Risk factors, driver behaviour and accident probability. The case of distracted driving.

Doctoral Dissertation

Panagiotis Papantoniou

Athens, August 2015
Objective

The analysis of the effect of road, traffic and driver risk factors on driver behaviour and accident probability at unexpected incidents, with particular focus on distracted driving.

The development of risky driver profiles and road characteristics leading to increased possibility of driver error.
Methodological steps

- Literature review
- Methodological review
- Research questions
- Driving simulator experiment
- Statistical analysis
- Methodological contributions
- Key research findings
- Further research
Several literature reviews were implemented in the following research topics:

- Driver behaviour and road safety
- Driver distraction
- Types of assessing driver distraction
- Driving simulator characteristics
- Driving simulator studies on driver distraction
Driver behaviour and road safety

• **Human factors** are the basic causes in 65-95% of road accidents (Salmon et al., 2011).

• Human factors involve a large number of specific factors that may be considered as accident causes, including (Yannis et al., 2013):
  • **Driver injudicious action** (speeding, traffic violations etc.)
  • **Driver error or reaction** (loss of control, failure to keep safe distances, sudden braking etc.)
  • **Behaviour or inexperience** (aggressive driving, nervousness, uncertainty etc.)
  • **Driver distraction** (cell phone use, conversation with passenger etc.)
  • **Driver impairment** (alcohol, fatigue etc.)
Driver Distraction

- **Driver distraction** is generally defined as “a diversion of attention from driving, because the driver is temporarily focusing on an object, person, task or event not related to driving, which reduces the driver’s awareness, decision making ability and/or performance, leading to an increased risk of corrective actions, near-crashes, or crashes”

- Driver distraction may include four different types: *physical* distraction, *visual* distraction, auditory *distraction* and cognitive *distraction* (Ranney et al., 2000)

- Driver distraction factors can be subdivided into those that occur outside the vehicle (external) and those that occur inside the vehicle (in-vehicle).
Types of assessing driver distraction

The following **experiment types** of assessing driving behaviour and driver distraction exist (Papantoniou et al., 2015):

- Field tests
- Naturalistic driving experiments
- Driving simulator experiments
- In-depth accident investigation
- Stated preference surveys
Driving simulators allow for the examination of a range of driving performance measures in a controlled, relatively realistic and safe driving environment.

**Advantages**
- safe environment
- greater experimental control
- large range of test conditions (e.g., night and day, weather conditions, road environments)

**Disadvantages**
- data generally include the effect of learning
- feeling of safety
- simulator sickness
• Most experiments are based on **small samples**, limited to rural road environment and no explicit (if at all) simulation of ambient traffic

• Participants in almost all driving simulator experiments implemented a practice scenario, but no specific performance measures were used to assess the driver’s **familiarization**

• No pattern could be identified as regards the selection of **number and duration of trials**

• In 30% of studies no **counterbalancing** in the different trials was reported
Methodological review

Two targeted literature reviews took place in order to investigate:

• the key driving performance parameters examined in driver distraction research

• the statistical analyses implemented in the scientific field of driver distraction
Driving performance parameters

- Driver distraction is a **multidimensional phenomenon** which means that no single driving performance measure can capture all effects of distraction.

- A lot of different methods and **measures** exist for evaluating driving performance the most common of which include lateral control, longitudinal control, reaction time, gap acceptance, eye movement and workload measures.

- The selection of the specific measures should be guided by the nature of the task examined as well as the specific **research questions**.
• 5% of the examined studies perform only **descriptive statistics** tests aiming to gain general information regarding different performance measures.

• In more than half of the examined studies the main statistical analysis is **repeated measures ANOVA**.

• Latent model analysis and especially **structural equation models** have never been implemented in the field of driver distraction.
Research questions

1. Design and implementation of a simulator experiment aiming to deal with the majority of **limitations** that have been noted in the assessment of the examined simulator studies. The basic limitations found in the literature are the following:
   - Large and representative sample
   - Randomisation of trials
   - Adequate practice drive
   - Investigation of an optimum number of driving factors

2. Need to demonstrate a composite driving performance measure in order to examine driver distraction as a **multidimensional phenomenon**

3. Development and application of an **innovative statistical analysis methodology**

4. Estimation of the **combined effect** of distraction sources, driver as well as road and traffic environment characteristics directly on driving performance.
A common simulator experiment in the framework of two research projects:

- **Distract** - Analysis of causes and impacts of driver distraction

- **DriverBrain** - Analysis of the performance of drivers with cerebral diseases

An **interdisciplinary** research team:

- Dpt. of Transportation Planning and Engineering NTUA
- Dpt. of Neurology of the University of Athens Medical School, UoA
- Dpt. of Psychology, School of Philosophy, Pedagogy and Psychology, UoA
Driving simulator characteristics

Technical characteristics
• Foerst Driving Simulator FPF
• 3 LCD wide screens 40”
• total angle view 170°
• driving position and support base

Driving simulator validation
• Research in order to compare the driving performance of young drivers in normal and simulation driving conditions
• 31 young drivers aged 20-30 participated in an experimental process including driving both in a driving simulator as well in real traffic conditions

Results
• Absolute values of drivers' performance vary between simulated and real driving conditions
• Relative differences remain mostly the same
Exclusion criteria

Driving criteria
• Have a valid driving license
• had driven for more than 3 years
• had driven more than 2500km during the last year
• had driven at least once a week during the last year
• had driven at least 10km/week during the last year

Neurological criteria
• not had important psychiatric history for psychosis
• not had any important kinetic disorder that prevent them from basic driving moves
• not had dizziness or nausea either as a driver or as a passenger
• not be pregnant
• not be an alcoholic or had any other drug addiction
• not had any important eye disorder that prevent him from driving safely
• not had any disease of the Central Nervous System
The design of the driving scenarios is a central component of the present PhD thesis and includes:

**Area type**
- Rural area
  - 2.1 km long, single carriageway and the lane width was 3m, with zero gradient and mild horizontal curves
- Urban area
  - 1.7km long, lane width 3.5m, separated by guardrails

**Distraction conditions**
- No distraction
- Cell phone use
- Conversation with the passenger
Traffic scenarios
- Low traffic
  Αντιστοιχεί σε μέσο κυκλοφοριακό φόρτο $Q_L=300$ vehicles/hour
- High traffic
  Αντιστοιχεί σε μέσο κυκλοφοριακό φόρτο $Q_H=600$ vehicles/hour

Unexpected incidents
- Child crossing the road
- Sudden appearance of an animal

Randomisation
- The purpose of randomisation is to remove bias and other sources of extraneous variation, which are not controllable

<table>
<thead>
<tr>
<th></th>
<th>Urban area</th>
<th>Rural area</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>$Q_L$</td>
<td>$Q_H$</td>
</tr>
<tr>
<td>No distraction</td>
<td>√</td>
<td>√</td>
</tr>
<tr>
<td>Cell phone use</td>
<td>√</td>
<td>√</td>
</tr>
<tr>
<td>Conversation with the passenger</td>
<td>√</td>
<td>√</td>
</tr>
</tbody>
</table>
During the **familiarization** with the simulator, the participant practiced in:

- handling the simulator (starting, gears, wheel handling etc.)
- keeping the lateral position of the vehicle
- keeping stable speed, appropriate for the road environment
- Braking and immobilization of the vehicle

During this practice drive, two **unexpected incidents** took place.

The following criteria must be verified (there is **no time restriction**) before the participant moves on to the next phase of the experiment:
Procedure of experiment

Researcher – coordinator of the experiment:
• the oral briefing and the delivery of the instructions to the participant
• assisting the participant during their familiarization drive
• assisting the participant to fill in the Self-assessment and Memory questionnaire
• filling a checklist (see Annex) for the control of the experiment with any comments related to anything remarkable regarding the driving of the participant
• the monitoring for and handling of simulator sickness
• the accomplishment of the driving simulator experiment
• assisting the participant in any other issue

Researcher responsible for:
• performing the distraction tasks during the experiment: the conversation task and the phone call with the participant
• assisting for any other secondary issues during the experiment
• organizing the files generated from the participants’ driving and editing statistically the data
Driving behaviour questionnaire
- Driving experience - car use
- Self-assessment of the older driver
- Distraction-related driving habits
- Emotions and behaviour of the driver
- Anger expression inventory during driving
- History of accidents, near misses, and traffic violations

Self-Assessment and memory questionnaire
- Memory
- Self assessment
- Driving skills
The sample of the analysis consists of **95 participants**

- 28 young drivers aged 18-34 years old
- 31 were middle aged drivers aged 35-54 years old
- 36 older driver aged 55-75 years old
Statistical analysis methodology

Data collected from the driving simulator experiment and the respective questionnaires are analysed by means of a dedicated statistical analysis method:

1. **Descriptive analysis**  
   (correlation table, boxplots)

2. **Regression analysis**  
   (6 general linear mixed models)

3. **Factor Analysis**  
   (2 factor analysis)

4. **Latent analysis**  
   (4 structural equation models)
Descriptive analysis

- **Database development**
  - Type of variable
  - Min, max, average value

- Several **boxplots** are presented in order to explain the effect of specific driver, road and traffic parameters as well as the examined distraction sources on selected driving performance measures.

- A **correlation table** is investigating any of a broad class of statistical relationships between driving simulator variables.
Within the framework of regression analysis, **6 general linear mixed models** are developed in order to identify several sets of explanatory variables that covary with specific driving performance measures of the driving simulator dataset.

- Average speed
- Reaction time
- Lateral position
- Average headway
- Speed variability
- Lateral position variability
Factor analysis

- Two factor analysis are developed in order to investigate which observed variables are most highly correlated with the common factors of driving performance and driver error and how many common factors are needed to give an adequate description of the data.

- Regarding driving performance, 5 factors are best fitted in the specific database. The interpretation of the results revealed that the five factors are: lateral measures, speed measures, vehicle direction measures, headway as well as vehicle revoluation.

- The variables that tend to explain better the “Driver Error” factor are: numbers of Outside Road Lines, Sudden Brakes and High Rounds per Minute.
Structural Equation Models (1/2)

• Structural Equation Modeling is a very general, powerful multivariate analysis technique that includes several analysis methods.

• SEM involves the evaluation of two models:
  
  **Measurement Model**
  - The part of the model that relates indicators to latent factors
  - The measurement model is the factor analytic part of SEM

  **Path model**
  - This is the part of the model that relates variable or factors to one another (prediction)
  - If no factors are in the model then only path model exists
Structural Equation Models (2/2)

Path diagram

- Latent variables
- Observed variables
- Single-headed arrow →
  Regression Coefficient or factor loading
- Double headed arrow ↔
  Correlations

Goodness-of-fit measures

- Standardized Root Average Square Residual (SRMR) < 0.08
- Root Average Square Error of Approximation (RMSEA) < 0.08
- Comparative Fit Index (CFI) > 0.90
- Tucker Lewis Index (TLI) > 0.90
Latent analysis overview

Risk Factors

Driver Error

Driving Performance

- Age
- Gender
- Experience
- Education
- Traffic
- Area type

Outside Road Lane
Hit of Side Bars
High Round per Minute

Time To Line Crossing
Stdev Lateral Position
Average Speed
Average Gear

Accident Probability at unexpected incidents
The latent variable reflects the underlying driving performance and the objective is the quantification of the impact of distraction, driver characteristics as well as road and traffic environment on driving performance.
### SEM regarding driving performance (2/3)

| Latent Variable                        | Est. | Std.err | t value. | P(>|z|) |
|----------------------------------------|------|---------|----------|--------|
| Driving Performance                    |      |         |          |        |
| Average Speed                          | 1,000| -       | -        | -      |
| Stdev Lateral Position                 | -0,085| 0,004   | -23,909  | 0.000  |
| Average Gear                           | 0,048| 0,002   | 21,887   | 0.000  |
| Time to Line Crossing                  | -0,109| 0,005   | -19,972  | 0.000  |

### Regressions

| Regression                                | Est.  | Std.err | t value. | P(>|z|) |
|-------------------------------------------|-------|---------|----------|--------|
| Distraction – Cell phone                  | -1,099| 0,342   | -3,213   | 0.001  |
| Area - Urban                              | -15,596| 0,467   | -33,410  | 0.000  |
| Traffic - Low                             | 1,123 | 0,285   | 3,943    | 0.000  |
| Gender - Female                           | -1,154| 0,303   | -3,802   | 0.000  |
| Age                                       | -0,155| 0,027   | -5,755   | 0.000  |
| Experience                                | 0,083 | 0,032   | 2,630    | 0.009  |

### Summary statistics

| Minimum Function Test | 305,74 |
| Degrees of freedom   | 20     |

### Goodness of fit

| SRMR         | 0,061  |
| RMSEA        | 0,136  |
| CFI          | 0,867  |
| TLI          | 0,809  |
• The effect of **cell phone** on driving performance is definitely negative
• **Conversation with the passenger** does not have a statistically significant effect
• Risk factors that affect driving performance include driver characteristics (**age**, **gender**, **driving experience**), **area type** and **traffic conditions**
SEM regarding driver error (1/3)

The latent variable reflects the underlying driver error and the objective is the quantification of the impact of distraction, driver characteristics as well as road and traffic environment on driving error.
SEM regarding driver error (2/3)

| Latent Variable                  | Est. | Std.err | t value. | P>|z|< |
|----------------------------------|------|---------|----------|-------|
| Driving Error                    |      |         |          |       |
| Hit Of Side Bars                 | 1.000| -       | -        | -     |
| Outside Road Lanes               | 0.741| 0.257   | 2.887    | 0.004 |
| High Rounds Per Minute           | 0.680| 0.243   | 2.803    | 0.005 |

<table>
<thead>
<tr>
<th>Regressions</th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Driver Errors</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Gender - Female</td>
<td>0.359</td>
<td>0.076</td>
<td>4.739</td>
<td>0.000</td>
</tr>
<tr>
<td>Age</td>
<td>0.031</td>
<td>0.009</td>
<td>3.393</td>
<td>0.001</td>
</tr>
<tr>
<td>Area - Urban</td>
<td>-0.393</td>
<td>0.062</td>
<td>-6.383</td>
<td>0.000</td>
</tr>
<tr>
<td>Experience</td>
<td>-0.030</td>
<td>0.010</td>
<td>-3.050</td>
<td>0.002</td>
</tr>
<tr>
<td>Education</td>
<td>-0.021</td>
<td>0.010</td>
<td>-2.167</td>
<td>0.030</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Summary statistics</th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Minimum Function Test</td>
<td>62.19</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Degrees of freedom</td>
<td>10</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Goodness of fit</th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>SRMR</td>
<td>0.032</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>RMSEA</td>
<td>0.096</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>CFI</td>
<td>0.823</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>TLI</td>
<td>0.682</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Risk Factors</th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Accident Probability</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>at unexpected incidents</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Driving Performance</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Neither conversing with a passenger nor talking on the cell phone has a statistical significant impact on driver error.

Risk factors that affect driver error include gender, age, experience, education and area type.
Two latent variables are created regarding driving performance and driver error while the objective of this analysis is the quantification of the impact of driving errors, distraction, driver characteristics as well as road and traffic environment on driving performance.
## SEM regarding driving performance and driver error (2/3)

### Latent Variable 1
**Driver Errors**
- **Hit Of Side Bars**: 1,000<br>
  Std.err: -<br>
  t value: -<br>
  P(>|z|): -
- **Outside Road Lanes**: 0,547<br>
  Std.err: 0,214<br>
  t value: 2,559<br>
  P(>|z|): 0,010
- **High Rounds Per Minute**: 0,950<br>
  Std.err: 0,276<br>
  t value: 3,436<br>
  P(>|z|): 0,001

### Latent Variable 2
**Driving Performance**
- **Average Speed**: 1,000<br>
  Std.err: -<br>
  t value: -<br>
  P(>|z|): -
- **Stdev Lateral Position**: -0,085<br>
  Std.err: 0,004<br>
  t value: -23,117<br>
  P(>|z|): 0,000
- **Average Gear**: 0,049<br>
  Std.err: 0,002<br>
  t value: 22,043<br>
  P(>|z|): 0,000
- **Average TTL**: -0,108<br>
  Std.err: 0,005<br>
  t value: -20,114<br>
  P(>|z|): 0,000

### Regression 1
**Driving Performance**
- **Driver Errors**: -51,016<br>
  Std.err: 11,417<br>
  t value: 4,468<br>
  P(>|z|): 0,000
- **Gender – Female**: -16,739<br>
  Std.err: 3,799<br>
  t value: -4,407<br>
  P(>|z|): 0,000
- **Age**: -2,244<br>
  Std.err: 0,681<br>
  t value: -3,297<br>
  P(>|z|): 0,001
- **Experience**: 2,103<br>
  Std.err: 0,694<br>
  t value: 3,031<br>
  P(>|z|): 0,002

### Regression 2
**Driver Errors**
- **Gender - Female**: 0,311<br>
  Std.err: 0,076<br>
  t value: 4,068<br>
  P(>|z|): 0,000
- **Age**: 0,042<br>
  Std.err: 0,010<br>
  t value: 4,125<br>
  P(>|z|): 0,000
- **Area - Urban**: -0,300<br>
  Std.err: 0,068<br>
  t value: -4,395<br>
  P(>|z|): 0,000
- **Experience**: -0,040<br>
  Std.err: 0,011<br>
  t value: -3,815<br>
  P(>|z|): 0,000
- **Education**: 0,004<br>
  Std.err: 0,001<br>
  t value: 3,174<br>
  P(>|z|): 0,002

### Summary statistics
- **Minimum Function Test**: 608,01
- **Degrees of freedom**: 40

### Goodness of fit
- **SRMR**: 0,088
- **RMSEA**: 0,158
- **CFI**: 0,793
- **TLI**: 0,711

![Diagram showing the relationships between driver errors, driving performance, risk factors, and accident probability]
• Driver error is a crucial factor that negatively affects driving performance
• Neither **road characteristics** (area type, traffic conditions) nor the **distraction sources** examined (cell phone use, conversation with a passenger) have a significant impact on this model
The latent variable reflects again the underlying driving performance of the participants and the objective is the quantification of the impact of driving performance, distraction, driver characteristics as well as road and traffic environment directly on **accident probability at unexpected incidents**
SEM regarding accident probability (2/3)

| Latent Variable          | Est.  | Std.err | t value. | P(>|z|) |
|--------------------------|-------|---------|----------|---------|
| Driving Performance      |       |         |          |         |
| Average Speed            | 1,000 | -       | -        | -       |
| Stdev Lateral Position   | -0,085| 0,004   | -23,803  | 0.000   |
| Average Gear             | 0,048 | 0,002   | 21,836   | 0.000   |
| Average TTL              | -0,109| 0,005   | -20,046  | 0.000   |

<table>
<thead>
<tr>
<th>Regression</th>
<th>Accident</th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Driving Performance</td>
<td>-0,007</td>
<td>0,002</td>
<td>-3,119</td>
<td>0.002</td>
</tr>
<tr>
<td>Gender - Female</td>
<td>0,074</td>
<td>0,034</td>
<td>2,198</td>
<td>0.028</td>
</tr>
<tr>
<td>Traffic – Low</td>
<td>0,104</td>
<td>0,033</td>
<td>3,142</td>
<td>0.002</td>
</tr>
<tr>
<td>Distraction – Cell phone</td>
<td>0,081</td>
<td>0,033</td>
<td>2,463</td>
<td>0.014</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Regression</th>
<th>Driving Performance</th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Gender - Female</td>
<td>-1,147</td>
<td>0,307</td>
<td>-3,737</td>
<td>0.000</td>
</tr>
<tr>
<td>Area - Urban</td>
<td>-15,614</td>
<td>0,468</td>
<td>-33,386</td>
<td>0.000</td>
</tr>
<tr>
<td>Distraction – Cell phone</td>
<td>-1,099</td>
<td>0,343</td>
<td>-3,208</td>
<td>0.001</td>
</tr>
<tr>
<td>Traffic - Low</td>
<td>1,131</td>
<td>0,286</td>
<td>3,956</td>
<td>0.000</td>
</tr>
<tr>
<td>Age</td>
<td>-0,156</td>
<td>0,028</td>
<td>-5,593</td>
<td>0.000</td>
</tr>
<tr>
<td>Experience</td>
<td>0,083</td>
<td>0,032</td>
<td>2,557</td>
<td>0.011</td>
</tr>
</tbody>
</table>

Summary statistics
- Minimum Function Test: 352.62
- Degrees of freedom: 31

Goodness of fit
- SRMR: 0.061
- RMSEA: 0.136
- CFI: 0.867
- TLI: 0.807
• **Cell phone** use has a negative effect on accident probability
• Drivers self-regulate their driving performance better while conversing with a passenger
• **Female** drivers at **low traffic** are more prone to accidents at unexpected incidents
Scientific contributions

1. Risky driving profiles
2. Estimation of driving performance and safety
3. Innovative statistical analysis methodology
4. Large and rigorous driving simulator experiment
Design and implementation of a large and rigorous driving simulator experiment

The basic limitations found in the literature that the present experiment tackled are the following:

- Large and representative sample
- Randomisation of trials
- Adequate practice drive
- Investigation of an optimum number of driving factors
Methodological contributions (2/2)

• Development and application of an **innovative** statistical analysis methodology

• **Latent analysis** through Structural Equation models is implemented for the first time in the field of driving performance and traffic safety

• Estimation of the **combined effect** of distraction sources, driver as well as road and traffic environment characteristics directly on driving performance
• Results regarding the effect of **driver distraction** indicate the different effect on driving performance between cell phone use and conversation with the passenger.

• **Driver characteristics** play the most crucial role in driving performance (gender, age, experience).

• Driving performance is worst in urban areas and high traffic conditions probably due to the complex **driving environment**.
Development of risky driver profiles regarding driver error and accident probability at unexpected incident. Results indicate that:

• more likely to commit driving errors are **young or old female drivers at urban areas**

• more likely to be involved in an accident at an unexpected incident are **female drivers in low traffic conditions while talking on the cell phone**
Further research

• Investigation of the effect of other parameters such as alcohol, fatigue etc. on driving performance through latent analysis

• Development of Structural Equation Model on different experimental methods (Naturalistic experiments, field test etc.)

• Further investigation of the parameters that affect the compensatory behaviour of the driver

• Investigation of different types of cell phone use such as a hands-free, bluetooth, typing an sms etc.)
Risk factors, driver behaviour and accident probability. The case of distracted driving.

Doctoral Dissertation

Panagiotis Papantoniou

Athens, August 2015