*(30).

Martínez-Ruiz, V., Jiménez-Mejías, E., Luna-Del-Castillo, J. D. D., García-Martín, M., Jiménez-Moleón, J. J., & Lardelli-Claret, P. (2014). Association of cyclists' age and sex with risk of involvement in a crash before and after adjustment for cycling exposure. *Accident Analysis and Prevention*, *62*. https://doi.org/10.1016/j.aap.2013.10.011

Matviienko, A., Ananthanarayan, S., Borojeni, S. S., Feld, Y., Heuten, W., & Boll, S. (2018). Augmenting bicycles and helmets with multimodal warnings for children. *MobileHCI 2018 - Beyond Mobile: The Next 20 Years - 20th International Conference on Human-Computer Interaction with Mobile Devices and Services, Conference Proceedings*. https://doi.org/10.1145/3229434.3229479

McDonald, J. D. (2008). Measuring personality constructs: The advantages and disadvantages of self-reports, informant reports and behavioural assessments. *Enquire*, *1*(1).

McKeown, D., & Isherwood, S. (2007). Mapping candidate within-vehicle auditory displays to their referents. *Human Factors*, *49*(3). https://doi.org/10.1518/001872007X200067

Mohebbi, R., Gray, R., & Tan, H. Z. (2009). Driver reaction time to tactile and auditory rear-end collision warnings while talking on a cell phone. *Human Factors*, *51*(1). https://doi.org/10.1177/0018720809333517

Møller, M., & Haustein, S. (2017). Anger expression among Danish cyclists and drivers: A comparison based on mode specific anger expression inventories. *Accident Analysis and Prevention*, *108*. https://doi.org/10.1016/j.aap.2017.09.016

Morgan, D. L., & Krueger, R. A. (2014). When to Use Focus Groups and Why. In *Successful Focus Groups: Advancing the State of the Art*. https://doi.org/10.4135/9781483349008.n1

Morgan, D. L., Krueger, R. A., & King, J. A. (1998). The focus group kit, Vols. 1–6. In *The focus group kit, Vols. 1–6.*

Murata, A., Kanbayashi, M., & Hayami, T. (2013). Effectiveness of automotive warning system presented with multiple sensory modalities. *Lecture Notes in Computer Science (Including Subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*. https://doi.org/10.1007/978-3-642-39173-6_11

Napitupulu, D., Abdel Kadar, J., & Kartika Jati, R. (2017). Validity testing of technology acceptance model based on factor analysis approach. *Indonesian Journal of Electrical Engineering and Computer Science*, *5*(3). https://doi.org/10.11591/ijeecs.v5.i3.pp697-704

National Bike Month: May 2022. (2022, May). https://www.census.gov/newsroom/stories/bike-month.htm

Naujoks, F., Kiesel, A., & Neukum, A. (2016). Cooperative warning systems: The impact of false and unnecessary alarms on drivers' compliance. *Accident Analysis and Prevention*, *97*. https://doi.org/10.1016/j.aap.2016.09.009

Neurauter, M. L. (2005). Multimodal warnings: Curve-warning design. *Proceedings of the Human Factors and Ergonomics Society*, *1999*, 1945–1949. https://doi.org/10.1177/154193120504902213

Newsholme, G. (2007). Too much data, too little information. Some thoughts on the explosion at Milford Haven in 1994. *Loss Prevention Bulletin*, *193*.

Ngwu, O. L., Rimu, A., Deb, S., Lummus, J. R., Kan, C., & Zhou, Y. (2022a). Teenage bicyclists' perceptions toward automated vehicles and their associated traffic infrastructure: A focus group discussion. *Transportation Research Part F: Traffic Psychology and Behaviour*, *89*, 371–384. https://doi.org/10.1016/J.TRF.2022.07.006

Ngwu, O. L., Rimu, A., Deb, S., Lummus, J. R., Kan, C., & Zhou, Y. (2022b). Teenage bicyclists' perceptions toward automated vehicles and their associated traffic infrastructure: A focus group discussion. *Transportation Research Part F: Traffic Psychology and Behaviour*, *89*, 371–384. https://doi.org/10.1016/J.TRF.2022.07.006

NHTSA. (2019). *Bicyclist and Pedestrian Safety*. National Highway Traffic Safety Administration. https://www.nhtsa.gov/sites/nhtsa.dot.gov/files/documents/14046-pedestrian_bicyclist_safety_resources_030519_v2_tag.pdf.

NHTSA. (2020). *Traffic Safety Facts*. https://crashstats.nhtsa.dot.gov/Api/Public/ViewPublication/812850

NHTSA. (2022). *Traffic Safety Facts - 2020 Data: Bicyclists and Other Cyclists*.

Nielsen, J. (1994). Enhancing the explanatory power of usability heuristics. *Conference on Human Factors in Computing Systems - Proceedings*. https://doi.org/10.1145/191666.191729

Nielsen, J. (1995). Nielsen Norman Group 10 Usability Heuristics for User Interface Design Author. *Designing Web Interfaces*.

Niepel, M., Rudolph, U., Schiltzwohl, A., & Meyer, W. U. (1994). Temporal Characteristics of the Surprise Reaction Induced by Schema-discrepant Visual and Auditory Events. *Cognition and Emotion*, *8*(5). https://doi.org/10.1080/02699939408408951

Nov, O., & Ye, C. (2008). Personality and technology acceptance: Personal innovativeness in IT, openness and resistance to change. *Proceedings of the Annual Hawaii International Conference on System Sciences*. https://doi.org/10.1109/HICSS.2008.348

NSC. (2021). *Bicycle Deaths*. https://injuryfacts.nsc.org/home-and-community/safety-topics/bicycle-deaths/

Oh, C., Oh, J. S., & Ritchie, S. G. (2005). Real-time hazardous traffic condition warning system: Framework and evaluation. *IEEE Transactions on Intelligent Transportation Systems*, *6*(3). https://doi.org/10.1109/TITS.2005.853693

O'hern, S., Estgfaeller, N., Stephens, A. N., & Useche, S. A. (2021). Bicycle rider behavior and crash involvement in australia. *International Journal of Environmental Research and Public Health*, *18*(5). https://doi.org/10.3390/ijerph18052378

O'Hern, S., Stephens, A. N., Young, K. L., & Koppel, S. (2020). Personality traits as predictors of cyclist behaviour. *Accident Analysis and Prevention*, *145*. https://doi.org/10.1016/j.aap.2020.105704

Oikawa, S., Hirose, T., Aomura, S., & Matsui, Y. (2016). Traffic Accidents Involving Cyclists Identifying Causal Factors Using Questionnaire Survey, Traffic Accident Data, and Real-World Observation. *Stapp Car Crash Journal*, *60*.

Ozili, P. K. (2023). The acceptable R-square in empirical modelling for social science research. In Social research methodology and publishing results: A guide to non-native english speakers (pp. 134-143). IGI Global.

Paulhus, D. L. (1991). Measurement and Control of Response Bias. In *Measures of Personality and Social Psychological Attitudes*. https://doi.org/10.1016/b978-0-12-590241-0.50006-x

Plummer-D'Amato, P. (2008). Focus group methodology Part 1: Considerations for design. *International Journal of Therapy and Rehabilitation*, *15*(2). https://doi.org/10.12968/ijtr.2008.15.2.28189

Politis, I., Brewster, S., & Pollick, F. (2013). Evaluating multimodal driver displays of varying urgency. *Proceedings of the 5th International Conference on Automotive User Interfaces and Interactive Vehicular Applications, AutomotiveUI 2013*. https://doi.org/10.1145/2516540.2516543

Pravossoudovitch, K., Cury, F., Young, S. G., & Elliot, A. J. (2014). Is red the colour of danger? Testing an implicit red-danger association. *Ergonomics*, *57*(4). https://doi.org/10.1080/00140139.2014.889220

Pucher, J., Komanoff, C., & Schimek, P. (1999). Bicycling renaissance in North America?: Recent trends and alternative policies to promote bicycling. *Transportation Research Part A: Policy and Practice*, *33*(7–8), 625–654. https://doi.org/10.1016/S0965-8564(99)00010-5

Qi, W., Shen, B., & Yu, Y. (2019). Assessment model on cyclists' road behaviors in University of China: A case of simplified cycling behavior questionnaire. *CICTP 2019: Transportation in China - Connecting the World - Proceedings of the 19th COTA International Conference of Transportation Professionals*. https://doi.org/10.1061/9780784482292.029

Rahman, T., de Santiago, B., Rimu, A., Deb, S., Pande, A., & Islam, M. (2022). A Virtual Reality Program to Improve Child Pedestrians' Safety at Street-Crossing Scenarios. *Human Factors and Systems Interaction*, *52*. https://doi.org/10.54941/ahfe1002150

Rattray, J., & Jones, M. C. (2007). Essential elements of questionnaire design and development. In *Journal of Clinical Nursing* (Vol. 16, Issue 2). https://doi.org/10.1111/j.1365-2702.2006.01573.x

Rubie, E., Haworth, N., Twisk, D., & Yamamoto, N. (2020). Influences on lateral passing distance when motor vehicles overtake bicycles: a systematic literature review. *Transport Reviews*, *40*(6). https://doi.org/10.1080/01441647.2020.1768174

Šabić, E., Chen, J., & MacDonald, J. A. (2021). Toward a Better Understanding of In-Vehicle Auditory Warnings and Background Noise. *Human Factors*, *63*(2). https://doi.org/10.1177/0018720819879311

Sayed, T., Zaki, M. H., & Autey, J. (2013). Automated safety diagnosis of vehicle-bicycle interactions using computer vision analysis. *Safety Science*, *59*. https://doi.org/10.1016/j.ssci.2013.05.009

Schaaff, K., & Adam, M. T. P. (2013). Measuring emotional arousal for online applications: Evaluation of ultra-short term heart rate variability measures. *Proceedings - 2013 Humaine Association Conference on Affective Computing and Intelligent Interaction, ACII 2013*. https://doi.org/10.1109/ACII.2013.66

Schumacker, R. E., & Lomax, R. G. (2004). A Beginner's Guide to Structural Equation Modeling. In *A Beginner's Guide to Structural Equation Modeling*. https://doi.org/10.4324/9781410610904

Senders, J. W., Kristofferson, a. B., Levison, W. H., Dietrich, C. W., & Ward, J. L. (1967). The attentional demand of automobie driving. In *Transportation Research Board*.

Shinar, D., & Gurion, B. (2019). Crash causes, countermeasures, and safety policy implications. *Accident Analysis and Prevention*, *125*. https://doi.org/10.1016/j.aap.2019.02.015

Simons, D., Clarys, P., de Bourdeaudhuij, I., de Geus, B., Vandelanotte, C., & Deforche, B. (2014). Why do young adults choose different transport modes? A focus group study. *Transport Policy*, *36*. https://doi.org/10.1016/j.tranpol.2014.08.009

Simpson, C. A., & Marchionda-Frost, K. (1984). Synthesized speech rate and pitch effects on intelligibility of warning messages for pilots. *Human Factors*, *26*(5). https://doi.org/10.1177/001872088402600503

Sivaji, A., Abdullah, A., & Downe, A. G. (2011). Usability testing methodology: Effectiveness of heuristic evaluation in E-government website development. *Proceedings - AMS 2011: Asia Modelling Symposium 2011 - 5th Asia International Conference on Mathematical Modelling and Computer Simulation*. https://doi.org/10.1109/AMS.2011.24

Smith, H., Howard, C., & Foord, T. (2003). Alarms management. Priority, floods, tears or gain? Introduction to the "problem." *Measurement and Control*, *36*(4). https://doi.org/10.1177/002029400303600402

Sorkin, R. D., Kantowitz, B. H., & Kantowitz, S. C. (1988). Likelihood alarm displays. *Human Factors*, *30*(4). https://doi.org/10.1177/001872088803000406

Stanton, N. A., Salmon, P. M., Rafferty, L. A., Walker, G. H., Baber, C., & Jenkins, D. P. (2013). Human factors methods: A practical guide for engineering and design, 2nd edition. In *Human Factors Methods: A Practical Guide for Engineering and Design, 2nd Edition*. https://doi.org/10.1080/00140139.2014.948659

Stephens, A. N., O'Hern, S., Young, K. L., Chambers, R., Hassed, C., & Koppel, S. (2020). Self-reported mindfulness, cyclist anger and aggression. *Accident Analysis and Prevention*, *144*. https://doi.org/10.1016/j.aap.2020.105625

Stewart, D. W., & Shamdasani, P. N. (2014). *Focus groups: Theory and practice* (Vol. 20). Sage publications.

Stewart, T. (2023a). *Overview of Motor Vehicle Traffic Crashes in 2021*. https://crashstats.nhtsa.dot.gov.

Stewart, T. (2023b). *Overview of Motor Vehicle Traffic Crashes in 2021*.

Stothart, C., Mitchum, A., & Yehnert, C. (2015). The attentional cost of receiving a cell phone notification. *Journal of Experimental Psychology: Human Perception and Performance*, *41*(4). https://doi.org/10.1037/xhp0000100

Straub, D., & Gefen, D. (2004). Validation Guidelines for IS Positivist Research. *Communications of the Association for Information Systems*, *13*. https://doi.org/10.17705/1cais.01324

Strohaeker, E. H., Moia, A., Steinmann, J., & Hagemeister, C. (2022). How do warnings influence cyclists' reaction to conflicts? Comparing acoustic and vibro-tactile warnings in different conflicts on a test track. *Transportation Research Part F: Traffic Psychology and Behaviour*, *90*, 151–166.

Sun, M., Ren, X., & Cao, X. (2010). Effects of multimodal error feedback on human performance in steering tasks. *Journal of Information Processing*, *18*. https://doi.org/10.2197/ipsjjip.18.284

Tabachnick, B. G., & Fidell, L. S. (2007). Using multivariate statistics (5th ed.). Boston, MA: Allyn & Bacon/Pearson Education. *Using Multivariate Statistics*.

Tavakol, M., & Dennick, R. (2011). Making sense of Cronbach's alpha. In *International journal of medical education* (Vol. 2). https://doi.org/10.5116/ijme.4dfb.8dfd

Taylor, D., & Davis, W. J. (1999). Review of basic research in bicycle traffic science, traffic operations, and facility design. *Transportation Research Record*, *1674*. https://doi.org/10.3141/1674-14

The League of American Bicyclists. (1880). *STATE BIKE LAWS*. https://bikeleague.org/bike-laws/state-bike-laws/?bikelaw-by-state=florida

Then, K. L., Rankin, J. A., & Ali, E. (2014). Focus group research: what is it and how can it be used? *Canadian Journal of Cardiovascular Nursing = Journal Canadien En Soins Infirmiers Cardio-Vasculaires*, *24*(1).

Thomas, J. R., Martin, P., Etnier, J., & Silverman, S. J. (2022a). *Research methods in physical activity*. Human kinetics.

Thomas, J. R., Martin, P., Etnier, J., & Silverman, S. J. (2022b). *Research methods in physical activity*. Human kinetics.

Tingvall, C., & Haworth, N. (1999). Vision Zero - An ethical approach to safety and mobility. *6th ITE International Conference Road Safety & Traffic Enforcement: Beyond 2000, Melbourne, 6-7 September 1999*.

Useche, S. A., Alonso, F., Boyko, A., Buyvol, P., Castañeda, I., Cendales, B., Cervantes, A., Echiburu, T., Faus, M., Feitosa, Z., Gene, J., Gonzalez-Marin, A., Gonzalez, V., Gnap, J., Ibrahim, M. K., Janstrup, K. H., Javadinejad, A., Makarova, I., McIlroy, R., … Montoro, L. (2022). Cross-culturally approaching the cycling behaviour questionnaire (CBQ): Evidence from 19 countries. *Transportation Research Part F: Traffic Psychology and Behaviour*, *91*. https://doi.org/10.1016/j.trf.2022.10.025

Useche, S. A., Alonso, F., Montoro, L., & Esteban, C. (2018). Distraction of cyclists: How does it influence their risky behaviors and traffic crashes? *PeerJ*, *2018*(9). https://doi.org/10.7717/peerj.5616

Useche, S. A., Montoro, L., Alonso, F., & Tortosa, F. M. (2018). Does gender really matter? A structural equation model to explain risky and positive cycling behaviors. *Accident Analysis & Prevention*, *118*, 86–95. https://doi.org/10.1016/J.AAP.2018.05.022

Useche, S. A., Montoro, L., Tomas, J. M., & Cendales, B. (2018). Validation of the Cycling Behavior Questionnaire: A tool for measuring cyclists’ road behaviors. *Transportation Research Part F: Traffic Psychology and Behaviour*, *58*. https://doi.org/10.1016/j.trf.2018.08.003

Useche, S. A., Philippot, P., Ampe, T., Llamazares, J., & de Geus, B. (2021). “Pédaler en toute sécurité”: The Cycling Behavior Questionnaire (CBQ) in Belgium – A validation study. *Transportation Research Part F: Traffic Psychology and Behaviour*, *80*. https://doi.org/10.1016/j.trf.2021.04.014

Useche, S., Montoro, L., Alonso, F., & Oviedo-Trespalacios, O. (2018). Infrastructural and human factors affecting safety outcomes of cyclists. *Sustainability (Switzerland)*. https://doi.org/10.3390/su10020299

Utriainen, R. (2020). Characteristics of commuters’ single-bicycle crashes in insurance data. *Safety*. https://doi.org/10.3390/safety6010013 van der Schalk, J., Hawk, S. T., Fischer, A. H., & Doosje, B. (2011). Moving Faces, Looking Places: Validation of the Amsterdam Dynamic Facial Expression Set (ADFES). *Emotion*, *11*(4). https://doi.org/10.1037/a0023853

van Schijndel-Speet, M., Evenhuis, H. M., van Wijck, R., van Empelen, P., & Echteld, M. A. (2014). Facilitators and barriers to physical activity as perceived by older adults with intellectual disability. *Intellectual and Developmental Disabilities*, *52*(3). https://doi.org/10.1352/1934-9556-52.3.175

von Sawitzky, T., Grauschopf, T., & Riener, A. (2022). "Attention! A Door Could Open."—Introducing Awareness Messages for Cyclists to Safely Evade Potential Hazards. *Multimodal Technologies and Interaction*, *6*(1). https://doi.org/10.3390/mti6010003

von Stülpnagel, R., & Binnig, N. (2022). How safe do you feel? – A large-scale survey concerning the subjective safety associated with different kinds of cycling lanes. *Accident Analysis and Prevention*, *167*. https://doi.org/10.1016/j.aap.2022.106577

Waard, de, Toffetti, A., Wiczorek, R., Sonderegger, A., Röttger, S., Bouchner, P., Franke, T., Fairclough, S., Noordzij, M., Brookhuis, K., Dettmann, A., & Bullinger, A. C. (2016). Spatially distributed visual, auditory and multi modal warning signals- a comparison. *Proceedings of the Human Factors and Ergonomics Society Europe Chapter*.

Wang, C., Zhang, W., Feng, Z., Sze, N. N., Xu, J., Zhang, X., Wang, K., Huang, W., & Luo, Y. (2019). Aberrant behaviours in relation to the self-reported crashes of bicyclists in China: Development of the Chinese Cycling Behaviour Questionnaire. *Transportation Research Part F: Traffic Psychology and Behaviour*, *66*. https://doi.org/10.1016/j.trf.2019.08.022

Wang, C., Zhang, W., Feng, Z., Wang, K., & Gao, Y. (2020a). Exploring Factors Influencing the Risky Cycling Behaviors of Young Cyclists Aged 15–24 Years: A Questionnaire-Based Study in China. *Risk Analysis*, *40*(8). https://doi.org/10.1111/risa.13499

Wang, C., Zhang, W., Feng, Z., Wang, K., & Gao, Y. (2020b). Exploring Factors Influencing the Risky Cycling Behaviors of Young Cyclists Aged 15–24 Years: A Questionnaire-Based Study in China. *Risk Analysis*, *40*(8). https://doi.org/10.1111/risa.13499

Wang, C., Zhang, W., Feng, Z., Wang, K., & Gao, Y. (2020c). Exploring Factors Influencing the Risky Cycling Behaviors of Young Cyclists Aged 15–24 Years: A Questionnaire-Based Study in China. *Risk Analysis*, *40*(8). https://doi.org/10.1111/risa.13499

Watanabe, M., & Sakai, K. (2021). Identifying tractor overturning scenarios using a driving simulator with a motion system. *Biosystems Engineering*, *210*. https://doi.org/10.1016/j.biosystemseng.2021.08.010

Welle, B., Sharpin, A. B., Adriazola-Steil, C., Bhatt, A., Alveano, S., Obelheiro, M., Imamoglu, C. T., Job, S., Shotten, M., & Bose, D. (2018). *Sustainable and safe: A vision and guidance for zero road deaths*.

Whitmire, J., Morgan, J. F., Oron-Gilad, T., & Hancock, P. A. (2011). The effect of in-vehicle warning systems on speed compliance in work zones. *Transportation Research Part F: Traffic Psychology and Behaviour*. https://doi.org/10.1016/j.trf.2011.04.001

Wierda, M., & Brookhuis, K. A. (1991). Analysis of cycling skill: A cognitive approach. *Applied Cognitive Psychology*, *5*(2). https://doi.org/10.1002/acp.2350050205

Wogalter, M. S., Kalsher, M. J., Frederick, L. J., Magurno, A. B., & Brewster, B. M. (1998). Hazard Level Perceptions of Warning Components and Configurations. *International Journal of Cognitive Ergonomics*, *2*(1–2).

Wolfe, E. S., Arabian, S. S., Breeze, J. L., & Salzler, M. J. (2016). Distracted biking: An observational study. *Journal of Trauma Nursing*, *23*(2). https://doi.org/10.1097/JTN.0000000000000188

Wood, J. M., Black, A. A., & Tyrrell, R. A. (2022). Increasing the conspicuity of cyclists at night by using bicycle lights and clothing to highlight their biological motion to oncoming drivers. *Transportation Research Part F: Traffic Psychology and Behaviour*, *90*. https://doi.org/10.1016/j.trf.2022.09.005

Xiao, Y., & Seagull, F. J. (1999). An Analysis of Problems with Auditory Alarms: Defining the Roles of Alarms in Process Monitoring Tasks. *Proceedings of the Human Factors and Ergonomics Society Annual Meeting*, *43*(3). https://doi.org/10.1177/154193129904300327

Yang, X., Huan, M., Si, B., Gao, L., & Guo, H. (2012). Crossing at a red light: Behavior of cyclists at urban intersections. *Discrete Dynamics in Nature and Society*, *2012*. https://doi.org/10.1155/2012/490810

Yannis, G., Nikolaou, D., Laiou, A., Stürmer, Y. A., Buttler, I., & Jankowska-Karpa, D. (2020). Vulnerable road users: Cross-cultural perspectives on performance and attitudes. *IATSS Research*, *44*(3), 220–229. https://doi.org/10.1016/J.IATSSR.2020.08.006

Yun, H., & Yang, J. H. (2020). Multimodal warning design for take-over request in conditionally automated driving. *European Transport Research Review*. https://doi.org/10.1186/s12544-020-00427-5

Zafeiriou, S., Zhang, C., & Zhang, Z. (2015). A survey on face detection in the wild: Past, present and future. *Computer Vision and Image Understanding*, *138*. https://doi.org/10.1016/j.cviu.2015.03.015

Zahabi, M., Abdul Razak, A. M., Shortz, A. E., Mehta, R. K., & Manser, M. (2020). Evaluating advanced driver-assistance system trainings using driver performance, attention allocation, and neural efficiency measures. *Applied Ergonomics*. https://doi.org/10.1016/j.apergo.2019.103036

Zhang, J., Johnson, T. R., Patel, V. L., Paige, D. L., & Kubose, T. (2003). Using usability heuristics to evaluate patient safety of medical devices. *Journal of Biomedical Informatics*, *36*(1–2). https://doi.org/10.1016/S1532-0464(03)00060-1

Zhang, Y., Yan, X., Wu, J., & Duan, K. (2020). Effect of Warning System on Interactive Driving Behavior at Unsignalized Intersection under Fog Conditions: A Study Based on Multiuser Driving Simulation. *Journal of Advanced Transportation*. https://doi.org/10.1155/2020/8871875

Zhao, X., Du, Y., Mao, E., Zhu, Z., & Song, Z. (2022). Agricultural tractor dynamics simulation research based on a driving simulator with real-time human-machine interaction. *Computers and Electronics in Agriculture*, *203*, 107443. https://doi.org/10.1016/J.COMPAG.2022.107443

Zoellick, J. C., Kluy, L., Rössle, S., Witte, J., Schenk, L., Kuhlmey, A., & Blüher, S. (2021). I'm curious, I'm open to it, I test it, I trust it! A focus groups study to understand a-priori trust in automated buses. *Transportation Research Part F: Traffic Psychology and Behaviour*, *81*, 55–64. https://doi.org/10.1016/J.TRF.2021.05.016

Zuki, F. S. M., & Sulaiman, S. (2016). Using multimodal warning as a method of optimizing an in-vehicle warning system. *2016 4th International Conference on User Science and Engineering (i-USEr)*, 91–94. https://doi.org/10.1109/IUSER.2016.7857940

APPENDIX A

The following questionnaire was developed using CBQ, CCBQ and individual knowledge of road rules as mentioned in section 2.3.1.

A. CYCLISTS' BEHAVIOR QUESTIONNAIRE

Demographics

1. Gender (Male, Female, other)

2. Age

3. What kind of transportation mode do you normally use (walking, bicycling, driving, other)

Next page

4. Duration of average weekly bicycling trip (<15 min, 15-30 min, 31-45 min, >45 minutes)

5. Tell us about your bicycling experience (Less than a year, 2-5 years, more than 5 years, not applicable)

6. What is your reason for bicycling (commute to school/work, exercise, other, not applicable)

7. Where do you typically ride your bicycle (urban area, residential neighborhood, park, other)

8. How many times have you crashed into other non-motor vehicles in the past three-year period while cycling? (0, 1-5 times, more than 5, not applicable)

9. How many times have you crashed into motor vehicles in the past three-year period while cycling? (0, 1-5 times, more than 5, not applicable)

10. Are you aware of all traffic rules for bicycling? (not at all, somewhat well, moderately well, most of them, very well, not applicable)

CBQ item rated on 6-point scale (1-very infrequently or never, 2-quite infrequently, 3-infrequently, 4-frequently, 5-quite frequently, 6-very often or always.)

Violations:

11. I cycle against traffic (the wrong way)

12. Even though there is an exclusive bicycle lane nearby, I cycle on the vehicular lane or on the sidewalk

13. I cycle under the influence of alcohol and / or other drugs or hallucinogens

14. I zigzag between vehicles when I am using a mixed lane to go faster

15. I cross the road when it appears to be a clear crossing, even if the traffic light is red

16. I carry potentially obstructive objects while riding the bicycle (food, packs, cigarettes, etc)

Error:

17. I do not brake on a “stop” or “yield” sign and come close to colliding with other vehicles or pedestrians

18. I misjudge a turn and hit the curb on the road

19. I try to brake, but I am not able to do so because of poor hand or foot positioning (for brakes) or a slippery surface

20. When using crosswalks, I stay on my bicycle and ride across, instead of getting off my bicycle and walking

21. I try to overtake vehicles that had previously used indicators to signal that they were about to turn

22. I sometimes mistake a traffic signal for another one, and maneuver according to the latter

Positive Behavior

23. I stop and look both sides before crossing a corner or intersection

24. I try to move at an appropriate speed to avoid sudden collision or braking

25. I usually keep a safe, recommended distance from vehicles and other road users

26. I avoid cycling under poor weather conditions (heavy rain, sleet, hail, high winds, etc.)

27. I use the helmet for cycling

28. When I travel at night, I use the necessary safety equipment (lights, vest, and reflectors)

Aggressive violations

29. I change course (such as turning, avoiding obstacles, passing pedestrians) without giving any signal to other road users, making a sudden sharp turn

30. I yell at other road users if they do not follow the rules

31. I make rude gestures (hand, face gestures) to other road users if they do not follow the rules

32. I cycle around other vehicles, cyclists, or pedestrians and “cut them off”, forcing them to brake or stop

33. I have races with other cyclists or drivers

34. I hit other road users if they slow down and block my way

Distractions and Forgetfulness

35. I get distracted and unintentionally hit a parked vehicle

36. I cycle with one hand and execute other actions with the other hand (holding up umbrellas, eating, using phones, etc.)

37. I listen to audio (news or music) while cycling and do not hear audible cues

38. I get forgetful and think about other things while cycling

39. I get distracted and do not see that there is an object or parked vehicle on the road and get close to hitting them

40. I get distracted and forget to gesture that I am turning left or right

Scenario-based survey (ordinal response a being the most conservative and c being the most risky)

Violation

41. You have been cycling on a two-way urban road with no designated bicycle lane for a while to reach your destination. You are riding in the right lane. As you approach the lights, you realize you have to turn right to reach your destination, but the traffic light is red. There is a “No Right Turn on Red” sign posted by the traffic lights. You don’t see any other vehicle or pedestrian at the intersection. You look further to make sure there are no approaching vehicles or pedestrians either. What would you do?

a. I will wait for the lights to turn green

b. I will get on the sidewalk to turn right to avoid any approaching vehicle

c. I will turn right anyway since there is no other vehicle at the intersection

Error

42. You are cycling in the designated bicycle lane on the right side of a two-way urban road with a bicycle lane on each side. You are about to turn right at an intersection. The traffic light is green, so it is now your turn to proceed. As you are about to turn into the street, you realize there is a pedestrian who is about to cross even though the light for the pedestrian is red. As a cyclist, what would you do?

a. I will come to a full stop and let the pedestrian cross

b. I will swerve around the pedestrian to get out of their way

c. I will continue turning because the pedestrian should not cross on a red light

Positive Behaviors

43. You are cycling in the designated bicycle lane on the right side of a very busy two-way urban road with a bicycle lane on each side. You are approaching a four-way intersection with 'All Way' stop signs. As you stop at the intersection, you realize you have arrived at the intersection first. As you are about to start crossing the intersection, you realize the car on your left has also been moving forward. What would you do in this situation?

a. I will stop and let the car pass first

b. I will ring my bell to alert the driver and I will be prepared to stop

c. I will continue cycling because I know I have right-of-way

Aggressive Violations

44. You are cycling on a two-way urban road that has no designated lane for bicyclists. The speed limit is 30 mph. You are cycling behind a very slow vehicle that is going at a speed of 15 mph. You get frustrated because they are going much slower even though they don't have any other vehicle in front of them. There is no one else on the lane next to your left (wrong-way traffic). The road has a solid yellow line separating the lanes of traffic. Based on the scenario, what will you do?

a. I will keep following the vehicle patiently

b. I will use the oncoming traffic lane to overtake the vehicle

c. I will ring my bell and show rude gestures at the driver to express my frustration

Distractions and Forgetfulness

45. You are cycling on a busy urban road that has two lanes in each direction, but no designated lanes for bicyclists. You have your phone mounted on your bicycle for using Google Maps. You are approaching a busy intersection. You have a green light and the right-of-way to proceed straight. As you are about to approach the intersection, you notice there are some people gathered with posters on the sidewalk on your right. At the same time, you hear a notification on your phone. What would you do?

a. I will first cross the intersection safely before considering the phone or the crowd

b. I will check the notification on my phone as I cross the intersection

c. I will divert my attention to the crowd to see what is going on as I cross the intersection

Check Questions.

46. I am asked to choose quite infrequently

47. I am asked to choose quite frequently

APPENDIX B

The following questionnaire was used in both Study I and Study II to determine the fitness of the participants to continue the study.

A. SIMULATION SICKNESS QUESTIONNAIRE

Simulation Sickness Questionnaire (SSQ) from Kennedy et al., 1993

Circle how much each symptom below is affecting you right now.

1. General discomfort	None	Slight	Moderate	Severe
2. Fatigue	None	Slight	Moderate	Severe
3. Headache	None	Slight	Moderate	Severe
4. Eyestrain	None	Slight	Moderate	Severe
5. Difficulty focusing	None	Slight	Moderate	Severe
6. Increased salivation	None	Slight	Moderate	Severe
7. Sweating	None	Slight	Moderate	Severe
8. Nausea	None	Slight	Moderate	Severe
9. Difficulty concentrating	None	Slight	Moderate	Severe
10. Fullness of head	None	Slight	Moderate	Severe
11. Blurred vision	None	Slight	Moderate	Severe
12. Dizziness (eyes open)	None	Slight	Moderate	Severe
13. Dizziness (eyes closed)	None	Slight	Moderate	Severe
14. Vertigo	None	Slight	Moderate	Severe

15. Stomach awareness	None	Slight	Moderate	Severe
16. Burping	None	Slight	Moderate	Severe

APPENDIX C

The following questionnaire was used in Study II to determine the willingness of cyclists to try new technology.

C. PERSONAL INNOVATIVENESS QUESTIONNAIRE

Personal Innovativeness – adapted from Agarwal and Prasad (1998)

1. If I heard about a new technology, I would look for ways to experiment with it.
2. Among my peers, I am usually the first to try out new technologies.
3. In general, I am hesitant to try out new technologies. [reverse-scaled]
4. I like to experiment with new technologies.

APPENDIX D

The following questionnaire is the final 11-item questionnaire that explains the behavior of the US cyclists. This was obtained from the survey results in Study I in section 2.4.2.

D. 11-ITEM CYCLISTS' BEHAVIOR QUESTIONNAIRE

1. I cycle against traffic (the wrong way)
2. I zigzag between vehicles when I am using a mixed lane to go faster
3. I cross the road when it appears to be a clear crossing, even if the traffic light is red
4. I change course (such as turning, avoiding obstacles, passing pedestrians) without giving any signal to other road users, making a sudden sharp turn
5. get distracted and forget to gesture that I am turning left or right
6. I yell at other road users if they do not follow the rules
7. I make rude gestures (hand, face gestures) to other road users if they do not follow the rules
8. I sometimes mistake a traffic signal for another one, and maneuver according to the latter
9. I get distracted and do not see that there is an object or parked vehicle on the road and get close to hitting them
10. I use the helmet for cycling
11. When I travel at night, I use the necessary safety equipment (lights, vest, and reflectors)

APPENDIX E

E. SAS outputs

The following SAS outputs show the results explained in section 3.4.2.1. The Modified Levene's test, followed by the multiple comparison tests for the significant interaction effects and main effects are shown for the different in mean reaction time. The same outputs for the difference in mean heart rate are then shown.

The TTEST Procedure

Variable: d

group	Method	N	Mean	Std Dev	Std Err	Minimum	Maximum
1		16	0.5235	0.4243	0.1061	0.0908	1.5070
2		16	0.4482	0.3046	0.0761	0.0930	0.8640
Diff (1-2)	Pooled		0.0754	0.3693	0.1306		
Diff (1-2)	Satterthwaite		0.0754		0.1306		

group	Method	Mean	95% CL Mean		Std Dev	95% CL Std Dev	
1		0.5235	0.2975	0.7496	0.4243	0.3134	0.6567
2		0.4482	0.2859	0.6105	0.3046	0.2250	0.4714
Diff (1-2)	Pooled	0.0754	-0.1913	0.3420	0.3693	0.2951	0.4937
Diff (1-2)	Satterthwaite	0.0754	-0.1925	0.3432			

Method	Variances	DF	t Value	Pr > |t|
Pooled	Equal	30	0.58	0.5681
Satterthwaite	Unequal	27.216	0.58	0.5686

Equality of Variances				
Method	Num DF	Den DF	F Value	Pr > F
Folded F	15	15	1.94	0.2108

Figure 19: Modified Levene's Test output for transformed difference in mean reaction time

The GLM Procedure
Least Squares Means
Adjustment for Multiple Comparisons: Tukey

speec_1	fortyfive_1	sqplustwo LSMEAN	LSMEAN Number
0	0	1.14326250	1
0	1	1.85162500	2
1	0	3.18698750	3
1	1	2.56530000	4

Least Squares Means for effect speec_1*fortyfive_1 Pr > |t| for H0: LSMean(i)=LSMean(j) Dependent Variable: sqplustwo				
i/j	1	2	3	4
1		0.3106	0.0001	0.0076
2	0.3106		0.0129	0.3044
3	0.0001	0.0129		0.4224
4	0.0076	0.3044	0.4224	

speec_1	fortyfive_1	sqplustwo LSMEAN	95% Confidence Limits	
0	0	1.143263	0.561112	1.725413
0	1	1.851625	1.269474	2.433776
1	0	3.186988	2.604837	3.769138
1	1	2.565300	1.983149	3.147451

Least Squares Means for Effect speec_1*fortyfive_1				
i	j	Difference Between Means	Simultaneous 95% Confidence Limits for LSMean(i)-LSMean(j)	
1	2	-0.708363	-1.807124	0.390399
1	3	-2.043725	-3.142486	-0.944964
1	4	-1.422037	-2.520799	-0.323276
2	3	-1.335363	-2.434124	-0.236601
2	4	-0.713675	-1.812436	0.385086
3	4	0.621688	-0.477074	1.720449

Figure 20: Tukey test output for the interaction effect of Audible*Speed for the difference in mean reaction time

The GLM Procedure
Least Squares Means
Adjustment for Multiple Comparisons: Tukey

day_1	sqplustwo LSMEAN	H0:LSMean1=LSMean2 Pr > |t|
0	1.47992500	0.0003
1	2.89366250	

day_1	sqplustwo LSMEAN	95% Confidence Limits	
0	1.479925	1.026889	1.932961
1	2.893663	2.440627	3.346698

Least Squares Means for Effect day_1				
i	j	Difference Between Means	Simultaneous 95% Confidence Limits for LSMean(i)-LSMean(j)	
1	2	-1.413738	-2.054397	-0.773078

Figure 21: Tukey test output for time of day for the difference in mean reaction time

The GLM Procedure
Least Squares Means
Adjustment for Multiple Comparisons: Tukey

red_1	sqplustwo LSMEAN	H0:LSMean1=LSMean2 Pr > |t|
0	2.81958125	0.0007
1	1.55400625	

red_1	sqplustwo LSMEAN	95% Confidence Limits	
0	2.819581	2.366545	3.272617
1	1.554006	1.100970	2.007042

Least Squares Means for Effect red_1				
i	j	Difference Between Means	Simultaneous 95% Confidence Limits for LSMean(i)-LSMean(j)	
1	2	1.265575	0.624915	1.906235

Figure 22: Tukey test output for visual for the difference in mean reaction time

The GLM Procedure
Least Squares Means
Adjustment for Multiple Comparisons: Tukey

red_1	fortyfive_1	hr LSMEAN	LSMEAN Number
0	0	23.3587500	1
0	1	14.8112500	2
1	0	32.8962500	3
1	1	36.6837500	4

Least Squares Means for effect red_1*fortyfive_1
Pr > |t| for H0: LSMean(i)=LSMean(j)

Dependent Variable: hr

i/j	1	2	3	4
1		0.3131	0.2280	0.0539
2	0.3131		0.0066	0.0011
3	0.2280	0.0066		0.8587
4	0.0539	0.0011	0.8587	

red_1	fortyfive_1	hr LSMEAN	90% Confidence Limits	
0	0	23.358750	17.486309	29.231191
0	1	14.811250	8.938809	20.683691
1	0	32.896250	27.023809	38.768691
1	1	36.683750	30.811309	42.556191

Least Squares Means for Effect red_1*fortyfive_1

i	j	Difference Between Means	Simultaneous 90% Confidence Limits for LSMean(i)-LSMean(j)	
1	2	8.547500	-3.249983	20.344983
1	3	-9.537500	-21.334983	2.259983
1	4	-13.325000	-25.122483	-1.527517
2	3	-18.085000	-29.882483	-6.287517
2	4	-21.872500	-33.669983	-10.075017
3	4	-3.787500	-15.584983	8.009983

Figure 23: Tukey test output for the interaction effect of Visual*Speed for the difference in mean heart rate

The GLM Procedure
Least Squares Means
Adjustment for Multiple Comparisons: Tukey

day_1	hr LSMEAN	H0:LSMean1=LSMean2 Pr > |t|
0	33.6868750	0.0008
1	20.1881250	

day_1	hr LSMEAN	90% Confidence Limits	
0	33.686875	29.534432	37.839318
1	20.188125	16.035682	24.340568

Least Squares Means for Effect day_1				
i	j	Difference Between Means	Simultaneous 90% Confidence Limits for LSMean(i)-LSMean(j)	
1	2	13.498750	7.626308	19.371192

Figure 24: Tukey test output for time of day for the difference in mean heart rate

The GLM Procedure
Least Squares Means
Adjustment for Multiple Comparisons: Tukey

front_1	hr LSMEAN	H0:LSMean1=LSMean2 Pr > |t|
0	29.9725000	0.0899
1	23.9025000	

front_1	hr LSMEAN	90% Confidence Limits	
0	29.972500	25.820057	34.124943
1	23.902500	19.750057	28.054943

Least Squares Means for Effect front_1				
i	j	Difference Between Means	Simultaneous 90% Confidence Limits for LSMean(i)-LSMean(j)	
1	2	6.070000	0.197558	11.942442

Figure 25: Tukey test output for collision type for the difference in mean heart rate

The GLM Procedure
Least Squares Means
Adjustment for Multiple Comparisons: Tukey

speec_1	hr LSMEAN	H0:LSMean1=LSMean2 Pr > |t|
0	22.5843750	0.0190
1	31.2906250	

speec_1	hr LSMEAN	90% Confidence Limits	
0	22.584375	18.431932	26.736818
1	31.290625	27.138182	35.443068

Least Squares Means for Effect speec_1				
i	j	Difference Between Means	Simultaneous 90% Confidence Limits for LSMean(i)-LSMean(j)	
1	2	-8.706250	-14.578692	-2.833808

Figure 26: Tukey test output for audible for the difference in mean heart rate

The TTEST Procedure

Variable: d

group	Method	N	Mean	Std Dev	Std Err	Minimum	Maximum
1		18	4.8156	3.7638	0.8871	0.2050	11.1550
2		14	6.6314	4.1737	1.1155	0.9150	11.2950
Diff (1-2)	Pooled		-1.8159	3.9467	1.4064		
Diff (1-2)	Satterthwaite		-1.8159		1.4252		

group	Method	Mean	95% CL Mean		Std Dev	95% CL Std Dev	
1		4.8156	2.9438	6.6873	3.7638	2.8243	5.6425
2		6.6314	4.2216	9.0413	4.1737	3.0258	6.7241
Diff (1-2)	Pooled	-1.8159	-4.6881	1.0564	3.9467	3.1538	5.2754
Diff (1-2)	Satterthwaite	-1.8159	-4.7427	1.1109			

Method	Variances	DF	t Value	Pr > |t|
Pooled	Equal	30	-1.29	0.2065
Satterthwaite	Unequal	26.53	-1.27	0.2137

Equality of Variances				
Method	Num DF	Den DF	F Value	Pr > F
Folded F	13	17	1.23	0.6781

Figure 27: Modified Levene’s Test output for difference in mean heart rate

APPENDIX F

F. Experiment setup

This is the experiment set up for the design is explain in section 3.4.1.3.

red	speech	day	front	45	trial	control	control	trial	trial	cnotrol	trial	control	trial	control
red	speech	day	right hook	30	control	trial	trial	control	control	trial	control	control	trial	trial
red	speech	night	front	30	trial	control	control	trial	trial	cnotrol	trial	control	trial	control
red	speech	night	right hook	45	control	trial	trial	control	control	trial	control	trial	trial	control
red	tone	day	front	30	trial	control	control	trial	trial	control	trial	control	control	trial
red	tone	day	right hook	45	control	trial	trial	control	control	trial	control	control	trial	trial
red	tone	night	front	45	trial	control	control	trial	trial	cnotrol	trial	control	trial	control
red	tone	night	right hook	30	control	trial	trial	control	control	trial	control	trial	trial	control
yellow	speech	day	front	30	trial	control	control	trial	trial	control	trial	control	control	trial
yellow	speech	day	right hook	45	control	trial	trial	control	control	trial	control	control	trial	trial
yellow	speech	night	front	45	trial	control	control	trial	trial	cnotrol	trial	control	trial	control
yellow	speech	night	right hook	30	control	trial	trial	control	control	trial	control	trial	trial	control
yellow	tone	day	front	45	trial	control	control	trial	trial	control	trial	control	control	trial
yellow	tone	day	right hook	30	control	trial	trial	control	control	trial	control	control	trial	trial
yellow	tone	night	front	30	trial	control	control	trial	trial	cnotrol	trial	control	trial	control
yellow	tone	night	right hook	45	control	trial	trial	control	control	trial	control	trial	trial	control
red	speech	day	front	45	trial	control	control	trial	trial	control	trial	control	control	trial
red	speech	day	right hook	30	control	trial	trial	control	control	trial	control	control	trial	trial
red	speech	night	front	30	trial	control	control	trial	trial	cnotrol	trial	control	trial	control
red	speech	night	right hook	45	control	trial	trial	control	control	trial	control	trial	trial	control
red	tone	day	front	30	trial	control	control	trial	trial	control	trial	control	control	trial
red	tone	day	right hook	45	control	trial	trial	control	control	trial	control	control	trial	trial
red	tone	night	front	45	trial	control	control	trial	trial	cnotrol	trial	control	trial	control
red	tone	night	right hook	30	control	trial	trial	control	control	trial	control	trial	trial	control
yellow	speech	day	front	30	trial	control	control	trial	trial	control	trial	control	control	trial
yellow	speech	day	right hook	45	control	trial	trial	control	control	trial	control	control	trial	trial
yellow	speech	night	front	45	trial	control	control	trial	trial	cnotrol	trial	control	trial	control
yellow	speech	night	right hook	30	control	trial	trial	control	control	trial	control	trial	trial	control
yellow	tone	day	front	45	trial	control	control	trial	trial	control	trial	control	control	trial
yellow	tone	day	right hook	30	control	trial	trial	control	control	trial	control	control	trial	trial
yellow	tone	night	front	30	trial	control	control	trial	trial	cnotrol	trial	control	trial	control
yellow	tone	night	right hook	45	control	trial	trial	control	control	trial	control	trial	trial	control

Design Table:

	A	B	C	D	E
1	-	-	-	-	+
2	-	-	-	+	-
3	-	-	+	-	-
4	-	-	+	+	+
5	-	+	-	-	-
6	-	+	-	+	+
7	-	+	+	-	+
8	+	+	+	+	-
9	+	-	-	-	-
10	+	-	-	+	+
11	+	-	+	-	+
12	+	-	+	+	-
13	+	+	-	-	+
14	+	+	-	+	-
15	+	+	+	-	-
16	+	+	+	+	+

Design Generator: E = ABCD

Alias structure:

I + ABCDE

A + BCDE

B + ACDE

C + ABDE

D + ABCE

E + ABCD

AB + CDE

AC + BDE

AD + BCE

AE + BCD

BC + ADE

BD + ACE

BE + ACD

CD + ABE

CE + ABD

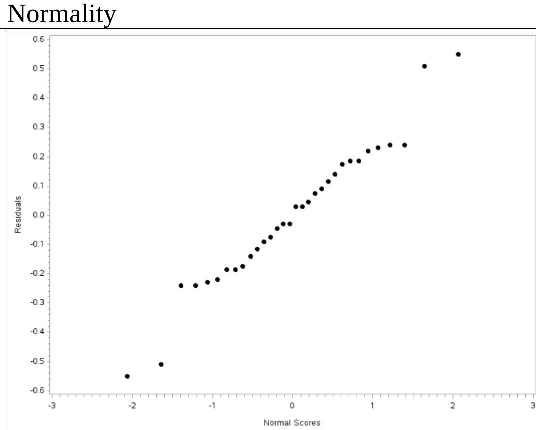
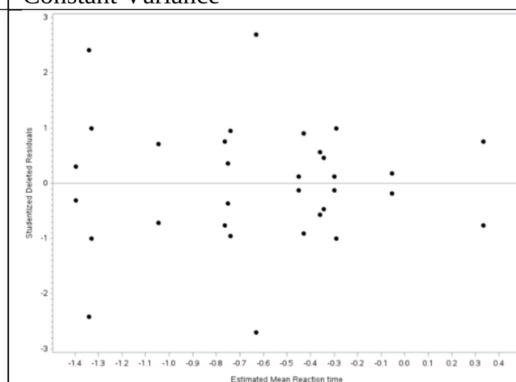
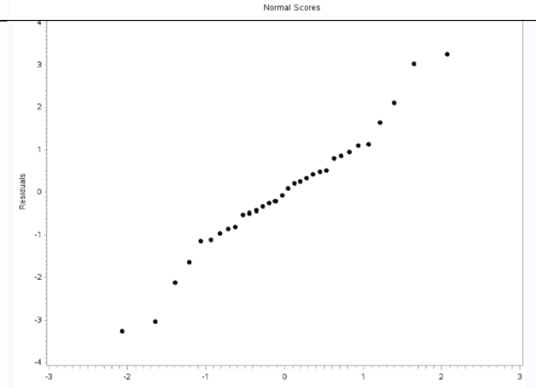
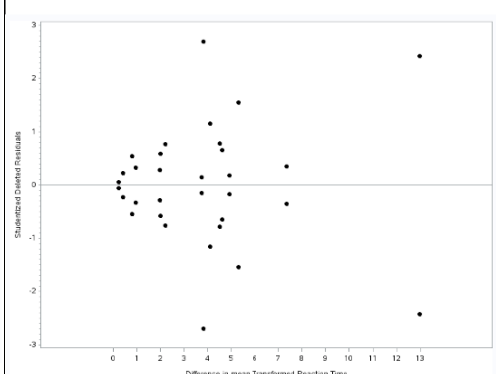
DE + ABC

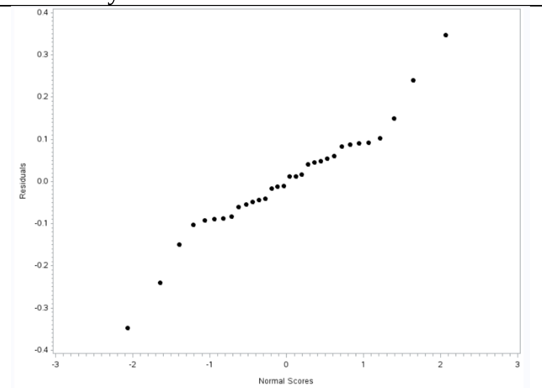
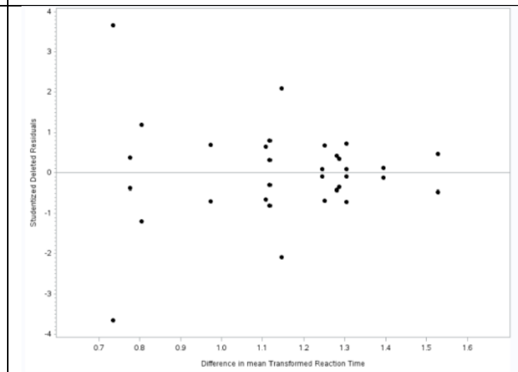
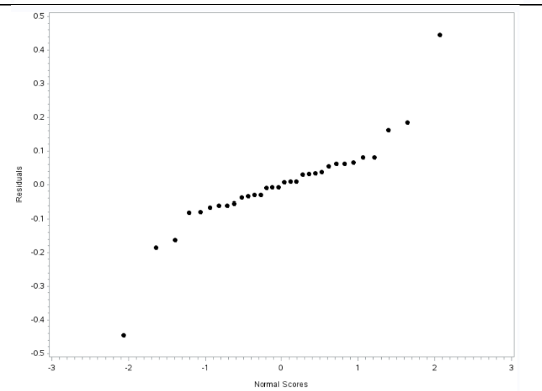
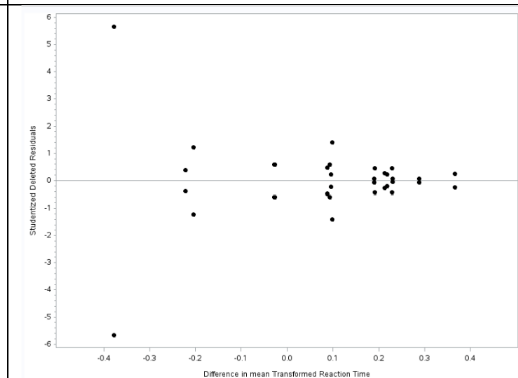
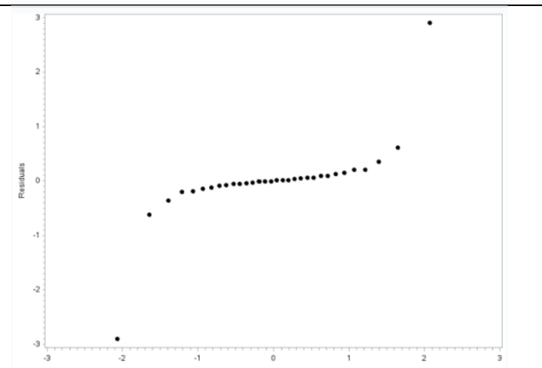
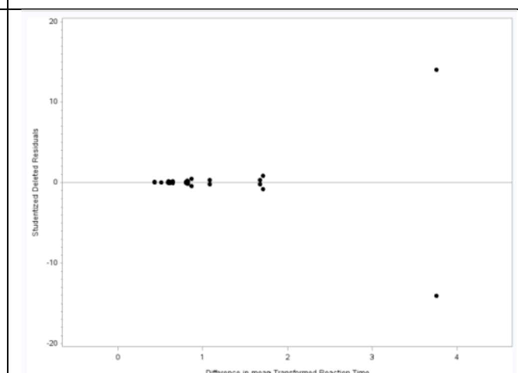
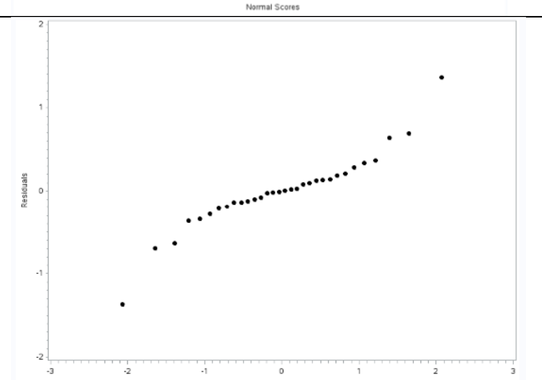
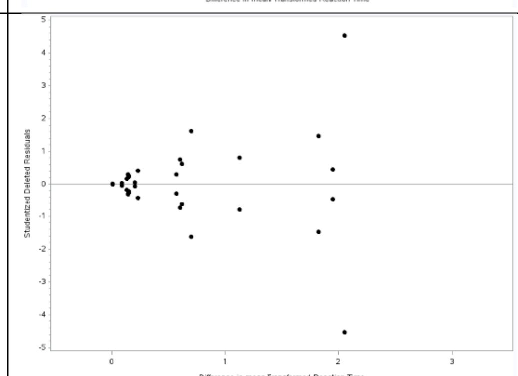
APPENDIX G

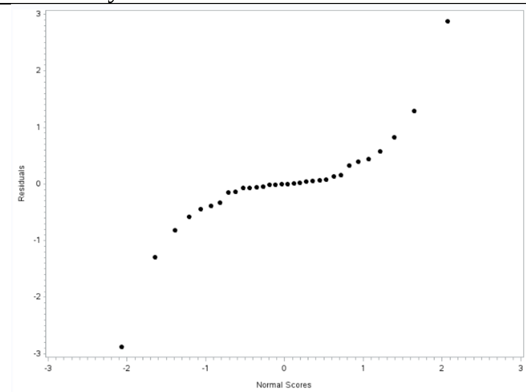
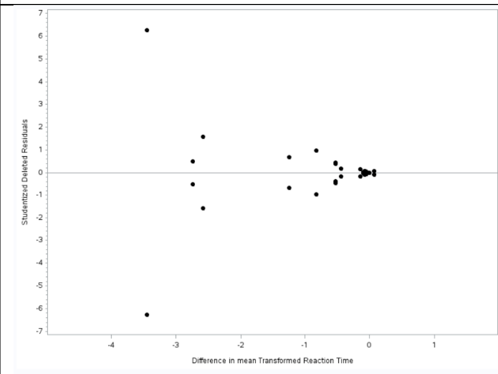
Transformation of difference in mean reaction time

The difference in mean reaction time had a violation of the constant variance assumption, as well as the normality assumption. Therefore, several variance stabilizing transformations were conducted on the difference in mean reaction time. The transformation results are summarized below in table 31 along with the respective normality and constant variance plots. The final transformation that was used was $(RT\text{-}RT_0+2)^2$ and the results are described in section 3.4.2.1.

Table 31: Transformation attempts for difference in mean reaction time

	Normality	Constant Variance
$RT\text{-}RT_0$		
$(RT\text{-}RT_0+2)^3$		

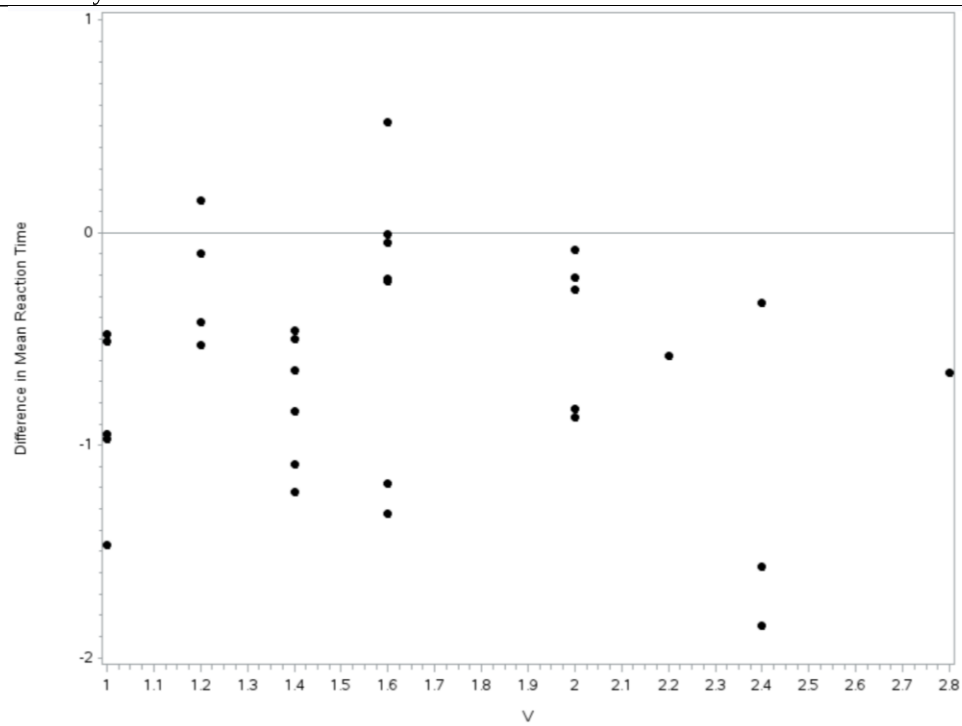
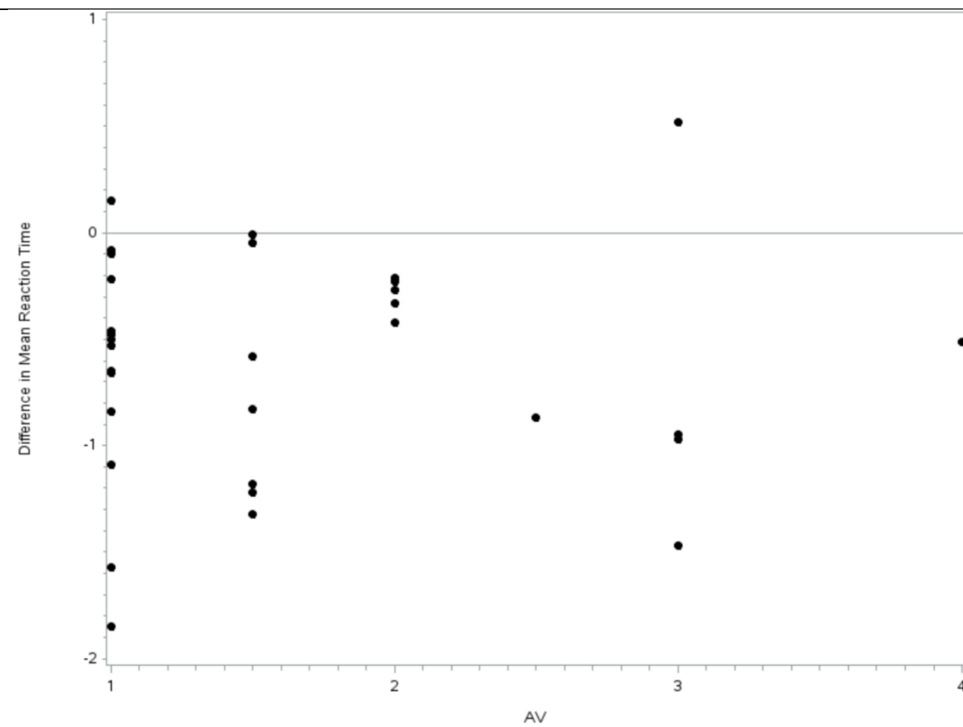
	Normality	Constant Variance
$\sqrt{\mathrm{RT}-\mathrm{RT0}}$		
$\mathrm{Lg}(\mathrm{RT}\text{-}\mathrm{RT}_0)$		
$\frac{1}{\mathrm{RT}-\mathrm{RT0}}$		
$(\mathrm{RT}\text{-}\mathrm{RT}_0)^2$		

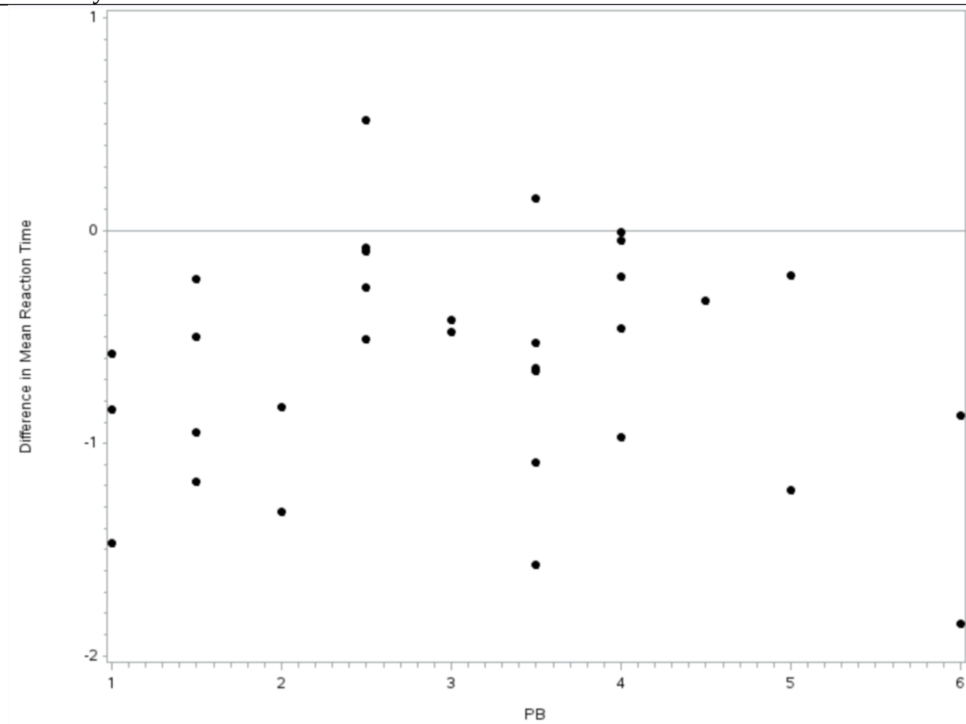
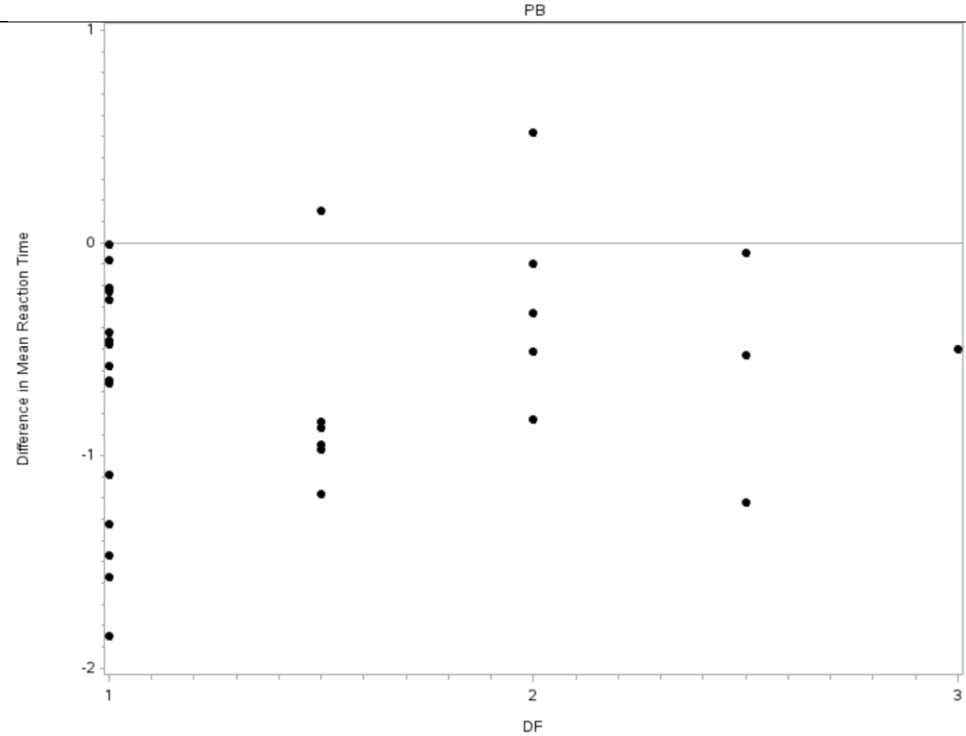
	Normality	Constant Variance
$(RT\text{-}RT_0)^3$		

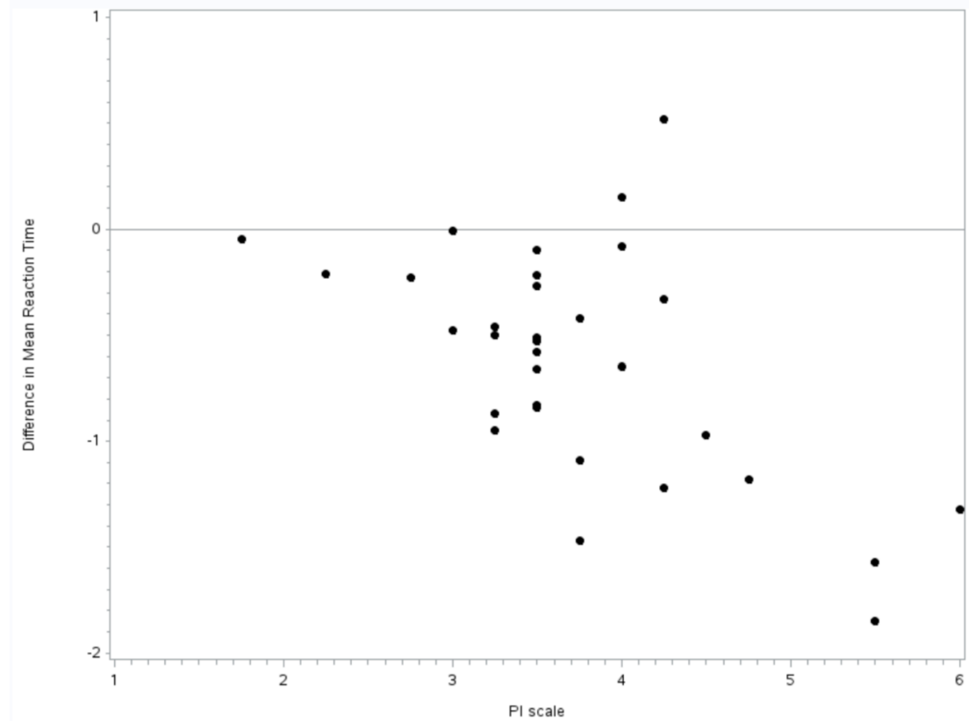
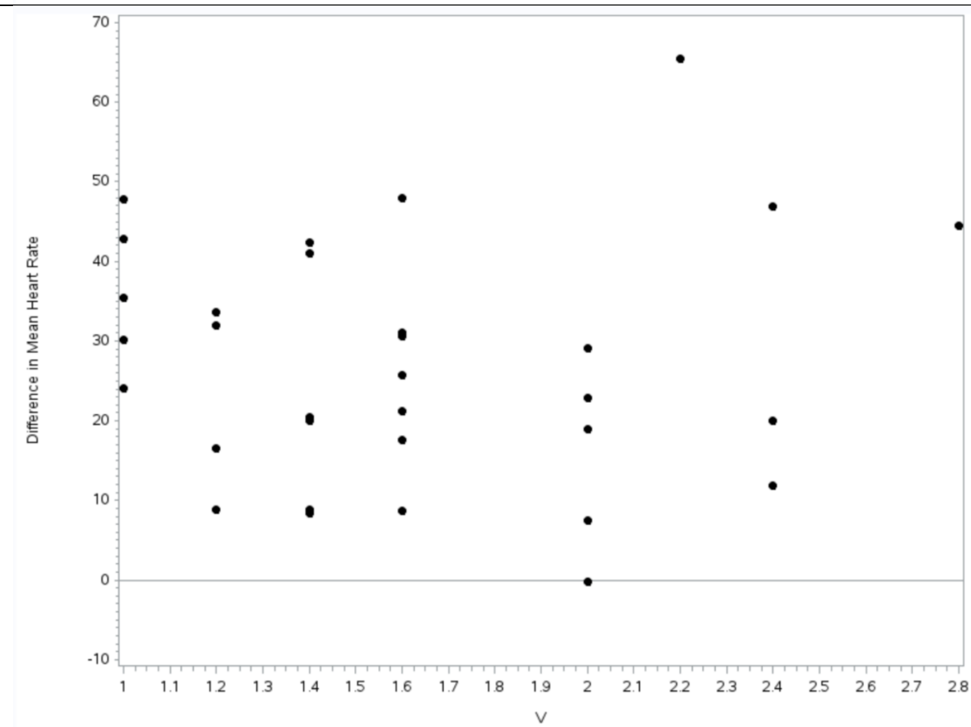
Influence of cyclists' behavior questionnaire and personal innovativeness on the warning system

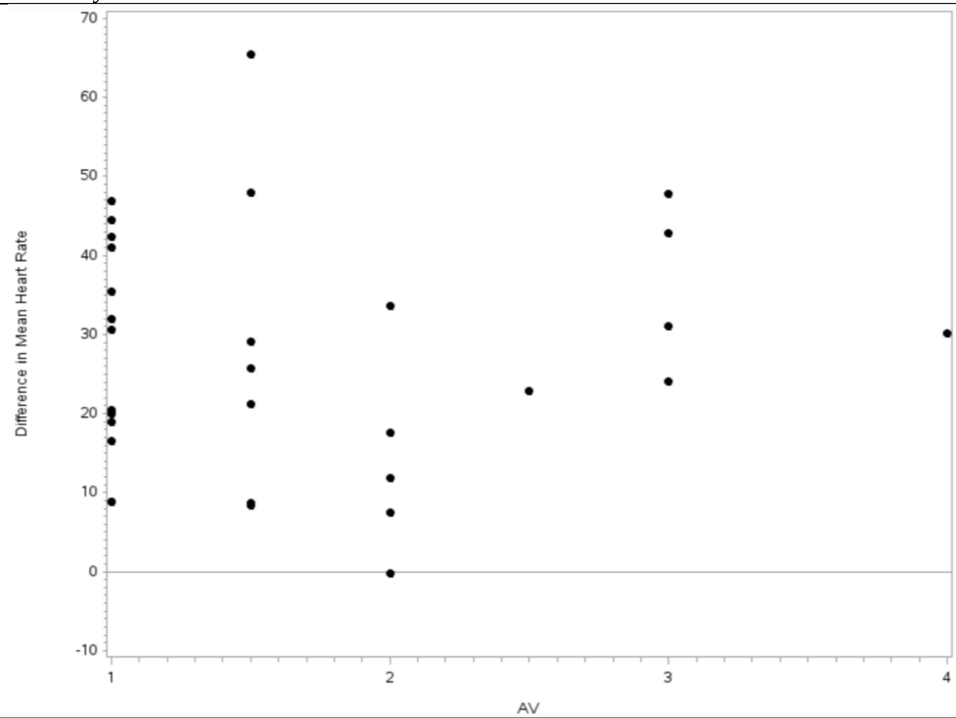
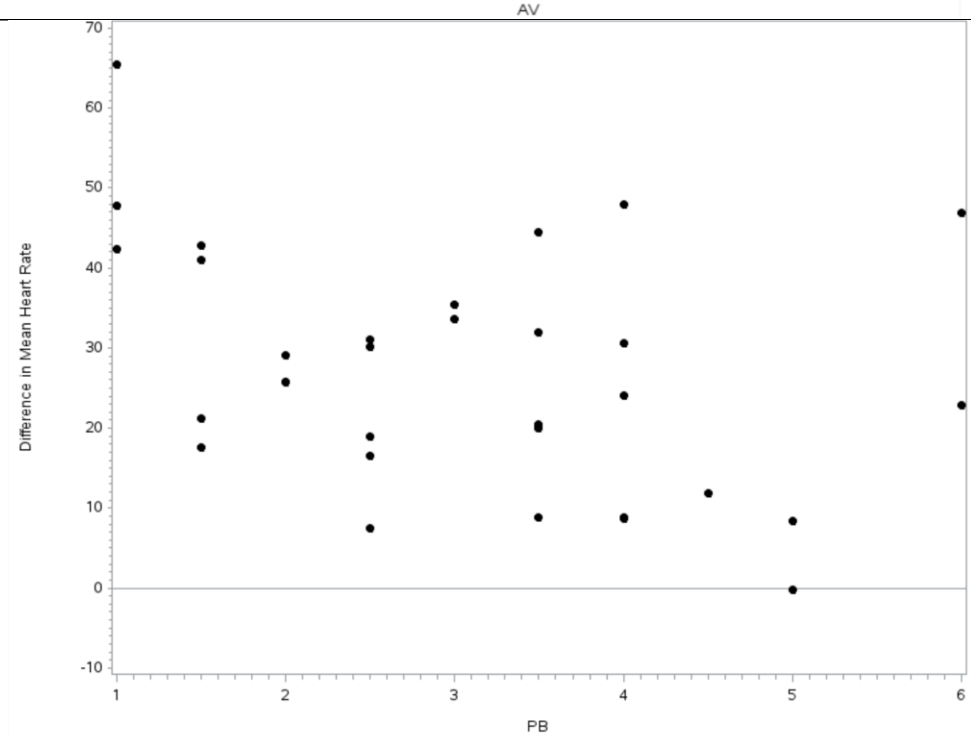
The scatterplots below represent the plots of the difference in mean reaction time ($RT\text{-}RT_0$) and the difference in mean heart rate ($HR\text{-}HR_0$) against the cyclists' behavior subscales- violation (V), aggressive violation (AV), positive behavior (PB), and distraction and forgetfulness (DF). The scatterplots of difference in mean reaction time ($RT\text{-}RT_0$) and the difference in mean heart rate ($HR\text{-}HR_0$) and the personal innovativeness (PI) questionnaire are also provided in the table. The correlation analyses based on these plots are mentioned in section 3.4.2.2.

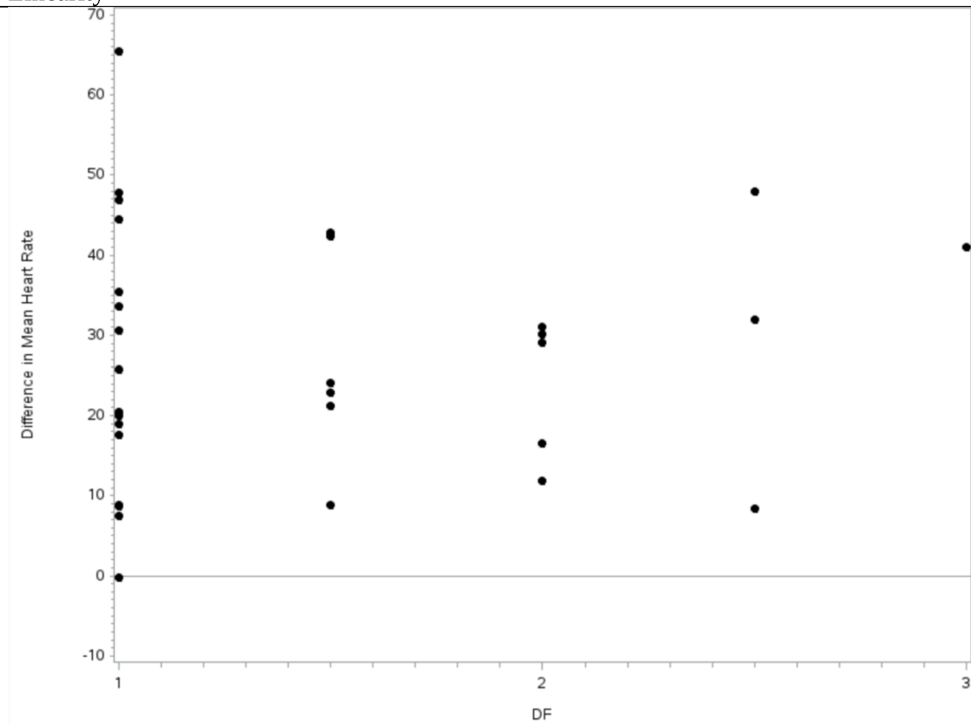
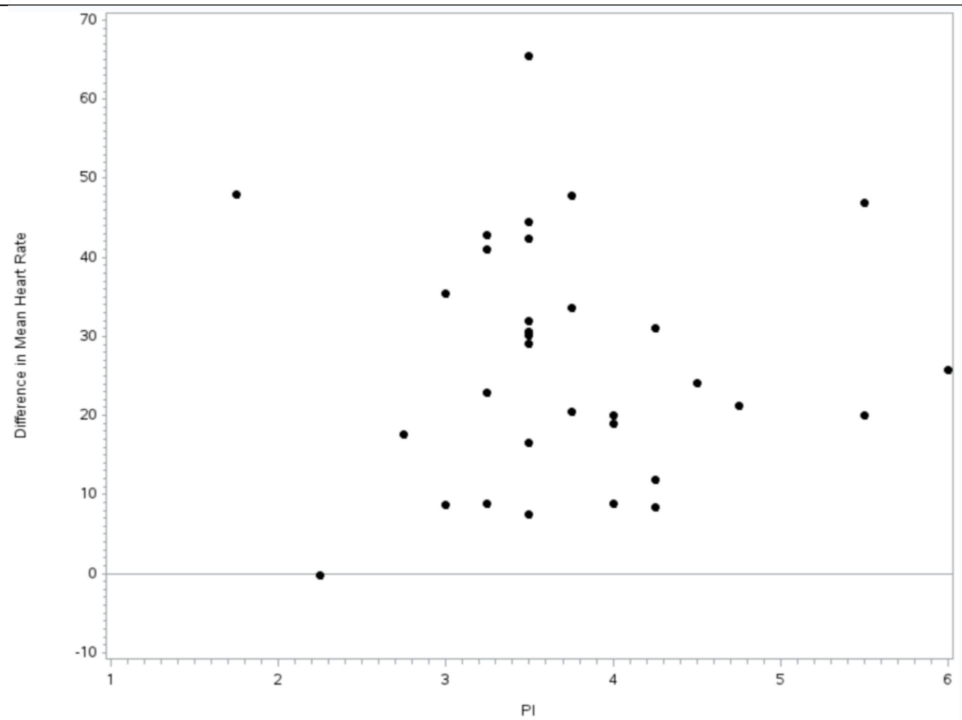
Table 32: Plots for difference in mean reaction time and difference in mean heart rate

	Linearity
($RT-RT_0$) vs V	
($RT-RT_0$) vs AV	

	Linearity
$(RT-RT_0)$ vs PB	
$(RT-RT_0)$ vs DF	

	Linearity
($RT-RT_0$) vs PI	
($HR-HR_0$) vs V	

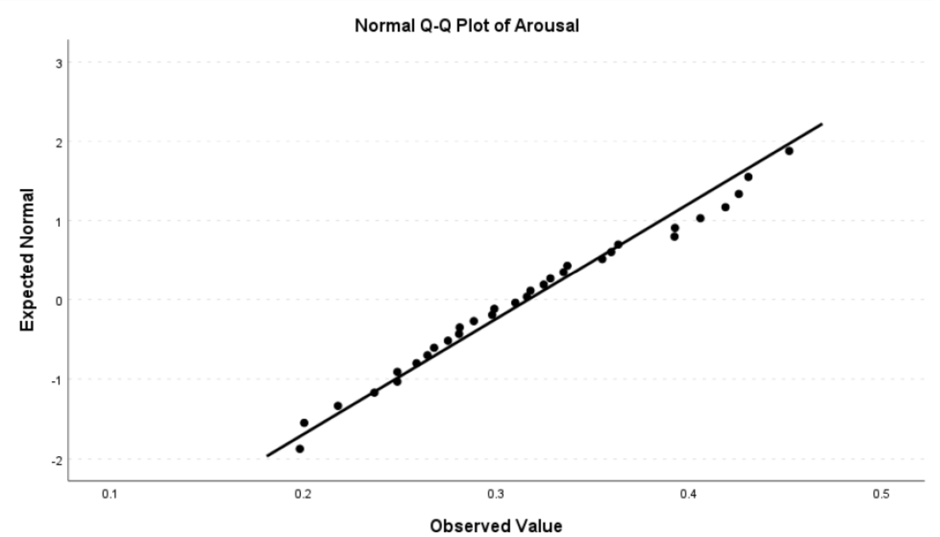
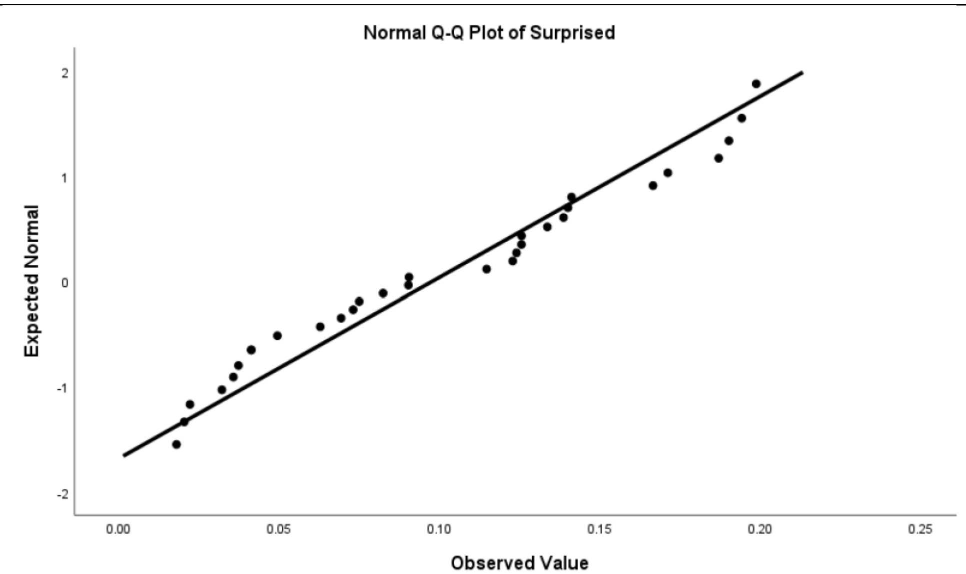
	Linearity
$(HR-HR_0)$ vs AV	
$(HR-HR_0)$ vs PB	

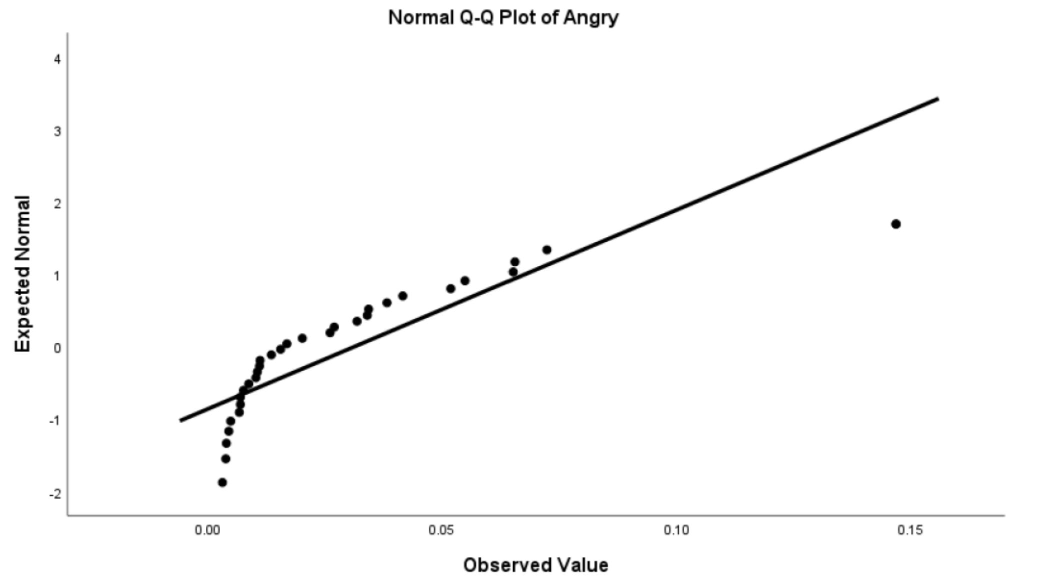
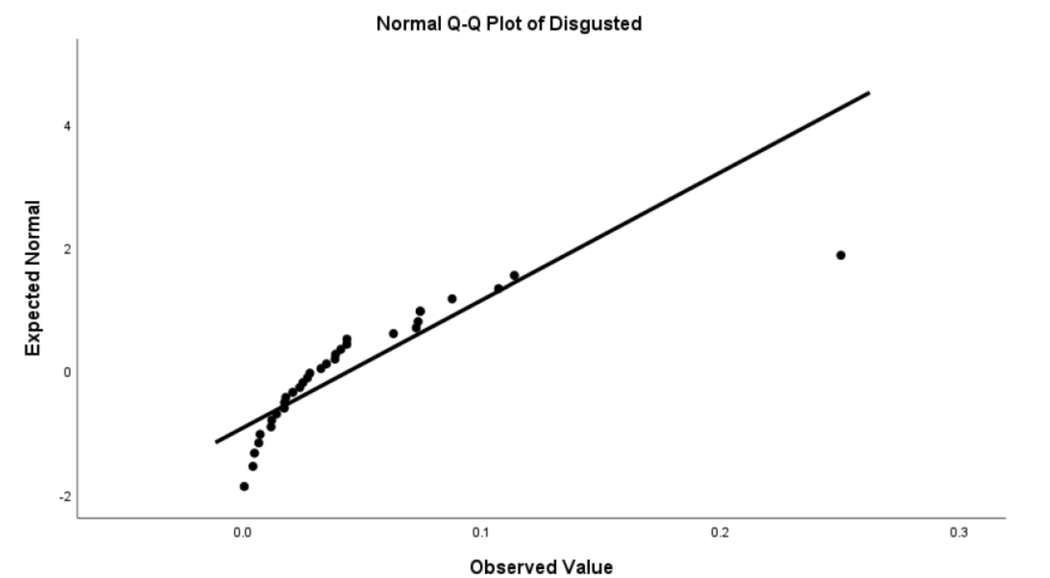
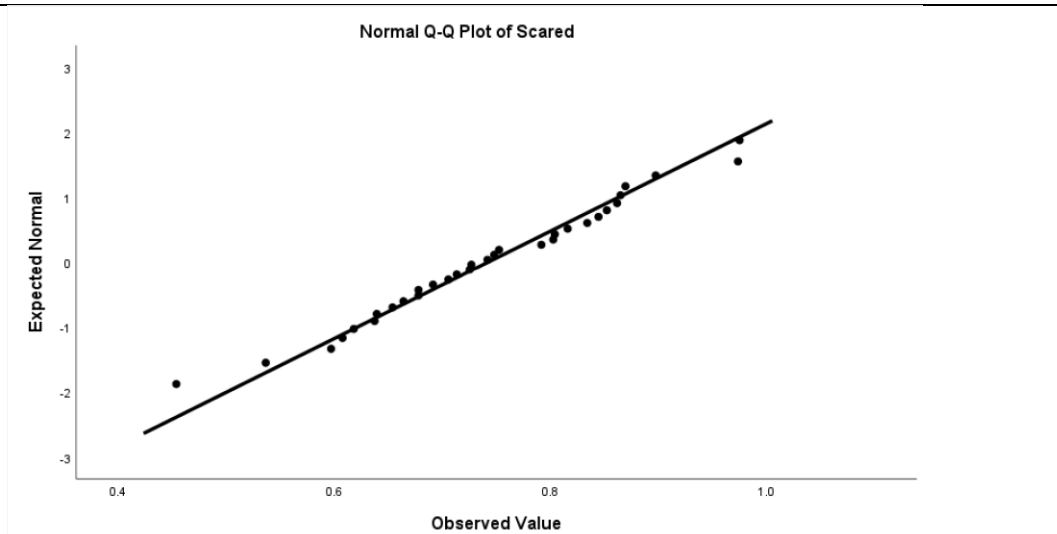
	Linearity
$(HR-HR_0)$ vs DF	
$(HR-HR_0)$ vs PI	

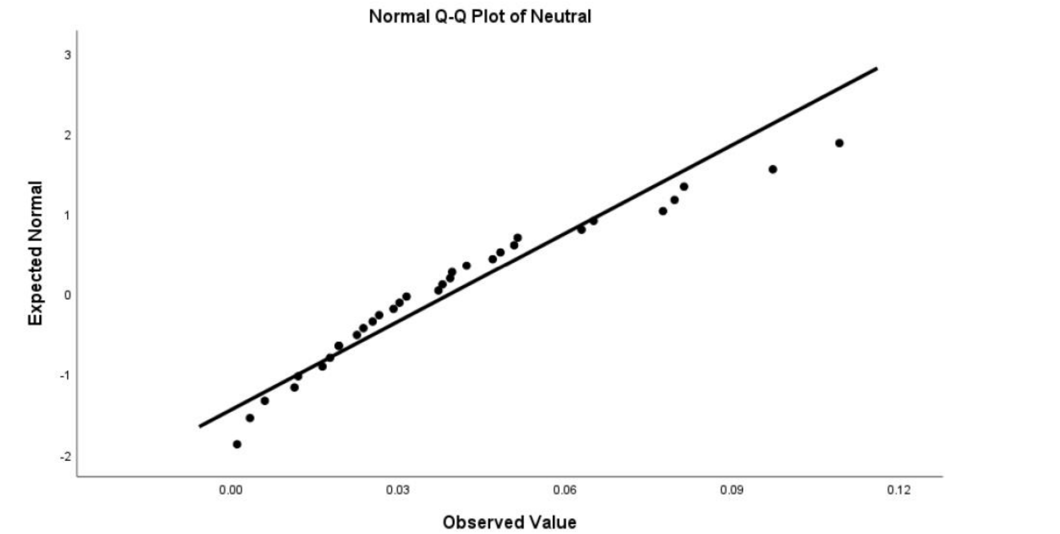
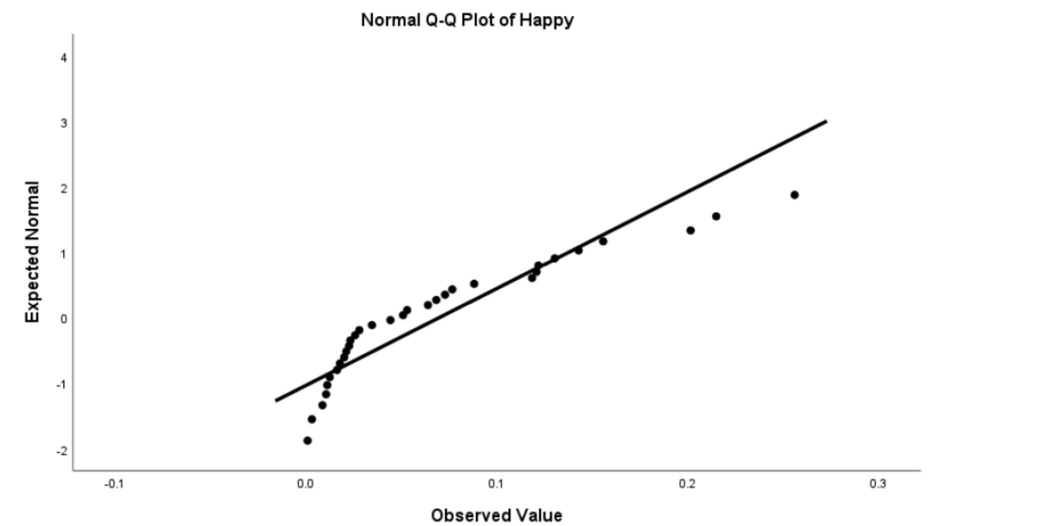
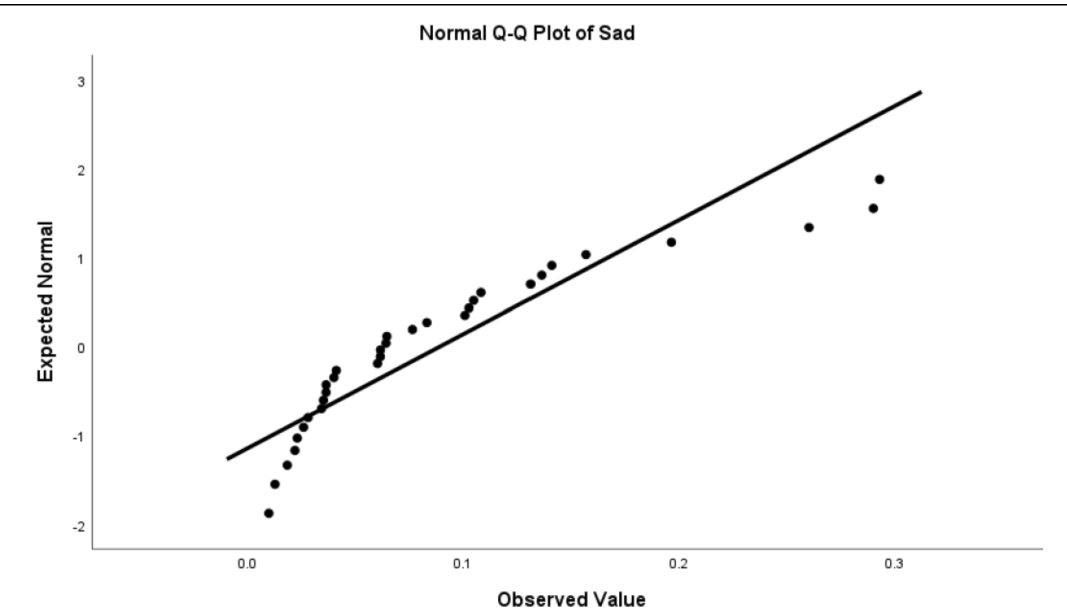
Influence of the warning signals on the emotional variables

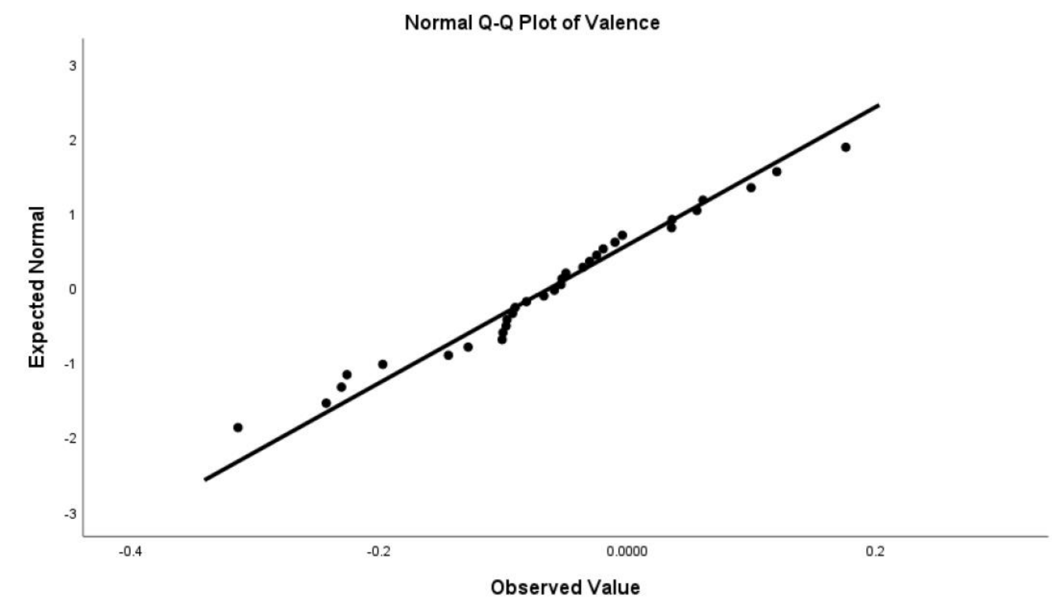
The Q-Q plots of each of the emotional variables- arousal, surprised, angry, disgusted, scared, neutral, happy, sad, and valence were analyzed to check the assumption of normality for conducting t-test in section 3.4.3.3. Table 28 summarizes the normality plots.

Table 33: Normality plots for emotional variables

Variable	
Arousal	
Surprised	

<table>
<tr><td>Angry</td><td>
</td></tr>
<tr><td>Disgusted</td><td>
</td></tr>
<tr><td>Scared</td><td>
</td></tr>
</table>

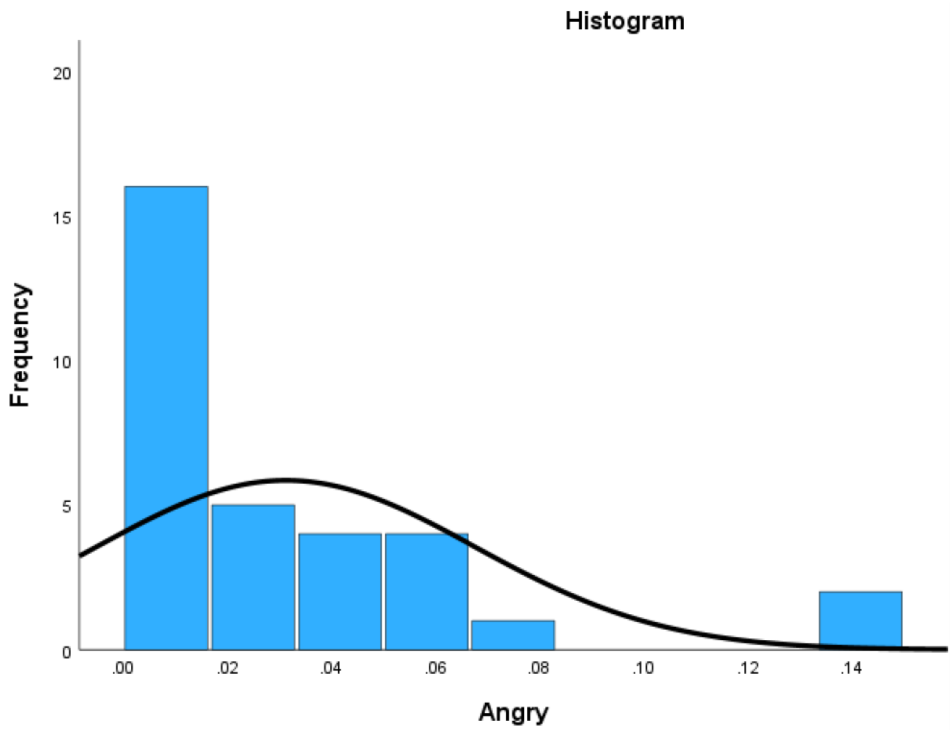
Neutral	
Happy	
Sad	

Valence	

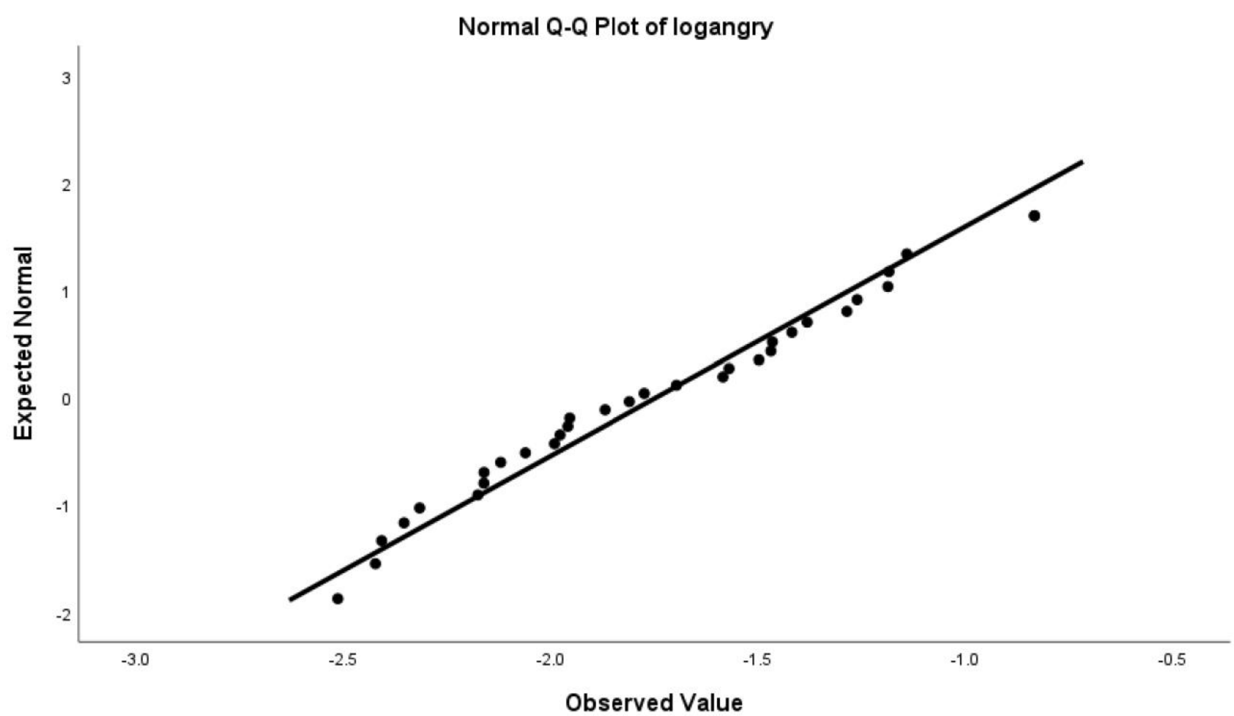
Transformation of the emotion variables:

Table 33 shows angry, disgusted, happy, and sad severely violated normality. Therefore, transformation was explored to get the distribution closer to normality. The transformation is explained in section 3.4.3.3 as well.

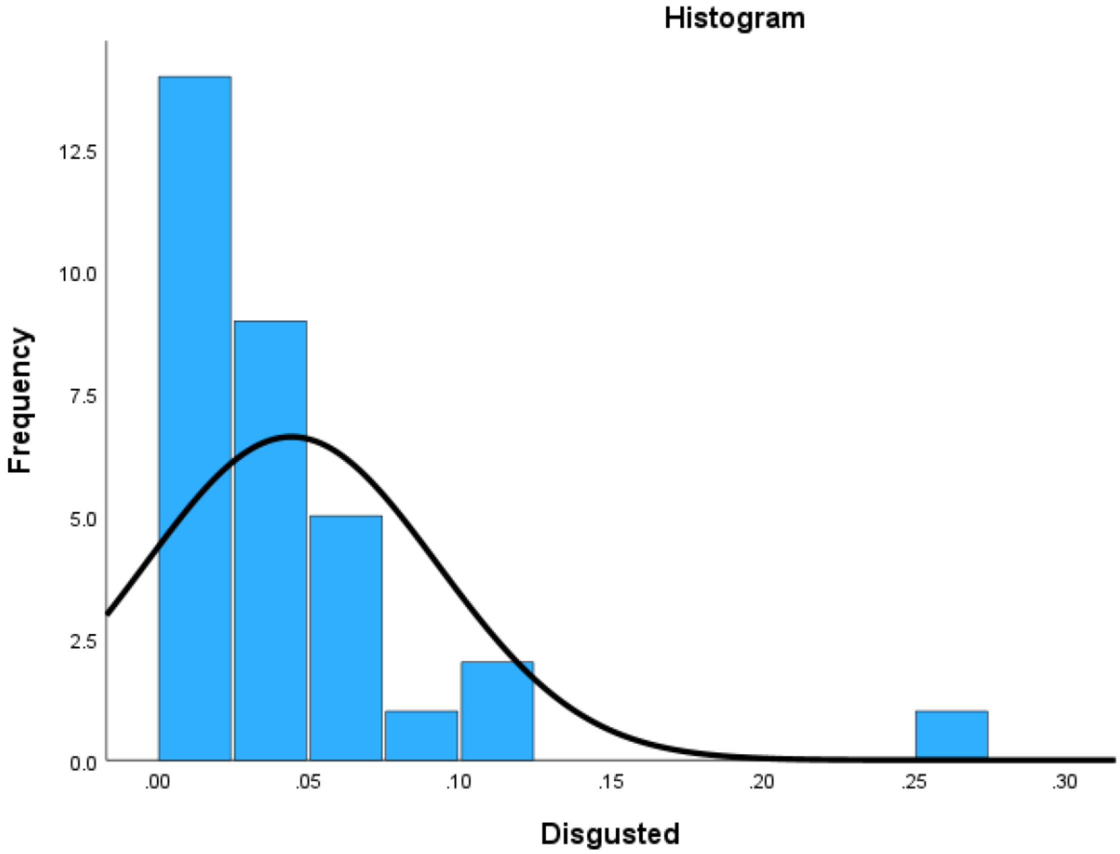
1. Angry:



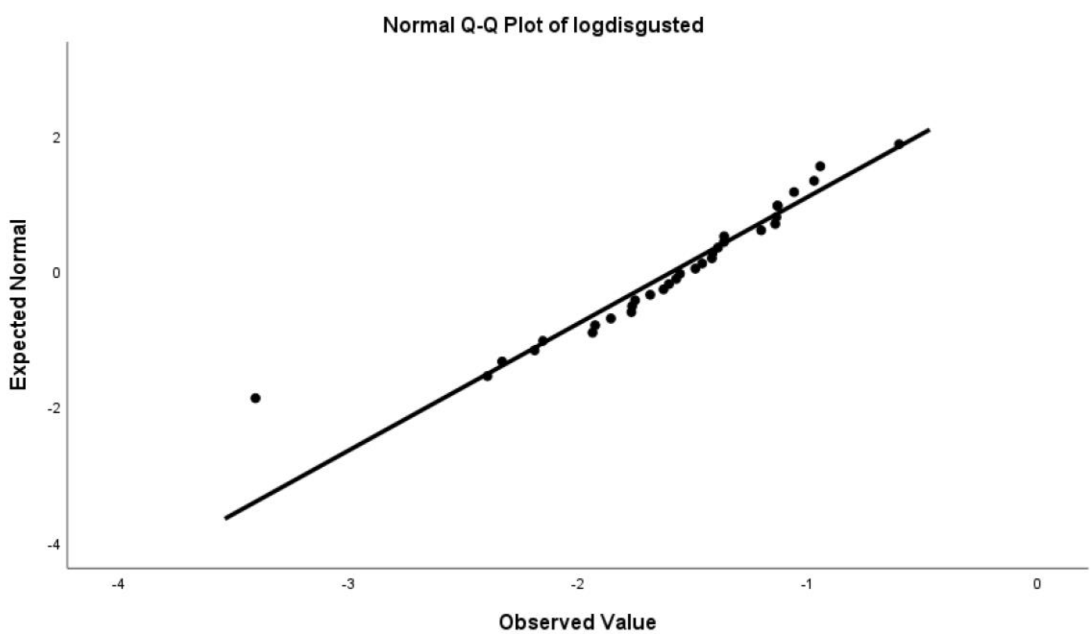
Log Transformation:



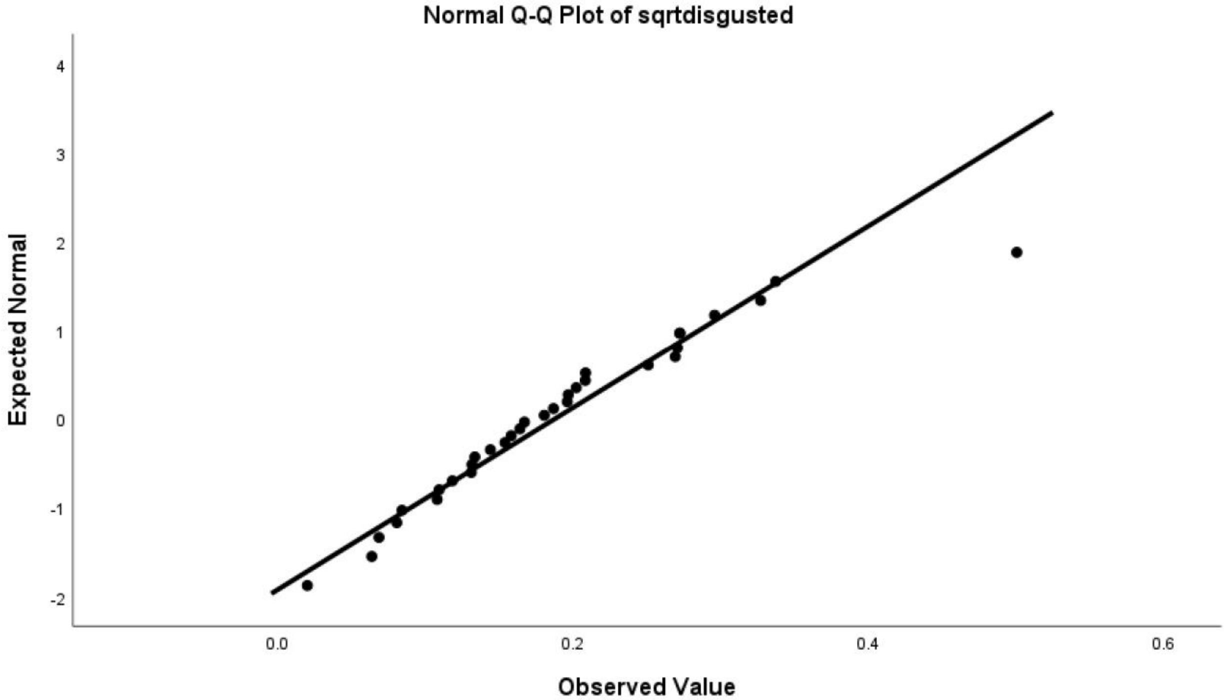
2. Disgusted:



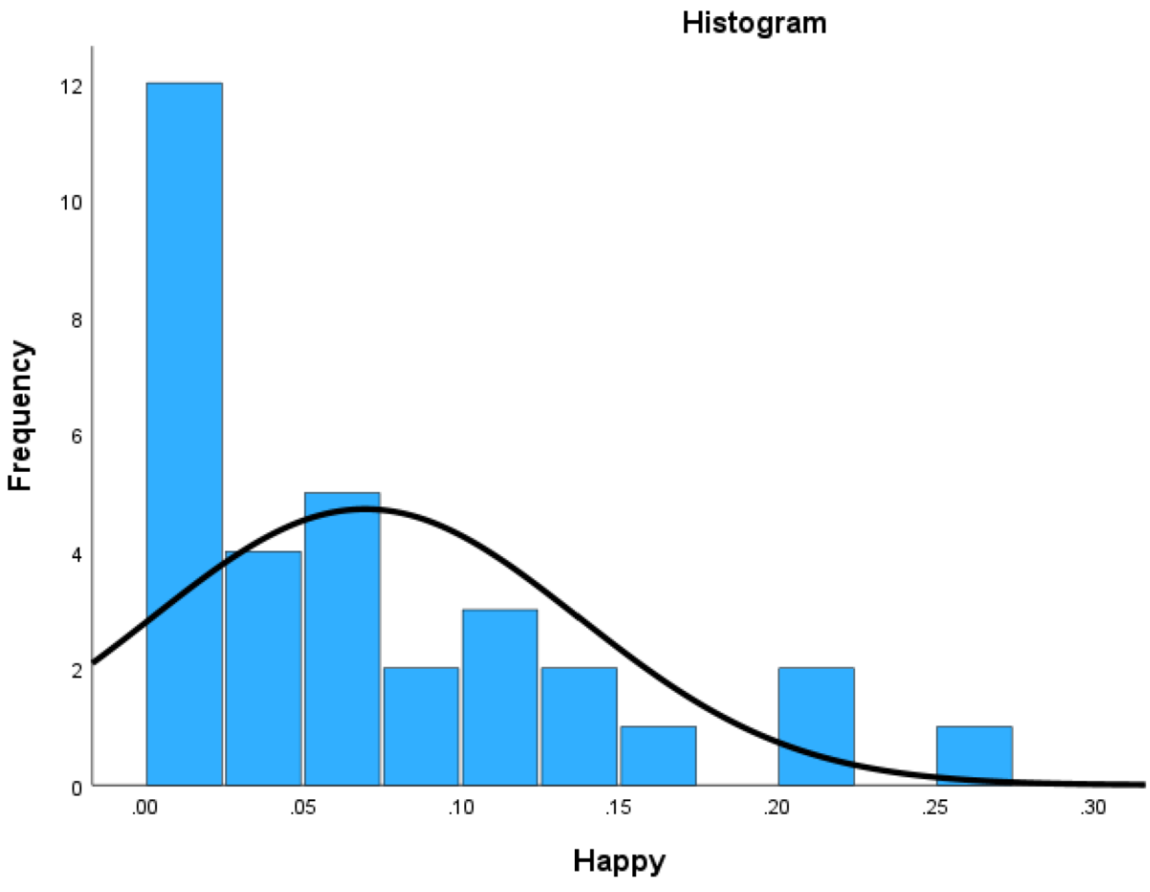
Log Transformation:



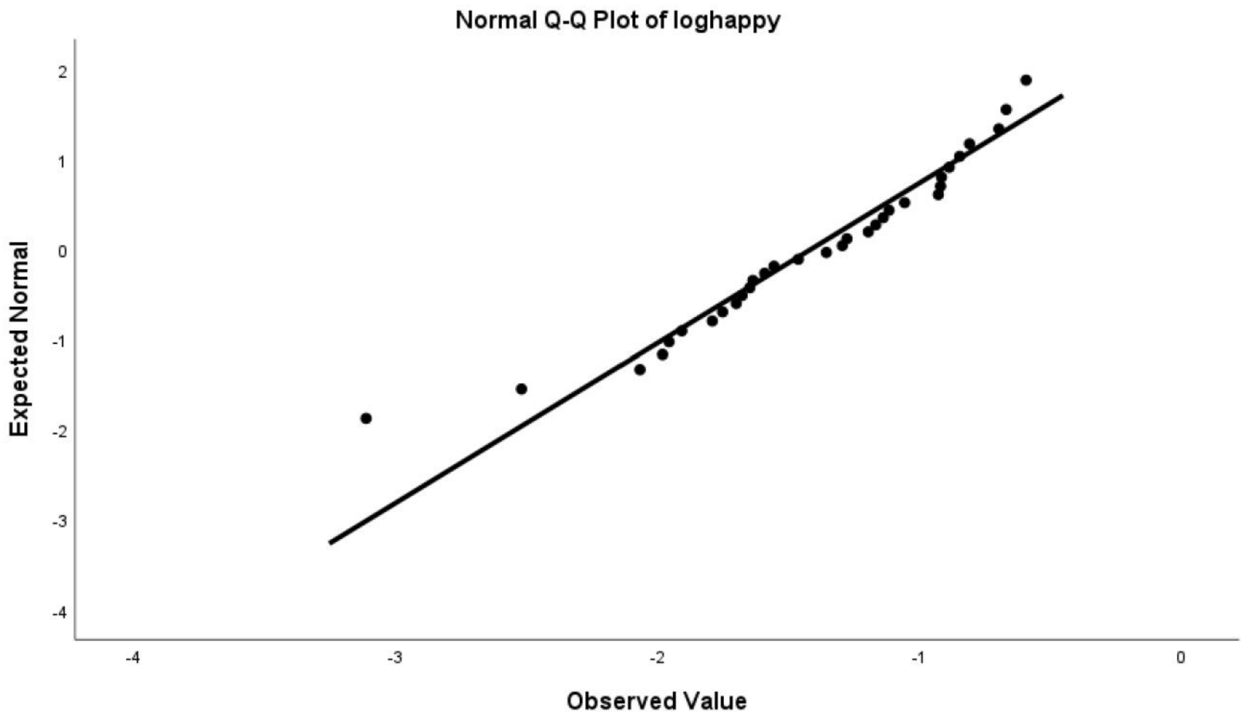
Square Root Transformation:



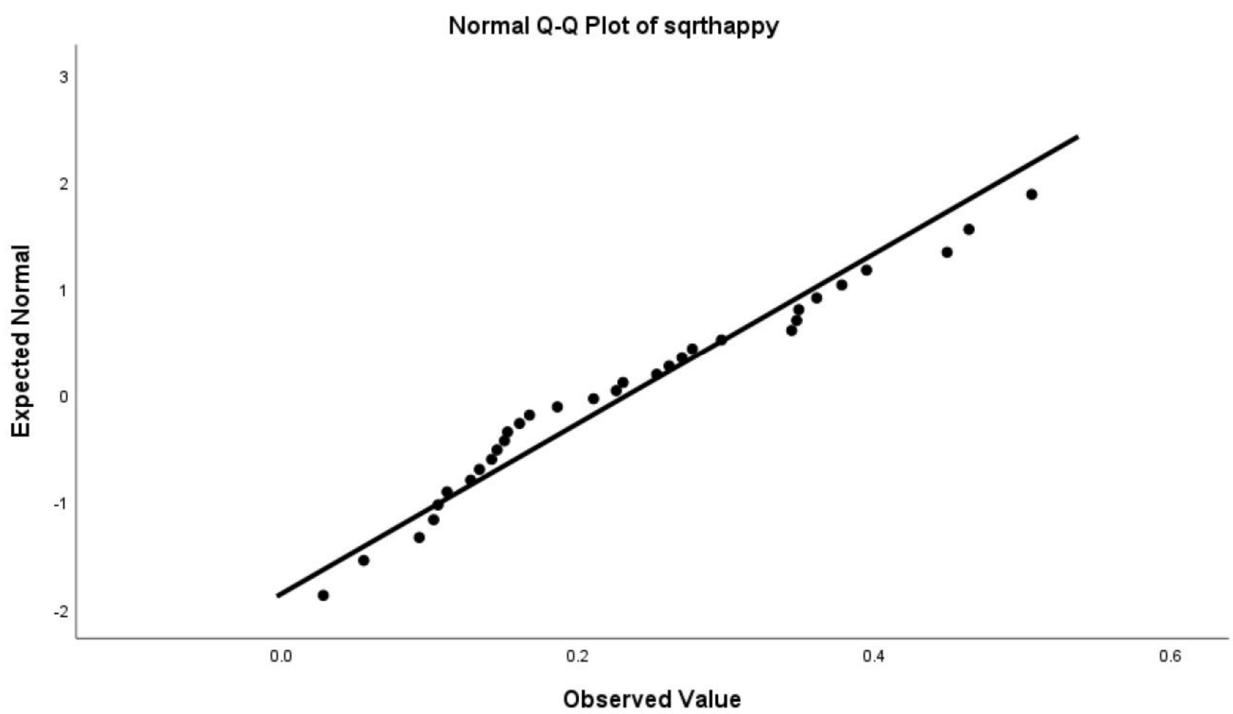
3. Happy



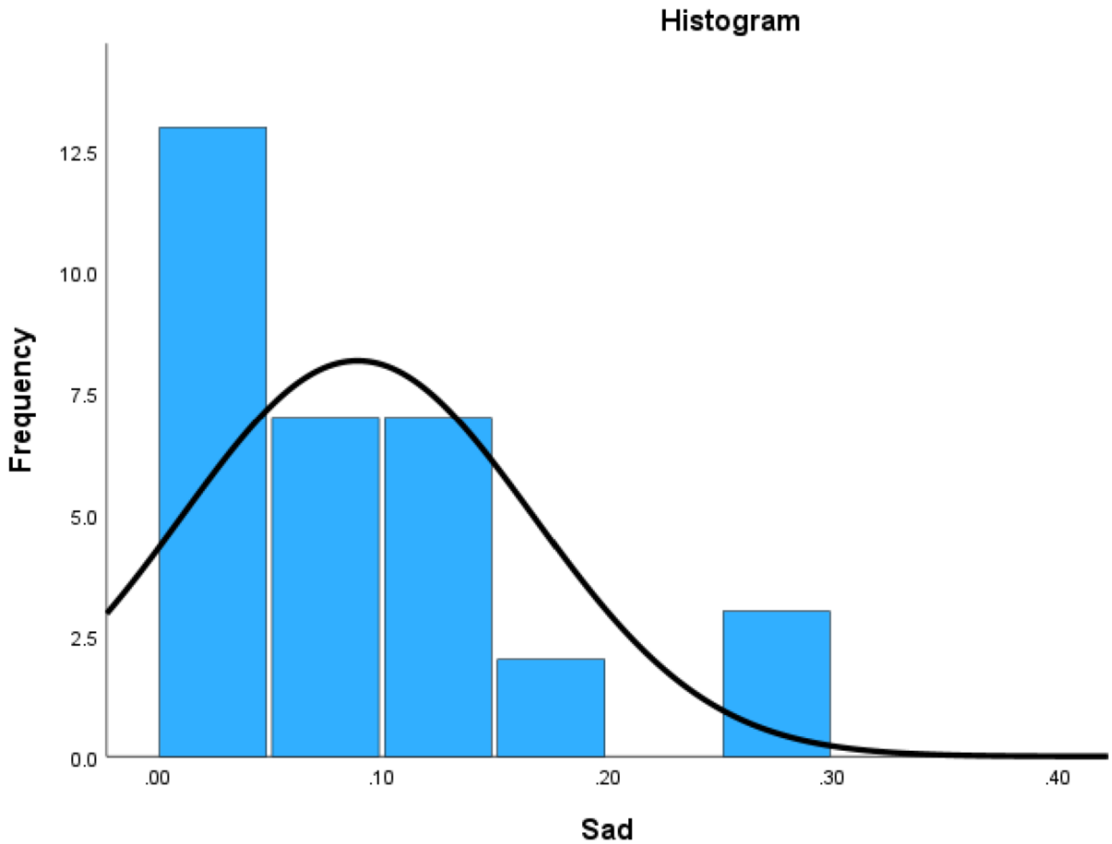
Log Transformation:



Square Root Transformation:



4. Sad



Log Transformation: