AN EXPLORATORY ANALYSIS OF FACTORS LEADING TO PREVALENCE OF DISTRACTED DRIVING BEHAVIOR IN PAKISTAN

By

Qamar Muneer

(00000330702)

Department of Transportation Engineering

NUST Institute of Civil Engineering

School of Civil and Environmental Engineering (SCEE)

National University of Sciences and Technology (NUST)

Sector H-12, Islamabad, Pakistan

December 2023

AN EXPLORATORY ANALYSIS OF FACTORS LEADING TO PREVALENCE OF DISTRACTED DRIVING BEHAVIOR IN PAKISTAN

By

Qamar Muneer

NUST2020-330702

A thesis submitted to the National University of Sciences and Technology,

Islamabad, in partial fulfillment of the requirements for the degree of

Master of Science in Transportation Engineering

Supervisor: Dr. Arshad Hussain

NUST Institute of Civil Engineering

School of Civil and Environmental Engineering (SCEE)

National University of Sciences and Technology (NUST)

Sector H-12, Islamabad, Pakistan

THESIS ACCEPTANCE CERTIFICATE

It is certified that final copy of MS thesis written by **Mr. QAMAR MUNEER (Registration No.00000330702)** of MS Transportation SCEE (NICE), has been vetted by undersigned, found **co**mplete in all respects as per NUST Statutes / Regulations, is free of plagiarism, errors, and mistakes and is accepted as partial fulfillment for award of MS degree. It is further certified that necessary amendments as pointed out by GEC members of the scholar have also been incorporated in the said thesis.

Dr. Arshad Hussain
Associate Professor
Head of Transportation Engineering Department
School of Civil & Environmental Engineering
NUST, Sector H-12, Islamabad

Signature: ______________________

Name of Supervisor: **Dr. Arshad Hussain**

Date: 03- 11- 2023

Signature (**HOD**): HoD Transportation Engineering
NUST Institute of Civil Engineering
School of Civil & Environmental Engineering
National University of Sciences and Technology

Date: 03- 11- 2023

Signature (**Associate Dean**): ______________________

Date: 04/12/2023

Dr. S. Muhammad Jamil
Associate Dean
NICE, SCEE, NUST

Signature (**Principle & Dean**): ______________________

Date: 04 DEC 2023

PROF DR MUHAMMAD IRFAN
Principal & Dean
SCEE, NUST

Form TH-4

National University of Sciences and Technology

MASTER'S THESIS WORK

We hereby recommend that the dissertation prepared under our Supervision by: (Student Name Qamar Muneer & Regn No.00000330702) Titled: **An Exploratory Analysis of Factors Leading to Prevalence of Distracted Driving Behavior in Pakistan** be accepted in partial fulfillment of the requirements for the award of **Master of Science** degree with (B+ Grade).

Examination Committee Members

1. Name: Dr. Muhammad Asif Khan — Signature: ____________

2. Name: Dr. Kamran Ahmed — Signature: ____________

Supervisor's name: Dr. Arshad Hussain — Signature: ____________

Date: 08-11-2023

____

Head of Department

HoD Transportation Engineering
NUST Institute of Civil Engineering
School of Civil & Environmental Engineering
National University of Sciences and Technology

Date: 08-11-2023

COUNTERSIGNED

Date: 04 DEC 2023

Principal & Dean

PROF DR MUHAMMAD IRFAN
Principal & Dean
SCEE, NUST

CERTIFICATE OF APPROVAL

This is to certify that the research work presented in this thesis, entitled "An Exploratory Analysis of Factors Leading to Prevalence of Distracted Driving Behavior in Pakistan" was conducted by Mr. Qamar Muneer under the supervision of Dr. Arshad Hussain.

No part of this thesis has been submitted anywhere else for any other degree. This thesis is submitted to the NUST Institute of Civil Engineering in partial fulfillment of the requirements for the degree of Master in Science in Field of Transportation Engineering at National University of Science and Technology.

Student Name: Qamar Muneer — Signature: ____________

Examination Committee:

a) GEC Member 1: Dr. Muhammad Asif Khan — Signature: ____________
Assistant Professor

b) GEC Member 2: Dr. Kamran Ahmed — Signature: ____________
Assistant Professor

Supervisor Name: Dr. Arshad Hussain — Signature: ____________

Name of HOD: Dr. Arshad Hussain — Signature: ____________

HoD Transportation Engineering
NUST Institute of Civil Engineering
School of Civil & Environmental Engineering
National University of Sciences and Technology

Name of Associate Dean: Dr. Syed Muhammad Jamil — Signature: ____________

Dr. S. Muhammad Ja
Associate Dean
NICE, SCEE, NUST

Name of Principal & Dean: Prof. Dr. Muhammad Irfan — Signature: ____________

PROF DR MUHAMMAD IRFAN
Principal & Dean
SCEE, NUST

AUTHOR'S DECLARATION

I Qamar Muneer hereby state that my MS thesis titled **An Exploratory Analysis of Factors Leading to Prevalence of Distracted Driving Behavior in Pakistan** is my own work and has not been submitted previously by me for taking any degree from this University, National University of Sciences and Technology (NUST) or anywhere else in the country/ world.

At any time if my statement is found to be incorrect even after I graduate, the university has the right to withdraw my MS degree.

Name of Student: Qamar Muneer

Date: December 5, 2023

PLAGIARISM UNDERTAKING

I solemnly declare that research work presented in the thesis titled "An Exploratory Analysis of Factors Leading to Prevalence of Distracted Driving Behavior in Pakistan" is solely my research work with no significant contribution from any other person. Small contribution/ help wherever taken has been duly acknowledged and that complete thesis has been written by me.

I understand the zero-tolerance policy of the HEC and National University of Sciences and Technology (NUST) towards plagiarism. Therefore, I as an author of the above titled thesis declare that no portion of my thesis has been plagiarized and any material used as reference is properly referred/cited.

I undertake that if I am found guilty of any formal plagiarism in the above titled thesis even after award of MS degree, the University reserves the rights to withdraw/revoke my MS degree and that HEC and the University has the right to publish my name on the HEC/University website on which names of students are placed who submitted plagiarized thesis.

Student/Author Signature: ____________________

Name: Qamar Muneer

DEDICATION

I dedicate this research work to my beloved parents and my thesis supervisor and co-supervisor whose constant support and guidance made it possible to accomplish this difficult task.

ABSTRACT

Distracted driving behavior is a major contributing factor to road accidents. To understand its impact in Pakistan, a case study was conducted to examine how different road types and conditions, trip time, law enforcement, and passenger type influence distracted driving, and this these five were main latent constructs in the analysis. This research also looked into the mediating effect of driving experience. Around 590 responses were gathered through online and face-to-face surveys, and 501 of them were considered for further analysis. The most of the respondents were males aged between 18-30 (40%) and 30-50 (27%). After conducting preliminary data testing, including data normality checks and descriptive analysis, the study employed Partial Least Squares Structural Equation Modeling (PLS-SEM) to test the hypotheses. The results revealed that all the factors had varying degrees of positive influence on distracted driving behavior. Trip timing had the most significant effect (0.196), indicating that drivers were more distracted during the daytime, possibly due to the ease of driving, allowing them to engage in non-driving activities. Road conditions (0.194), age (0.168), and law enforcement (0.161) also had significant effects on distracted driving behavior. Surprisingly, the presence of passengers did not significantly impact distracted driving in this study, contrary to some prior research findings. Older drivers exhibited more distracted behavior. Gender and education had no direct impact, but they showed significant indirect effects through driving experience, suggesting mediation. These results underscore the importance of implementing appropriate interventions involving various stakeholders (such as policymakers, police, mental health experts, advocates, etc.) to raise awareness, change behaviors, and increase risk perception related to distracted driving behavior and its dangers. Such interventions can play a crucial role in curbing distracted driving incidents and improving road safety.

Keywords: Distracted Driving, PLS-SEM, Latent Construct, Mediation, Intervention

ACKNOWLEDGEMENT

I am extremely grateful to Allah Almighty the Omnipotent for bestowing me with the strength, courage, wisdom and patience to finish my research. I wish to show my utmost thankfulness to my advisor Dr. Arshad Hussain and Co Advisor Dr. Muhammad Asif Khan, whose supervision, endless inspiration and support enabled me to finish this research. I also sincerely thank my GEC members for providing guidance and motivation throughout the process. I acknowledge and deeply appreciate the faculty of the NICE PG program for their assistance and direction during this research endeavor.

I am tremendously and honestly appreciative of all my friends for their help, support and well wishes. Finally, I owe my deepest thanks and admiration to my parents for their firm support, heartfelt prayers and best wishes that were integral to the successful completion of this research.

(Engr. Qamar Muneer)

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

LIST OF ABBREVIATIONS

Abbreviation	Complete Form
WHO	World Health Organization
CB	Covariance Based
PLS	Partial Least Square
SEM	Structure Equation Modeling
NIPLAS	Nonlinear Iterative Partial Least square
PCR	Principal Components Regression
CNN	Convolutional Neural Networks
ANOVA	Analysis of Variance
CMB	Common Method Bias
VIF	Variance Inflation Factor
AVE	Average Variance Extracted
HTMT	Heterotrait-Monotrait
SRMR	Standardized Root Mean Square
SPSS	Statistical Package for Social Sciences
RTC	Road Type and Conditions
TT	Trip Timing
LE	Law Enforcement
PT	Passenger Type
D	Distracted Driving Behavior

CHAPTER 1: INTRODUCTION

1.1 Background

As published by WHO in its recent report of 2020, Deaths caused by Road crashes in Pakistan reached 28,170, that is 1.93 percent of the total deaths in Pakistan (World Health Organization, 2020). Back in 2016, statistics of WHO regarding Pakistan revealed 27,582 deaths were because of road crashes and approximated cost of Serious Injuries and Fatalities summed up to $12,550 million which accounted for 4.5 percent of Pakistan GPD in 2016 (*Road Safety in Pakistan | Traffic Accidents, Crash, Fatalities & Injury Statistics | GRSF*, 2016).

Obtrusive reasons for accidents are easily observed and perceived where as some insidious causes for specific road crashes which are purely caused by distracted driving, are usually ruled out. Distracted Driving divert the attention of drivers while driving vehicle, thereby compromising their cognitive, emotional and physical capabilities required by the careful driving. The distracted include but not limited to talking, eating, music playing, using phone, making calls, looking at objects of interest etc. Some previous studies revealed that Cell Phone use and younger drivers are two bigger concerns areas for distracted driving (Atchley et al., 2012; Hancock et al., 2003; Neyens & Boyle, 2008).

Changing the behavior of drivers could be the key to reducing frequency or severity of the road crashes mainly caused by distracted driving. In some foreign countries like USA Traffic Survey was carried out at national level which was focused on studying attitudes, knowledge, opinions, experiences of the drivers. Studies revealed that it was also felt by drivers that distracted driving was an immense issue. Some states had expanded the programs and had been keeping the data base updated regarding distracted driving behaviors which assisted them in better planning, decision making, formulating counter policies, mitigation measures etc. regarding safety issues.

Unfortunately, Pakistan lacks such kinds of studies let alone upgrading the modern threat detection tools in vehicle, deploying Artificial Intelligence for better and safer driving.

This study intends to point out the reasons of distracted driving, what is the impact of the particular distractions upon drivers when drivers indulge in them and then determining the relative prevalence of the distracting activities across various classes of highway, while considering road type, vehicle type and trip type as well.

1.2 Problem Statement

Distracted Driving has been pointed out as the one of the major reasons for road crashes specially among young drivers. Driving under influence or driving while indulging in other non-primary driving tasks compromise decision making, awareness of surrounding and perception-reaction time and concomitant measures. Besides, due to vehicle proliferation, advanced gadgets e.g. phone; drivers are distracted now than ever. The need is to carry out the pertinent studies and find out insidious factors that compel the drivers for being involved in distracted driving behavior, their pattern and impact upon driving.

1.3 Research Objective

The objectives of the Research are given as follows:

- Finding the underlying factors that elicit the distracted driving
- Measure the relative impact of each factor upon distracted driving behavior
- Find the mediation role of driving experience
- Inculcate the demographics factor e.g. age, gender, education throughout the study

1.4 Scope of Thesis

To study the effect of various factors namely Road Type & Characteristics, Trip Timing, Law Enforcement, Passenger Type on Distracted Driving Behavior of the drivers. In addition to this the effect of Gender, Age and Education is also measured. The effect of demographic variables (age, gender, education) will be measured directly upon Distracted Driving Behavior and also indirectly in the presence of another variable “Driving Experience” known as mediator variable. Thus, mediation analysis is also performed.

1.5 Organization of Thesis

The following organization of the chapters of thesis has been made:

Chapter 1

Explains what is distracted driving and its impact upon road safety, drivers, economy. Then outlines the general finding of the germane studies conducted abroad and in Pakistan. The chapter presents an overview for aims of the research, receded by Problem statement, Research scope and organization of the chapters of the thesis.

Chapter 2

The chapter 2 contains apropos and detailed literature review explaining the previous and current studies performed in the fields of distracted driving which include but not limited to finding causes for driving distractions, their impacts, detection of distracting activities and counter measures etc. Besides, it comprises the research gap as well.

Chapter 3

This chapter focuses on the methodologies involved in the research i.e. PLS-SEM, t-value, p-value etc. It succinctly explains the PLS-SEM its benefits and the particular software required to perform the intended analyses. It provides an insight into SmartPls which is the main software for the analysis.

Chapter 4

The chapter 4 proffers the results obtained by the Analysis of the data gathered by Questionnaire and meticulously interpret the results as well.

Chapter 5

The final chapter consists of the conclusion drawn by intricate analysis, results and findings of the study. In addition to that it also states the plausible additional recommendations for future works.

CHAPTER 2: LITERATURE REVIEW

2.1 Introduction

In the recent times distracted driving seems to be one of the major responsible factors for traffic hazards, especially among drivers of less ages. The National Highway Safety Administration approximated that being distracted during driving accounts for 10 percent of all deathly accidents in the USA (Tison et al., 2011). Distracted Driving can be explained as involvement in activities that disturb or take away the substantial amount of due attention of the drivers. Distractions can affect person cognitively, visually and manually, for instance while using phone driver’s attention is distributed thus he’s cognitively impaired, same way looking at phone will disturb driver visually as he’ll use phone/music player etc. Besides, while performing these task driver’s hand or hands will be occupied rendering the drivers manually distracted. Distraction are in-vehicle and out-of-vehicle. In vehicle means the non-primary driving task inside the vehicle e.g. phone use, make-up, making hair, talking to other passenger, taking care of child etc pulling down window, setting side mirrors, fantasizing during driving etc. A person can be involved in multiple distractions at the same time e.g. taking to other passenger and adjusting music player pari passu. (Prat et al., 2017; Strayer et al., 2013; Wen et al., 2019). Out-of-vehicle are like looking at bill boards or at objects of interest or enjoying weather too much etc.

2.2 Background

Being distracted during driving drivers' surrounding and situational awareness, performance of driving and decision-making are negatively affected because of diverted attention from driving to other non-driving tasks.

The issue of distraction driving is being exacerbated by the technological advancement and their provision in vehicle, e.g. provision of GPS, satellite radios, OS on dashboard mini pc etc. in the vehicle (Lee et al., 2001; Strayer et al., n.d.).

Now a day’s vehicles (Ebnali et al., 2021) are loaded with technologies such as climate control, a navigator, weather updates, parking amenities, multimedia displays and a lot more. Indubitably drivers are being benefited from such disruptive innovations (Ahangari et al., 2019). But it's indispensable for drivers to refrain from being distracted and shall pay due attention towards driving tasks.

It is well conceded that there is a colossal detrimental effect of increasing trends of road accidents on the economy of a country thereby affecting its GDP. Besides, as revealed by a study carried out by Chen et al (S. Chen et al., 2019), it is approximated that road crashes will incur a cost of nearly as much as 1.8 trillion dollars in the next 15 years. The need of the hour is to acquire pertinent information germane to the fatal road incidents and devise methods to minimize them.

It has been found out that Driver Distraction (DD) is a majorly responsible for exacerbating the already dwindling situation of road safety (Dingus et al., 2016; Oviedo-Trespalacios & Regan, 2021)Also, with the invention and application of Level 2 Automated Driving Systems (ADS), the responsibility of the driver is more central in monitoring such systems and technologies that demand unhindered attention of drivers.

Even with the provided facility of level 3 ADS, the driver role as a backup or fallback solution, is crucial with respect to safety. It is needed to consider the driver's availability to overtake the steering task. Use of Artificial intelligence and automated system in driver monitoring systems, or occupant monitoring in the absence of vehicle driver are becoming new and necessary trends in Europe for newly designed vehicles laced with automated facilities, whose protocols are commensurate with the one laid out by the NCAP.

Driver distraction can be thought of as engagement with task or activities that will distract or avert the drivers' attention away from what is inevitable for efficient and healthy driving (Regan et al., 2008).

2.3 Driving Distractions

All the task or activities in which drivers indulge themselves deliberately or un-intentionally which avert their attention from driving, be it physically, mentally or visually can be referred to as Driving Distractions or shortly abbreviated as DD.

There're numerous types of activities in which drivers can be observed. NHTSA has pointed out some of the tasks which are frequent and common as well. They are given as follows:

- Conversing to other passengers in the auto-mobile

- Drinking or eating or taking medicine
- Calling on phone and holding the phone
- Calling on phone while using hand-free
- Reading news, feeds or pod cast etc.
- Reading text or emails on phone
- Forwarding/sending text messages or e-mails
- Taking care of children in the vehicle
- Doing make-up, hair making, shaving or looking in the
- Adjusting the music player
- Singing along with the song/music being played
- Using a computer e.g. laptop or PDA
- Watching a film
- Setting the GPS for directions

Besides, the survey was also carried out in order to affirm that what was the reason for indulging in a particular or set of particular activities and what the drivers would do after being involved in distracted driving.

Following reasons/impacts were put forwarded to the respondents to have their view regarding causes and impact of driving distractions in which they find themselves.

- Importance of the call
- Depends upon who is calling
- Phone availability
- Call is related to work
- Call is social or personal
- Call is as per routine
- Call is unprecedented
- Call is from someone familiar
- Call is from unknown number
- Exquisite weather condition
- Low traffic on road
- Time of the day or night

- Try to travel at low speed
- To avoid boredom or dozing
- Law is not strict
- Other people do it as well
- I need directions

2.4 Introduction to PLS-SEM

As the word PLS-SEM suggests that this technique consists two methodologies involved in its root. First is PLS which stands for Partial Least Square and the 2nd term is SEM which means Structure Equation Modeling. That's why next portion in the writing will explain these two things separately. SEM is technique which uses regression and other statistical tools to explore the connection between variables. In order to do that it requires computation of coefficients for system of equations involved in regression, those coefficients are measured by Partial Least Square (PLS) Method.

SEM is actually a set of statistical tools and methods which are applied combine to achieve the purpose of the study. It measures the variable which we can't directly calculate. The dependent variable or the variable which can't be measured directly are called "Lateral Constructs" and the variable explaining the dependent variable are called "observed variables". The technique also offers effects of a third variable upon the prediction of a dependent variable. In simple words change in relation between independent variable and dependent variable in the presence of a 3rd variable e.g. whether the presence of the 3rd variable weakens or strengthens the relationships among independent and dependent variable or not, which in term of SEM is called "moderation analysis". Then it's also checked whether 3rd variable is of significance or not meaning by, does the independent variable affect the dependent variable directly or it affect indirectly through the involvement of 3rd variable. This is called "mediation analysis". As far as the word Partial Least Square is concerned this is simply a technique to reduce multi-collinearity among independent variables when they have large dimensions. This improves the prediction of dependent variables. Another method to reduce collinearity is Principal Component Analysis but use of this method makes results precarious.

Following topic expatiates about PLS and SEM; their history, their creator, how they evolved over the period of time and how do they differ from other techniques offering the same solutions, some mathematical background for coefficient computation and involved algorithms. Cessation contains name and some description of the software used in this field (Mateos-Aparicio, 2011).

2.4.1 Partial Least Squares (PLS) Methods

Partial Least Square is a method used for computing coefficients for different system of equations involved in regression. The effect of independent variables is observed on dependent variables. To reduce the collinearity between the independent variable and to better explain the dependent variables, PLS methods is used. The next headings explain its origin, mathematical background, purpose and effectiveness etc.

2.4.1.1 Origins of the Partial Least Squares (PLS) Methods

The origin of the PLS methods can be attributed back to the Swedish Professor Herman Wold. He was mentor of Karl Jöreskog who founded Structure Equation Modeling techniques. Colleagues of Herman mostly employed maximum verisimilitude in their analysis, on the contrary, Herman always went for the method based upon least square approaches. He invented Fixed Point Algorithm that he used to measure the coefficient of system of simultaneous equation. Later he modified it and extended the method to compute the principal components and canonical correlations using iterative procedures. This work was the basis that led Herman to the creation of PLS techniques as highlighted by Wynne Chin (Chin, 1998). These two methods laid foundation for the Nonlinear Iterative Partial Least square (NIPLAS) algorithm in which Wold employed Ordinary Least Square method to compute principal components by an ordered iteration of regressions. Besides, he applied NIPLAS to compute canonical relations by an ordered iteration of multiple regressions.

2.4.1.2 PLS Regression Vs PCR (Principal Components Regression)

In regression while computing coefficients of a model, sometimes independent or explanatory variables contains extreme dependence relationship among them which is called multi-collinearity. This multi-collinearity renders the computed coefficients insignificant leading to faulty or erroneous prediction of the variable being interpreted.

To get rid of multi-collinearity the dimensionality is decreased of the explanatory variables and this is done by getting a set of new explanatory variable which are free from multi-collinearity. The Principal Components Regression (PCR) method was widely famous to reduce the dimensionality and multi-collinearity. The PCR carious out principal component analysis and find a set of new explanatory variable which are free from multi-collinearity. But using this method is like choosing between Scylla and Charybdis as the method reduces dimensionality but it compromises the prediction of dependent variables which make this method precarious.

Therefore, new technique was required. Here came PLS-R. It was an extension of PCR but with additional regression step. This method chose the principal components to define covariance between X and Y. In simple words this method aims to extract the explanatory variable while having most of the variation of real explanatory variable and the chosen set can also be used to model behavior of dependent variable. That’s why the PLS regression technique is relatively more suitable for prophetic task (Chin et al., 2003). Similarly, Barcely (Barclay et al., 1995) also conceded that PLS-R is a good recommendation for predictive studies.

In short both techniques PCR and PLS-R intends to reduce the multi-collinearity by choosing another new variable who have less multi-collinearity but the former technique compromises prediction of dependent variables whereas latter one doesn’t.

Initially the method was confined to Social Sciences only but in 1983 son of Herman Wold expanded the idea and they applied it in another field as well. Rather, he along with Harald Martens broadened the PLS-R and came up with new techniques for analytical chemistry(Valencia et al., 2003).

Benefits:

- It solves problem of multi-collinearity in models of regression
- It doesn’t compromise the interpretation of dependent variable
- It can work with small amount of data.
- PLS does not require the data to be normally distributed (Falk & Miller, 1992).

2.4.1.3 PLS Regression and PLS-Path Modeling

PLS Regression is technique of multivariable nature and its purpose is to remove the problem of multi-collinearity issue for explanator variables in regression. It perform the required action by reducing the dimensionality of independent or explanatory variables and at the same time not maring the prediction accuracy for dependent variables.

PLS Path modeling is use of PLS for SEM and in computes the coefficients for structural equations by emplying partial least square method rahter than ordinarly least square method. It assumes the structural model to be linear therby validatign use of regression methods to calculate coefficients.

2.4.2 Algorithms for PLS Methods

There exist two famous algorithms for modeling via partial least square methods. These were invented by Wold and later on his successor modified and extended its applications. The history, refinement, application, benefits over PCA, advantages and extensiveness are explained before an further technical explanation is given in the following sections.

2.4.2.1 Algorithms for PLS Regression

There exist two methods for PLS Regression based upon nature of dependent variable. First method called PLS-1 is used when dependent variable ‘q’ is univariate meaning by one dependent variable has to be interpreted by set of explanatory or independent variables ‘p’ and therefore used univariate regression methods. Second method PLS-2 involves prediction of multi-variant dependent variables and thus involves multivariate regression techniques. In case of PLS-2 $q > 1$ means we’ve more than one dependent variable to be explained by p -explanatory variables.

PLS-1 is quite simple it consists extraction of first component which leads to computation of remaining components from first component while maintaining the orthogonality. Orthogonality means absence of collinearity.

Succinct explanation of this method is described below (Valencia and Diaz-Llanos, 2003). The 1st component can be described as follows:

$$t_1 = w_{11}x_1 + w_{12}x_2 + \cdots + w_{1p}x_p = \sum_{j=1}^{p} w_{1j}x_j$$

Here x_j refer to explanatory variables and y is the variable to be interpreted. The w_{ij} coefficients can be given as:

$$\omega_{1j} = \frac{cov(x_j, y)}{\sqrt{\sum_{j=1}^{p} cov^2(x_j, y)}} = \frac{(x_j, y)}{\sqrt{\sum_{j=1}^{p} (x_j, y)^2}} \quad with\ j = 1, 2, 3, \ldots \ldots \ldots \text{p}$$

In order to compute coefficients we need to find dot products (x_j , y) for each j value from 1 to p.

If we consider that single component model is not enough, the need for second component becomes inevitable. This second component can be shown by t_2 and can be thought of a linear combination which contains regression error of x_j variables on first constituent. Component orthogonality is also confirmed this way. So, that's why we need to compute residues too, which for single component regression can be given as follows:

$$e_1 = y - \hat{y} = \hat{\beta}_1 \ldots t_1 \ with\ \hat{\beta}_1 = \frac{(y, t_1)}{||t_1||^2}$$

Moving towards 2nd component which can be calculated by following equation:

$$t_2 = w_{21}e_{11} + w_{22}e_{12} + \cdots + w_{2p}e_{1p}$$

Where coefficient are as follows:

$$\omega_{2j} = \frac{cov(e_{1j}, e_1)}{\sqrt{\sum_{j=1}^{p} cov^2(e_{1j}, e_1)}} = \frac{(e_1, e_{1j})}{\sqrt{\sum_{j=1}^{p} (e_1, e_{1j})^2}} \quad with\ j = 1, 2, 3, \ldots \ldots \ldots \text{p}$$

As for as residues e_{ij} are concerned, simple regression of x_j upon t_1 i.e. $x_j^* = \hat{\alpha} t_1 (j = 1, \ldots, p)$ gives us following residues.

$$e_{1j} = x_j - x_j^* = x_j - \hat{\alpha}_j \,.\, t_1$$

And calculation of regression coefficients has also been done thus

$$\hat{\alpha}_j = \frac{(x_j, t_1)}{||t_1||^2}$$

Dot product of (e_1, e_{ij}) for j=1…p leads us to computation of t_2.

The following procedure is iterated until calculation of significant no of components that need to be retained.

PLS2 algorithm is just expansion of PLS-1 but it requires further use of PCA in order to reduce multi-collinearity.

2.4.2.2 Algorithms for PLS Path-Modeling

Indubitably, NIPLAS is the most beneficial tool for PLS-Path Modeling which was created by Professor Wold back in 1966. The method is versatile in nature as it doesn't need hardcore condition to work with like large data samples, data normality. It performs number of iterations in order to compute the required coefficients without requiring normalized and enormous data.

There exist essentially two steps in the Path Modeling algorithm of PLS: the computation of the measurement model and the calculation of structural model parameters. Latent variables are obtained by linear combination of weighted attributes given by the computation of the measurement model, whereas, relation among latent variables is revealed or interpreted by coefficients which are calculated by estimation of the structural parameters. Most algorithm employ linear regression (which uses least square method) to measure structural coefficients. But if profusion of latent variables is present, we need to use PLS-Regression to cater for the issue of multi-collinearity.
Fornell algorithm is the simplest and useful technique for most researches and a basic representation of PLS-Path modeling.

2.4.3 Structural Equation Modeling (SEM)

SEM is a statistical method that studies the associations among various variables in a simultaneous way.

It is not a single procedure or technique but instead a family of pertinent statistical methods. It examines the relationship among dependent and independent variables, effect of other variable upon dependent and independent variable relations and prediction. It's a versatile technique and let the dependent variable be independent variable for further analysis or prediction. SEM is somewhat like multiple regression but much more flexible and inclusive.

2.4.3.1 Brief History of SEM

Bollen (Bollen, 1989)asserts that the development of SEM was based on three key analytical advancements: (1) path analysis, (2) latent variable modelling, and (3) general covariance estimation techniques. Sewall Wright created the fundamental path analysis in 1934, which calculated the correlation matrix of observed data to assess the link between variables. Later, this approach was expanded to social sciences, including psychology, sociology, and economics.

The exploratory factor analysis, which is regarded as another significant origin component for SEM, was created by psychologist Charles Spearman. Interdisciplinary integration led to the first generation of SEM which was accomplished by Wiley, Keesling and Joreskog. Further the concept of Maximum Likelihood Estimation was also introduced into SEM which assisted in computation error and direct effects (Golob, 2003; Kline, 2011; Mateos-Aparicio, 2011).

As (Matsueda, 2012) explains, SEM went through mainly four stages for its development

i. Path Analysis Development and the concept was later expanded.
ii. Inter disciplinary growth between psychology, sociology and economics which produced empirical SEM applications.
iii. Development of methods to cater for ordinal, discrete and limited dependent variables
iv. In fourth stage statistical approaches were inculcated into SEM network

2.4.3.2 Theories behind Structural Equation Modeling (SEM)

Two renowned analysis methods, path analysis and confirmatory factor analysis are basically types of Structure Equation Modeling (Tabachnick & Fidell, 1996). Lei & Wu (Lei & Wu, 2007) suggested that SEM can be thought of an extension of general linear modeling (GLM).

2.4.3.3 Model Specification in SEM

Usually, SEM contains two components, a structural model and a measurement model. The measurement model is designed to check the reliability and validity of the constructs. Construct is the dependent variable which can't be measured directly and that's why it's studied or examined through set of explanatory variables called Observed variables. The structural model on the other hand help to assess the relationship among lateral constructs. Besides,

SEM offers graphical representation by means of a path diagram. Figure 1 shows some common symbols used in SEM. Figure 2 illustrates SEM models briefly.

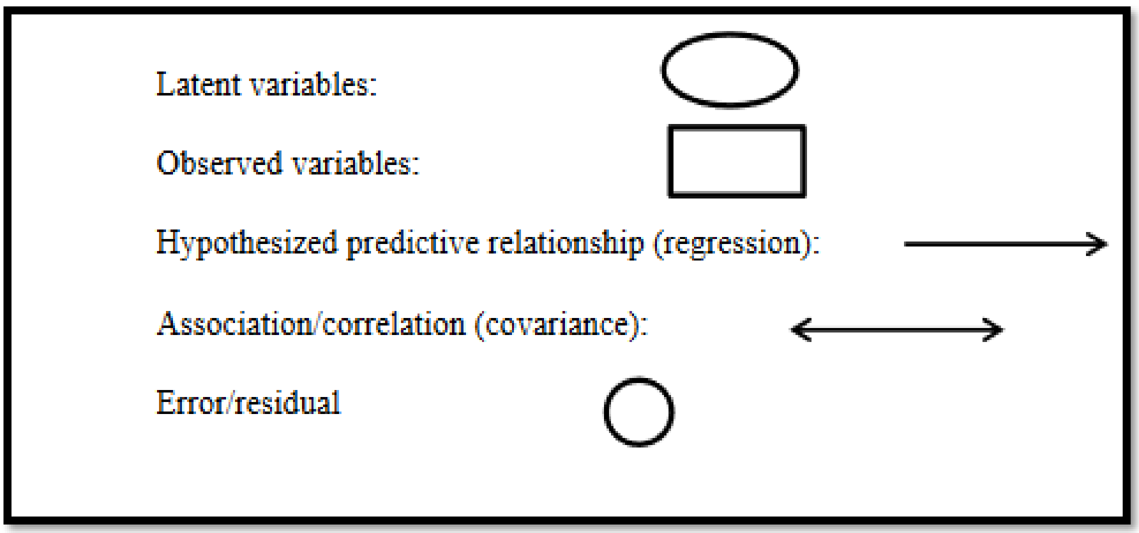


Figure 2-1: Common Symbols used in SEM with their purpose

The *Figure 2-1* shows a schematic representation of SEM where we've 6 variables called "observed variables" shown in rectangular boxes and then we've error related to computation of those variable which are shown in circle and range from e1 to e6. In the Oval we have dependent variables (DV), Factor 1 and Factor 2. These two are "lateral constructs". Lateral Construct "Factor 2" is being measuring by observed variables 4, 5 and 6. Whereas, Factor 1 is computed by Variable 1, 2 and 3; also, Factor is being associated with Factor 2 as well. The arrows between variables and factors are called paths.

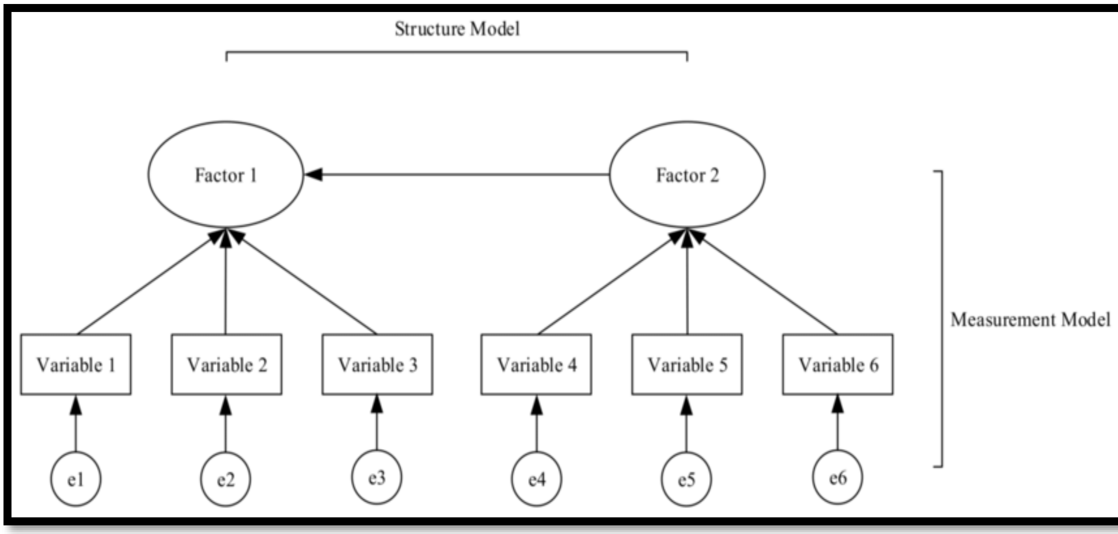


Figure 2-2: A schematic representation of SEM based model

The fundamental structure of SEM can be considered of a theorized model that has a number of hidden parameters that respond to (1) the regression coefficients, and (2) the covariance and variance of the independent variables (IV) in that model (Bentler, 2006). The simple regression equation can be stated as:

$$y = \beta_1 x_1 + \beta_2 x_2 + \beta_3 x_3 + \cdots . \beta_n x_n + \varepsilon$$

'y' is DV and x and ε are IV, ε is actually error term and β is coefficient.

Generally, in matrix algebra form it can be written as

$$\eta = \beta\eta + \xi\gamma + \varepsilon$$

If q is assumed to be number of dependent variable and r is assumed to be no of independent variable, then η is a $q \ x \ l$ vector of dependent variables, β is $q \ x \ q$ matrix of regression coefficient among dependent variables, γ will represent a $q \ x \ r$ a matrix for regression coefficient among independent and dependent variables. And ξ represent *r x l* vector for independent variables and ε shows regression error matrix.

The later step in the SEM is estimating the parameters which can be calculated by various method as explained earlier. And some other methods are un-weighted least squares (ULS), Generalized Least Square, Asymptotically distribution free and Browne's method. I'll restrict to my method of PLS and will not explain the afore-mentioned in details as these methods out of the scope of the thesis but meticulous data can be found in (Kline, 2011; Raykov & Marcoulides, 2012; Tabachnick & Fidell, 1996)

The next and the last step in SEM involves checking goodness-of-fit or fitness of our model. There are numerous criteria to check the intended objects. Mostly different kinds of indices are employed. They include p-value inter alia t-value, chi-square ($\chi 2$) test, Tucker-Lewis index, Normal fit index etc.

2.4.4 Partial Least Square Structural Equation Modeling (PLS-SEM)

Along with development of PLS-R method there was extension of PLS methods. This 2nd line of PLS technique was based upon method of Joreskog. It was also structure equation modeling but it was based upon covariance and it would measure coefficients for set of equations by altering and adjusting the covariance matrix. Best fit for theoretical covariance matrix was achieved where theoretical covariance matrix was obtained from the model and preliminary empirical covariance matrix. But this technique was hard core as it required large data sets and multivariate normality.

On the contrary SEM based upon PLS an approach put forwarded by Professor Herman Wold was soft core, its requirements for data were low. PLS techniques use partial least square method to compute coefficient for system of equations and the same time not demanding data normality and provision of large sample.

PLS-SEM gained popularity after creation of software for this peculiar yet robust and efficient technique. It was designed by Lohmoller in 1984 and was called LVPLS ver1.6

2.4.5 Availability of Software for the PLS and PLS-SEM Models

SmartPls: Smart Pls is the latest and most unique and versatile software. It is user friendly and doesn’t require data normality. It offers structural and measurement model in the same model. Provides old and new checks for construct validity & reliability and discriminant validity, mediation vs moderation analysis and lot more. It employs PLS in its analysis for SEM. The study will use this software for research purpose.

R package OpenMx: It has graphical capabilities featuring R and it is open source and has inherent ability to optimize user-specified objective functions.

LISERAL: Provides General linear modeling and multilevel modeling, good for measuring residuals in SEM and factor analysis.

Mplus: can handle variety of data types, e.g. continuous, categorical and has latest missing data handling techniques, also has Bayesian SEM.

EQS: It is more capable for exploratory analysis, estimation of reliability, can handle missing values.

SPSS: It's very common and prevalent software, provides factor analysis, correlation, regression and analysis of variance for SEM. Besides, it has graphical user interface. LS-GUI, PLS-Graph, SPAD-PLS, XLSTA, LVPLS are also among the list.

2.5 Previous Studies on Distracted Driving and SEM

A lot of researches have been performed in the area reaching different conclusion as to what causes the distracted behavior and what is the most prevalent type of distracted activity among drivers of different age, salary, driving experience, gender and profession etc. M Rezapour (Mahdi & Khaled, 2022) identified the responsible causes for bad driving. MNL was applied and it was pointed out that the driving behavior is conflated with demographic characteristics, e.g., having car of different type, age and experience, etc.
(Mahdi & Khaled, 2022).

Cai et al studied effect of deprived sleep on driving. The study included working people, young and old. The young drivers shown more lane departures than old drivers and an enhanced amount of risk of near crash events as compared to the old drivers (Cai et al., 2021). In a research conducted by (Brodsky, 2001) the effect of paper music tempo on the performance of drivers was studied. (Horberry et al., 2006) studied the impact of in-vehicle entertainment facilities and hands-free on driving behavior of drivers of various ages; and it was observed that distracted drivers tended to reduce their speed and were less aware of any dangers. The effect thickened when drivers were distracted visually. (Fitch et al., 2013) also noticed resembling results. Similarly, the bearing of phone conversations and passengers on drivers' performance was analyzed by (Drews et al., 2008). Higher lane position variance was observed for drivers talking on the phone; besides, increased distance headway was seen for distracted drivers. The drivers texting on phone were found in slower driving conditions overall and increased braking time for such drivers was observed. Also, an enhanced risk of crash was highly possible (Drews et al., 2009; Yannis et al., 2010). (Owens et al., 2011) carried out a research to observe the impact of texting while using handheld as opposed to an integrated system. The study concluded that texting drivers glance away longer, leading to increased steering variance and velocity, indicating difficulty maintaining lateral position, requiring quick, large steering wheel adjustments. The aftermath of texting under varying environmental and weather conditions was studied by (Yannis et al., 2014), and these conditions proved to have an effect on driving performance. (Ortiz et al., 2018) extended the work and found out that while texting both the number and duration of lane excursions increase, resulting in a higher standard deviation of lane position. Older drivers show a greater tendency to drive out of their lanes for longer periods compared to younger drivers. Additionally, they exhibit reduced lateral control and a higher likelihood of being involved in crashes. (Ahangari et al., 2020) performed extensive and inclusive research and studied the effect of various distraction activities on speed and lane changing while being on different classes of highways such as freeway, local roads, etc. The author noticed that the distracted behavior was more pronounced when the driver indulged in eating or drinking. While several researchers have focused on the impact of distracted driving on driver's performance and crash probability, some have extended the concept to construct prediction and detection models. For instance, (Murphy-Chutorian et al., 2007) developed Machine Learning models to identify distracted driving behavior based upon certain movements of the drivers; whereas, (Li et al., 2013) also developed models for evaluations of distractions subjectively. Similarly, (Tango & Botta, 2013) created models using Support Vector Machine (SVM) to detect distracted driving based upon some input variables including speed, time lane change and collisions, steering, position of the brake and accelerator. (Ahangari et al., 2019) created a model for predicting distracted driving via employing machine learning techniques. The very author then refined the concept to observe various distracted driving patterns across different classes of highway (Ahangari et al., 2021). In Pakistan (Javid & Faraz, 2017) conducted a study to measure the effect of driving distractions on crash risks

2.6 Research Methodologies, Experimental Setup and Measures in Previous Studies

This section describes the different methods employed in the area of distracted driving studies and what were the environment/set-up for the studies and what was being measured. The methodologies they used, the results they found and what they referred and concurred. To some extent the limitations and constraints are also defined.

2.6.1 Experimental Environment

There are two major classifications for the environments of experiments real and simulated. In the real environment people actually drive on roads are observers observe the drivers and their behavior either by sitting with them or by recording them.

The real data are collected via the following way (Papantoniou et al., 2017; Shahverdy et al., 2020)

- Sitting with a driver and noting his or her distractions
- Road side observations
- From CCTV cameras or video recorded by observers
- Camera installed in the vehicle

On the other hand, Simulated environment involves volunteers and a driving simulator. The driver or participants provides demographic survey and then drives the vehicle in the simulator and all the apropos parameters are calculated/observed manually or the simulator is programmed to record them automatically and later on the data is extracted from the simulator (Papantoniou et al., 2015).

2.6.2 Data Collection Methods

Over the course of times various researches have employed various methods in order to get the relevant data for their research.

Some of them are given as follows:

- **Sensor**

There exists a vast variety of sensors that can be of use for the detection of Driving Distraction. This can be via gather data on performance measures. For example, one of the sensors that had been in use is a smartphone's accelerometer, gyroscope, and magnetometer to get performance measure data (Shahverdy et al., 2020; Yu et al., 2018).

A lot of researches have involved programmed cameras in order to

detect visual features that could help in detection of Distracted Driving (J.-C. Chen et al., 2020; Owens et al., 2011).

- **Simulator extracted Data**

In Simulated environment there are volunteers and a driving simulator. The driver drives the vehicle in the simulator and all the apropos parameters are calculated/observed manually or the simulator is programmed to record them automatically and later on the data is extracted from the simulator (Ahangari et al., 2019, 2020, 2021; Cassidy & MacDonald, 2009).

- **Questionnaire**

The reason a lot of researches have involved questionnaire in their research because it's cost effective and less hectic, data can be obtained directly in the desired format and can be easily manipulated as per analysis. Besides, it requires no volunteers and it's relatively safer as no physical driving is performed (Brodsky & Slor, 2013; Cassidy & MacDonald, 2009).

2.6.3 Analysis Methods

Different studies new and old have used different method for the required analysis.

Some of them are given below:

SEM, PLS, Multiple Logistic Regression, two-way ANOVA test, ANOVA test, p-value test, t-test, Multiple Linear Regression, Liner regression are used for analysis and hypothesis testing whereas the art of state new methods of Machine learning inter alia CNN, ANN and SVM have been used to detect the driving distractions (Ahangari et al., 2019, 2021; Cai et al., 2021; Mahdi & Khaled, 2022).

2.6.4 Experimental Measures

Various studies (Ahangari et al., 2019, 2020, 2021; Brodsky & Slor, 2013; Cassidy & MacDonald, 2009; Murphy-Chutorian et al., 2007; Owens et al., 2011; Son & Park, 2021; Torres et al., 2019; Yu et al., 2018) based upon purpose, methodology and data; have computed or observed different measures, which are as follows:

- Measuring brake and its effect as a response towards the action of a distracted activity
- Effects of Distraction upon speed of the vehicle
- Lane position and lateral control of the vehicles
- Steering of the vehicles
- The increased or decreased headway of the vehicles and longitudinal control
- Driver's physiological and psychological responses
- Factors affecting driver's behaviors

2.7 Driving Distractions Trends in The World

In developed countries like Australia, studies have revealed that one person in every two drivers are found engaged in driving distractions e.g. phone use, browsing etc. (Oviedo-Trespalacios et al., 2017). Similarly, UK has 22-30% of drivers who indulge in visual and manual distractions, making or receiving phone calls while driving on daily basis (Sullman et al., 2018).

In Colombia (Oviedo-Trespalacios et al., 2017) researches approximated that nearly 78% of drivers with age range of 15-25 years old are occasionally found using phone while driving vehicles. (Truong et al., 2016) performed in Vietnam shown that out of 26,300 riders monitored at 12 road sites that as much as 8% drivers were found using a mobile phone.

In state of Iowa, from 2001 to 2010, around fifty-five hundred motor vehicle accidents resulted mainly from cell phone use as reported by GTSB. Theoretically, distracted driving is an activity of set of activities that could can turn away the driver's attention from primary task of driving; those distraction activities posse a greater risk of driving error and crash involvement.

The most frequent and major driving distraction actions as defined by NHTSA are texting inter alia grooming, reading, using navigation system, watching movie, talking to other people in vehicle, enjoying weather scenery, looking at people or objects of interest, eating and drinking, singing along being played songs. According to the studies from national phones surveys on distracted driving behavior and attitude, conducted in 2011 the most frequent distraction is conversing with passenger while driving, almost eighty percent of drivers do it. Followed by adjusting the radio activity which is found in 65% of people who are driving. Around 25 percent of the drivers were involved in using phone. Young drivers e.g. age of 25 years were three time more prone to reading or sending text and emails (Tison et al., 2011). Back in 2005, a 100-car naturalistic driving study was conducted by the Virginia PISU, Virginia DOT and NHTSA, where around exceeding time of 42,000 hours of data was recorded which accounted around over two million vehicles the study revealed that 80 percent of the road accidents are because of lack of driving attention and the behavior was more pronounced among younger drivers (Hanowski et al., 2006).

Figure 2-3 shows the overall %age of unfocussed drivers who met deadly crashes in USA countrywide. It depicts that most drivers who had fatal crashes were less than 20 years old (Vermette, 2010)

Whereas Figure 2-4 presents chronological deaths as revealed by NHTSA for time span of 2010-2019.

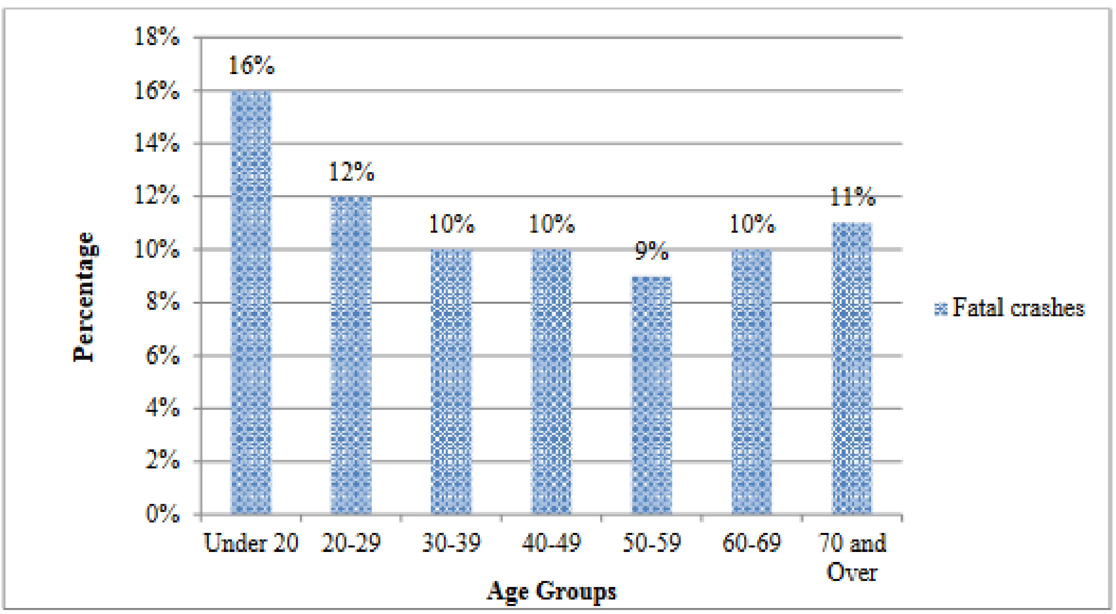


Figure 2-3:Age wise distribution of distracted drivers who died in road crashes
Image courtesy: Vermette, E. (2010).Curbing distracted driving 2010 survey of state safety programs. Washington, DC: GHSA.

Figure 2-4: NHTSA deaths statistics caused by distracted driving
https://www.cdc.gov/transportationsafety/distracted_driving/index.html

2.8 Driving Distractions Trends in Pakistan

Overall, there has not been done enough work in the area of distracted driving behaviors in Pakistan like in other countries. A lot of apropos studies and research are required to fill the lacuna in this very field and in order to circumspect the security threats faced by modern days' innovations with respect to vehicles and driving. Some of the related researches which have been done in the field of distracted driving are given below.

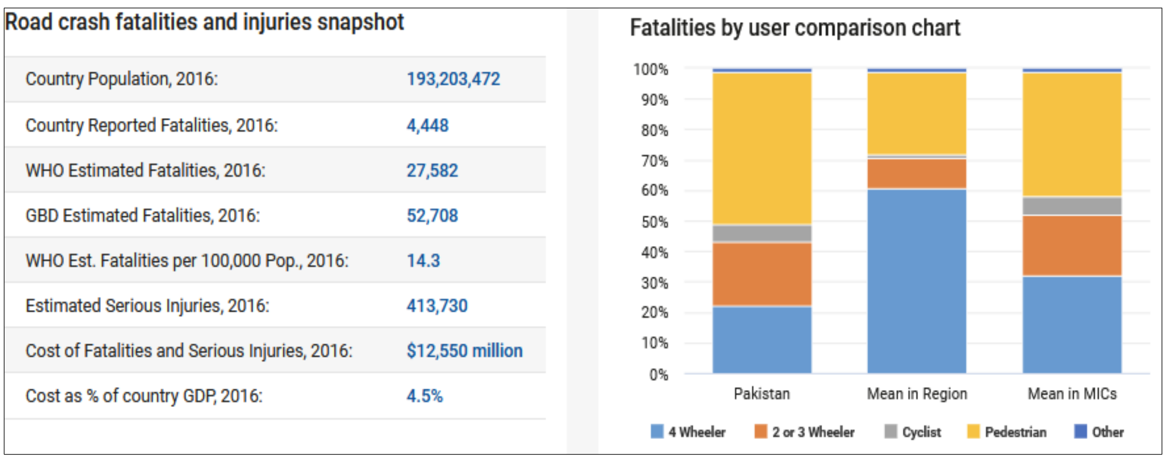
(Khan et al., 2020) interviewed young drivers who were university students and had experienced road crashes.

Analysis of the collected data revealed various driving distraction e.g. weather condition, sleep deprivation etc. causing road crashes in Pakistan. Besides, the effect was more accentuated in the youth. It was concluded that proper education was required to tackle the situation.

Another research performed in Lahore by Javid (Javid & Faraz, 2017) shows that economic development in Lahore is directly linked with proliferation of vehicles thereby increasing crash probability. Data regarding distracted driving was gathered with the help of questionnaire and its analysis revealed that young drivers indulge mostly in looking out of vehicles, music listening and using phone. These activities resulted in dangerous aftermaths. Formulation of proper policies and enforcement was provided as a solution the problem of distracted driving.

Ajmal Khan (Khoso, 2019) showed that lack of enforcement of law on N-5 has made the drivers indulge mostly in distracting activities and not following safety protocol as well, e.g. not wearing helmet and over-speeding.

As published by WHO in its recent report of 2020, in Pakistan Deaths due to road crashes reached 28,170, that is 1.93 percent of the total deaths in Pakistan. Back in 2016, statistics of WHO regarding Pakistan revealed 27,582 deaths were because of road crashes and approximated cost of Serious Injuries and Fatalities summed upto $12,550 million which accounted for 4.5 percent of Pakistan GPD in 2016. Figure 2-5 explains the stance with respect to facts and figures.



Road crash fatalities and injuries snapshot

Country Population, 2016:	193,203,472
Country Reported Fatalities, 2016:	4,448
WHO Estimated Fatalities, 2016:	27,582
GBD Estimated Fatalities, 2016:	52,708
WHO Est. Fatalities per 100,000 Pop., 2016:	14.3
Estimated Serious Injuries, 2016:	413,730
Cost of Fatalities and Serious Injuries, 2016:	$12,550 million
Cost as % of country GDP, 2016:	4.5%

Figure 2-5:Road crashes statistics revealed by WHO and GBD

Image courtesy: https://www.roadsafetyfacility.org/country/pakistan

2.9 Research Gap and Hypotheses Development

The extensive examination of the aforementioned studies primarily centers around two main aspects. Firstly, it involves observing and quantifying the effects of various distraction activities (such as phone usage, conversing with passengers, eating, etc.) on a driver's driving performance, as well as their physical and psychological conditions, and how these activities contribute to an increased risk of accidents. The second main focus is on developing AI models to identify and predict instances of unfocused driving behavior.

However, the core issue of distracted driving, which leads to numerous negative consequences, has not been thoroughly investigated. Specific details, such as the factors linked to distracted driving behavior and the circumstances in which drivers are more likely to remain inattentive, remain underexplored. Only limited literature exists that specifically analyzes the relationship between various road and environmental factors and distracted driving. For instance, Foss and Goodwin (2014) conducted an analysis specifically aimed at identifying the factors responsible for distracted driving behavior in adolescent drivers, along with the relative prevalence of different distraction activities. Similarly, Ahangari et al. (2021) studied how distracted driving is associated with different classes of highways. Nonetheless, a comprehensive and thorough study is still needed to determine the impact of certain factors on distracted driving, unlike other studies that primarily focus on the effects of distracted driving on other parameters.

Therefore, following hypotheses were formulated to delve into the research and fill the lacuna.

i) H1: Road Type and Conditions have Significant Impact on Distracted Driving
ii) H2: Driving Timing Significant Impact on Distracted Driving
iii) H3: Law Enforcement has Significant Impact on Distracted Driving
iv) H4: Passenger Type has Significant Impact on Distracted Driving
v) H5: Driving Experience Mediates the Relationship between Age and Distracted Driving
vi) H6: Driving Experience Mediates the Relationship between Gender and Distracted Driving
vii) H7: Driving Experience Mediates the Relationship between Education and Distracted Driving

The following Figure 2-6Figure 2-6 illustrates the hypotheses to be analyzed.

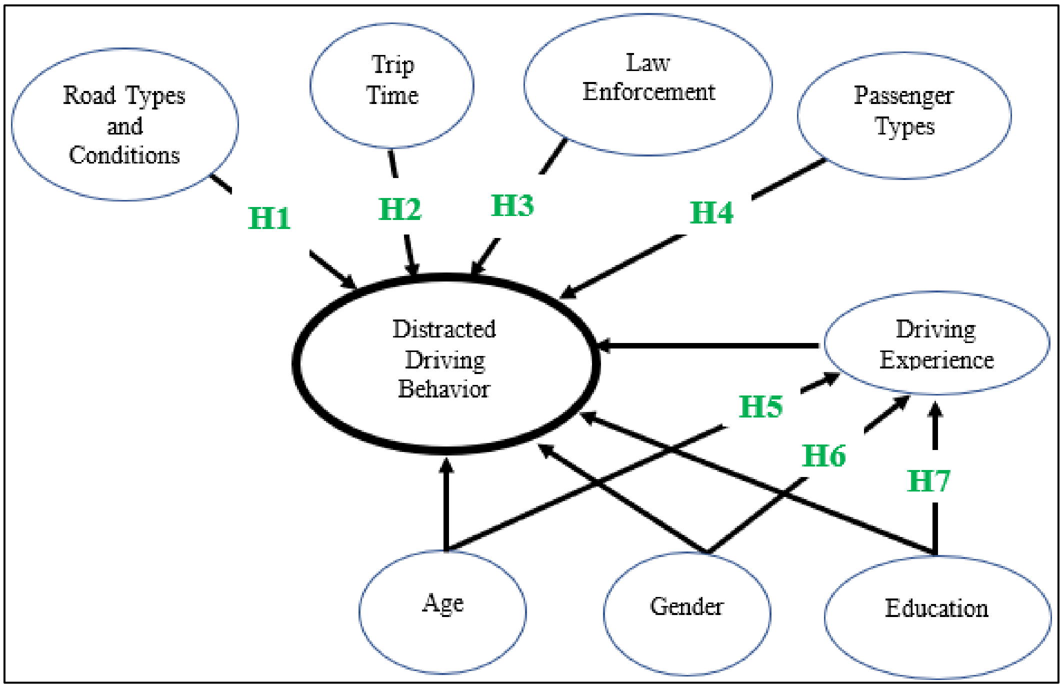


Figure 2-6:The Proposed Hypotheses

2.10 Summary

The chapter initially discussed the background of the research and provided brief introduction as well. Later on, Distraction were detailed along with most frequent types of driving distractions and their impacts once driver indulge in the distracted activities. Then discussion included some previous researches related to our field which explained about different methodologies for the research, various data extraction methods used in studies, variety of experimental environments in which studies were conducted and various parameter/end results which were measured in them. Followed by the researches, results and distraction trends in the other countries especially USA. Finally, the related studies conducted in Pakistan were briefly explained along with the result, conclusion and recommendation. Also, the research gap was highlighted which led to this research; therefore, hypotheses were formulated to perform the research. Besides, PLS-SEM was discussed separately and its use in previous researches.

CHAPTER 3: RESEARCH METHODOLOGY

3.1 Introduction

This chapter offers the methodology used to assess and tests the hypotheses developed after literature review in chapter 2. This is a questionnaire bases research which gathers responses from people of varying demographics, the responses are then run through some basic test which produce the final responses to be utilized to proceed with the research's next stages. This is followed by descriptive analysis and then use of SmartPLS for developing model to investigate the intended purposes, where a series of tests are performed again and compared against specified set of criteria which help assess the model reliability and to concur the analyses. The chapter will expatiate the steps involved from cradle-to-grave for carrying out the research.

3.2 Research Methodology

The next sections discuss in details the methodology and stratagem put into effect to carry out the research. It starts with questionnaire preparation its initial testing, finalization, data collection and sorting, then organizing the data. It is followed by basic and primary analysis and then comes main analysis of Structure Equation Modeling, where hypotheses are tested and the relations are explored.

3.2.1 Designing the Questionnaire

To examine factors influencing distracted driving behaviors, I designed an original questionnaire after extensively reviewing prior literature in this domain. Based on review of over 20 published studies on distracted driving in various countries, key variables were identified, including demographics, driving experience, road conditions, trip timing, passengers, and law enforcement. With these constructs in mind, I compiled a pool of closed-ended questions with Likert scale (1-5) response options that could effectively capture data on these variables.

The language of the questionnaire was English & Urdu. Participation was open to all drivers with adequate experience, not limited by demographics. The questionnaire had three sections. The first section obtained respondent demographics like gender, age, education, income, driving experience, vehicle type, and crash history. The crash severity scale KABCO from the Federal Highway Administration (FHWA) was used to categorize past crash levels. The second and largest section asked about distraction likelihood under various conditions, using a 5-point Likert scale. This covered highway types, day versus night driving, passenger types, and law enforcement contexts. The third section evaluated specific distracted behaviors like phone use while driving. It also used a 5-point frequency scale.

3.2.2 Pre-Testing

After developing the research questionnaire by adapting items from prior studies, the next step was assessing content validity to ensure the instrument adequately measures the intended variables. As (Sekaran & Bougie, 2016) note, content validity evaluates how representative and informative the scale items are for capturing the constructs of interest.

To ascertain content validity, two processes were followed. First, language experts proofread the questionnaire to improve phrasing. Second, field experts and professionals pre-tested the survey to evaluate understandability and construct representation. The experts agreed the items exhibited sufficient comprehension and reflected the core concepts in the questionnaire. Their recommendations were incorporated to refine the research instrument. These validity assessment procedures helped verify the final questionnaire has appropriate content validity before administration to collect study data.

3.2.3 Data Collection

The survey form was distributed using Google Forms and was made available both online and physically. Before participating, respondents were asked for their consent at the beginning of the survey. The majority of responses were collected physically to ensure a diverse range of answers. To achieve this, the survey was conducted at various locations such as Universities, Libraries, Shopping Centers, Bus Stations, and Cafeterias. Participants were given clear guidance to address any confusion they might have had regarding the questionnaire. To accommodate individuals who were illiterate or had difficulty understanding the research purpose or questions, translations were provided in the vernacular language. Their responses were considered as well. Due to the extensive scale of the survey, data collection took a significant amount of time.

3.2.4 Preliminary Tests

Some tests need to be performed after gathering the responses from the respondents. These tests organize the data in better way and help to reduce the unreliability of data and later on the research done using that data. This enhances validity and reliability of the research. The succeeding sections shed light on these tests, what are their thresholds and how they are done.

3.2.4.1 Missing Value Analysis

The most effective way to handle missing data is to prevent it from occurring in the first place (Dong & Peng, 2013). There are various techniques to manage data and minimize missing values. In this study, the researcher used an online smart application (Google Forms) for data collection, which required respondents to complete the entire questionnaire before submission. This restriction on submitting partial responses significantly reduced the incidence of missing data. Utilizing web-based forms with required response settings is an efficient way to curb missing values, relative to post-survey follow-ups or statistical imputation. Overall, a key takeaway is that proactive prevention of missing data through careful questionnaire design and delivery is preferable to reactive approaches. Besides, IBS SPSS 23 was also later on applied to check for missing values.

3.2.4.2 Data Normality Test

Various methods exist to assess data normality. One well-known and reliable approach widely employed is to examine the skewness and kurtosis of the data distribution (Cooper et al., 2003). In this research, this method was used to check the normality of the data. Skewness tells about asymmetry in the data distribution around its mean and can be either positive or negative, indicating an imbalance in the distribution. On the other hand, kurtosis indicates the degree of peakness or flatness of the distribution curve, reflecting how the data is clustered around the central values (Tabachnick et al., 2013).

Previous works have suggested common guidelines for acceptable ranges of skewness and kurtosis. The suggested values generally fall within the ranges of -2.58 to +2.58 (Hair et al., 2006) for skewness and -3 to +3 (Hair et al., 2011) for kurtosis.

3.2.4.3 Common Method Bias (CMB)

CMB is a potential problem in cross-sectional studies when attitudes and behaviors are measured simultaneously, as it can artificially inflate correlations between variables (Lindell & Whitney, 2001). To check for collinearity issues that may indicate CMB, variance inflation factor (VIF) values were examined in this study. As per (Hair Jr et al., 2021)VIF values above 3.000 signify high collinearity or potential CMB.

3.2.5 Descriptive Analysis

It was carried out to organize, arrange, and interpret the characteristics of the collected data. As emphasized by (Sekaran, 2003) descriptive statistics are essential for structuring data, generating meaningful summaries, and enhancing understanding. The frequency distribution, arithmetic mean, and standard deviation were calculated to achieve the research objectives.

3.2.6 Structure Equation Modeling

SEM is commonly used to examine cause-effect associations between latent variables. It examines directional links between observed and latent variables by incorporating both exogenous and endogenous factors. This study used PLS-SEM, as compared to covariance-based (CB) SEM, PLS-SEM offers several advantages over CB-SEM. Social science data often violates multivariate normality, which can underestimate standard errors and overestimate model fit in CB-SEM (Lei & Lomax, 2005). In contrast, PLS-SEM transforms non-normal data based on the central limit theorem, eliminating the normality requirement (Cassel et al., 1999). Additionally, PLS-SEM handles complex models well even with small samples (Reinartz et al., 2009).

Due to these benefits, PLS-SEM has become very popular in transportation research for behavioral modeling (Nguyen-Phuoc et al., 2019; Zhang et al., 2019). Given the non-normal social science data in this study and the complex conceptual model, PLS-SEM was chosen as the most suitable SEM approach.

3.2.6.1 Measurement Model Evaluation

In Partial Least Square Structure Equation Modeling first of all the evaluation of measurement model is performed. This assesses the reliability and validity of the constructs. It assesses the quality of the constructs or the main variables that are being tested in the model. Its evaluation is subjected to fulfilment of the factors explained in following sections:

3.2.6.1.1 Factor Loading and Internal Consistency Reliability

Factor loading shows the correlation between indicators and their constructs. High values mean the indicator strongly reflects that construct. During the evaluation of the model all the indicator should have loading higher than 0.60 (Gefen & Straub, 2005), these values are given by SmartPLS software. Internal consistency reliability checks if a construct's indicators are measuring that construct reliably and consistently. After performing the factor loading analysis, the next part of assessing the measurement model is reliability analysis, which includes calculating composite reliability. The recommended threshold for composite reliability is 0.70 (Ringle et al., 2020).

3.2.6.1.2 Convergent Validity

It means how well the indicators of a specific construct converge. It is one way of establishing the validity of a construct. Average variance extracted (AVE) is one method to quantify convergent validity. It measures the amount of variance in the indicators that is taken by the underlying latent construct, relative to variance due to measurement error. An AVE value of 0.50 or higher supports adequate convergent validity.

3.2.6.1.3 Discriminant Validity

It tells the extent to which a construct is truly discrete from other constructs in a structural equation model (SEM). It indicates that a construct measures a phenomenon that is unique and captures phenomena not shown in model by other constructs (Cheung & Wang, 2017).

The two widely applied method to verify discriminant validity include Fornell-Larcker Criterion and HTMT. The Fornell-Larcker criterion assesses discriminant validity by comparing the AVE value for each construct with the squared correlations between constructs (Barclay et al., 1995) . For adequate discriminant validity, the AVE values should exceed the squared correlations. The HTMT ratio is another assessment of discriminant validity. As per (Henseler et al., 2015), HTMT ratios should be under 0.90 for distinct constructs.

In summary, both the Fornell-Larcker criterion and HTMT assessments provided evidence of adequate discriminant validity, signifying the constructs differ sufficiently from each other in the model.

3.2.6.2 Model Fit

One commonly used method to gauge model fit is the standardized root mean square residual (SRMR). This study utilized SRMR to see the fitness of model based upon its data. SRMR computes the average discrepancy between observed correlations and correlations implied by model (Pavlov et al., 2021). It is based on comparing the actual correlation matrix from the data to the estimated correlation matrix from the structural model. The lesser the value of SRMR the better the fitness of model is. A SRMR below 0.10 is generally considered an acceptable fit (Henseler et al., 2015). This means the model-implied correlations do not differ greatly on average from the empirically observed correlations.

3.2.6.3 Structural Model Evaluation

The structural model is a key component of structural equation modeling (SEM). Its main purpose is to specify the theorized associations or paths between the latent constructs in the research model. In particular, the structural model aims to:

- Visually represent the theoretical/conceptual relationships between constructs that are assumed based on prior research and knowledge.
- Quantify the strengths of the relationships between constructs by estimating path coefficients along the connections.
- Identify how certain constructs directly or indirectly influence other constructs in the model.

- Test the conceptual model to evaluate how well it explains the covariance between the constructs.

Its results are utilized to support or reject the hypotheses and also to observe the mediation.

3.2.6.4 Mediation Analysis

Mediation analysis in structural equation modeling (SEM) examines indirect effects of an IV on a DV through mediator(s). Including mediators in SEM models provides insights into the underlying mechanisms and causal pathways between predictors and outcomes (Sarstedt et al., 2020).

The analysis involves estimating the direct effect, indirect effect, and total effect along with their significance levels. The total effect is the overall impact of one variable on another through all direct and indirect paths. The direct effect measures the relationship between two variables excluding mediators. The indirect effect represents the influence through intermediary variable(s). These effects elucidate the comprehensive relationships in SEM models, illuminating both direct connections and indirect chains of influence via mediators. Mediation analysis gives idea of the specific role of mediators in the causal system and helps unpack complex phenomena into component pathways. Overall, it is an invaluable statistical technique in SEM to dissect the mechanisms underlying observed relationships between variables.

3.3 Summary

The chapter explained the methodology of the research that will be used to carry out the study. An original questionnaire is required to examine factors influencing distracted driving, after reviewing prior literature to identify key variables like demographics, driving experience, road conditions, etc. The survey was developed in English and Urdu with closed-ended questions using 5-point Likert scales. It had 3 sections - demographics, distraction likelihood in different conditions, and engagement in distracting activities. Content validity will be assessed through expert review and pre-testing. An online form was used to prevent missing data. Preliminary tests included checking data normality via skewness and kurtosis and common method bias using VIF values. Descriptive analysis provided details on the sample characteristics. PLS-SEM is to be used for analysis due to its advantages with non-normal data and complex models. The measurement model will be evaluated by assessing factor loadings, composite reliability, convergent validity and discriminant validity. SRMR is used to evaluate model fit. The structural model tests hypothesized relationships between constructs. Mediation analysis will be examined through direct, indirect and total effects to provide insights into causal mechanisms.

CHAPTER 4: RESULTS AND DISCUSSION

4.1 Introduction

This chapter discusses the response collections' output, preliminary tests and later on intricate analyses performed on the data. It starts with describing the final questionnaire which is followed by discussion of the number of the responses collected and how many were finalized for further research. Then data is subjected to some basic test of normality distribution, common method bias. Finally, the results of main analyses along with discussion are offered. This includes the descriptive analysis and structure equation modeling which is subdivided into measurement and structural model. The model fit assessment is also proffered. Lastly, of mediation analysis is expatiated.

4.2 The Questionnaire

The questionnaire prepared after consulting many germane research papers, articles and reports published by authentic resources. Then the questionnaire was subjected to the testing and checking. Finally, it was finalized after incessant pertinent research and thorough pre-testing by the experts and peers. The final form of the questionnaire is attached in Annexure 1.

4.3 Data Collection

The data collection process was time-consuming because of the large number of responses gathered. The online survey yielded approximately 150 responses, 130 of which were used for the research. The in-person survey produced around 441 responses, with 371 being selected for the study. All of the in-person survey responses were carefully entered into an Excel spreadsheet. A counter check was also conducted to verify the accuracy of the data entry.

4.4 Preliminary Tests

These tests are done to organize the data, to remove the ambiguous or unnecessary data. Its removes inconsistent and missing responses which in turn improves the quality of the research being carried out using that data. The results obtained after performing the preliminary tests are presented in the succeeding sub sections.

4.4.1 Missing Value Analysis

The data analysis process involved identifying and addressing various issues in the responses. Outliers in the data were detected using SPSS and subsequently removed from the dataset. Similarly, responses with missing values were also eliminated. Additionally, inconsistent responses, such as respondents claiming to have 25+ years of driving experience at the age of 30, were excluded as they seemed implausible. After conducting a thorough analysis of missing and inconsistent values for both online and physical responses, a total of 501 valid responses (130 from the online survey and 371 from the physical survey) were selected to proceed with further research. These 501 responses were deemed reliable and appropriate for the study.

4.4.2 Data Adequacy

(Hair et al., 2010) provided a guideline that the sample size for a SEM (structural equation modeling) model must be 5 to 10 times the amount of questionnaire items. This particular study had 24 questionnaire indicators (items). Following the guideline, the minimum required sample size is 5 * 24 = 120 or 10 * 24 = 240. The total number of responses collected for this study was 501. Since 501 is more than the minimum required sample size of 240, the quantity of responses obtained is considered adequate for the SEM analysis as per the guideline.

4.4.3 Data Normality Test

The data normality was assessed using skewness and kurtosis measures. Acceptable ranges for skewness and kurtosis are commonly considered to be between -2.58 and +2.58 (Hair et al., 2006), and between -3 and +3 (Hair, Ringle, and Sarstedt, 2011) respectively. In Table 4-1, the values of skewness, kurtosis, mean, and standard deviation for the research items are presented. The skewness values ranged from -0.681 to 1.884, and the kurtosis readings were from -1.592 to 1.981, all of which fall within the acceptable limits. This indicates that the data approximates a normal distribution, which is essential for certain statistical analyses and model assumptions.

Table 4-1: The Mean, Std. Deviation, Skewness & Kurtosis of the Data

Constructs	Mean	Std. Deviation	Skewness	Kurtosis
Demographics				
Gender	1.158	.3648	1.884	1.556
Marriage	1.653	.4766	-.643	-1.592
Age	1.485	.8332	1.662	1.981
Education	3.513	.8777	-0.681	0.819
Driving Experience	1.842	1.0184	1.267	1.178
Road Types and Conditions (RTC)				
When driving on local roads (RTC1)	2.547	1.3492	.467	-.992
When driving on arterial road (RTC2)	3.128	1.2679	.013	-1.011
When driving on motorway/freeway (RTC3)	3.140	1.3815	-.097	-1.251
When driving on non-familiar road (RTC4)	3.144	1.3592	-.123	-1.206
When driving on a road, notorious for accidents or road is in bad condition (RTC5)	3.192	1.3591	-.144	-1.196
When there's too much traffic on the road (RTC6)	3.269	1.4315	-.227	-1.290
Trip Timing (TT)				
When driving during day time (TT1)	2.938	1.3705	.107	-1.156
When driving during night time (TT2)	2.952	1.2545	.109	-.944
Passenger Type (PT)				
When driving with parents or elder relatives (PT1)	3.387	1.2999	-.324	-.967
When driving with young kids or infants (PT2)	3.305	1.3147	-.222	-1.117
When driving a taxi with passenger (PT4)	3.242	1.2584	-.125	-.915
Law Enforcement (LE)				
When there's a fine on distraction activities (LE1)	3.417	1.2867	-.350	-.939
When there's a warden on the road to fine (LE2)	3.465	1.3013	-.374	-1.021
When there's no fine on distraction activities (LE3)	3.070	1.2653	.017	-.957
When no fine and no warden on the road (LE4)	3.086	1.2754	-.010	-1.022
Distracted Driving Behavior (D)				
Eating/drinking/taking medicine while driving (D1)	2.984	1.2553	.189	-1.017
Using phone while driving (D2)	2.892	1.3492	.177	-1.179
Doing make-up/shaving/looking in mirror while driving (D3)	3.521	1.4290	-.410	-1.252
Use dashboard for movie, music, GPS during driving (D4)	2.802	1.3187	.300	-1.053

4.4.4 Common Method Bias

The VIF values were used to examine the collinearity issues or common biasness. As per (Hair Jr et al., 2021) VIF values above 3.000 signify high collinearity or potential CMB. The values are presented in the Table 4-2 and it can be seen that all values are less than 3, hence there is no collinearity issue.

Table 4-2: Variance Inflation Factor Values

Constructs	**Age**	**Distracted Driving Behavior (D)**	**Driving Experience**	**Education**	**Gender**	**Law Enforcement (LE)**	**Passenger Type (PT)**	**Road Type and Conditions (RTC)**	**Trip Timing (TT)**
Age		1.583	1.025						
Distracted Driving Behavior (D)									
Driving_Experience		1.637							
Education		1.064	1.028						
Gender		1.043	1.005						
Law Enforcement (LE)		2.587							
Passenger Type (PT)		2.291							
Road Type and Conditions (RTC)		2.041							
Trip Timing (TT)		1.819							

4.5 Descriptive Analysis

The majority of respondents are male (84.2%). Females make up just 15.8% of respondents. This coincides with the general driver's population in Pakistan, as mostly drivers in Pakistan are male. Most respondents are single (65.3%) compared to married (34.7%). This aligns with the young age of the sample. The sample skews young, with

69.7% aged 18-30 years old. Only 15.8% are 30-40 years old. 11.4% are 40-50 years old. Very few respondents are over 50. Based on high inclination towards age bracket of 18-30 years, this research seems to be centered around studying driving behavior of young people. Almost 45.5% of respondents have a post-graduate degree, 35.7% have an undergraduate degree and only 9.4% have a PhD. This is a highly educated sample. The most common jobs are private sector (28.7%), students (35.1%), and government jobs (19.4%). Only 9% have their own business. Out of 501 samples, 36.1% earn 35,000-100,000 monthly, 21.2% earn 20,000-35,000. 19.2% earn 100,000-150,000. 13.2% earn over 200,000. This covers a wide income range. Precisely, 47.3% have 1-5 years of driving experience, 31.7% have 5-10 years. Only 13.4% have 10-15 years driving experience. Very few have over 15 years driving experience. This aligns with the young sample. Manual car is driven by 39.3% population whereas, 27.9% drive an automatic car, 24.4% drive a motorcycle. Few drive luxury cars or buses/vans. 50.9% drivers drive on daily basis, 25.7% drive randomly, 13% drive on alternate days, 10.4% drive 1-2 times per week. Most drive frequently. Descriptive analysis shows that 53.7% drivers have been in a crash due to distractions, 46.3% have not. Distractions seems to be a major crash factor. Out of the 57.3 who have faced accidents due to distractions, 33.5% experienced property damage only, 7.2% had a serious injury crash, 6.2% had a severe injury crash. Distractions lead to many minor crashes but also severe ones. In short, the sample of 501 respondents contains mostly young, educated males who drive frequently, with 84.2% male, 69.7% aged 18-30 years old, and 45.5% with a post-graduate degree. The majority (65.3%) are single and work in private jobs (28.7%) or are students (35.1%), with incomes ranging widely from 20,000 to over 200,000 monthly. Nearly half (47.3%) have 1-5 years of driving experience and drive manual cars (39.3%) or bikes (24.4%) daily (50.9%). Over half (53.7%) reported having a crash due to distractions, though most crashes only resulted in property damage (33.5%). The frequencies are presented in Table 4-3.

Table 4-3: Demographics of the Respondents

Variable	**Frequency**	**Percentage %**
Gender		
Male	422	84.2
Female	79	15.8

Marital Status		
Married	174	34.7
Single	327	65.3
Age		
18-30 Years	349	69.7
30-40 Years	79	15.8
40-50 Years	57	11.4
50-60 Years	14	2.8
60-70 Years	2	0.4
Education		
Matriculation	18	3.6
Intermediate	29	5.8
Undergraduate	179	35.7
Post-Graduate	228	45.5
Ph.D.	47	9.4
Job		
Govt. Job	97	19.4
Private Job	144	28.7
Own Business	45	9.0
Student	176	35.1
Jobless	39	7.8
Monthly Income		
20,000-35,000	106	21.2
35,000-100,000	181	36.1
100,000-150,000	96	19.2
150,000-200,000	52	10.4
>200,000	66	13.2
Driving Experience		
1-5 Years	237	47.3
5-10 Years	159	31.7
10-15 Years	67	13.4
15-20 Years	23	4.6
20-25+ Years	15	3.0
Vehicle Type		
Auto Transmission Car	140	27.9
Manual Transmission Car	197	39.3
Bike	122	24.4
Luxury Car	34	6.8
Bus/Van	8	1.6
Driving_Frequency		

Daily	255	50.9
Alternate Days	65	13.0
Once/Twice in a Week	52	10.4
Random	129	25.7
Crash_due_to_Distractions		
Yes	269	53.7
No	232	46.3
Crash_Nature		
No Accident	232	46.3
Fatal	14	2.8
Critical	20	4.0
Severe	31	6.2
Serious	36	7.2
Property Damage Only	168	33.5

4.6 Structure Equation Modeling

As described earlier in the research methods portion, PLS-SEM techniques has been used in this study unlike CB-SEM. The next discussion explains the results of the research. Thus, the model was developed via SmartPLS 4 and the schematic depiction of the model is given in Figure 4-1.

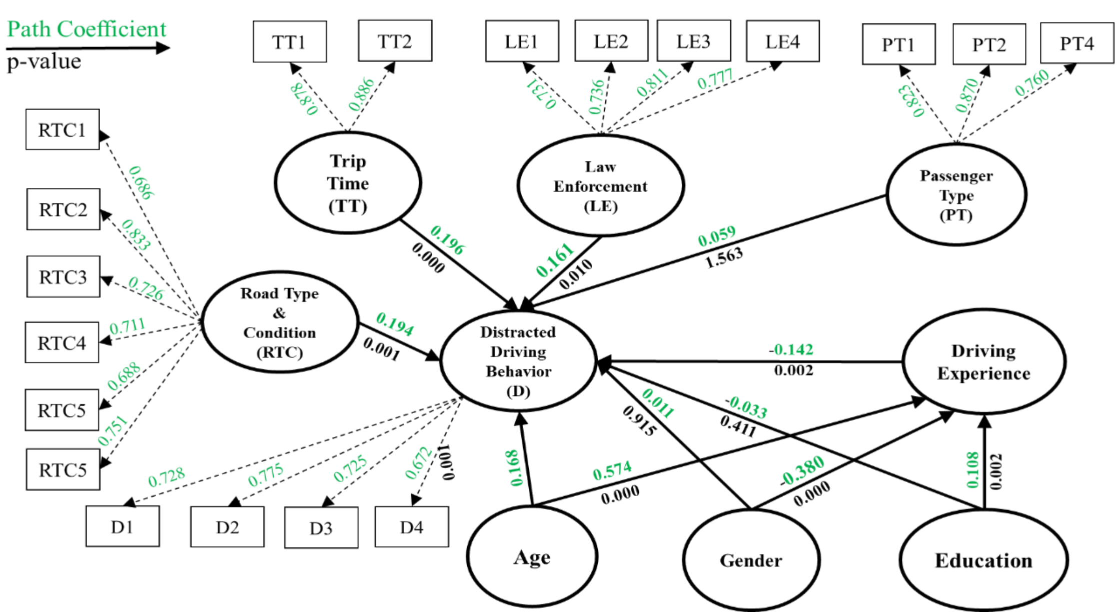


Figure 4-1: The PLS-SEM Model

4.6.1 Measurement Model Evaluation

This is the first step in the structure equation modeling. In measurement model the quality of the latent constructs or the dependent variables was thoroughly checked. It is done by means of various parameters. First one is factor loading, then there is internal consistency reliability. It is followed by validity check where we have convergent validity and discriminant validity. AVE is used to check the convergent validity whereas, the Fornell Larcker criteria helps checking in discriminant validity.

4.6.1.1 Factor Loading and Internal Consistency Reliability

The measureent model shows the link between the consructs and their indicaor variales. All representing or indicating attributes under 0.60 loading were eliminated as recommended by (Gefen & Straub, 2005). Only one indicator, PT3, was removed for having the lowest loading. Table 4-4 displays all the factor loadings.

After analyzing the factor loadings, reliability analysis was done next by calculating the composite reliability for each construct. As stated by (Ringle et al., 2020) the composite reliability must meet the recommended edge of 0.70. As evident in Table 4-4, all constructs in the model meet this benchmark for composite reliability.

4.6.1.2 Convergent Validity

Convergent validity was checking by calculating the (AVE) for each construct. As stated by (Ringle et al., 2020), the AVE should be 0.50 or above to demonstrate adequate convegent validity. As shown in Table 3, all constructs in the model meet this criterion, with AVE values exceeding the 0.50 threshold. This indicates that all costruts demonstrate sufficent convergent validity based on their AVE scores.

4.6.1.3 Discriminant Validity

To examine the discriminant validity Fornell-Larcker’s criterion and the Heterotrait-Monotrait (HTMT) ratio were used. Table 4-5 shows the results of the Fornell-Larcker analysis. According to this method, discriminnt validty is tested by comparing the AVE values to the squared correlations betwen construts (Barclay et al., 1995). To demonstrate disriminant validty, the squared correlation of inter-constructs needs to be less than its own

AVE values. As seen in Table 4-5, the AVE values exceed the squared correlations between that construct and other constructs. Additionally, HTMT ratios should be under 0.90 for adequate discriminant validity per (Henseler et al., 2015). As shown in Table 4-6, all HTMT ratios in this study are below 0.90, meeting this threshold. One PT value is slightly above 0.90 but since it is not excessively high and also passes the Fornell-Larcker criterion, this construct is not problematic. Overall, the results indicate satisfactory discriminant validity.

Table 4-4: The Factor Loadings, Cronbach Alpha, Composite Reliability and AVE

Item	Factor Loading	Cronbach's Alpha	Composite Reliability	AVE
Distracted Driving Behavior (D)		0.700	0.816	0.527
Eating/drinking/taking medicine while driving (D1)	0.728			
Using phone while driving (D2)	0.775			
Doing make-up/shaving/looking in mirror while driving (D3)	0.725			
Use dashboard for movie, music, GPS during driving (D4)	0.672			
Law Enforcement (LE)		0.764	0.849	0.585
When there's a fine on distraction activities (LE1)	0.731			
When there's a warden on the road to fine (LE2)	0.736			
When there's no fine on distraction activities (LE3)	0.811			
When no fine and no warden on the road (LE4)	0.777			
Passenger Type (PT)		0.753	0.859	0.670
When driving with parents or elder relatives (PT1)	0.823			
When driving with young kids or infants (PT2)	0.870			
When driving a taxi with passenger (PT4)	0.760			
Road Types and Conditions (RTC)		0.829	0.875	0.539
When driving on local roads (RTC1)	0.686			
When driving on arterial road (RTC2)	0.833			
When driving on motorway/freeway (RTC3)	0.726			
When driving on non-familiar road (RTC4)	0.711			
When driving on a road, notorious for accidents or road is in bad condition (RTC5)	0.688			

When there’s too much traffic on the road (RTC6)	0.751			
Trip Timing (TT)		0.714	0.875	0.778
When driving during day time (TT1)	0.878			
When driving during night time (TT2)	0.886			

Table 4-5: The Values for Fornell & Larker’s Criterion

Item	Age	Distracted Driving Behavior (D)	Driving Experience	Education	Gender	Law Enforcement (LE)	Passenger Type (PT)	Road Type and Conditions (RTC)	Trip Timing (TT)
Age	1								
Distracted Driving Behavior (D)	0.078	0.726							
Driving_Experience	0.595	-0.01	1						
Education	0.151	0.03	0.187	1					
Gender	-0.028	0.03	-0.148	0.059	1				
Law Enforcement (LE)	0.064	0.462	0.119	0.085	-0.011	0.765			
Passenger Type (PT)	-0.005	0.421	0.08	0.092	-0.028	0.713	0.819		
Road Type and Conditions (RTC)	0.004	0.462	0.037	0.095	0.045	0.643	0.622	0.734	
Trip Timing (TT)	-0.043	0.442	0.031	0.123	0.039	0.609	0.546	0.582	0.882

Table 4-6: HTMT Ratios

Item	Age	Distracted Driving Behavior (D)	Driving Experience	Education	Gender	Law Enforcement (LE)	Passenger Type (PT)	Road Type and Conditions (RTC)	Trip Timing (TT)
Age									
Distracted Driving Behavior (D)	0.14								
Driving_Experience	0.595	0.088							

Education	0.151	0.055	0.187						
Gender	0.028	0.049	0.148	0.059					
Law Enforcement (LE)	0.071	0.624	0.134	0.102	0.036				
Passenger Type (PT)	0.022	0.571	0.093	0.105	0.033	0.952			
Road Type and Conditions (RTC)	0.062	0.597	0.063	0.107	0.049	0.812	0.794		
Trip Timing (TT)	0.119	0.622	0.098	0.145	0.046	0.822	0.744	0.747	

4.6.2 Model Fit

One commonly used model fit test is the SRMR. The SRMR value for this model was 0.065. As stated by (Henseler et al., 2015), SRMR values below 0.1 indicate adequate model fit. Since the SRMR value of 0.065 is well below the 0.1 threshold, this suggests the model has an acceptable fit with the data overall.

4.6.3 Structural Model Evaluation

The hypothesized relationshis (paths) between the contructs are examined by structural model. The impact of all elements on distracted driving behavior was analyzed by examining the path-coefficients and their significance, which were generated through bootstrapping with 5000 subsamples. The results showed that road type & conditions (RTC) has a significant positive effect on distracted driving behavior (D) ($\beta = 0.194$, $p < 0.001$), supporting H1. Trip Timing (TT) also positively impacts distracted driving behavior (D) significantly ($\beta = 0.196$, $p = 0.000$), supporting H2. Law enforcement (LE) has a significant positive effect on distracted driving behavior (D) ($\beta = 0.161$, $p < 0.01$), supporting H3. However, the relationship between passenger type (PT) and distracted driving behavior (D) is insignificant ($\beta = 0.092$, $p > 0.05$), leading to the rejection of H4. In summary, the structural model analysis provides support for three hypothesized relationships (H1, H2, H3) and rejects one (H4), as presented in Table 4-7.

4.6.4 Mediation Analysis

This mediation analysis tested three hypotheses. Hypothesis H5 looked at whether driving experience mediates the relationship between age and distracted driving (D). The results gave that the direct effect of age on distracted driving behavior (D) was significant (β = 0.161, $p < 0.01$) when controlling for the mediator. Additionally, the indirct effect through the mediator was also substantial (β = -0.082, $p < 0.003$), indicating partial mediation. Overall, H5 is supported since driving experience partially mediates the relation of age and distracted driving behavior (D).

Hypothesis H6 examined if driving experience mediates the link of gender and distracted driving behavior (D). The direct link of gender on distracted driving behavior (D) was not substantial (β = 0.011, $p > 0.05$). However, the indirect effect was profound (β = 0.054, $p < 0.016$), showing full mediation. This suggests the gender and distracted driving behavior (D) relationship is entirely mediated through driving experience, supporting H6.

Finally, H7 looked at whether driving experience mediates the education-distracted driving behavior (D) linkage. The direct consequence of education on distracted driving behavior (D) was not profound (β = -0.049, $p > 0.05$). But, the indiret effect was substantial (β = -0.015, $p < 0.03$), indicating full mediation. So H7 is supported as driving experience fully mediates the education- distracted driving behavior (D) link.

In summary, driving experience acts as a mediator in the relationships between age-distracted driving behavior (partial mediation), gender- distracted driving behavior (full mediation), and education- distracted driving behavior (full mediation), as detailed in Table 4-8.

Table 4-7: Path Coefficients and P-Values to test the Hypotheses

Hypotheses	**Path Coefficients**	**Standard Deviation**	**T Statistics**	**P Values**
H1: Road Type & Condition -> Distracted Driving Behavior	0.194	0.060	3.215	0.001
H2: Trip Timing -> Distracted Driving Behavior	0.196	0.054	3.624	0.000
H3: Law Enforcement -> Distracted Driving Behavior	0.161	0.062	2.576	0.010
H4: Passenger Type -> Distracted Driving Behavior	0.092	0.059	1.563	0.118

Table 4-8: The Results for Mediation Analysis

Effects	Path Coefficients	Standard Deviation	T Statistics	P Values
Direct Effect				
Age -> Distracted Driving Behavior	0.168	0.050	3.375	0.001
Gender -> Distracted Driving Behavior	0.011	0.106	0.107	0.915
Education -> Distracted Driving Behavior	-0.033	0.041	0.822	0.411
Indirect Effect				
H5: Age -> Driving_Experience -> Distracted Driving Behavior	-0.082	0.027	3.000	0.003
H6: Gender -> Driving_Experience -> Distracted Driving Behavior	0.054	0.022	2.415	0.016
H7: Education -> Driving_Experience -> Distracted Driving Behavior	-0.015	0.007	2.168	0.030
Total Effects				
Age -> Distracted Driving Behavior	0.086	0.041	2.112	0.035
Gender -> Distracted Driving Behavior	0.065	0.106	0.616	0.538
Education -> Distracted Driving Behavior	-0.049	0.040	1.209	0.227

4.7 Discussion on the Results

The statistics in Table 4-4 are standardized load factors while Table 4-7 provides path coefficiets. The load coeficients in the measurment model indicate how much the observed variables reflect the latent variables. Table 4-4 shows the load coefficients were mostly above 0.6, meaning the observed variables adequately reflected their relevant latent variables. The path-coefficients indicate the degree of effect between the latent variables. The total effect was the dirct and indirect effects added together. In Table 4-7 and Table 4-8, the effects of different factors on driver Distracted Driving Behavior (D) were calculated. All factors except passenger type had significant effects on distracted driving behavior (D). The positive effects were: Road Type and Conditions (0.194), Trip Timing (0.196), Law Enforcement (0.161) and Age (0.168). The effect of Passenger type, Gender and Education wasn’t significant. Among all factors, Trip Timing had the greatest influence (0.196) on Distracted Driving Behavior, which shows that time of driving (night or day) can highly affect the behavior of distracted driving. Drivers were found more distracted during the day than at night. This explains, as during the day time driving is relatively easier due to enhanced visibility which makes drivers to indulge in other non-driving activities. Then, driving behavior was secondly most effected by Road Type and Condition. Better the road type and road conditions are, the more likely will the driver be distracted. The distracted driving behavior (D) was then most effected by law enforcement conditions. People reported highly distracted behavior in case of poor law enforcement. Thus, higher the lawlessness, higher the drivers will be distracted during driving. The effect of passenger on distracted driving behavior (D) was found insignificant in this research which is also supported by (Huisingh et al., 2019; Papantoniou et al., 2019). This might be because of the cultural aspects and drivers in Pakistan find passenger presence to have less effect on their driving and thus effect of passenger on driving behavior is not construed as a factor leading to distracted driving. Also, the passenger types in this study were parents, children, friends and customers (in case driving taxi). In all the scenario, except friends, people reported to be less distracted which explains the insignificant effect of passenger type on distracted driving (D). Thus, first 3 hypotheses were supported and H4 was not supported as concluded in Table 4-9.

The effect of age was also significant on distracted driving behavior (0.168). It was observed that people with higher age were more distracted, this affect could be due to the fact that they possessed more driving experience and relatively it's easier for them to indulge in non-primary driving tasks. In the population sample, around 70 % of the drivers were 18-30 years old and around middle-aged accounted for 27%, so it can be construed that middle-aged drivers were more distracted whereas, the young drivers are less the distracted ones. But some of the effect of age on driving behavior was also passing through driving experience, as the indirect effect of age on Driving behavior through Driving experience was also significant as shown in Table 4-8, which depicted partial mediation of driving experience. The direct effect of gender and education on distracted driving behavior was found insignificant but the indirect effect in the presence of driving experience was found significant (value presented in Table 4-8), therefore complete mediation was observed for gender and education where driving experience acted as mediator between gender, education and driving behavior.

Table 4-9: Overall Achievement of the Proposed Hypotheses

Hypotheses	Path Coefficients	Standard Deviation	T Statistics	P Values	Supported
H1: Road Type & Condition -> Distracted Driving Behavior	0.194	0.06	3.215	0.001	Yes

H2: Trip Timing -> Distracted Driving Behavior	0.196	0.054	3.624	0.000	Yes
H3: Law Enforcement -> Distracted Driving Behavior	0.161	0.062	2.576	0.010	Yes
H4: Passenger Type -> Distracted Driving Behavior	0.092	0.059	1.563	0.118	No
H5: Age -> Driving_Experience -> DDV	-0.082	0.027	3.000	0.003	Yes
H6: Gender -> Driving_Experience -> Distracted Driving Behavior	0.054	0.022	2.415	0.016	Yes
H7: Education -> Driving_Experience -> Distracted Driving Behavior	-0.015	0.007	2.168	0.030	Yes

4.8 Summary

This chaper offered the results of all the investigation performed on the data and the model construed out of the data. First of all, 501 right responses were filtered out for further research after carrying out missing value analysis and looking for inconsistent values. Then the data normality was checked and it was acceptable. This was followed by checking collinearity issue by dint of observing VIF which were all less than 3 indicating now common biasness. Also, descriptive analysis was executed for demographic information of the data. Finally, the model was designed using SmartPLS 4 and it had independent and dependent variables. The measuremen and structural model were evaluated. The measuremet model was evaluated by means of factor loading, Cronbach alpha, composite reliability, AVE, convergent and discriminant validity. All were within acceptable limits. Then structure model was also subjected to evaluation via path coefficients, t-value and p-value. All hypotheses were supported except effect of passenger on driving behavior (H4). Mediation analysis was also executed and driving experience was seen to be mediating the relationship between, age/gender/education and distracted driving behavior.

CHAPTER 5: CONCLUSIONS AND RECOMMENDATIONS

5.1 Summary

The lacuna in the previous studies regarding assessing effects of different factors on distracted driving behavior and especially a great deficiency of such research in Pakistan laid basis for the need of this research. As this is more of a behavioral research that falls at the intersection of transportation and psychology, therefore after rigorous studies of pertinent material some factors were finalized and a questionnaire was prepared to collect the responses from drivers. Almost five months were utilized to collect the data and finally an intricate analysis comprising a set of statistical tests along with the PLS-SEM techniques were applied to tests the following hypotheses.

i) H1: Road Type and Conditions have Significant Impact on Distracted Driving
ii) H2: Driving Timing Significant Impact on Distracted Driving
iii) H3: Law Enforcement has Significant Impact on Distracted Driving
iv) H4: Passenger Type has Significant Impact on Distracted Driving
v) H5: Driving Experience Mediates the Relationship between Age and Distracted Driving
vi) H6: Driving Experience Mediates the Relationship between Gender and Distracted Driving
vii) H7: Driving Experience Mediates the Relationship between Education and Distracted Driving

5.2 Conclusions

The following conclusion are drawn based upon the results and interpretation stated in chapter 4.

- Most drivers were male, educated, aged 18-30, male domination truly represents the actual culture of gender-based distribution in Pakistan.
- 5 main variables (road conditions, trip timing, law enforcement, passenger type, distracted driving behavior) with 24 observed variables were used for SEM to see which factors influence Distracted driving behavior most.

- All factors positively related to Distracted driving behavior and significant as suggested by p-value being less than 0.05 and t-value greater than 1.96, except for passenger type who had insignificant effect on distracted driving.
- Trip timing (TT) influenced Distracted driving behavior the most (0.196). And the drivers are more distracted during day than night, possibly due to ease of driving in daylight enabling non-driving activities.
- Second highest significant effect was seen for road conditions (0.194), followed by age (0.168) and law enforcement (0.161).
- Impact of passengers on Distracted driving behavior was insignificant. It might be due to cultural aspects and how Pakistani drivers view passenger effect on driving. Also, the passenger in the study were mainly children, parents or customers and in Pakistani culture people don't indulge in non-driving tasks in presence of such passenger.
- Age had also significant (0.168) effect on driving behavior. It was found that older drivers were more distracted, potentially due to greater experience making indulging in non-driving tasks easier.
- Some of the effect of age was also significantly passing through mediator driving experience, which indicated partial mediation.
- Direct effects of gender and education on Distracted driving behavior were insignificant.
- Indirect effects of gender and education were significant in presence of driving experience, indicating driving experience mediates between gender, education, and Distracted driving behavior.

5.3 Recommendations

Based on the observations and conclusions concurred upon, the following recommendations are offered:

- Implement distraction mitigation strategies during daytime driving when drivers are most susceptible to distraction, such as limiting use of electronic devices, keeping passengers from distracting the driver, and taking breaks.

- Improve road infrastructure and conditions to reduce mental workload for drivers, enabling them to focus more on driving. This could include adding rumble strips, widening lanes, and enhancing signage.
- Increase enforcement of distracted driving laws, as the research found higher law enforcement was associated with less distraction. This could involve penalties for phone use or other distractions.
- Develop education campaigns targeting older drivers about the risks of distraction and ways to minimize engagement in non-driving activities while driving.
- Conduct further research to understand the role of cultural factors in driver distraction, as passengers did not affect distraction in this study as expected. Customized interventions may be needed.
- Include driving experience more prominently in distracted driving research and interventions, given its mediating effects between demographics and distraction found in this study. Experience likely changes distraction risk.
- Focus on high-risk demographic groups identified here, such as younger male drivers, in terms of interventions to reduce distraction. Tailored approaches to these groups could improve outcomes.
- Increase data collection on driver distraction to better understand root causes and implement evidence-based mitigation strategies. More research is needed on this important issue.
- Overall, multi-faceted efforts are required to modify the unsafe driving conduct involving key actors across sectors and leveraging strategies like education and advertising.

REFERENCES

Ahangari, S., Jeihani, M., & Dehzangi, A. (2019). A machine learning distracted driving prediction model. *Proceedings of the 3rd International Conference on Vision, Image and Signal Processing*, 1–6.

Ahangari, S., Jeihani, M., Rahman, M. M., & Dehzangi, A. (2021). PREDICTING DRIVING DISTRACTION PATTERNS IN DIFFERENT ROAD CLASSES USING A SUPPORT VECTOR MACHINE. *International Journal for Traffic & Transport Engineering*, *11*(1).

Ahangari, S., Jeihani, M., Salahshour, B., & Ndegwa, M. (2020). A comprehensive analysis of distracted driving using a driving simulator. *International Journal for Traffic and Transport Engineering*, *10*(2), 252–265.

Atchley, P., Hadlock, C., & Lane, S. (2012). Stuck in the 70s: The role of social norms in distracted driving. *Accident Analysis & Prevention*, *48*, 279–284.

Barclay, D., Higgins, C., & Thompson, R. (1995). *The partial least squares (PLS) approach to casual modeling: personal computer adoption ans use as an Illustration*.

Bentler, P. M. (2006). EQS 6 structural equations program manual. Encino: Multivariate Software, Inc. Bentler, PM, & Chou, C.(1987). Practical issues in structural modeling. *Sociological Methods & Research*, *76*, 78–117.

Bollen, K. A. (1989). *Structural equations with latent variables* (Vol. 210). John Wiley & Sons.

Brodsky, W. (2001). The effects of music tempo on simulated driving performance and vehicular control. *Transportation Research Part F: Traffic Psychology and Behaviour*, *4*(4), 219–241.

Brodsky, W., & Slor, Z. (2013). Background music as a risk factor for distraction among young-novice drivers. *Accident Analysis & Prevention*, *59*, 382–393.

Cai, A. W. T., Manousakis, J. E., Singh, B., Kuo, J., Jeppe, K. J., Francis-Pester, E., Shiferaw, B., Beatty, C. J., Rajaratnam, S. M. W., & Lenné, M. G. (2021). On-road driving impairment following sleep deprivation differs according to age. *Scientific Reports*, *11*(1), 1–13.

Cassel, C., Hackl, P., & Westlund, A. H. (1999). Robustness of partial least-squares method for estimating latent variable quality structures. *Journal of Applied Statistics*, *26*(4), 435–446.

Cassidy, G., & MacDonald, R. (2009). The effects of music choice on task performance: A study of the impact of self-selected and experimenter-selected music on driving game performance and experience. *Musicae Scientiae*, *13*(2), 357–386.

Chen, J.-C., Lee, C.-Y., Huang, P.-Y., & Lin, C.-R. (2020). Driver behavior analysis via two-stream deep convolutional neural network. *Applied Sciences*, *10*(6), 1908.

Chen, S., Kuhn, M., Prettner, K., & Bloom, D. E. (2019). The global macroeconomic burden of road injuries: estimates and projections for 166 countries. *The Lancet Planetary Health*, *3*(9), e390–e398.

Cheung, G. W., & Wang, C. (2017). Current approaches for assessing convergent and discriminant validity with SEM: Issues and solutions. *Academy of Management Proceedings*, *2017*(1), 12706.

Chin, W. W. (1998). The partial least squares approach to structural equation modeling. *Modern Methods for Business Research*, *295*(2), 295–336.

Chin, W. W., Marcolin, B. L., & Newsted, P. R. (2003). A partial least squares latent variable modeling approach for measuring interaction effects: Results from a Monte Carlo simulation study and an electronic-mail emotion/adoption study. *Information Systems Research*, *14*(2), 189–217.

Cooper, D. R., Schindler, P. S., Cooper, D. R., & Schindler, P. S. (2003). *Business research methods*.

Dingus, T. A., Guo, F., Lee, S., Antin, J. F., Perez, M., Buchanan-King, M., & Hankey, J. (2016). Driver crash risk factors and prevalence evaluation using naturalistic driving data. *Proceedings of the National Academy of Sciences*, *113*(10), 2636–2641.

Dong, Y., & Peng, C.-Y. J. (2013). Principled missing data methods for researchers. *SpringerPlus*, *2*, 1–17.

Drews, F. A., Pasupathi, M., & Strayer, D. L. (2008). Passenger and cell phone conversations in simulated driving. *Journal of Experimental Psychology: Applied*, *14*(4), 392.

Drews, F. A., Yazdani, H., Godfrey, C. N., Cooper, J. M., & Strayer, D. L. (2009). Text messaging during simulated driving. *Human Factors*, *51*(5), 762–770.

Ebnali, M., Lamb, R., Fathi, R., & Hulme, K. (2021). Virtual reality tour for first-time users of highly automated cars: Comparing the effects of virtual environments with different levels of interaction fidelity. *Applied Ergonomics*, *90*, 103226.

Falk, R. F., & Miller, N. B. (1992). *A primer for soft modeling.* University of Akron Press.

Fitch, G. M., Soccolich, S. A., Guo, F., McClafferty, J., Fang, Y., Olson, R. L., Perez, M. A., Hanowski, R. J., Hankey, J. M., & Dingus, T. A. (2013). *The impact of hand-held and hands-free cell phone use on driving performance and safety-critical event risk*.

Gefen, D., & Straub, D. (2005). A practical guide to factorial validity using PLS-Graph: Tutorial and annotated example. *Communications of the Association for Information Systems*, *16*(1), 5.

Golob, T. F. (2003). Structural equation modeling for travel behavior research. *Transportation Research Part B: Methodological*, *37*(1), 1–25.

Hair, J. F., Anderson, R. E., Babin, B. J., & Black, W. C. (2010). *Multivariate data analysis: A global perspective (Vol. 7)*. Upper Saddle River, NJ: Pearson.

Hair, J. F., Black, W. C., Babin, B. J., Anderson, R. E., & Tatham, R. L. (2006). *Multivariate data analysis 6th Edition*. Pearson Prentice Hall. New Jersey. humans: Critique and reformulation ….

Hair, J. F., Ringle, C. M., & Sarstedt, M. (2011). PLS-SEM: Indeed a silver bullet. *Journal of Marketing Theory and Practice*, *19*(2), 139–152.

Hair Jr, J. F., Hult, G. T. M., Ringle, C. M., & Sarstedt, M. (2021). *A primer on partial least squares structural equation modeling (PLS-SEM)*. Sage publications.

Hancock, P. A., Lesch, M., & Simmons, L. (2003). The distraction effects of phone use during a crucial driving maneuver. *Accident Analysis & Prevention*, *35*(4), 501–514.

Hanowski, R. J., Olson, R. L., Hickman, J. S., & Dingus, T. A. (2006). *The 100-car naturalistic driving study: A descriptive analysis of light vehicle-heavy vehicle interactions from the light vehicle driver's perspective*. United States. Department of Transportation. Federal Motor Carrier Safety ….

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*, 115–135.

Horberry, T., Anderson, J., Regan, M. A., Triggs, T. J., & Brown, J. (2006). Driver distraction: The effects of concurrent in-vehicle tasks, road environment complexity and age on driving performance. *Accident Analysis & Prevention*, *38*(1), 185–191.

Huisingh, C., Owsley, C., Levitan, E. B., Irvin, M. R., MacLennan, P., & McGwin, G. (2019). Distracted Driving and Risk of Crash or Near-Crash Involvement among Older Drivers Using Naturalistic Driving Data with a Case-Crossover Study Design. *Journals of Gerontology - Series A Biological Sciences and Medical Sciences*, *74*(4), 550–555. https://doi.org/10.1093/gerona/gly119

Javid, M. A., & Faraz, N. S. (2017). Understanding the behaviour of young drivers in relation to traffic safety aspects in Lahore. *Pakistan Journal of Science*, *69*(1), 144.

Khan, K., Zaidi, S. B., & Ali, A. (2020). Evaluating the nature of distractive driving factors towards road traffic accident. *Civil Engineering Journal*, *6*(8), 1555–1580.

Khoso, A. K. (2019). Prevalence of nonuse of helmet and seat belt and distracted driving on national highway in Pakistan. *Journal of Injury and Violence Research*, *11*(4 Suppl 2).

Kline, R. B. (2011). Principles and practice of structural equation modeling (3. Baskı). *New York, NY: Guilford*, *14*, 1497–1513.

Lee, J. D., Caven, B., Haake, S., & Brown, T. L. (2001). Speech-based interaction with in-vehicle computers: The effect of speech-based e-mail on drivers' attention to the roadway. *Human Factors*, *43*(4), 631–640.

Lei, M., & Lomax, R. G. (2005). The effect of varying degrees of nonnormality in structural equation modeling. *Structural Equation Modeling*, *12*(1), 1–27.

Li, N., Jain, J. J., & Busso, C. (2013). Modeling of driver behavior in real world scenarios using multiple noninvasive sensors. *IEEE Transactions on Multimedia*, *15*(5), 1213–1225.

Lindell, M. K., & Whitney, D. J. (2001). Accounting for common method variance in cross-sectional research designs. *Journal of Applied Psychology*, *86*(1), 114.

Mahdi, R., & Khaled, K. (2022). Identification of factors associated with various types of impaired driving. *Humanities & Social Sciences Communications*, *9*(1).

Mateos-Aparicio, G. (2011). Partial least squares (PLS) methods: Origins, evolution, and application to social sciences. *Communications in Statistics-Theory and Methods*, *40*(13), 2305–2317.

Matsueda, R. L. (2012). *Key advances in the history of structural equation modeling.*

Murphy-Chutorian, E., Doshi, A., & Trivedi, M. M. (2007). Head pose estimation for driver assistance systems: A robust algorithm and experimental evaluation. *2007 IEEE Intelligent Transportation Systems Conference*, 709–714.

Neyens, D. M., & Boyle, L. N. (2008). The influence of driver distraction on the severity of injuries sustained by teenage drivers and their passengers. *Accident Analysis & Prevention*, *40*(1), 254–259.

Nguyen-Phuoc, D. Q., Nguyen, H. A., De Gruyter, C., Su, D. N., & Nguyen, V. H. (2019). Exploring the prevalence and factors associated with self-reported traffic crashes among app-based motorcycle taxis in Vietnam. *Transport Policy*, *81*, 68–74.

Ortiz, C., Ortiz-Peregrina, S., Castro, J. J., Casares-López, M., & Salas, C. (2018). Driver distraction by smartphone use (WhatsApp) in different age groups. *Accident Analysis & Prevention*, *117*, 239–249.

Oviedo-Trespalacios, O., King, M., Haque, M. M., & Washington, S. (2017). Risk factors of mobile phone use while driving in Queensland: Prevalence, attitudes, crash risk perception, and task-management strategies. *PLoS One*, *12*(9), e0183361.

Oviedo-Trespalacios, O., & Regan, M. A. (2021). Driver distraction. *International Encyclopedia of Transportation: Volume 7, Transport Psychology, Transport Sustainability and Health. Vol. 7.*, 113–120.

Owens, J. M., McLaughlin, S. B., & Sudweeks, J. (2011). Driver performance while text messaging using handheld and in-vehicle systems. *Accident Analysis & Prevention*, *43*(3), 939–947.

Papantoniou, P., Antoniou, C., Papadimitriou, E., Yannis, G., & Golias, J. (2015). Exploratory analysis of the effect of distraction on driving behaviour through a driving simulator experiment. *Proceedings of the 6th Pan-Hellenic Road Safety Conference, Hellenic Institute of Transportation Engineers, National Technical University of Athens, Athens*.

Papantoniou, P., Papadimitriou, E., & Yannis, G. (2017). Review of driving performance parameters critical for distracted driving research. *Transportation Research Procedia*, *25*, 1796–1805.

Papantoniou, P., Yannis, G., & Christofa, E. (2019). Which factors lead to driving errors? A structural equation model analysis through a driving simulator experiment. *IATSS Research*, *43*(1), 44–50.

Pavlov, G., Maydeu-Olivares, A., & Shi, D. (2021). Using the standardized root mean squared residual (SRMR) to assess exact fit in structural equation models. *Educational and Psychological Measurement*, *81*(1), 110–130.

Prat, F., Gras, M. E., Planes, M., Font-Mayolas, S., & Sullman, M. J. M. (2017). Driving distractions: An insight gained from roadside interviews on their prevalence and factors associated with driver distraction. *Transportation Research Part F: Traffic Psychology and Behaviour*, *45*, 194–207.

Raykov, T., & Marcoulides, G. A. (2012). *A first course in structural equation modeling*. routledge.

Regan, M. A., Lee, J. D., & Young, K. (2008). *Driver distraction: Theory, effects, and mitigation*. CRC press.

Reinartz, W., Haenlein, M., & Henseler, J. (2009). An empirical comparison of the efficacy of covariance-based and variance-based SEM. *International Journal of Research in Marketing*, *26*(4), 332–344.

Ringle, C. M., Sarstedt, M., Mitchell, R., & Gudergan, S. P. (2020). Partial least squares structural equation modeling in HRM research. *The International Journal of Human Resource Management*, *31*(12), 1617–1643.

Road Safety in Pakistan | Traffic accidents, crash, fatalities & injury statistics | GRSF. (2016). https://www.roadsafetyfacility.org/country/pakistan

Sarstedt, M., Hair Jr, J. F., Nitzl, C., Ringle, C. M., & Howard, M. C. (2020). Beyond a tandem analysis of SEM and PROCESS: Use of PLS-SEM for mediation analyses! *International Journal of Market Research*, *62*(3), 288–299.

Sekaran, U. (2003). A refresher on some statistical terms and test. *Research Methods for Business: A Skill Building Approach (4th Ed., Pp. 391–411). New York: John Wiley & Sons*.

Sekaran, U., & Bougie, R. (2016). *Research methods for business: A skill building approach*. john wiley & sons.

Shahverdy, M., Fathy, M., Berangi, R., & Sabokrou, M. (2020). Driver behavior detection and classification using deep convolutional neural networks. *Expert Systems with Applications*, *149*, 113240.

Son, J., & Park, M. (2021). The Effects of Distraction Type and Difficulty on Older Drivers' Performance and Behaviour: Visual vs. Cognitive. *International Journal of Automotive Technology*, *22*(1), 97–108.

Strayer, D. L., Cooper, J. M., Turrill, J., Coleman, J., Medeiros-Ward, N., & Biondi, F. (2013). *Measuring cognitive distraction in the automobile*.

Strayer, D. L., Drews, F. A., & Johnston, W. A. (n.d.). *Cell Phone Induced Failures of Visual Awareness During Simulated Driving*.

Sullman, M. J. M., Przepiorka, A. M., Prat, F., & Blachnio, A. P. (2018). The role of beliefs in the use of hands-free and handheld mobile phones while driving. *Journal of Transport & Health*, *9*, 187–194.

Tabachnick, B. G., & Fidell, L. S. (1996). *Using Multivariate Statistics, New York: HarperCollins College Publishers*.

Tabachnick, B. G., Fidell, L. S., & Ullman, J. B. (2013). *Using multivariate statistics* (Vol. 6). pearson Boston, MA.

Tango, F., & Botta, M. (2013). Real-time detection system of driver distraction using machine learning. *IEEE Transactions on Intelligent Transportation Systems*, *14*(2), 894–905.

Tison, J., Chaudhary, N., Cosgrove, L., & Group, P. R. (2011). *National phone survey on distracted driving attitudes and behaviors*. United States. National Highway Traffic Safety Administration.

Torres, R., Ohashi, O., & Pessin, G. (2019). A machine-learning approach to distinguish passengers and drivers reading while driving. *Sensors*, *19*(14), 3174.

Truong, L. T., Nguyen, H. T. T., & De Gruyter, C. (2016). Mobile phone use among motorcyclists and electric bike riders: A case study of Hanoi, Vietnam. *Accident Analysis & Prevention*, *91*, 208–215.

Valencia, J. L., Diaz-Llanos, F. J., & Sainz-Calleja, F. J. (2003). Regresión PLS en las ciencias experimentales. *Madrid: Complutense*, 47–55.

Vermette, E. (2010). *Curbing distracted driving: 2010 survey of state safety programs*.

Wen, H., Sze, N. N., Zeng, Q., & Hu, S. (2019). Effect of music listening on physiological condition, mental workload, and driving performance with consideration of driver temperament. *International Journal of Environmental Research and Public Health*, *16*(15), 2766.

World Health Organization. (2020). *Road Traffic Accidents in Pakistan*. https://www.worldlifeexpectancy.com/pakistan-road-traffic-accidents

Yannis, G., Laiou, A., Papantoniou, P., & Christoforou, C. (2014). Impact of texting on young drivers' behavior and safety on urban and rural roads through a simulation experiment. *Journal of Safety Research, 49*, 25-e1.

Yannis, G., Papadimitriou, E., Karekla, X., & Kontodima, E. (2010). Mobile phone use by young drivers: effects on traffic speed and headways. *Transportation Planning and Technology*, *33*(4), 385–394.

Yu, J., Chen, Y., & Xu, X. (2018). *Sensing Vehicle Conditions for Detecting Driving Behaviors*. Springer.

Zhang, F., Mehrotra, S., & Roberts, S. C. (2019). Driving distracted with friends: Effect of passengers and driver distraction on young drivers' behavior. *Accident Analysis & Prevention*, *132*, 105246.

ANNEXURE 1

SECTION-I: Demographic Survey					
Question	**Option 1**	**Option 2**	**Option 3**	**Option 4**	**Option 5**
What is your gender?	Male	Female	Other		
Are you Married?	Yes	No			
What is your age in years?	18-30	30-40	40-50	50-60	60-70
Level of Education?	Matric	Inter.	Undergrad	Post-grad	Ph.D
What is your job?	Govt. Job	Private Job	Own business	Student	Jobless
What is your monthly household income?	20,000-35,000	35,000-100,000	100,000-150,000	150,000-200,000	Above 200,000
How much is your Driving Experience in years?	1-5	5-10	10-15	15-20	20-25 +
What vehicle do you drive?	Auto Car	Manual Car	Bike	Luxury Car	Bus/Van
When do you drive?	Daily	Alternate Days	Once/twice in week	Random	
Have you ever been in crash due to distracted driving?	Yes	No			
What was the nature of **that** Crash?	Fatal (someone died)	Critical	Severe	Serious	Moderate (property damage only)

SECTION-II Driving Under Given Conditions					
Driving Under Given Conditions					
On the scale of 1-5 **how much would you be distracted (using phone, talking, eating) during driving, under following conditions**? Ř ȡŀƴ Ăǁnjźƽ ȡ ĭ ǈ źź ź 1-5ź Ő ŋ×ź źƽ ˝ ź ×ş źź (Ŕ ȡ ȿź ǈ ǈ ź ŧź)Őź	Very Likely ***1***	Likely ***2***	Neutral ***3***	Unlikely ***4***	Very Unlikely ***5***
When driving on Local roads (small roads near home)					
When driving on Arterial road (e.g. Kashmir highway, N5)					
When driving on Motorway/Freeway					
When driving on non-familiar road					
When driving on a road, notorious for accidents or road is in bad condition					
When there's too much traffic on the road					
When driving during day time					
When driving during night time					
When driving with parents or elder relatives					
When driving with young kids or infants					
When driving with friends					
When driving with passenger and you are an (Uber/Careem) driver.					
When there's a fine on distraction activities					
When there's a warden on the road to fine					
When there's no fine on distraction activities					
When no fine and no warden on the road					

Distracted Driving Behavior							
On the scale of 1-5 how much you do following activities while driving? ž ź [illegible] 1-5 ,ź ź Ő ź ź þ			Very Likely	Likely	Neutral	Unlikely	Very Unlikely
			1	***2***	***3***	***4***	***5***
DISTRACTION ACTIVITIES	1	Talking to other passengers in the vehicle					
	2	Taking care of child					
	3	Eating/drinking/taking medicine					
	4	Using Phone					
	5	Doing personal grooming, such as putting on make-up, shaving, looking in the mirror					
	6	Use Dashboard for movie, music, GPS					
	7	Looking at scenery, ads , crash scene etc					

FILLED SAMPLE RESPONSES

RESPONSE 1

Anything that turns away your attention from driving is called ***Driving Distraction*** e.g. talking, eating, using phone, looking out of car window etc.

361

Kindly tick ✓ Only 1 Option

SECTION-I : Demographic Survey					
Question	**Option 1**	**Option 2**	**Option 3**	**Option 4**	**Option 5**
What is your gender?	Male	Female	Other		
Are you Married?	Yes	No			
What is your age in years?	18-30	30-40	40-50	50-60	60-70
Level of Education?	Matric	Inter.	Undergrad	Post-grad	Ph.D
What is your job?	Govt. Job	Private Job	Own business	Student	Jobless
What is your monthly household income PKR?	20,000-35,000	35,000-100,000	100,000-150,000	150,000-200,000	Above 200,000
How much is your Driving Experience in years?	1-5	5-10	10-15	15-20	20-25 +
What vehicle do you drive?	Auto Car	Manual Car	Bike	Luxury Car	Bus/Van
When do you drive?	Daily	Alternate Days	Once/twice in week	Random	
Have you ever been in crash due to distracted driving?	Yes	No			
What was the nature of **that** Crash?	Fatal(someone died)	Critical	Severe	Serious	Moderate(property damage only)

SECTION-II Prevalence of Distractions					
Scale: 1=Very Likely, 2= Likely, 3=Neutral, 4=Unlikely, 5= Very Unlikely					
On the scale of 1-5 **how much would you be distracted (using phone, talking, eating) during driving, under following conditions?** نیچے دی گئی صورتحالوں میں، 5-1 کے سکیل پر، آپ اپنے آپ کو دوران ڈرائیونگ دوسرے کاموں (جیسے کہ باتیں کرنا، فون استعمال کرنا وغیرہ) میں کتنا مشغول رکھتے ہیں؟	**Very Likely**	**Likely**	**Neutral**	**Unlikely**	**Very Unlikely**
When driving on Local roads (small roads near home)	O	O	O	O	O
When driving on Arterial road (e.g. Kashmir highway, N5)	O	O	O	O	O
When driving on Motorway/Freeway	O	O	O	O	O
When driving on non-familiar road	O	O	O	O	O
When driving on familiar road	O	O	O	O	O
When driving on a road, notorious for accidents or road is in bad condition	O	O	O	O	O
When driving on a road that is well managed and accident severity is low	O	O	O	O	O
When approaching a U-turn on a highway	O	O	O	O	O
When there's too much traffic on the road	O	O	O	O	O
When driving an Automatic Transmission Vehicle	O	O	O	O	O
When driving a Manual Transmission Vehicle	O	O	O	O	O
When riding a Bike	O	O	O	O	O
	O	O	O	O	O

When driving for business/work trip	O	O	O ✓	O	O
When driving for leisure/recreation trip	O	O	O	O	O ✓
When driving during day time	O ✓	O	O	O	O
When driving during night time	O	O	O	O ✓	O
When driving with parents or elder relatives	O	O ✓	O	O	O
When driving with young kids or infants	O	O	O	O ✓	O
When driving with friends	O ✓	O	O	O	O
When driving with passenger and you are an (Uber/Careem) driver.	O	O	O ✓	O	O
When there's a fine on distraction activities	O	O	O	O	O ✓
When there's a warden on the road to fine	O	O	O ✓	O	O
When there's no fine on distraction activities	O ✓	O	O	O	O
When no fine and no warden on the road	O	O	O ✓	O	O
When driving during very hot/cold weather	O	O	O	O	O ✓
When driving during foggy weather condition	O	O ✓	O	O	O
When driving during rain	O	O	O	O ✓	O

Frequency of Distractions

Scale: 1=Very Likely, 2= Likely, 3=Neutral, 4=Unlikely, 5= Very Unlikely

On the scale of 1-5 how much you do following activities while driving? نیچے دیئے گئے کاموں میں سے ,5-1 کے سکیل پر، دوران ڈرائیونگ آپ کونسے کام زیادہ کرتے ہیں؟		**Very Likely**	**Likely**	**Neutral**	**Unlikely**	**Very Unlikely**
DISTRACTION ACTIVITIES	Talking to other passengers in the vehicle	O ✓	O	O	O	O
	Taking care of child	O	O	O	O ✓	O
	Eating/drinking/taking medicine	O	O ✓	O	O	O
	Using Phone	O	O	O	O	O ✓
	Doing personal grooming, such as putting on make-up, shaving, looking in the mirror	O ✓	O	O	O	O
	Use Dashboard for movie, music, GPS	O	O	O ✓	O	O
	Looking at scenery, ads , crash scene etc	O ✓	O	O	O	O

RESPONSE 2

Anything that turns away your attention from driving is called ***Driving Distraction*** e.g. talking, eating, using phone, looking out of car window etc.

Kindly tick ✓ Only 1 Option

SECTION-I : Demographic Survey

Question	Option 1	Option 2	Option 3	Option 4	Option 5
What is your gender?	Male	Female ✓	Other		
Are you Married?	Yes	No ✓			
What is your age in years?	18-30 ✓	30-40	40-50	50-60	60-70
Level of Education?	Matric	Inter.	Undergrad	Post-grad	Ph.D ✓
What is your job?	Govt. Job	Private Job	Own business	Student ✓	Jobless
What is your monthly household income PKR?	20,000-35,000	35,000-100,000 ✓	100,000-150,000	150,000-200,000	Above 200,000
How much is your Driving Experience in years?	1-5 ✓	5-10	10-15	15-20	20-25 +
What vehicle do you drive?	Auto Car	Manual Car	Bike	Luxury Car	Bus/Van
When do you drive?	Daily	Alternate Days	Once/twice in week	Random ✓	
Have you ever been in crash due to distracted driving?	Yes	No ✓			
What was the nature of **that** Crash?	Fatal(someone died)	Critical	Severe	Serious	Moderate(property damage only)

SECTION-II Prevalence of Distractions

Scale: 1=Very Likely, 2= Likely, 3=Neutral, 4=Unlikely, 5= Very Unlikely

On the <u>scale of 1-5</u> **how much would you be distracted (using phone, talking, eating) during driving, under following conditions?** نیچے دی گئی صورتحالوں میں، 5-1 کے سکیل پر، آپ اپنے آپ کو دوران ڈرائیونگ دوسرے کاموں (جیسے کہ باتیں کرنا، فون استعمال کرنا وغیرہ) میں کتنا مشغول رکھتے ہیں؟	Very Likely	Likely	Neutral	Unlikely	Very Unlikely
When driving on Local roads (small roads near home)	○	○	○	●	○
When driving on Arterial road (e.g. Kashmir highway, N5)	○	○	●	○	○
When driving on Motorway/Freeway	○	○	●	○	○
When driving on non-familiar road	○	○	○	●	○
When driving on familiar road	○	○	●	○	○
When driving on a road, notorious for accidents or road is in bad condition	○	●	○	○	○
When driving on a road that is well managed and accident severity is low	○	○	●	○	○
When approaching a U-turn on a highway	○	●	○	○	○
When there's too much traffic on the road	○	●	○	○	○
When driving an Automatic Transmission Vehicle	○	○	●	○	○
When driving a Manual Transmission Vehicle	○	●	○	○	○
When riding a Bike	○	○	●	○	○

When driving for business/work trip	○	●	○	○	○
When driving for leisure/recreation trip	○	○	●	○	○
When driving during day time	○	●	○	○	○
When driving during night time	○	○	●	○	○
When driving with parents or elder relatives	○	○	○	○	○
When driving with young kids or infants	○	●	○	○	○
When driving with friends	○	●	○	○	○
When driving with passenger and you are an (Uber/Careem) driver.	○	●	○	○	○
When there's a fine on distraction activities	○	○	●	○	○
When there's a warden on the road to fine	○	●	○	○	○
When there's no fine on distraction activities	○	○	●	○	○
When no fine and no warden on the road	○	●	○	○	○
When driving during very hot/cold weather	○	○	●	○	○
When driving during foggy weather condition	○	●	○	○	○
When driving during rain	○	○	●	○	○

Frequency of Distractions

Scale: 1=Very Likely, 2= Likely, 3=Neutral, 4=Unlikely, 5= Very Unlikely

On the scale of 1-5 how much you do following activities while driving? نیچے دیئے گئے کاموں میں سے ,5-1 کے سکیل پر، دوران ڈرائیونگ آپ کونسے کام زیادہ کرتے ہیں؟		Very Likely	Likely	Neutral	Unlikely	Very Unlikely
DISTRACTION ACTIVITIES	Talking to other passengers in the vehicle	○	●	○	○	○
	Taking care of child	○	○	○	●	○
	Eating/drinking/taking medicine	○	○	○	●	○
	Using Phone	○	○	○	●	○
	Doing personal grooming, such as putting on make-up, shaving, looking in the mirror	○	○	○	●	○
	Use Dashboard for movie, music, GPS	○	○	●	○	○
	Looking at scenery, ads , crash scene etc	○	○	●	○	○