InDev: In-Depth understanding of accident causation for Vulnerable road users

HORIZON 2020 - the Framework Programme for Research and Innovation

Deliverable 2.1 – part 3 of 5

Review of current study methods for VRU safety

Appendix 5 – Systematic literature review: Behavioural observations

Due date of deliverable: (30.08.2016)

Start date of project: 01.May 2015
Duration: 36 months

Organisation name of lead contractor for this deliverable:
(Warsaw University of Technology, Poland)

Revision 1.2

<table>
<thead>
<tr>
<th>Dissemination Level</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>PU</td>
<td>Public</td>
</tr>
<tr>
<td>PP</td>
<td>Restricted to other programme participants (including the Commission Services)</td>
</tr>
<tr>
<td>RE</td>
<td>Restricted to a group specified by the consortium (including the Commission Services)</td>
</tr>
<tr>
<td>CO</td>
<td>Confidential, only for members of the consortium (including the Commission Services)</td>
</tr>
</tbody>
</table>
Document information

Authors
Appendix 5: Wouter van Haperen (HU)

Project Coordinator
Aliaksei Laureshyn
Department of Technology and Society
Lund University
Box 118
221 00 Lund, Sweden

Phone: +46 46 222 91 31
Email: aliaksei.laureshyn@tft.lth.se

Coordinator of WP 2
Piotr Olszewski
Department of Civil Engineering
Warsaw University of Technology
Al. Armii Ludowej 16
00-637 Warsaw, Poland

Phone: +48 22 234 6331
Email: p.olszewski@il.pw.edu.pl

Project funding
Horizon 2020
Grant agreement No. 635895
Revision and History Chart

<table>
<thead>
<tr>
<th>Version</th>
<th>Date</th>
<th>Comment</th>
</tr>
</thead>
<tbody>
<tr>
<td>1.1</td>
<td>16-07-13</td>
<td>First draft</td>
</tr>
<tr>
<td>1.2</td>
<td>16-07-29</td>
<td>Second draft</td>
</tr>
</tbody>
</table>

This project has received funding from the *European Union’s Horizon 2020 research and innovation programme* under grant agreement No 635895.

This publication reflects only the authors’ view. Responsibility for the information and views expressed therein lies entirely with the authors. The European Commission is not responsible for any use that may be made of the information it contains.
# Table of Contents

Summary .................................................................................................................. 1

1. Introduction ........................................................................................................ 2

2. Scope .................................................................................................................. 4

3. Methodology ...................................................................................................... 6
   3.1. Review team ................................................................................................ 6
   3.2. Search protocol ......................................................................................... 6
   3.3. Screening .................................................................................................. 7
   3.4. Paper retrievals ...................................................................................... 7

4. Codebook ........................................................................................................... 9
   4.1. General Information .............................................................................. 9
   4.2. Research Topic ....................................................................................... 9
   4.3. Methodological information .................................................................. 11
   4.4. Indicators ............................................................................................... 11
       4.4.1. Infrastructural indicators .............................................................. 12
       4.4.2. Traffic Safety Aids ...................................................................... 12
       4.4.3. Distractions ............................................................................... 13
       4.4.4. Road user characteristics ............................................................ 13

5. Findings ............................................................................................................. 14
   5.1. Road user focus ....................................................................................... 14
       5.1.1. Countries ..................................................................................... 16
   5.2. Topics and indicators ............................................................................. 16
       5.2.1. Infrastructure ............................................................................. 19
       5.2.2. Topics ......................................................................................... 19
       5.2.3. Topics of interest ....................................................................... 24
   5.3. Purpose of road user behaviour observation ............................................ 25
   5.4. Methodological aspects .......................................................................... 27
       5.4.1. Research Design ......................................................................... 27
       5.4.2. Semi-controlled research ............................................................. 28
       5.4.3. Control groups ........................................................................... 29
       5.4.4. Data collection techniques ........................................................... 29
       5.4.5. Combination with other methodologies ....................................... 30
       5.4.6. Number of sites ........................................................................... 31
       5.4.7. Sample sizes ............................................................................... 32
       5.4.8. Observation period ..................................................................... 32

6. Discussion .......................................................................................................... 34
   6.1. Review process ......................................................................................... 34
       6.1.1. Bias ............................................................................................. 34
       6.1.2. Limitations ................................................................................... 35
   6.2. Strengths, weaknesses, opportunities and threats ..................................... 35
       6.2.1. Strengths ..................................................................................... 36
       6.2.2. Weaknesses ............................................................................... 36
       6.2.3. Opportunities .............................................................................. 37
       6.2.4. Threats ......................................................................................... 37
   6.3. Methodological considerations ................................................................. 38

7. Conclusion ......................................................................................................... 39
8. Acknowledgements .................................................................................................................40
9. References ...............................................................................................................................41
  9.1. References included in review ..............................................................................................41
  9.2. Additional references in report ...........................................................................................57
  9.3. Potential relevant but irretrievable references .................................................................58
ANNEX 1 .......................................................................................................................................60
List of Figures

Figure 1: The Swiss Cheese Model of Reason (2002). Figure retrieved from http://www.skybrary.aero. ........................................ 2
Figure 2: The flow chart of the screening process. .................................................................................................................. 8
Figure 3: The amount of studies including the specified road user types. On the left, VRU-studies are shown (n = 214) and driver-studies are depicted on the right (n = 477). ........................................................ 15
Figure 4: On the left, the evolution over time of research focusing on VRUs and drivers. On the right, the three most common road user types are plotted. Due to low numbers in earlier years, the graphs depict the period from 1965 till 2015. ........................................................ 15
Figure 5: The road user focus of the different countries conducted road user behavioural observations, sorted based on number of VRU-studies. .............................................................................. 17
Figure 6: The pie charts showing the countries with most research efforts for the defined road user types. ........................................................................................................................................................................... 18
Figure 7: The infrastructural elements of interest, sorted by number of VRU-studies. .......................................................... 19
Figure 8: Pie charts depicting the shares of topic-categories. ......................................................................................... 20
Figure 9: The research topics found in included references. Data is sorted per category based on the number of references for studies relating to VRU-types. ................................................................. 21
Figure 10: The indicator-categories. ............................................................................................................................ 22
Figure 11: The indicators of the included studies. Data is sorted per category based on the number of VRU-studies. ................................................................................................................................. 23
Figure 12: On top, the shares of the research goals. Below, their evolution over time. ......................................................... 26
Figure 13: On top, the shares of the research designs. Below, their evolution over time. ...................................................... 28
Figure 14: The number of studies including the defined data collection tools. ................................................................. 30
Figure 15: The cumulative distribution of the use of data sources for studies in which all road users are included. ................................................................................................................................. 30
Figure 16: The combination of behavioural observation studies with other methodologies. .............................................. 31
Figure 17: The number of sites related to research goal. ................................................................................................. 31
Figure 18: The observed sample sizes. ......................................................................................................................... 32
Figure 19: Reported observation periods. ...................................................................................................................... 33

List of Tables

Table 1: The most common safety evaluation methods used in scientific literature. ......................................................... 2
Table 2: The topic types defined in this review. .................................................................................................................... 10
Table 3: The infrastructural indicators found in the behavioural observation studies. .................................................... 13
Table 4: Semi-controlled research ........................................................................................................................................... 29
Table 5: The use of control groups in road user behavioural observation studies (n=21). ................................................ 29
Table 6: A SWOT analysis on road user behavioural observation studies. .............................................................................. 36

List of Abbreviations

VRU - Vulnerable Road User
Summary

Observation of road user behaviour has been reported since the 1930, but especially during recent years the number of (peer-reviewed) studies and reports is increasing rapidly. Several methodologies have been developed to study road user behaviour, of which behavioural observation studies aim at collecting naturalistic behavioural data. However, an overview of the current extent, range and nature of this type of research is lacking. Therefore, a scoping review was performed in order to identify how road user behavioural observation studies have been conducted, which topics have been covered and which research gaps still exist in literature, focusing on the evaluation of safety through measurements of behaviour. The aim of this report is to a) provide an overview of conducted road user behaviour observation studies, b) assess their usefulness, c) prevent duplicate research efforts, d) identify which indicators have been applied and e) indicate which areas of road user behaviour research needs further examination.

The review team, consisting of two members, carefully created and tested a search protocol to systematically retrieve relevant literature from three major online databases (ScienceDirect, Web of Science and TRID). The search term “Traffic Behaviour” AND “Safety OR Observation” was utilized and yielded more than 21,000 results. After the removal of duplicates and several screening rounds, 583 papers remained. Studies were excluded if they were published in any other language than English, if it only contained stated behaviour (e.g., questionnaires and focus groups) and if the data was collected with the participants’ awareness of being part of an experiment (e.g., naturalistic driving and driving simulator research). Based on subsamples of the included papers, a codebook was designed in order to extract relevant information. The publication years ranged from 1939 till 2015 and the majority of the studies were conducted in the USA (38%), Canada (8%) and China (8%). It was found that 36% of all included studies contained at least one vulnerable road user type (VRU-studies), while at least one other road user type was present in 82% of all studies (driver-studies). For both study categories, the main goal of behavioural observation is to simply observe what happens (> 50%), followed by the evaluation of safety improving treatments (around 30%) and the development of microsimulation models or automated video-analysis software. In total, 26 research topics were identified, of which VRU-studies mainly focused on crossing behaviour, yielding and red-light running, while for driver-studies the area of speeding is most often examined. Furthermore, the review identified 47 indicators used for behavioural observation analyses, of which red light running and yielding are most often used in VRU-studies and speed and headways in driver-studies.

Based on the findings of this review, it was concluded that road user behavioural observation studies are a useful tool to investigate underlying processes of (vulnerable) road users’ safety. The main strength of such studies is that naturalistic data is gathered without road users being aware that they are being observed for research purposes. It enables the observation and identification of behavioural and situational processes that contribute to crash occurrence. However, such studies are limited to what happens, since researchers cannot manipulate or control the driving environment. Because the use of video cameras to capture behavioural observations has become the major data collection technique in recent years, the current efforts to improve and further develop automated video-analysis software tools might prove to be a valuable asset in behavioural observation studies and traffic safety evaluation in general.
1. Introduction

The term ‘traffic safety’, or ‘road safety’, is well known by specialists and the general public, but a clear definition has never been formulated (Evans, 2004, p7; Elvik et al., 2009, p3). In its most basic form, safety means the absence of unintended harm to living creatures or inanimate objects. However, most studies and reports focusing on road safety consider only fatal and (severe) injury crashes when discussing traffic safety levels (e.g. WHO, 2009; EU, 2015), thereby ignoring crashes in which only minor or no injuries at all occur. These studies use crash data retrieved from police and hospital departments, even though problems regarding this form of data collection have been acknowledged. The most important limitation of crash data regards underreporting or incorrect reporting. Elvik and Mysen (1999) conducted a meta-analysis of 49 studies from 13 countries and found that in general underreporting rates were the highest for cyclists and lowest for car occupants, when hospital data and police reports were compared. However, the precise underreporting rates remains unclear, since it is very likely that accidents without physical or property damage are not reported to either source. Another important limitation of traditional crash data analysis regards the inability to capture behavioural and situational processes that have contributed to the occurrence of the crash. According to the Swiss Cheese model, Reason et al. (2006) argues that crashes are caused by a series of safety compromising events and that crashes will only occur if all these events occur at the same time (Figure 1). It is therefore important to identify which safety compromising events can contribute to crash occurrence, in order to develop efficient safety improving measures and programmes. Traditional crash data cannot offer such information.

![Figure 1: The Swiss Cheese Model of Reason (2002).](http://www.skybrary.aero)

Multiple methodologies to evaluate traffic safety that do not use traditional crash data have been proposed and applied in scientific literature (Table 1). These approaches might be considered as proxies for traffic safety, since they do not analyse or evaluate crashes itself, but focus on processes and events that precede them. These methodologies can be distinguished in naturalistic data collection (reflecting driving behaviour in the real environment) or controlled data collection (in which researchers have the ability to manipulate and control traffic events).

<table>
<thead>
<tr>
<th>Actual (objective) traffic safety</th>
<th>Proxy for actual (objective) traffic safety</th>
</tr>
</thead>
<tbody>
<tr>
<td>Naturalistic data collection</td>
<td>Controlled data collection</td>
</tr>
<tr>
<td>Revealed</td>
<td>Simulated</td>
</tr>
<tr>
<td>Crash data</td>
<td>Stated</td>
</tr>
<tr>
<td>Behavoural Observation*</td>
<td>Driving Simulator</td>
</tr>
<tr>
<td>Traffic conflict observation</td>
<td>Microsimulation</td>
</tr>
<tr>
<td>Naturalistic Driving</td>
<td>Questionnaires</td>
</tr>
<tr>
<td></td>
<td>Interviews</td>
</tr>
<tr>
<td></td>
<td>Focus groups</td>
</tr>
</tbody>
</table>

* The focus of this scoping review.
Controlled data collection techniques can be further divided into stated and simulated research methodologies. Stated research designs investigate attitudes and behaviours based on questioning participants, usually with the use of questionnaires, interviews or focus groups. These methodologies are limited in the sense that collected data is dependent on behaviour that is stated by respondents (e.g. attitudes, opinions and beliefs), rather than behaviour actually shown in a natural setting. The advantage of simulated studies is that different scenarios or new types of infrastructure can be tested, without the need of constructing them first. However, known issues regard behavioural adaptation during driving simulator research and issues regarding validity for both driving simulator research and microsimulation studies (Jenness, 2007; Bedard et al., 2010; Schectman et al., 2009; Y Wang et al. 2010).

Naturalistic data collection methodologies focus on the observation of behaviour that is shown in the real driving environment. A distinction can be made between studies in which road users are aware of being observed (naturalistic driving studies) and studies in which unobtrusive observations of road user behaviour are made (behavioural observation and traffic conflict studies). Although the concept of unobtrusive is not clearly defined in behavioural observation studies, it seems generally regarded as the avoidance of informing road users of their participation in the study. When road users know that they are being observed, they may change the behaviour of interest (Porter, 2011). Naturalistic driving studies enable the collection of vast amount of data during individuals’ trips, but the awareness of the sensors on their vehicle might cause behavioural adaptation effects. Therefore, we regard behavioural observation studies, and, as an extension traffic conflict studies, to most closely represent natural road user behaviour. The main difference between these two methodologies is that traffic conflict studies try to measure traffic safety in terms of the expected number of (injury) accidents, while road user behavioural observation studies focus on observing what happens, rather than to quantify traffic safety levels.

This scoping review is part of the Horizon 2020 project InDeV, in which several reviews are conducted in order to assess the usefulness of different methodologies to evaluate vulnerable road users’ safety. These reviews have as main objective to critically review the usefulness of different methodologies for accident causation studies with relevance to vulnerable road users in particular. This review focuses on the applicability of road user behavioural observation studies and aims to provide a descriptive insight into the current practice reported in scientific literature. Research topics, behavioural indicators and methodological aspects are examined in order to a) provide an overview of conducted road user behaviour observation studies, b) assess their usefulness, c) prevent duplicate research efforts, d) identify which methodological aspects and indicators are important and e) indicate which areas of road user behaviour research needs further examination.
2. Scope

Behavioural observation studies using naturalistic data was one of the first methods used in traffic safety research. Nearly a century ago, Dodge (1923) argued that studying human factors (behaviour) of road users is important to improve traffic safety. The oldest behavioural observation study found (and retrievable) by the authors dates back to 1934 (Greenshields et al., 1934), in which the method of taking pictures was presented as a new form of data collection for analysing road user behaviour. Since then, the use of behavioural observation studies has become common practice for various research aims, including testing the effectiveness of a certain countermeasure or to develop models for microsimulation software. The amount of published peer-reviewed journal articles is growing rapidly, but an overview of current research efforts is lacking. In order to guarantee the transparency and replicability and to clearly communicate the focus of this review, key concepts are defined as follows:

- **Road users** are all users of the road infrastructure that can move freely and are not constraint to guiding systems (e.g. trains on rails). Transportation modes that are guided are excluded since drivers/riders of those modes have very limited control over their own direction. Aviation and ships are excluded from the review, since they do not make use of the (public) road infrastructure.
- **Road user safety** is the absence of unintended harm to road users or damage to the vehicle of their mode of transportation. Road user safety in this research is independent of injury severity.
- **Road user behavioural observation studies** (from here referred to as behavioural observation studies) are studies observing road user behaviour, in which the road users observed are not informed (beforehand) of their participation in the research. These studies focus on how road users pass the observation site in a naturalistic setting and should be related to traffic safety aspects.
- **Vulnerable road users** (VRUs) are those road users that do not have a protective shell around them (Wegman and Aarts, 2006). These include, among others, pedestrians, cyclists and riders of powered-two-wheelers.
- **All road users that do not comply to the definition of vulnerable road users were labelled as drivers in this research. Examples are cars, trucks and busses.**

The amount of documents reporting on research of road user behaviour is extensive. A preliminary search on three major online databases yielded over 21,000 papers, reports and other research documents. In order to guarantee the feasibility of the review, it was therefore important to formulate exclusion criteria. The most important requirements for inclusion are listed below:

- **Document type**: Only peer-reviewed journal articles were included in this review, thereby discarding research reports, book sections and conference proceedings.
- **Publication year**: Road user behaviour has already been observed since the 1930s. Although problems regarding the accessibility of papers published before the 1990’s arise – those papers are not necessarily digitalized, nor might authors be still professionally active to be contacted – it is preferable to include those older studies as well, since they provide insight into how the behavioural observation techniques have evolved over time. Furthermore, exclusion of older publications might sketch a distorted perspective on the actual usefulness of behavioural observation studies. Therefore, no restriction on publication year was introduced.
• **Publication language:** Only articles in English were considered for inclusion, due to the limitation of the knowledge of (certain) foreign languages.

• **Involved road users:** Although the main focus of the InDeV project involves VRUs, it was decided to include all peer-reviewed articles regarding behavioural observations. This decision enables us to examine to what extent such studies have been applied to VRUs in particular and might identify research gaps or methodological challenges.
3. Methodology

Because the amount of published studies concerning behavioural observation studies increased rapidly in recent years, synthesizing available evidence is important. A review of the literature helps in summarising findings to identify relevant research opportunities. Most studies use narrative reviews, which use implicit processes to provide evidence (Garg et al., 2008). However, the reader cannot determine if this evidence is based on the author’s experience, how much literature was searched for and if certain studies were ignored due to contradicting findings. Studies described in a narrative review are mostly those that reinforce the ideas and research objectives of the study being conducted. In order to avoid subjectivity in the process of summarizing literature that is available on a certain topic, other reviewing techniques have been developed. Scoping reviews, for example, use a systematic approach for retrieving relevant articles. Such reviews aim to “map rapidly the key concepts underpinning a research area and the main sources and types of evidence available. They can be undertaken as stand-alone projects in their own right, especially where an area is complex or has not been reviewed comprehensively before” (Mays, Roberts, & Popay, 2001; Wilson, Lavis, & Guta, 2012). Guiding future research and reducing duplicate efforts are important objectives (Armstrong et al., 2011; Wee & Banister, 2016). An additional advantage is that such reviews can be used for many applications, even outside the authors’ intended purposes (Armstrong et al., 2011). The quality of a scoping review is determined by its clear definition of terms, the systematic retrieval of relevant literature, the transparency and replicability of the data extraction process and the acknowledgement of posed limitations. The following sections describe the important elements of the review process. The design of the codebook is discussed in chapter 4.

3.1. Review team

Although it was originally planned to include members from all partner countries into the review team, it was decided, based on the high amount of references, to limit the review team as much as possible. It is inevitable to have subjectivity issues when the review team is extensive as each member of the review team may interpret defined criteria differently (Mallet et al., 2012). Therefore, the review team was limited to the two members who created and tested the search protocol and designed the first version of the codebook. Regular discussions and multiple consistency checks were part of the entire review process.

3.2. Search protocol

Three major online databases were systematically searched for possible relevant journal articles: Web of Science, ScienceDirect and TRID. The authors believed that these three electronic databases are comprehensive enough to yield most relevant references regarding road user behaviour observation studies. After testing several combination of terms in the Web of Science and ScienceDirect databases, a search term was formulated and used in all three databases: Traffic AND (Behavio*r OR Safety). Several additional filters were set for the databases, which can be found in Annex 1. References were retrieved on the late afternoon at December 2nd 2015. Thirteen papers accepted for publication in 2016 were found, but in our review recoded as papers published in 2015. The three databases yielded 21,169, which were all imported into the Endnote referencing software. After the automatic and manual removal of duplicate entries, 12,121 references remained for screening.
3.3. Screening

One of the most important stages during a scoping review is the screening of the references found by the search strategy. During this stage, an assessment of the relevance of the found references is made. It is important that exclusion criteria are defined as clearly and unambiguously as possible, in order to limit the influence of selection bias (the extent to which different individuals in- or exclude references) and to guarantee the replicability of the review process for other researchers. In our review, three rounds of screening were used to identify the relevant references: selection, relevance and eligibility screening. A flow chart of the screening process can be found in Figure 2.

- **Selection screening**: The first screening round was used to remove all references that were not peer-reviewed journal articles published in English. Examples include conference proceedings, non-peer-reviewed journal articles, book sections and research reports. After this screening round, 7,007 references remained. Unlike the other steps of the screening process, the selection screening was performed by only one member of the review team.

- **Relevance screening**: The second round of screening evaluated the relevance of found references. Both members of the research team checked the titles and abstracts in order to determine if the articles regarded the unobtrusive observation of road user behaviour. Three exclusion criteria were formulated: not relating to (road) traffic, no collection of uninformed observed behaviour (e.g. driving simulator, questionnaires, crash data analyses) and being a traffic conflict study only. During this stage, the differences in in- or exclusion between the research members were discussed. In case of doubt, references were kept for eligibility screening.

- **Eligibility screening**: Almost 700 full papers were examined for data extraction. At the start of this screening round, papers were coded into the codebook by both review members, until a satisfactory level of consistency was reached (after around 75 references). Then, the papers were divided based on publication year (even vs odd years). Additionally, a subsample of papers was coded by both review members as well, in order to continue to check for consistency. Whenever a reviewer was not sure about certain aspects of the extracted information or if an article did not seem to be eligible for information extraction, a notation was made and the references was checked by the other review member as well. Some papers were found that used the same data (the NGSIM data-set), however, since their application and use of the data differed, all those papers were kept.

3.4. Paper retrievals

The automatic text retrieval function in Endnote was used to collect full text articles. Papers that could not be found were searched for manually through google scholar and Research Gate. A list of missing articles was then formulated and sent to the involved project partners. Finally, the library of Hasselt University was consulted in order to try to acquire the missing articles. From the 620 references, 37 publications could not be located. The majority of these papers were published before 1985 (56%).
Figure 2: The flow chart of the screening process.
4. Codebook

A codebook was created to structure the information of interest. A subsample of references was used to compose and test both the completeness of the codebook and the consistency of coding between the members of the review team. Around 55 papers were used before the final working version of the codebook was created. After the information extraction process, the codebook was revised again in order to include elements that were difficult to categorize based on definitions used in the working version. The following sections describe the elements that were included in the final version.

4.1. General Information

The first section of the codebook contained the most general and basic information regarding the included journal articles. The following information was extracted:

- Unique identifier and the member of the review team who initially coded the article
- Research goal: monitoring, evaluation, model development, software development
- Research focus: traffic safety, mobility or both
- Full reference
- Exclusion (ineligible or irretrievable)

4.2. Research Topic

The second section of the codebook collected information regarding the main topic. In order to structure the data, a categorization for classifying research topics was proposed, based on the examination of the first 55 papers (Table 2). The topics were defined as follows:

- Infrastructural-intersection: Phase change warning systems and dilemma zones both relate to driver behaviour during the final stages of the green or amber phase. The difference between them was defined as the presence of a certain system that warned approaching drivers for the oncoming phase change. Studies that specifically focused on red light running were considered separately. With regard to the topics of yielding and crossing, the distinction was defined as yielding representing the studies in which the process of negotiation between road users who shall cross first was observed, while crossing studies represented the actual crossing and elements related to it (e.g. waiting position and waiting time). Finally, shared space was considered as an infrastructural-intersection topic. Studies investigating this topic focused on situations in which road users have no allocated position, forcing road users to interact and communicate with each other.

- Infrastructural-road section: Speeding, car following and merging are topics that can be allocated to the category of road sections, since they reflect behaviour that is typically shown on stretches of road. Other topics include overtaking/passing (defined as one road user overtaking another or two road users passing each other while traveling in the opposite direction), lane changes, gap acceptance and dedicated infrastructure (observing behaviour regarding the use of annotated road user positions, e.g. bicycle lanes).

- Situational: Research regarding work-zone safety, the influence of weather conditions and violations other than violations directly related to the other defined topics (e.g. red-light running) were allocated to this category. A few studies were also found that focused on emergency vehicle warning systems, which alerted drivers that an emergency vehicle was approaching or being approached.
• Personal: The use of safety systems, like seatbelts, child restraints and protective clothing were considered as topics relating to personal characteristics. The topics of turn indicator use, mobile phone use, driver distraction, driver aggression, drunk driving and risky driving behaviour were also allocated to this category.

• Other: Finally, a category was created to capture all studies of which the topic could not be allocated to one of the defined topics. Examples include the effect of road lighting, hand positions on the steering wheel and the effect of speed bumps.

In order to keep the list of topics as structured as possible while keeping a sufficient amount of detail, no specific categories for these topics were created.

Table 2: The topic types defined in this review.

<table>
<thead>
<tr>
<th>Category</th>
<th>Topic</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Infrastructural</strong></td>
<td>Intersection</td>
</tr>
<tr>
<td></td>
<td>Phase Change Warning Systems</td>
</tr>
<tr>
<td></td>
<td>Dilemma Zone</td>
</tr>
<tr>
<td></td>
<td>Red Light Running</td>
</tr>
<tr>
<td></td>
<td>Crossing</td>
</tr>
<tr>
<td></td>
<td>Yielding</td>
</tr>
<tr>
<td></td>
<td>Shared space</td>
</tr>
<tr>
<td></td>
<td>Speeding</td>
</tr>
<tr>
<td></td>
<td>Car-following</td>
</tr>
<tr>
<td></td>
<td>Merging</td>
</tr>
<tr>
<td></td>
<td>Overtaking / Passing</td>
</tr>
<tr>
<td></td>
<td>Lane change</td>
</tr>
<tr>
<td></td>
<td>Gap acceptance</td>
</tr>
<tr>
<td></td>
<td>Dedicated infrastructure</td>
</tr>
<tr>
<td><strong>Situational</strong></td>
<td>Work-zone Safety</td>
</tr>
<tr>
<td></td>
<td>Weather conditions</td>
</tr>
<tr>
<td></td>
<td>Emergency Vehicle Warning</td>
</tr>
<tr>
<td></td>
<td>Other violations</td>
</tr>
<tr>
<td><strong>Personal</strong></td>
<td>Seatbelt</td>
</tr>
<tr>
<td></td>
<td>Turn Indicator</td>
</tr>
<tr>
<td></td>
<td>Child restraint</td>
</tr>
<tr>
<td></td>
<td>Protective clothing</td>
</tr>
<tr>
<td></td>
<td>Mobile phone</td>
</tr>
<tr>
<td></td>
<td>Drink driving</td>
</tr>
<tr>
<td></td>
<td>Aggression</td>
</tr>
<tr>
<td></td>
<td>Distraction</td>
</tr>
<tr>
<td></td>
<td>Risky driving behaviour</td>
</tr>
<tr>
<td><strong>Other</strong></td>
<td></td>
</tr>
</tbody>
</table>

In addition to the specification of the research topic, a notion of the type of infrastructure was made as well. A categorization of four types was made:

• **Intersections**: Included are signalized, priority controlled, priority-from-the-right (road users arriving from the right have the right-of-way) and stop-controlled (stop-signs at all legs) intersections and roundabouts;

• **Crossing facilities**: pedestrian crossings (not part of an intersection), cyclist crossings (not part of an intersection), channelized right-turn lanes and railroad crossings;

• **Road sections**: straight roads, curves and work zones;

• **Other**: shared space, parking lots and other locations (which were specified in an additional variable).
4.3. Methodological information

The third section of the codebook focused on methodological aspects. Several types of information were extracted and structured, including on which road user types the studies focused, whether data was collected on all involved or a pre-selected road user type (e.g. yielding studies focusing on both drivers and pedestrians or only focusing on pedestrians) and if the behavioural observations were combined with other methodologies (e.g. the traffic conflict technique or the driving simulator). Other important information extracted regarded:

- **Research design:** Four types were distinguished:
  - Single observation: a single site or multiple comparable sites are monitored. No comparison between sites is made.
  - Before-after: At the same site(s), measurements are made before and after the implementation of a safety improving treatment. The effectiveness of the treatment is evaluated.
  - With-without: At the same site(s), measurements are made with the treatment activated and the treatment deactivated (e.g. countdown timers). This type of research design is most often applied when the treatment has been in place for a long time or if the treatment is part of a series of different treatments to be tested at the same location.
  - Cross-sectional: Two or more sites are compared to test the effectiveness of a safety improving treatment. The difference with with-without research designs is that in cross-sectional research comparisons between sites are made.

For before-after and with-without studies, it was also investigated if a control group was used to address confounding factors such as regression-to-the-mean and natural variability problems. A notion was also made when a semi-controlled design was used, in which instructed road users provoked traffic events of interest.

- **Data collection technique:** The means of data collection was coded. The following categories were defined: human observers, cameras, vehicles equipped with cameras (observing other road users), detectors, (handheld) speed guns and other sensors.

- **Data collection characteristics:** Finally, information regarding the collected data itself was extracted. This information focused on the number of testing sites, their geographical location (country), the sample size and the time of day data was collected (peak or off-peak, day or night, week or weekend). Peak hours were determined based on the statement of the authors that volumes of road users of interest were high or not at the selected time of data collection. It should be noted that this section of the codebook solely focused on the characteristics of the data that was actually analysed, because not all studies have analysed all the data that were collected.

4.4. Indicators

A wide variety of indicators was used in the 583 included studies, which relate to different aspects of traffic safety events, situational aspects and road users characteristics. A categorization was applied to structure the 47 identified indicators, closely related to the categorization created for the research topics. It is important to note that this proved to be a difficult task since multiple indicators were difficult to allocate to only one category (e.g. speed and gap acceptance). Only indicators with a connection to traffic safety were
included (road user characteristics being an exception). Indicators related to mobility (e.g. start-up lost time, delay time) were ignored. The following categories were defined:

- Infrastructure indicators
- Traffic Safety Aids
- Distractions
- Personal characteristics

4.4.1. Infrastructure indicators

The majority of the identified indicators (n = 30) can be related directly to infrastructural elements (Table 3). The following categorization was applied:

- **Intersection-Dilemma Zone**: Five indicators were found to describe behaviour during dilemma zone scenarios. The most common ones included the stop-or-go decision and yellow- and red-light running. Two other indicators tried to describe the severity of yellow- or red-light running by recording the distance-to-stop-line during the onset of the amber or red phase and to measure the intersection-entry-time (the time after the onset of the amber or red phase that a vehicle passes the stopping line).

- **Intersection-Yielding**: Four indicators were identified that can be used to examine yielding behaviour. The most basic indicator, yielding, just indicates if a road user stopped to let the other one pass first. Other indicators include looking behaviour, yielding distance (describing the distance at which yielding occurs) and the occurrence of evasive actions (sudden changes in speed or direction to avoid a collision).

- **Intersection-Crossing**: Indicators related to the crossing manoeuvre can broadly be distinguished into pre-crossing and during-crossing indicators. The former includes waiting position, waiting time, gap acceptance, gap size and stop sign compliance. Behavioural indicators describing the process of the crossing manoeuvre itself include crossing time, crossing path and jaywalking (road users do not cross at the crossing facility).

- **Intersection-Railroad Crossing**: A special type of intersection is the railroad crossing. Several studies investigated the extent to which road users respect the traffic rules, focusing on stop sign compliance or adhering to signalling lights.

- **Road section**: Indicators used to describe behaviour at road sections include measurements that do not necessarily require interaction (speed, space and time headway, lateral position and lane choice) and measurements in which two or more road users need to interact with each other (overtaking, overtaking attempts, merging, merging distance and lane changing).

- **Violations**: Finally, indicators relating to violations were identified. They included wrong-way driving and violations that could not be allocated to the other defined indicators.

4.4.2. Traffic Safety Aids

Seven indicators were identified that relate to the use of traffic safety aids. Three of them are exclusive for vehicles and regard the use of turn indicators, seatbelts and child restraints, while the use of protective clothing and pedestrian push buttons are only applicable for VRUs. Finally, the use of (head)lights was encountered in the literature. An additional category was included, indicating if measurements were used that could not be allocated to one of the previously mentioned indicators.
4.4.3. Distractions

In total, five indicators were identified that relate to (driver) distraction. A sixth indicator was defined to capture measurements that could not be allocated to the five identified indicators. The indicators regarded personal behaviours (mobile phone use, smoking and carrying items) and situational characteristics (the presence of a passenger for vehicles and the group size of crossing road users).

4.4.4. Road user characteristics

The final category of indicators consists of road user characteristics. It includes age, gender and ethnicity. Some studies also focused on following certain road users over space and time or relating shown behaviour to crash/violation history. For these studies, the license plate or vehicle registration number was collected. These indicators are not considered to be behavioural, but important to include into the review since literature has shown that for example men are more willing to take risks than woman (e.g. Buss, 2004; Yagil, 1998).

Table 3: The infrastructural indicators found in the behavioural observation studies.

<table>
<thead>
<tr>
<th>Type (Element)</th>
<th>Indicator</th>
</tr>
</thead>
<tbody>
<tr>
<td>Intersection</td>
<td>Dilemma zone</td>
</tr>
<tr>
<td></td>
<td>Yielding</td>
</tr>
<tr>
<td></td>
<td>Yielding distance</td>
</tr>
<tr>
<td></td>
<td>Looking</td>
</tr>
<tr>
<td></td>
<td>Evasive action</td>
</tr>
<tr>
<td></td>
<td>Waiting position</td>
</tr>
<tr>
<td></td>
<td>Waiting time</td>
</tr>
<tr>
<td></td>
<td>Crossing time</td>
</tr>
<tr>
<td></td>
<td>Crossing path</td>
</tr>
<tr>
<td></td>
<td>Jaywalking</td>
</tr>
<tr>
<td></td>
<td>Gap acceptance</td>
</tr>
<tr>
<td></td>
<td>Gap size</td>
</tr>
<tr>
<td></td>
<td>Stop-sign compliance</td>
</tr>
<tr>
<td>Railroad crossing</td>
<td>Sign compliance</td>
</tr>
<tr>
<td></td>
<td>Light-compliance</td>
</tr>
<tr>
<td>Road section</td>
<td>Speed</td>
</tr>
<tr>
<td></td>
<td>Headway</td>
</tr>
<tr>
<td></td>
<td>Lateral position</td>
</tr>
<tr>
<td></td>
<td>Lance choice</td>
</tr>
<tr>
<td></td>
<td>Overtaking</td>
</tr>
<tr>
<td></td>
<td>Overtaking Attempts</td>
</tr>
<tr>
<td></td>
<td>Merging</td>
</tr>
<tr>
<td></td>
<td>Merging Distance</td>
</tr>
<tr>
<td></td>
<td>Lane change</td>
</tr>
<tr>
<td></td>
<td>Wrong-way driving</td>
</tr>
<tr>
<td></td>
<td>Other</td>
</tr>
</tbody>
</table>
5. Findings

The process of data retrieval showed that the amount of peer-reviewed journal articles relating to road user behavioural observation studies is extensive: 583 relevant and retrievable articles were found. The previous chapter described how information from these articles was extracted and structured. With the use of descriptive analyses, this chapter describes the main findings and observation of the current application of behavioural observation studies published in scientific literature. In order to acquire a better view on the applicability of such studies with regard to VRUs' safety, a distinction between studies including at least one VRU road user type (hereafter referred to as VRU-studies) and studies including at least one non-VRU road user type (hereafter referred to as driver-studies) was made. Based on this categorization, 214 of the 583 studies (37%) were labelled as VRU-studies and 477 of the 583 studies (82%) as driver-studies. It is important to note that this categorization does not take into account if behavioural data of these road user types is collected. For example, studies investigating yielding behaviour between car drivers and pedestrians in which data for both road users was collected and studies only focusing on the crossing pedestrian were both labelled as VRU- and driver-studies. The following sections describe the descriptive analyses of the information coded in the codebook. The findings are structured as follows: first an overview of road user focus is given, then topics and behavioural indicators are described, afterwards the purpose of behavioural observation studies is explained and finally methodological aspects are discussed.

5.1. Road user focus

Figure 3 provides an overview of the specified road user types' involvement in behavioural observation studies, shown in percentages on the y-axes and in absolute numbers on the data labels. The numbers of the different road user types do not add up to 100%, since studies can include multiple VRU or driver types. The VRU-studies show that the majority of research interest has focused on pedestrians (67%) and, to a lesser extent, cyclists (25%). The area of powered-two-wheelers' behaviour is less addressed: only 33 studies (15%) were found. This might seem surprising, since riders of powered-two-wheelers are highly represented in crash statistics worldwide (WHO, 2009). However, our experience during the selection process indicated that motorcyclists’ safety might mainly be addressed by investigating attitudes, using driver behaviour questionnaires (e.g. Özkan et al., 2011, Elliott et al., 2007). Other road user types included in VRU-studies are staged pedestrians and cyclists, which are used to investigate driver behaviour at certain events of interest (e.g. yielding or lateral overtaking distance). Finally, behaviour of e-bikes has been observed during recent years.

Car drivers were included in almost all driver-studies. Determining the involvement of heavy vehicles was difficult, since many studies did not formulate clear statements regarding their in- or exclusion. Several articles only reported the inclusion of heavy vehicles by describing them in result tables, a few studies clearly stated whether or not they were included, but many studies made no remark at all. Therefore, it is rather difficult to formulate any meaningful observations regarding heavy vehicles' involvement in behavioural observation studies.

Only three studies were found that included road users that could not be allocated to one of the formulated road user types. These studies examined behaviour of car drivers while overtaking or passing a mobile home (Hall & Harkley, 1999), violations by taxi drivers (Frederiksen, Frank & Freeman, 1939) and behaviour of busses and commuters near
bus-stops in Pakistan (Mirza et al, 1999). The latter study was also the only research found that was categorized as a ‘driver’ study, but that did not include car drivers.

![Figure 3](image-url)  
**Figure 3:** The amount of studies including the specified road user types. On the left, VRU-studies are shown (n = 214) and driver-studies are depicted on the right (n = 477).

The number of published studies on road user behaviour observation studies has been growing rapidly since the first journal article was published in 1939. In order to examine how these publications have evolved over time, the cumulative distributions of both VRU- and driver-studies were plotted (Figure 4, left). The first study including VRUs was found in 1973, while drivers have already been subject of behavioural observation studies since 1939. The cumulative distributions for both VRUs and drivers show an exponential growth. It can be seen that the majority of the studies (>50%) have been carried out within the last five and last ten years for VRU and drivers respectively. When considering the three most examined road user types (Figure 4, right), it can be observed that studies involving car drivers and pedestrians have grown exponentially. For cyclists however, the exponential growth seems to have reached it limits: the increasing rate of research efforts declined slightly from 2013.

![Figure 4](image-url)  
**Figure 4:** On the left, the evolution over time of research focusing on VRUs and drivers. On the right, the three most common road user types are plotted. Due to low numbers in earlier years, the graphs depict the period from 1965 till 2015.
5.1.1. Countries

For all 583 articles it was noted down in which countries data was collected. In total, 51 different countries were identified. The majority of the studies took place in America and Europe. When considering country level, most studies were conducted in the USA (38%), Canada (8%), China (8%), the UK (4%) and Israel (4%). For almost half (n = 24) of the identified countries, only one or two studies were found. A few countries were found that invested more research efforts into VRUs than drivers (Austria, Bosnia-Herzegovina, Denmark, Egypt, India, Malaysia, Poland and Vietnam), but it should be noted that the total amount of behavioural observation studies in these countries is rather limited. A complete overview of the amount of VRU and driver studies can be found in Figure 5.

Several studies (n = 16) were found that used testing sites in multiple countries. Main reasons for such a decision were to enlarge the dataset (e.g. Rudloff et al., 2011) or to compare different driving cultures (e.g. Marczak et al., 2013). From these studies, three quarters included testing locations in two countries, but four studies were found that made comparisons between three countries.

Pie charts were plotted to investigate which countries invested most efforts in observing the behaviour shown by the defined road user types (Figure 6). Each chart shows the countries with the most studies relating to the defined road user type. One should be careful during interpretation of these figures, since those do not relate to the research efforts per country: the fact that the USA has the highest score for almost all road user types is the result of the USA having conducted most studies. For example, when looking at pedestrians, the USA has actually the lowest share of studies including pedestrians (21% of all studies conducted in the USA) compared to the other four ‘top’ countries (29%, 29%, 42% and 42% for Canada, China, UK and Israel respectively). However, based on this reasoning, there is an important observation regarding the USA’s share in powered-two-wheeler and electric bike research. The pie charts show that for these road user types, most studies have been conducted in Asia and that currently behavioural observation research into e-bikes is limited to China, Singapore and Taiwan only.

Due to the limited amount of studies in most countries, sorting road user types based on research efforts per country will not provide a meaningful perspective on the allocation of research efforts. Furthermore, it is important to note that the results of this review regard behavioural observation studies only and that it is possible that other research methodologies might be more applicable to investigate, for example, powered-two-wheelers’ behaviour.

5.2. Topics and indicators

This section provides an overview of the topics and indicators that have been applied in the scientific literature. First, the infrastructural elements that were the focus of conducted studies as defined in section 4.2 are discussed. Then, research topics and indicators are briefly described. Finally, topics and corresponding indicators are discussed in more detail.
Figure 5: The road user focus of the different countries conducted road user behavioural observations, sorted based on number of VRU-studies.
Figure 6: The pie charts showing the countries with most research efforts for the defined road user types.
5.2.1. Infrastructure

Figure 7 shows the use of infrastructural elements in studies observing VRUs’ and drivers’ behaviour. It can be seen that VRU-studies mainly focus on locations where interaction between road users is required. Mainly signalized intersections (39%) and pedestrian crossings (31%) have been extensively researched. Driver-studies on the other hand mainly focused on road sections (34%), where interaction with other road users is not necessarily required. Furthermore, the infrastructural elements of railroad crossings and stop-controlled intersections mainly relate to driver studies.

![Diagram showing infrastructure elements and their number of studies](image)

**Figure 7**: The infrastructural elements of interest, sorted by number of VRU-studies.

5.2.2. Topics

The pie charts in Figure 8 show the distribution of the different topic categories. The most prominent area of research has been ‘Intersection’, with almost 38% of all driver- and almost 80% of all VRU-studies. For driver studies, the share of ‘road sections’ is comparable (36%), followed by ‘personal’ and ‘situational’ topics. The same order of shares of research topics can be found for VRU-studies. From 34 studies (of which 9 included VRU-studies and 32 driver-studies) the topics could not be allocated to one of the specified categories. Examples include the observation of hand positions on the steering wheel (Fourie, 2011; Walton, 2005), validating driving simulators (Yan, 2008) and cycling on the sidewalk (Okinaka & Shimazaki, 2011).
A full list of VRU- and driver-studies related to the defined research topics can be found in Figure 9. It shows that the topics of crossing (40%) and yielding (23%) have received most attention in VRU studies. Red light running and overtaking follow, but their share is rather limited (10% and 4% respectively). The other topics are researched quite infrequently, not exceeding more than seven out of 214 studies per topic.

The amount of topics represented by driver studies is slightly higher compared to VRU studies, since topics like seatbelt and child restraint usage can only apply for drivers. This is reflected in the shares of the research topics: the most popular one, speeding, accounts for ‘just’ 16% of all driver studies, followed by yielding, crossing and seatbelt usage (13%, 13% and 7% respectively). Drivers' behaviour during the dilemma zone is a well-known topic: the combination of Red-Light-Running, Dilemma Zone and Phase Change Warning Systems adds up to 53 studies (11%). Surprisingly, studies regarding violations like mobile phone use and drink driving only formed a very small part of topic focus, while they are recognized as very important road safety problems (e.g. the EU states that yearly around 600 deaths and hospital admissions in the Netherlands are caused by mobile phone use while driving, while 2600 deaths and 330,000 serious injuries occur in the USA because of this distraction (2009)). A possible explanation could be that such studies are only published as research reports or that different methodologies are used to investigate these safety issues. For example, driver behaviour questionnaires are very useful to analyse attitudes, opinions, beliefs, emotions and behavioural and cognitive processes when considering driving under the influence of alcohol (Drew et al, 2010). Only three out of 27 topics had more VRU- than driver-studies (crossing behaviour, red light running and the use of protective clothing), even though the amount of driver studies was more than twice as high (212 versus 476), suggesting that these areas of research are very useful for observing behaviour of VRUs.
Figure 9: The research topics found in included references. Data is sorted per category based on the number of references for studies relating to VRU-types.
Indicators

The bar chart in Figure 10 depicts the distribution of indicators. With regard to VRU-studies, it shows that the behavioural indicators belonging to the categories of crossing (87%), road section (48%) and yielding (45%) are most often measured. The observation that indicators relating to road sections are well represented is somewhat contradicting to the observation that mainly intersections have been the topic of research for VRU studies. Possible explanations for this finding might be the classification of ‘speed’ as an indicator relating to road sections and that for studies involving both VRU’s and drivers, the indicators relating to road sections measure the behaviour shown by drivers. For driver-studies, most measured behavioural indicators relate to road sections (79%), followed by crossing (39%), yielding (27%) and dilemma zones (27%). Personal characteristics are also well-reported for both study types: 43% in VRU-studies and 33% for driver-studies.

A full list of VRU- and driver-studies measuring the defined indicators can be found in Figure 11. It shows that speed is most often measured for both study types: in 44% of the VRU-studies and 59% of the driver-studies speed data was collected. Other popular indicators for VRU studies are red light running (35%), yielding (32%) and looking behaviour (22%). For driver studies, headways (18%), yielding (14%) and gap acceptance (13%) are behavioural indicators measured often. It was further found that twelve out of 43 indicators were more often measured in VRU-studies, most of them belonging to the category of ‘crossing’ indicators. With regard to personal characteristics, it can be seen that age and gender belong to the top ten of most collected indicators for both study types, although they do not directly relate to aspects of traffic safety. Furthermore, especially VRU-studies considered the influence of group size, defined as the amount of road users present during the behavioural action of interest (19%).
Deliverable D2.1 “Review of current study methods for VRU safety – part 3“

<table>
<thead>
<tr>
<th>INDICATORS [#]</th>
<th>VRU</th>
<th>Driver</th>
</tr>
</thead>
<tbody>
<tr>
<td>Intersection</td>
<td></td>
<td></td>
</tr>
<tr>
<td>RLR</td>
<td>5</td>
<td>21</td>
</tr>
<tr>
<td>YLR</td>
<td>4</td>
<td>35</td>
</tr>
<tr>
<td>Stop/Go</td>
<td>16</td>
<td>7</td>
</tr>
<tr>
<td>Distance to Stop Line</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Intersection Entry Time</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Yielding</td>
<td>23</td>
<td>67</td>
</tr>
<tr>
<td>Looking</td>
<td>34</td>
<td>77</td>
</tr>
<tr>
<td>Evasive Action</td>
<td>21</td>
<td></td>
</tr>
<tr>
<td>Yielding Distance</td>
<td>6</td>
<td></td>
</tr>
<tr>
<td>Crossing</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Crossing Path</td>
<td>23</td>
<td>38</td>
</tr>
<tr>
<td>Jaywalking</td>
<td>17</td>
<td>36</td>
</tr>
<tr>
<td>Waiting Time</td>
<td>30</td>
<td>61</td>
</tr>
<tr>
<td>Gap Acceptance</td>
<td>13</td>
<td>28</td>
</tr>
<tr>
<td>Waiting Position</td>
<td>14</td>
<td>33</td>
</tr>
<tr>
<td>Crossing Time</td>
<td>16</td>
<td></td>
</tr>
<tr>
<td>Gap Size</td>
<td>14</td>
<td></td>
</tr>
<tr>
<td>Stop-Sign Compliance</td>
<td>6</td>
<td></td>
</tr>
<tr>
<td>Railroad Crossing</td>
<td>Light Compliance</td>
<td>3</td>
</tr>
<tr>
<td></td>
<td>Sign Compliance</td>
<td>12</td>
</tr>
<tr>
<td>Road Section</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Speed</td>
<td>280</td>
<td></td>
</tr>
<tr>
<td>Lateral Position</td>
<td>58</td>
<td></td>
</tr>
<tr>
<td>Headway</td>
<td>10</td>
<td>85</td>
</tr>
<tr>
<td>Lane Change</td>
<td>17</td>
<td></td>
</tr>
<tr>
<td>Overtaking</td>
<td>7</td>
<td>36</td>
</tr>
<tr>
<td>Merging</td>
<td>17</td>
<td>20</td>
</tr>
<tr>
<td>Merging Distance</td>
<td>10</td>
<td></td>
</tr>
<tr>
<td>Overtaking Attempts</td>
<td>5</td>
<td></td>
</tr>
<tr>
<td>Lane Choice</td>
<td>5</td>
<td></td>
</tr>
<tr>
<td>Violations</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Wrong-way Driving</td>
<td>9</td>
<td></td>
</tr>
<tr>
<td>Other Violations</td>
<td>6</td>
<td></td>
</tr>
<tr>
<td>Traffic Safety Aids</td>
<td>Protective Clothing</td>
<td>2</td>
</tr>
<tr>
<td></td>
<td>Pedestrian Push Button</td>
<td>12</td>
</tr>
<tr>
<td></td>
<td>Turning Indicator</td>
<td>2</td>
</tr>
<tr>
<td></td>
<td>Lights</td>
<td>3</td>
</tr>
<tr>
<td></td>
<td>Seatbelt</td>
<td>48</td>
</tr>
<tr>
<td></td>
<td>Child Restraint</td>
<td>4</td>
</tr>
<tr>
<td>Distractions</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Groupsize</td>
<td>16</td>
<td>41</td>
</tr>
<tr>
<td>Carrying Items</td>
<td>14</td>
<td></td>
</tr>
<tr>
<td>Mobile Phone</td>
<td>10</td>
<td>22</td>
</tr>
<tr>
<td>Passengers</td>
<td>6</td>
<td>23</td>
</tr>
<tr>
<td>Other distractions</td>
<td>6</td>
<td></td>
</tr>
<tr>
<td>Smoking</td>
<td>2</td>
<td></td>
</tr>
<tr>
<td>Characteristics</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Gender</td>
<td>849</td>
<td></td>
</tr>
<tr>
<td>Age</td>
<td>70</td>
<td></td>
</tr>
<tr>
<td>Ethnicity</td>
<td>77</td>
<td></td>
</tr>
<tr>
<td>License Plate</td>
<td>15</td>
<td></td>
</tr>
<tr>
<td>VRU Driver</td>
<td>13</td>
<td></td>
</tr>
</tbody>
</table>

**Figure 11:** The indicators of the included studies. Data is sorted per category based on the number of VRU-studies.
5.2.3. Topics of interest

This subsection looks closer into the connection between topics and indicators. Unlike the other analyses, findings discussed in this chapter relate to all studies, unless it is explicitly stated that observations are based on the VRU- or driver-study level.

5.2.4.1 Dilemma Zone

The most common safety indicators used during dilemma zone situations are the stop-or-go decision and whether road users run the red light or not. A few studies also included the indicator 'yellow light running', which states if the driver passes the stop bar after the onset of the amber light, but before the light has turned to red. The added value of such an indicator on its own is difficult to determine, since the meaning of the amber light is fully dependent on the approaching driver: in general the amber light states that passing is not allowed, unless one cannot stop safely anymore. Consequently, running an amber light might not necessarily be regarded as an unsafe behaviour. Therefore, several studies included the measurement of the distance between the driver and the stop line during the onset of the amber/red phase. A further extension of these indicators, able to interpret some form of severity, is the 'intersection entry time', which describes at what time after the onset of the amber/red phase a road user is (illegally) entering the intersection. Although relevant for safety evaluation, this indicator has only been found seven times. When looking into more detail at the indicator of red and yellow light violations, it can be observed that the majority of the studies had monitoring purposes (66%), while some studies were found that tested the effectiveness of, for example, red light cameras (Polders et al., 2015). Studies specifically focusing on red-light violations have been applied for both VRU- and driver-studies, but the indicator itself has also been used for other topics (e.g. crossing in general).

5.2.4.2 Yielding

In total, 67 studies were found that examined yielding behaviour of drivers and VRUs. Most of them were conducted at pedestrian crossing facilities that were not part of an intersection (n = 33) or priority controlled intersections (n = 23). The remaining studies examined the effect of unprotected left- and right-turn signal phasing in which turning drivers need to yield to drivers, cyclists and/or pedestrians going straight, evaluated stop-sign compliance and monitored crossing facilities on roundabouts. The majority of the studies tested the effectiveness of a certain measure to increase drivers' yielding towards cyclists and pedestrians. Examples include advanced yield markings, additional signage and push buttons. Other studies were limited to monitoring purposes only, in which instructed pedestrians or cyclists were used to observe, for example, the effect of race (Goddard et al., 2015), staring (Guegen et al., 2015) and hand gestures (Zhuang, 2014).

Yielding is mainly being evaluated as whether a road user stops to enable another road user to go first. Many times this has been coded as a binary (yes/no) variable, but few studies included the manner of yielding, distinguishing in a hard or soft yield (e.g. Samuel et al., 2013; Schroeder et al., 2013). Another yielding-related indicator, mainly applied in earlier years, is if road users performed evasive actions (e.g. hard breaking or swerving) to avoid a collision (e.g. Van Houten et al., 1985).

5.2.4.3 Crossing

Studies focusing on crossing behaviour (n = 126, 22%) have mainly focused on signalized intersections (n = 47), pedestrian crossing facilities (n = 30) and railroad crossings (n = 19). Most of these studies monitored behaviour (n = 69), rather than to test the effectiveness of a safety treatment (n = 36). Both VRUs and drivers have been subject of
crossing research. The most applied indicator for VRU-studies is the crossing path, while for driver-studies gap-acceptance is most often considered. Few studies provide more detail on the aspect of gap acceptance by examining the actual size of accepted and rejected gaps. Although railroad crossings form a major safety hazard across the EU (Eurostat, 2016), only 21 studies (3%) examining crossing behaviour were found. Most of these studies only considered drivers and the extent to which they complied to lights and signs. A few studies were found that observed pedestrians (Lobb et al., 2003; Siques, 2002) and cyclists (Cobey et al., 2013).

5.2.4.4 Road sections

The most often used indicator in driver-studies is speed, most likely because speeding has been identified as one of the biggest problems in road safety as speeding greatly increases the risk of accident (Elvik, 2004). Speed is most often measured in terms of driving speed in km/h, but several studies used the pace of driving/walking during crossing (e.g. in a hurry, relaxed) as an indicator (e.g. Walmsley & Lewis, 1989; Zeedyk et al., 2002). Research solely focusing on speeding aims at evaluating traffic control devices (e.g. Yang et al., 2015) or determining the effect of enforcement like point-to-point measurements (e.g. Montella et al., 2015). Other indicators that have been used on topics relating to road sections focus mainly on drivers and include, among others, following distance (in space or time), lateral position and lane changing.

5.2.4.5 Traffic safety aids

The use of seatbelts belongs to the top five of most researched topics (n = 49). Five studies were found to evaluate the effectiveness of a certain treatment to increase seatbelt use, while the majority of the studies only monitored seatbelt use. Combinations with questionnaires are not uncommon (n = 16, 33%). Further analyses shows that in the early 80s studies in the USA and Israel were conducted to investigate the effect of the implementation of the mandatory seatbelt use law (Hakkert et al., 1981; Matthews, 1982; Lund et al., 1984). Observation of the use of other traffic safety aids is limited and regards turn indicator use, protective clothing of motorcyclists and the use of push buttons at traffic signals by pedestrians.

5.2.4.6 Distractions

The area of distractions is not extensively researched. The main indicator used are group size and passenger presence for drivers and motorcyclists, describing whether the road user is traveling alone or not. A few studies also included mobile phone use, smoking, carrying items or other forms of distractions like eating and drinking.

5.3. Purpose of road user behaviour observation

The assessment of the available literature showed that road user behaviour observation studies can broadly be used for four different research goals:

- **Monitoring** (e.g. Walker, I., 2007 and Goddard et al., 2015)
  - One location or multiple (identical) locations are observed to monitor the behaviours performed by road users. The goal of such studies is to ‘look what happens’ and not to examine the influence of a certain (infrastructural) intervention or safety improving measure.
- **Evaluation** of certain measures (e.g. Zhang et al., 2015 and Polders et al., 2015)
  - Using a before-after, with-without or cross-sectional research design, the effect of an (infrastructural) intervention of safety improving measure is evaluated. Behavioural measures (like hand gestures or staring by staged
pedestrians) are not considered to be studies testing the effectiveness of certain measures, since these regard the effect of shown behaviour rather than a measure.

- **Model development** (e.g. Li et al., 2014 and Shiomi et al., 2015)
  - Real world data is used to develop/calibrate/validate predictive models (e.g. microsimulation models). Studies developing explaining models are not regarded as model development.

- **Automated video-analysis software development** (e.g. St-Aubin et al., 2015 and Zaki et al., 2012)
  - Video data of traffic events is used to develop and test automated video-analysis tools (e.g. tracking algorithms and road user classification). This type of research goal is an extension of monitoring. Because the development of automated video analysis software is relatively new, publications with this research goal are limited and can only be found in recent years.

As can be seen in Figure 12, the proportions or research goals are similar between the VRU- and driver-studies. The majority of studies have focused on monitoring of road user behaviour, followed by research evaluating the effect of safety improving measures. Model development and testing automated video-analysis software tools only reflect a small portion of behavioural observation studies. When considering the research aims in function of time, again no major differences between VRU- and driver-studies can be found. The different research goals seem to be growing exponentially, although testing automated video-analysis software tools seems to be an emerging topic, since those studies were only found in the last five years.

**Figure 12:** On top, the shares of the research goals. Below, their evolution over time.
Most monitoring studies focus on the simple observation of road user behaviour, using a pre-specified protocol of indicators of interest. Examples include the examination of seatbelt and child restraint use, yielding behaviour or crossing behaviour at railroad crossings. A few studies were found to be somewhat unique in their approach to monitor behaviour:

- Manan et al (2015) collected data on motorcyclists’ behaviour at priority-controlled T-intersections, in which the motorcyclist should yield. In their research design, human observers and video cameras were used to monitor road user behaviour without a pre-defined list of indicators of interest. During data collection, human observers described the behaviour shown by the motorcyclists by speaking into a recording device. Based on these spoken descriptions, relevant indicators were identified and categorized. This approach is unique in the sense that the research did not depart from a list of pre-specified indicators, but that the aim of the research was to create a list of indicators. Such an approach is important to identify which types of behaviour need to be considered for safety evaluation practices.

- Jonasson (1999) used a similar approach, in which he just observed what happened in order to answer his research questions regarding informal traffic rules. Although not much attention in literature is given to informal traffic rules, further research is required since it indicates that the official rules are violated at many occasions. This can lead to safety issues when two conflicting road users adhere to the different rule types (formal versus informal).

- Rosenbloom et al. (2007) observed driver behaviour from inside the car. Female drivers were observed while driving in familiar and unfamiliar surroundings while a friend was recording violations from the passenger seat. This study is unique in the sense that observations were carried out inside the car, but without the awareness of the driver that she was being observed or participated in an experiment.

- Papadimitriou et al. (2011) modelled pedestrian crossing behaviour during a trip. Pedestrians were randomly selected and followed during their walking trip through Athens, Greece. Video recordings were made without violating the privacy of the participants or the awareness of the pedestrians that they were being followed. A similar approach was used by Gates et al. (2011), which followed car drivers when they drove through a work zone.

### 5.4. Methodological aspects

This section provides an overview of the most important methodological aspects regarding the use of behavioural observation studies. Elements like research design, control groups, data collection techniques, observation periods and sample sizes are briefly discussed.

#### 5.4.1. Research Design

The pie charts in Figure 13 show the proportions of the use of the different research designs. More than 50% of both VRU- and driver-studies used a single observation design to measure road users’ behaviour, followed by before-after research. The proportions of with-without and cross-sectional research designs are comparable for both study types and only account for around 20% of all applied research designs. The two graphs below the evolution over time of the use of the different research designs. It can be seen that all of them grow exponentially, but that the use of cross-sectional research designs for VRU-studies have only been used during the last ten years.
5.4.2. Semi-controlled research

Several studies (n = 33) were found that used an instructed road user to provoke traffic events of interest (Table 4). In this approach, the road user received instructions to behave in a certain way, enabling the collection of data of the other road user during specific events. The majority of these studies used a VRU as instructed road user, in which either yielding towards pedestrians or cyclists (n = 14) or drivers overtaking cyclists (n = 8) was investigated. The yielding studies focused on aspects like ethnicity (Goddard et al., 2015), staring behaviour (Guéguen et al., 2015) and hand gestures (Zhuang and Whu, 2014), while the overtaking studies focused on the lateral distance between drivers and cyclists during the overtaking manoeuvre. The remaining eleven studies made use of instructed car drivers that provoked overtaking manoeuvres (n = 9), tested the position of braking lights (n = 1) and evaluated drivers’ aggression (n = 1). The latter one was the only study in which the confederate of the research team hindered other road users in the sense that there was no possibility to pass (Ellison et al., 1995).

Semi-controlled research designs are mainly used to monitor traffic events (64%) or to evaluate the effectiveness of a safety treatment (33%). Our findings suggest that mainly VRU-studies benefit from this approach, since 10% of the VRU-studies used this research design compared to only 2% for driver-studies. Another interesting observation is that of these 10%, almost 64% of the studies were conducted during the last five years.
Deliverable D2.1 “Review of current study methods for VRU safety – part 3”

Table 4: Semi-controlled research

<table>
<thead>
<tr>
<th>Instructed Road User</th>
<th>Topic</th>
<th>Number of Studies</th>
</tr>
</thead>
<tbody>
<tr>
<td>Pedestrians</td>
<td>Yielding behaviour of drivers</td>
<td>13</td>
</tr>
<tr>
<td>Cyclists</td>
<td>Overtaking behaviour of drivers</td>
<td>8</td>
</tr>
<tr>
<td></td>
<td>Yielding behaviour of drivers</td>
<td>1</td>
</tr>
<tr>
<td></td>
<td>Overtaking</td>
<td>9</td>
</tr>
<tr>
<td></td>
<td>Braking lights position</td>
<td>1</td>
</tr>
<tr>
<td></td>
<td>Driver aggression</td>
<td>1</td>
</tr>
</tbody>
</table>

5.4.3. Control groups

In the field of road user behavioural observation studies, the use of control groups is quite rare. Only 21 out of 581 references (4%) used some sort of mechanism to address confounding factors like temporal variability or regression-to-the-mean bias (Table 5). The type of control group can broadly be categorized into sites (n = 17) and road users (n = 4). The studies reported that results needed to be modified, based on the data of the control groups. This could indicate that for evaluation purposes also the domain of road user behavioural observation is in need of control groups. Even more, Islam et al. (2014) emphasized the use of control groups since they found that it affected their results significantly.

Table 5: The use of control groups in road user behavioural observation studies (n=21).

<table>
<thead>
<tr>
<th>Control group</th>
<th>Element</th>
<th>n</th>
<th>%</th>
</tr>
</thead>
<tbody>
<tr>
<td>Type</td>
<td>Infrastructure type</td>
<td>17</td>
<td>81</td>
</tr>
<tr>
<td></td>
<td>Road user group</td>
<td>4</td>
<td>19</td>
</tr>
<tr>
<td>Research Goal</td>
<td>Monitoring</td>
<td>2</td>
<td>10</td>
</tr>
<tr>
<td></td>
<td>Effectiveness testing</td>
<td>19</td>
<td>90</td>
</tr>
<tr>
<td>Effect</td>
<td>High</td>
<td>17</td>
<td>81</td>
</tr>
<tr>
<td></td>
<td>Low</td>
<td>4</td>
<td>19</td>
</tr>
</tbody>
</table>

5.4.4. Data collection techniques

Several protocols and devices can be used to collect data. Six types of data collection tools were identified. Figure 14 gives an overview of the amount of studies that used the defined data collection tools. For both VRU- and driver-studies, cameras (either photo or video) have been used most often, followed by human observers. Vehicles equipped with cameras were regarded as separate category, since they enable researchers to observe road user behaviour in time rather than in space (fixed location). The camera equipped vehicles consisted of cars (n = 17), bicycles (n = 4) and helicopters (n = 4). For cars and bicycles, the main use of the cameras was to observe overtaking and passing behaviour, while the helicopters focused on merging manoeuvres. With regard to the use of helicopters, it was pointed that although this technology offers the opportunity to capture data without the risk of occlusion by other road users, it is a relative expensive means of data collection (Polus, 1985).

The results show that speed measurements are mainly collected using speed guns (their share is higher compared to detectors or other sensors). Of ten studies the data collection tools could not be allocated to one of the defined categories. These studies used traffic fines (n = 4) or did not specify the data collection tools (n = 6).
Figure 14: The number of studies including the defined data collection tools.

The cumulative distribution of the use of the data collection tools over time is for both VRU and Driver-studies depicted in Figure 15. The most important observation is that for a long time human observers were the main source of data collection, but that for both study types the use of cameras has become more common during the last five years. It also shows the (slow) increase of all data collection tools over the years.

Figure 15: The cumulative distribution of the use of data sources for studies in which all road users are included.

With regard to data collection using cameras, one of the main concerns nowadays is the privacy legislation. In many countries current regulations prohibit to film inside vehicles to examine for example looking behaviour. Most probably these regulations have influenced data collection over the years, since only one study was found that positioned cameras in such a manner that behaviour inside cars could be observed (Tenkink & Van der Horst, 1990). Other studies observing inside vehicles made use of human observers.

5.4.5. Combination with other methodologies

Several studies were found that combined the use of behavioural observation studies with another safety evaluation method (Figure 16). Of these, the use of stated behaviour methodologies (questionnaires, interviews and focus groups), was found in 80 studies (13.7%). Their main use is to determine to what extent self-reported behaviour (and
attitudes, beliefs and opinions) resemble observed behaviour (e.g. Geller, 1980; Hakkert 1981). The inclusion of crash data was found in 44 (7.5% of all) studies, mostly applied to identify locations of interest for behavioural observations. The combination with microsimulation tools was found in 27 studies, followed by Traffic Conflict Techniques with 20 studies. Combinations with driving simulators or naturalistic driving techniques were found rarely. Finally, five studies were found that complemented observed behaviour with police records (violations and traffic fines).

Figure 16: The combination of behavioural observation studies with other methodologies.

5.4.6. Number of sites

The number of (testing) sites used in behavioural observation studies ranges from one single location to 548 locations. Figure 17 provides an overview of the use of study sites related to the research goal (due to a low number of studies (n = 5), the research goal of software development was omitted), in which the range between one and 30 sites is shown. It can be seen that for all research goals the use of only one site is most common and that for the goal of model development almost half of all studies used one location. The graph further shows that the number of locations used decreases exponentially. It can also be seen that for the goal of effectiveness testing almost 70% of all studies used four or less testing locations. The highest number of testing locations were found for monitoring studies observing seatbelt use (Eustace & Bartel, 2002; Russo et al., 2014) and mobile phone use (Wenners et al., 2013; Cooper et al., 2013).

Figure 17: The number of sites related to research goal.
5.4.7. Sample sizes

The bar chart in Figure 18 shows the distribution of the sample sizes of the included behavioural observation studies. The majority of the studies used a sample size between 100 and 5,000 records for both VRU- and driver-studies. Studies with less than 100 records were rare and mainly related to monitoring and model development purposes. Studies with more than 100,000 records were also uncommon. These studies mainly focused on speeding using (loop) detectors as data collection tool (e.g. Montella et al., 2015; De Pauw et al., 2014). An exception is the study of Savolainen et al. (2011), who observed 103,047 pedestrians within the urban environment during five stages of enforcement. The lowest sample size was found for the study of Manan & Varhelyi (2015), which focused on identifying relevant behavioural indicators for merging behaviour of motorcyclists. An important remark is that around 18% of the 583 included studies did not report about the sample size. Most articles mentioned the observation period, but did not specify how many data-entries were analysed.

![Sample Sizes Chart](chart.png)

**Figure 18:** The observed sample sizes.

5.4.8. Observation period

The observation period was examined based on three aspects, namely peak or off-peak, daytime or night-time and week or weekend (Figure 19). The findings show that 69 studies (12%) did not report on any of these three aspects at all, making it difficult to compare findings. The element of day- or night-time observations is most often reported (68%), followed by week or weekend (56%) and peak or off-peak (49%) observations. From the information that is being reported, it can be derived that most research takes place during the day on weekdays. Only ten studies were found that only observed during night-time hours, focusing mainly on work zone safety and the effect of road lighting. It can further be seen that only eight pedestrian studies included night-time observations, but always in combination with daytime observations. With regard to peak- and non-peak-hours, no clear statements could be formulated. It is dependent on the research aims if high traffic volumes or free flowing traffic is needed for observing specific road user behaviour.
Figure 19: Reported observation periods.
6. Discussion

This scoping review gives an overview of the road user behavioural observation studies that have been published in English in peer-reviewed journals. In total, 583 papers were found eligible and retrievable and were included in this study, from which information regarding topic, indicators and methodological aspects was extracted. This chapter discusses the findings. The chapter is structured as follows: first some remarks regarding the review process will be mentioned, then a SWOT analysis is conducted and finally methodological considerations are discussed.

6.1. Review process

Although scoping reviews use a systematic approach for data retrieval, the quality and conclusions of such reviews are subject to bias and limitations. The following two sections describe which issues should be taken into account while interpreting the results.

6.1.1. Bias

As with any other study, retrieving information from the vast amount of available research is most likely subject to some form of bias. For scoping reviews, bias can be introduced during several stages of the review process. Following is a brief description of the types of bias that might have influenced our results:

- **Database bias:** When incorrect databases are used for information retrieval, the number of references found does not accurately reflect the total amount of references that can be found. This review only considered the databases of Science Direct, Web of Science and TRID, yielding 2,530, 1,920 and 16,719 results respectively. The assumption was made that these databases are sufficient enough to capture most relevant references, since these databases include the major peer-reviewed journals in transportation and traffic safety sciences.

- **Publication bias:** Publication bias regards the exclusion of research that has not been published. According to Rothstein et al. (2005) “publication bias is the term for what occurs whenever the research that appears in the published literature is systematically unrepresentative of the population of the completed studies”. Published studies in medical sciences tend to have more positive results than unpublished studies, which have lower effects or no significant findings. Evidence on publication bias shows that the bias has more to do with the significance of the results than the quality of the study. In our review, we choose to only include studies that have been published in the scientific literature, thereby most likely discarding many research that is linked with road user behavioural observation, but has not been published. However, since the aim of this study is to provide the current state-of-the-art, we believe that by using this criterion we capture only references that should be of higher quality. Furthermore, since the aim of this study is not to synthesize all available evidence about a specific topic (e.g. the influence of a safety improving measure), we feel it is justified to only include peer-reviewed journal articles as a first step to gain insights into the usefulness of behavioural observation studies.

- **Article selection bias:** In order to minimise the risk of losing relevant references, several search terms were tested and compared. We found that a rather broad formulation of terms was needed, resulting in a high number of references. During the first round of screening, only titles and abstracts were used to formulate an in-
or exclusion decision. It might be possible that due to unclear titles and abstracts relevant articles were excluded, but it is difficult to determine the extent of this bias. Several steps were taken minimize this form of bias:

- Whenever the abstract did not provide a clear indication as to whether to include or exclude the reference, the methodology section was examined;
- In case of doubt, the reference was kept for full-paper screening;
- The entire screening process was carried out by both members of the review team, i.e. both members evaluated all references. Different decisions were discussed and if no consensus was reached, the reference was kept for full-paper screening.

### 6.1.2. Limitations

The inclusion of selection-criteria unavoidably leads to limitations. However, in order to guarantee the feasibility, transparency and replicability of the review, restraints were needed. The following aspects might be considered as limitations for this study:

- The most important limitation posed in this research regards the inclusion of only peer-reviewed journal articles. This choice is justified by the objective to present the current state of the art in behavioural observation studies. Since peer-reviewed articles follow an extensive review process by experts in the field, we assume that only studies with higher quality are included. Furthermore, limiting the scope to peer-reviewed journal articles only is not uncommon in the field of scoping reviews (Pham et al., 2014). As a result, book sections, conference proceedings and research reports were not taken into account, even though they might contain important information regarding current the use of behavioural observation studies. For example, the InDeV-questionnaire (chapter 3, part 1) showed that several studies regarding seatbelt and child restraint usage were published as research reports and therefore were not included in this review. It might indicate that the current practice and the state-of-the-art in road user behavioural observation studies differ.

- Only articles in English were considered for inclusion, due to the limitation of the knowledge of (certain) foreign languages. It should be noted that most peer-reviewed articles are published in English anyway: English is the predominant language in contemporary research. Researchers outside the English speaking world who want recognition of their work have little choice but to publish their results in English (Egger et al., 1997).

- The process of reference retrieval was carried out December 2nd, 2015. All peer-reviewed studies that met the criteria and were published or accepted for publication were included. Publications scheduled to be released in 2016 were recoded as 2015 (n = 2). Several papers were not retrievable. The majority of these papers regard publication from before 1985, a time period in which journal articles were not necessarily digitized yet. In order to acquire as much references as possible, a list of missing references was sent out to the project partners, authors were contacted through Research Gate and the UHasselt library was consulted. Despite these efforts, 37 papers could not be located and might influence our results slightly.

### 6.2. Strengths, weaknesses, opportunities and threats

In total, 583 behavioural observation studies were analysed. Based on the extracted information and the findings of the analysis, a SWOT-analysis was conducted pointing...
out the advantages and disadvantages of using behavioural observation studies (Table 6). Following is a brief description of the identified elements.

**Table 6: A SWOT analysis on road user behavioural observation studies.**

<table>
<thead>
<tr>
<th>Internal factors</th>
<th>Positive</th>
<th>Negative</th>
</tr>
</thead>
<tbody>
<tr>
<td>Strength</td>
<td>Natural driving behaviour</td>
<td>Weakness</td>
</tr>
<tr>
<td></td>
<td>Behavioural and situational processes</td>
<td>Control of traffic events</td>
</tr>
<tr>
<td></td>
<td>Data quality</td>
<td>Control groups</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Data processing</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Bias</td>
</tr>
<tr>
<td>External factors</td>
<td>Opportunities</td>
<td>Threats</td>
</tr>
<tr>
<td></td>
<td>Amount of data</td>
<td>Privacy legislation</td>
</tr>
<tr>
<td></td>
<td>Control group</td>
<td>Validity</td>
</tr>
</tbody>
</table>

### 6.2.1. Strengths

The main strength of road user behavioural observation studies is that naturalistic data is gathered without road users’ knowledge that they are being observed for research purposes, thereby limiting the effects of behavioural adaptation. In surveys, for example, respondents try to answer questions in such a way that makes them look good (Paulhus, 1984). Natural settings in which road users are unaware of being observed reduces this bias and may lead to risky and aggressive behaviour while driving (Shinar, 1998). Notion should be made of the trend during recent years to install permanent traffic camera’s for monitoring purposes, but we believe that road users have grown accustomed to them and therefore do not adapt their behaviour anymore. Furthermore, using these ‘unobtrusive’ observations enables researchers to examine natural behaviour of road users, with the opportunity to identify behavioural and situational processes that lead to traffic safety issues. This is important, since other forms of data collection techniques fail to include such information. Current studies have led to the creation of microsimulation models and suggestions as to why road users behave in a certain manner (e.g. informal traffic rules).

Another strength of behavioural observation studies regards the quality of the data. Although earlier studies using human observers reported frequently on inter-observer agreement rates, the application of video cameras enables to watch traffic events of interest as many times as needed. This improves the quality of the data, since real world driving behaviour can be examined multiple times with different perspectives (for example once for identifying looking behaviour and once for examining situational circumstances like weather and road conditions).

### 6.2.2. Weaknesses

One of the major drawbacks of behavioural observation studies is that researchers are dependent on what happens. Unlike for example driving simulator research or microsimulation, traffic events of interest cannot be triggered without the ability to collect data of all road users involved. Semi-controlled research designs partly overcome this issue by using instructed road users to provoke certain traffic events (e.g. yielding), but are still constrained by the limitation that only data can be collected of the road user that is being ‘trapped’ in the traffic event of interest.

The findings of the review indicate that the main aim of behavioural observation studies is to observe what happens, rather than to perform specific safety evaluations. If, however, the effects of a safety treatment were tested (n = 202), only sixteen studies (8%) were found that included some sort of a control group. Those studies found that results needed to be adjusted to control for natural variability. This suggests that, as for many other domains, behavioural studies also need control groups when the effectiveness of a...
safety treatment is tested. However, the review showed that control groups are rarely applied.

Other weaknesses regard the form of data collection. When human observers are used, one should be careful to guarantee the objectivity of data collection and keep track of the inter-observer agreement rates between trained observers. As already mentioned video data can partly address this issue, but one should be aware that technical problems and occlusions by for example other road users can limit its quality. Furthermore, video cameras enable the continuous collection of large samples of data, but the labour and costs that are needed to analyse the data are quite extensive. Current efforts in developing automated video analysis software tools might provide to be a valuable asset for data analysis. However, current accuracy of these software tools are not always able to capture the quality needed for traffic safety evaluation and observation.

Finally, sampling bias can influence the results of behavioural observation studies (and other research methodologies). Behaviour observed is only a small sample of all behaviours that can occur and it can never be concluded that the same behaviours would be conducted by people who have not been observed (Porter, 2011). Although a proper sampling design can minimize this form of bias, it cannot be eliminated.

6.2.3. Opportunities

During recent years the streetscape is changing. Cameras are installed rapidly, with as main purpose to monitor the current state of traffic and the environment. Opportunities exist to use the video footage from these cameras to observe road user behaviour. However, issues regarding the data-storage, privacy legislation and data quality should be taken into account. Combined with the current efforts to improve and develop automated video-analysis software tools, huge opportunities exist to ease safety evaluation practices. Current software applications are already able to classify road users and trace them through the video image, but once (tracking) accuracy is improved further such software might be exploited to observe crossing and yielding behaviour in more detail.

It was found that almost 14% of the studies combined the observation of behaviour with other methodologies. Such combinations offer the opportunity to compare behaviour in the real driving environment with stated or simulated behaviour. Even more, behavioural observation data might be used as validation tool, since it reflects natural road user behaviour.

6.2.4. Threats

The most important threat for road user behaviour observation studies regards privacy legislations when video cameras are used for data collection. Personal experiences with video observation studies showed that strict rules regarding the collection of personal data exists. Permits are required to be allowed to make video recordings, on which license plates or faces should not be recognizable. This prohibits to observe inside drivers’ vehicles and as a result excludes certain research topics (e.g. seatbelt use, mobile phone use and looking behaviour). These data can be gathered with human observers, but as explained earlier, more risks of human subjectivity exists.

Another important threat regards the validity of the behavioural indicators. The review shows that a wide variety of indicators is used to describe behaviour, but the relation between these behaviours and safety is rarely validated. Literature has shown a relation between speed and safety (Elvik, 2004), but for other indicators no such link has yet been
proven. It is generally assumed that the used behavioural indicators are a valid proxy for traffic safety.

6.3. Methodological considerations

One of the main observations regarding the research design is that road user behavioural observation studies rarely include control groups, when a safety evaluation is being performed. The sixteen studies that used one indicated that results needed to be adjusted due to confounding factors. Although the use for cross-sectional research designs might be less useful or feasible (if locations to be compared are as similar as possible), studies using a before-after and with-without research design should include some type of mechanism to control for confounding factors.

When looking at the number of testing sites, it was found that the majority of the studies only looked at one or two locations, even for studies evaluating the effect of a certain safety improving treatment. For these latter studies, the share of studies limiting themselves to two or less testing sites amounted to 48%. The question is, however, to what extent these limited amount of selected observation sites are representative and, as a result, if conclusions can be generalized. Therefore, it should be encouraged to increase the amount of testing locations. This study found that 20% of the evaluation studies used more than eight locations, showing that the inclusion of multiple testing sites is not uncommon practice.

An important consideration regards the interest of road user type behaviour. Although behavioural observation studies are dependent on what happens, semi-controlled research designs have been used to force traffic events of interest to occur. Although this methodology prevents data collection of the instructed road user, data collection efforts can be used efficiently, observation periods can be reduced and one can control the sample size. The review showed that especially for VRU-studies such designs are useful for monitoring purposes.

Finally, efforts should be allocated into validating the use of behavioural indicators. The review showed that many indicators have been used, but their added value with regard to safety is rarely discussed. It should be noted that it might be difficult to do so, since not all road user behaviour observation studies have as main goal to evaluate traffic safety.
7. Conclusion

Road user behavioural observation studies are mainly used to monitor traffic events and to simply observe what happens (> 50%). Evaluation of safety improving measures account for one third of all studies, suggesting that behavioural observation can also be used to directly evaluate traffic safety aspects. From the 583 studies included in this review, 36% included VRUs. Especially during the last fifteen years, an increase in research efforts of VRU safety is observed. Behavioural observations seem very useful for examining how road users interact with other road users or navigate through a crossing. Almost all studies involving a VRU were found to take place at some sort of crossing (e.g. intersection, railroad crossing, pedestrian crossing). For studies involving drivers on the other hand, most current reported research efforts have focussed on road stretches, in which interaction with other road users was not necessarily required.

An important notion to be made is that the lack of current reporting issues might limit the value of journal articles. This review encountered difficulties with determining heavy goods vehicles. Furthermore, taking into account transparency and replicability and to enable comparisons between studies, information about the observation period should be clearly stated. It was found that 12% of all studies did not provide any information and that information regarding peak- of off-peak, daytime or night-time and week or weekend observations was missing in 51%, 44% and 32% of the studies respectively. Additionally 18% of all studies neglected to specify the sample size.

This scoping review limited itself to peer-reviewed journal articles published in English only. Although it was not the aim of this review to depict the current practice everywhere (one is limited in resources), this research suggests that certain topics have not been subject of research much. For example, rail road crossings are considered to be a major safety hazard across the EU (Eurostat, 2016) and powered-two-wheelers are highly represented in crashes (WHO, 2004), but both were only found sporadically in the review. However, the InDeV-questionnaire hinted that many relevant studies might be published as research reports only. Therefore, it is advised to extend the boundaries of this scoping review in future research.
8. Acknowledgements

This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 635895 (InDeV- In-Depth-understanding of accident causation for Vulnerable road users). This publication reflects only the authors’ views. The European Commission is not responsible for any use that may be made of the information it contains.
9. References

9.1. References included in review


Deliverable D1.2 “Review of current study methods for VRU safety – part 3”


Charlton, S. G. (2003). Restricting intersection visibility to reduce approach speeds. Accident Analysis & Prevention, 35(5), 817-823. doi:http://dx.doi.org/10.1016/S0001-4575(02)00522-0


Deliverable D2.1 “Review of current study methods for VRU safety – part 3”


Dell'Acqua, G. (2014). Behavioural effects of fixed speed cameras on motorways: Overall improved speed compliance or kangaroo jumps? Accident Analysis and Prevention, 73, 132-140. doi:10.1016/j.aap.2014.08.019


Eckes, H. K., Tao, R., & Mangum, B. C. (2004). EVALUATION OF PEDESTRIAN COUNCITOWN SIGNALS IN MONTGOMERY COUNTY, MARYLAND. Maryland Transportation Research Record(1878), pp. 36-41.

Deliverable D.1 “Review of current study methods for VRU safety – part 3”


Faw, H. W. (2013). To signal or not to signal: That should not be the question. Accident Analysis & Prevention, 59, pp 374-381.


- 44 -
Deliverable D2.1 “Review of current study methods for VRU safety – part 3”


Deliverable D2.1 “Review of current study methods for VRU safety – part 3”


Lipovac, K., Teslic, M., Marin, B., & Deric, M. (2015). Self-reported and observed seat belt use – A case study: Bosnia and Herzegovina. Accident Analysis & Prevention, 84, 74-82. doi:http://dx.doi.org/10.1016/j.aap.2015.08.010


Liu, B. (2007). Association of intersection approach speed with driver characteristics, vehicle type and traffic conditions comparing urban and suburban areas. Accident Analysis & Prevention, 39(5), pp 216-227.


Deliverable D2.1 “Review of current study methods for VRU safety – part 3”


Deliverable D2.1 “Review of current study methods for VRU safety – part 3”


Ng, O. K., & Saccamanno, F. (2010). Speed Reduction Profiles Affecting Vehicle Interactions at Level Crossings with No Trains. Transportation Research Record(2149), 108-114. doi:http://dx.doi.org/10.3141/2149-14


Papaioannou, P. (2007). Driver behaviour, dilemma zone and safety effects at urban signalised intersections in Greece. Accident Analysis and Prevention, 39(1), 147-158. doi:http://dx.doi.org/10.1016/j.aap.2006.06.014


Deliverable D2.1 “Review of current study methods for VRU safety – part 3”

Polderman, E., Comu, J., De Courten-My湲ins, T., Daniels, S., Brisk, J., Brisk, T., ... Wets, G. (2015). Drivers’ behavioural responses to combined speed and red light cameras. Accident Analysis & Prevention, 81, pp 153-166.

- 51 -
Deliverable D.1 “Review of current study methods for VRU safety – part 3”


Walker, I. (2007). Drivers overtaking bicyclists: Objective data on the effects of riding position, helmet use, vehicle type and apparent gender. Accident Analysis & Prevention, 39(2), 417-425. doi:http://dx.doi.org/10.1016/j.aap.2006.08.010
Deliverable D2.1 “Review of current study methods for VRU safety – part 3”

Walker, I., Garrard, I., & Jowitt, F. (2014). The influence of a bicycle commuter’s appearance on drivers’ overtaking proximities: An on-road test of bicyclist stereotypes, high-visibility clothing and safety aids in the United Kingdom. Accident Analysis & Prevention, 64, 69-77. doi:http://dx.doi.org/10.1016/j.aap.2013.11.007


Wesley, P. (1994). Speed as a measure of driver risk: Observed speeds versus driver and vehicle characteristics. Accident Analysis & Prevention, 16(2), 89-103. doi:http://dx.doi.org/10.1016/0001-4575(84)90344-4


Wong, J., Xue, S., Yang, Y., Yan, X., & Qu, X. (2015). In-depth analysis of drivers’ merging behavior and rear-end crash risks in work zone merging areas. Accident Analysis & Prevention, 77, pp 51-61.


9.2. Additional references in report


9.3. Potential relevant but irretrievable references


Deliverable D2.1 “Review of current study methods for VRU safety – part 3”

ANNEX 1

In order to guarantee the feasibility and the effectiveness of the search protocol, several options in the databases were selected prior to the reference retrieval process. The table below shows how the search terms were defined and which filters were used.

<table>
<thead>
<tr>
<th>Search term: Traffic Behavior AND &quot;Observation OR Safety&quot;</th>
<th>TRID</th>
<th>Web of Science</th>
<th>Science Direct</th>
</tr>
</thead>
<tbody>
<tr>
<td>“Result Type”</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Only articles and papers</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>“Subject Category”</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Design</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Highways</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Pedestrians and Bicyclists</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Policy</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Research</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Transportation (General)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Vehicles and equipment</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>“Languages”</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>English</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>“Research Areas”</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Transportation Sciences</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Behavioral Sciences</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>“Publication Title”</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Accident Analysis &amp; Prevention</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Journal of Safety Research</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Journal of Transport Geography</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Procedia – Social and Behavioral Sciences</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Safety Science</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Transportation Research Part A</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Transportation Research Part B</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Transportation Research Part C</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Transportation Research Part F</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Transport Policy</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>“Topic”</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Unite state</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Driver</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Vehicle</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Traffic</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Network</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Europe</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Road</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Human</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Accident</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>