



Project No. 635895 — InDeV

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

HORIZON 2020 - the Framework Programme for Research and Innovation

Deliverable 2.1 – part 4

Review of current study methods for VRU safety

Appendix 6 – Scoping review: surrogate measures of safety in site-based road traffic 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)

Version 1.0 (2016-09-30)

Dissemination Level		
PU	Public	X
PP	Restricted to other programme participants (including the Commission Services)	
RE	Restricted to a group specified by the consortium (including the Commission Services)	
CO	Confidential, only for members of the consortium (including the Commission Services)	

Document information

Appendix 6: Aliaksei Laureshyn (Lund University, Sweden)
Carl Jonsson (Lund University, Sweden)
Tim De Ceunynck (Hasselt University, Belgium)
Åse Svensson (Lund University, Sweden)
Maartje de Goede (TNO, the Netherlands)
Nicolas Saunier (Polytechnique Montréal, Canada)
Paweł Włodarek (Warsaw University of Technology, Poland)
Richard van der Horst (TNO, the Netherlands)
Stijn Daniels (Hasselt University, Belgium)

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

Version	Date	Comment
1	2016-09-30	First official version

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

1. Introduction.....	1
2. Background	3
2.1. Historical perspective	3
2.2. Limitations of accident data.....	3
2.3. The “pyramid” concept	4
2.4. What is severity?.....	5
2.5. Relation between accidents and conflicts	5
2.6. Validity	8
2.7. Reliability.....	9
3. Methodology.....	12
4. Results	15
4.1. Overview of the literature	15
4.2. Surrogate safety indicators	17
4.3. Traffic conflict techniques.....	35
4.4. Surrogate safety studies of VRU.....	50
4.5. Study design	53
5. Discussion	58
5.1. Validity of applied techniques and indicators	58
5.2. Safety continuum and continuous safety indicators	60
5.3. Outcome severity in case of an accident.....	61
5.4. Subjective component in safety studies	61
5.5. Surrogates safety measures for VRUs.....	62
5.6. Strengths and limitations (of this study)	64
6. Conclusions.....	66
7. Acknowledgements	67
8. References.....	68

1. Introduction

Surrogate measures of safety are meant to be tools to investigate traffic safety. The term “surrogate” describes that these measures do not rely on accident data and instead are meant to be an alternative or a complement to analyses based on accident records. The “traffic safety” can generally be considered as “the absence of unintended harm to living creatures or inanimate objects” (*Evans, 2004*). Vision Zero set the highest priority in traffic safety work on reduction (and ultimately elimination) of the risk for fatalities and serious injuries in the road transport system (*Johansson, 2009*).

The approach based on surrogate safety measures has advantages compared to accident-based analysis being pro-active (and thus more ethical as there is no need to wait for accidents to happen), and in some conditions more time-efficient, informative and even accurate (*Svensson, 1992, Hydén, 1987*). There are, however, still many unresolved issues when it comes to selection of the appropriate indicators, their validity, data collection and analysis procedures, etc. The traditional reliance on human observers and related costs turned out to be a serious hinder for wide acceptance of the surrogate safety analysis in everyday practice.

The last decade has been characterized by great improvements in sensor technologies that can be applied for collection of traffic data in general and surrogate safety measures in particular (*Tarko et al., 2017, Laureshyn, 2010, Saunier et al., 2010*). This opens new opportunities but also poses new questions and challenges.

Researchers new to the surrogate safety measures struggle to gain a clear overview of the current state of the field. The literature in this domain is vast and diverse; while many of the old (but still highly relevant) research effort reports exist only in paper format and are hard to reach, the publications related to the technical improvements in this area seem to grow in number very rapidly. Even for more experienced researchers there is a risk of losing track of the critical points of attention. The lack of holistic overview seems to lead to “reinventing the wheel” and errors from the past being repeated.

Therefore, this report presents a scoping review of the current state in the field of surrogate safety measures, providing a descriptive insight into the literature that is available, and identifying critical challenges and opportunities that should be central in future research. A particular focus of this report is the application of the surrogate safety measures for studies of vulnerable road users (VRUs). Most of the earlier research was performed with “cars in mind” and the existing methods are often not directly applicable on pedestrians or cyclist. Therefore, we aimed at putting together the scattered knowledge about existing application of surrogate safety measures on VRUs as well as highlighting the knowledge gaps that still have to be covered in the future.

It has been a challenge to reach some of the publications, particularly the old ones. To save the efforts for the interested researchers, we address them to the library of the ICTCT (International Co-operation on Theories and Concepts in Traffic safety - *ICTCT*,

2016) where some of the most important and influential papers can be found (www.ictct.org/ictct-library/tct-library).

2. Background

2.1. Historical perspective

The first attempts to study safety based on observation of critical situations in traffic (traffic conflicts) emerged in the 1950-60s (*Perkins & Harris, 1967, Forbes, 1957*). After successful first attempts, the approach rapidly gained popularity, leading to the development of numerous new conflict observation techniques in the 1970s. Countries that developed their own techniques included Sweden, the Netherlands, France, Germany, Finland, Canada, the United Kingdom, the United States, Austria, etc. (*Asmussen, 1984*). Since all techniques used different indicators and protocols to select traffic conflicts, comparability of the findings as well as limited validation for each individual technique were identified as significant issues. Therefore, in the 1980s and early 1990s, research efforts emphasized aspects such as validity (*Hauer & Gårder, 1986, Williams, 1981*), comparability (*Grayson, 1984*) and reliability of observations and assessments (*Kruysse, 1991, Hydén, 1987, Lightburn & Howarth, 1979*).

The International Committee on Traffic Conflict Techniques (ICTCT) deserves mentioning in this context. Built by a group of researchers working with traffic conflicts in 1977 (*ICTCT, 2016*), this group organised a number of workshops and international calibration studies and developed solid theoretical grounds for the use of surrogate safety measures.

However, the traffic conflict techniques were applied less frequently after the early 1990s. The most important reasons for this languish were the significant costs in time and efforts required to collect such data in the field by human observers. Recent improvements in advanced video analysis techniques (*Laureshyn, 2010, Saunier et al., 2010*) and other sensor technologies (*Tarko et al., 2017*) have led to a renewed interest in the use of traffic conflicts – or other relevant surrogate indicators - to assess road safety.

2.2. Limitations of accident data

Traditionally, road safety is described in terms of number of accidents or injuries that occur in traffic. While such indicators have the most direct connection with the subject studied, they have a number of serious limitations:

- Traffic accidents are random events and the number of accidents registered every year at the same place is not the same, even if the traffic situation is unchanged. From this perspective, the actual accident number is also an indirect measure, while the “true” safety characteristic is the “expected number of accidents” that cannot be measured but only estimated based on the accident history or using some other methods (*Hauer, 1997*).
- Traffic accidents are rare events and it takes a long time to collect a sufficient amount of accidents data to produce reliable estimates of the expected number of accidents. During that time the traffic conditions might (and usually do) change. There is also an ethical problem in that one has to wait for sufficient number of accidents to occur and thus for people to suffer before anything can be said about the (un)safety.
- Not all accidents are reported. The level of underreporting depends on the accidents severity and types of road users involved. This is particularly a problem

for the vulnerable road users (*Elvik et al., 2009, Amoros et al., 2006, Alsop & Langley, 2001*).

- The actual process of the accidents is often unclear as only the final accidents outcome can be observed during the registration phase. accidents reconstructions and in-depths investigations are usually costly and not always possible to perform; doing it with a help of witnesses and those involved in the accidents presents a great risk of bias. Without information about the process preceding the accident it is very difficult to understand the link between (contributing) behaviour and accident and thus limits the possibilities to propose effective counter-measures in order to change/reduce this behaviour.

For these reason, there is an interest to use some other, indirect measures for traffic safety (also called surrogate safety measures). Indirect in this context means that these measures are not based on accidents, but rather on other occurrences in traffic that are causally related to accidents or injuries, can indicate safety performance and help to understand the process that leads to accidents.

2.3. The “pyramid” concept

The basic concept that the traffic conflicts theory is based on is that traffic process can be seen as a number of elementary events. These events differ in their degree of severity (unsafety) and there exists some relation between the severity and frequency of events of that severity. *Hydén (1987)* illustrated the concept with a “safety pyramid” (see Figure 1). The base of the pyramid represents the undisturbed passages which of course are very safe and occur most of the time. At the other end, the very top of the pyramid consists of the most severe events such as fatal or injury accidents and that are very infrequent compared to the total number of the events. If the form of the relation between the severity and frequency of the events is known, it is theoretically possible to calculate the frequency of the very severe but infrequent events (accidents) based on known frequency of the less severe, but more easily observable events (conflicts).

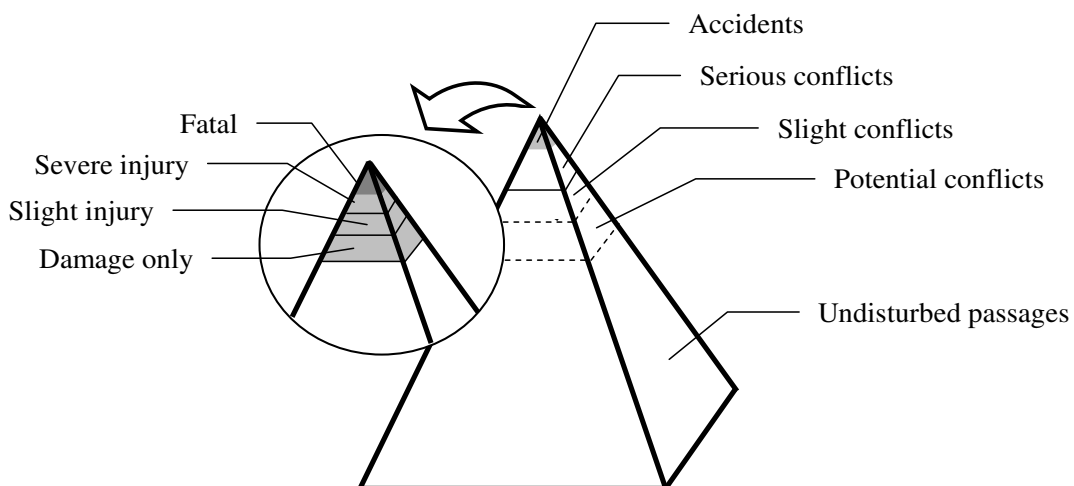


Figure 1. “Safety pyramid” (adopted from *Hydén, 1987*).

Svensson (1998) limited the events to be included in the safety hierarchy to those in which there actually was a risk of a collision (thus, for example, all single passages were excluded). Elaborating further on the meaning of the distribution shape, she pointed out that the most frequent events are not necessarily the safest ones (“severity

“diamond” model, see Figure 2a). The suggested explanation was that road users optimise their behaviour in respect to safety and travel time, preferring to be involved in events of moderate severity but saving some time through keeping higher speeds and accepting smaller gaps.

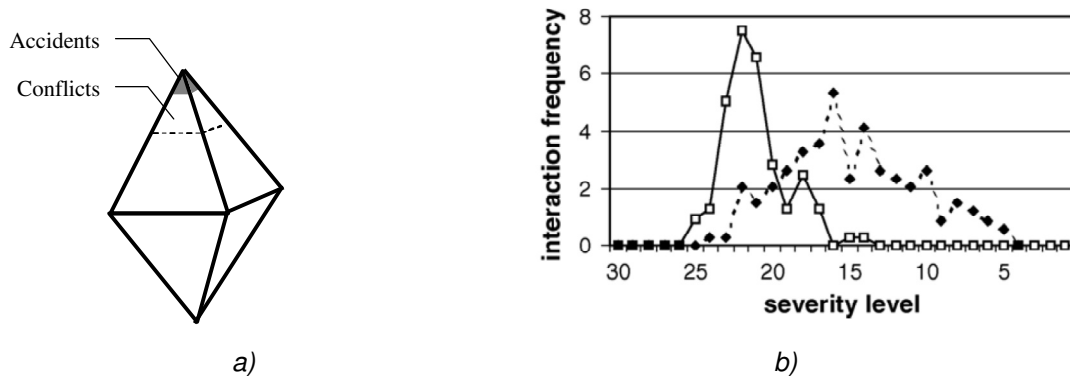


Figure 2. "Severity diamonds" (Svensson, 1998): a) conceptual representation; b) observed distributions of events with different severity levels (according to the Swedish traffic conflict technique) at two sites with different forms of intersection control.

Moreover, comparing different types of road environment, *Svensson (1998)* showed that the shape of the distribution varies depending on factors such as regulation form, road design, frequency of interactions, type of manoeuvre and road users involved, etc. (Figure 2b).

2.4. What is severity?

The concept of “severity” of an event also requires clarification. Most traffic conflict indicators that have been developed over the last decades express the severity of an event as its proximity to a collision in terms of time or space (*Zheng et al., 2014c*). However, the proximity to a collision is only one dimension “severity”. The potential consequences in case a collision had taken place is another dimension of “severity” that should preferably be taken into account in some way as well (*Laureshyn, 2010*). Following the goals set by Vision Zero in road safety – “no one will be killed or seriously injured within the road transport system” (*Johansson, 2009*) – an appropriate definition for the severity can be “a nearness to a serious personal injury” (*Laureshyn et al., 2017*). The potential consequences of an event are dependent on the type of road users involved and their vulnerability, speed, mass, type of collision, collision angle, etc. Recently, some attempts to combine these two dimensions into one indicator have been done (*Laureshyn et al., 2017, Bagdadi, 2013*).

The choice of the surrogate safety measure(s) that are used to operationalise the severity concept has a direct effect on which events are selected for inclusion into the “safety diamond”, as well as their position in the distribution (and thus the shape of the distribution).

2.5. Relation between accidents and conflicts

How the events of different severity are related has a direct effect on whether there are theoretical grounds to extrapolate the knowledge from the less severe events to the more severe ones (accidents).

Two alternative models relating traffic conflicts and accidents have been described by *Güttinger (1982)*. In the first, the conflict is defined as a set of initial conditions that depending on the successfulness of the evasive action either develop further into a collision or resolve without any consequences (see Figure 3a). Defined in this way, conflicts and collisions belong to the same dimension as a conflict always precedes a collision and with a certain probability may develop into a collision. The indicators describing the severity in this case should be based on the initial conditions and not on the intensity of the evasive action or the final outcome of the event.

In the alternative model (Figure 3b), it is the evasive action that results either in a collision or “an avoided collision” – conflict. In this definition, conflicts and collisions exist “in parallel”.

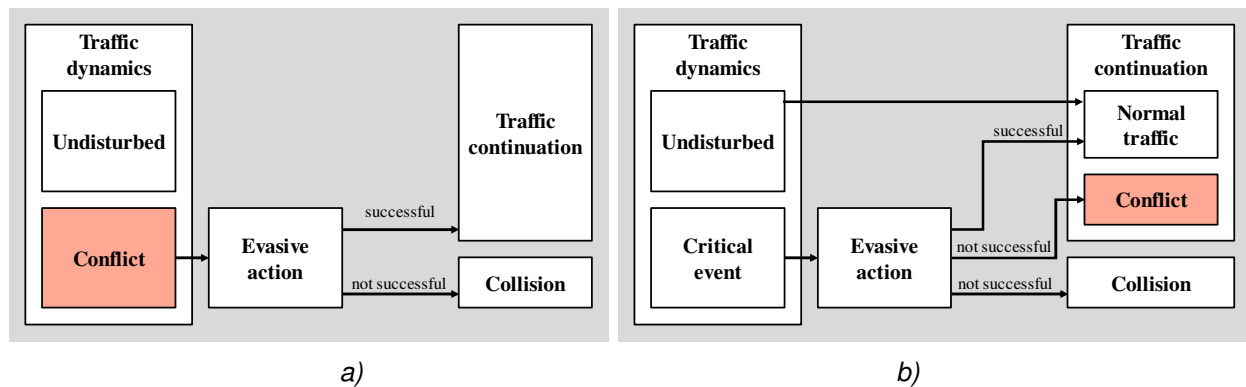


Figure 3. Two models of relation between traffic conflicts and accidents (adopted from *Güttinger, 1982*): a) conflict precedes a collision; b) conflict is “parallel” with a collision.

Which model lies behind a surrogate safety methodology is crucial for the validity of the methodology. If conflicts and accidents do not belong to the same continuum, the use of one type of events to predict frequency of the other type of events is not well-motivated. The causal chain is broken. For example, there might be some factors always present in collision situations and absent in conflict situations (or vice versa) that are crucial for whether the situation is resolved successfully or not.

The model in Figure 3a can be interpreted in a way that a conflict is an opportunity that either result in an accident or not. In this case, the relation between conflicts and accident should take a form (*Hauer & Gårder, 1986, Hauer, 1982*):

$$\lambda = c \cdot \pi ,$$

where λ – number of accidents expected to occur on an entity during a certain period of time;

c – expected number of conflicts on an entity in that time;

π – accident-to-conflict ratio (or probability of a conflict to develop into an accident).

Hauer & Gårder (1986) pointed out that such representation is very simplistic as conflicts vary in degree of their severity and the probability to develop into an accident for conflicts of different severity is most probably not the same. Therefore, the more correct representation would be:

$$\lambda = \sum c_i \cdot \pi_i ,$$

c_i, π_i – expected number of conflicts and accident-to-conflict ratio for each severity category i .

In practice, one often has to work with a mixture of conflicts of varying severity which creates a problem of identifying the conflicts groups that have relatively stable π -values (Davis et al., 2011). As an alternative, the probability of a collision might be presented as a continuous function of the severity measure. If the chosen severity measure treats conflict and accidents as events belonging to a same continuous severity dimension (Tarko, 2012), a collision is then an event with severity above a certain (quite high) threshold. Special statistical methods can be applied to estimate the frequency of the events with severity above the threshold based on frequencies of the more frequent events (e.g. extreme-value theory, Tarko, 2012, Songchitruksa & Tarko, 2006). For example, if severity is defined by a parameter such as Time-to-Collision (TTC), Post-Encroachment Time (PET), available braking distance/time, etc. a value of zero is an obvious threshold at which a physical contact takes place meaning it is a collision.

Attractive from a theoretical point of view, such approach is not problem-free when practically applied. For some definitions of severity, particularly those attempting to reflect the risk of injuries rather than simply risk of a collision, the choice of a threshold is not always obvious (Laureshyn et al., 2017). Even the postulate of the “continuity” is sometimes questioned (Campbell et al., 1996). It appears that more severe traffic conflicts have more similarities with accidents (Svensson & Hydén, 2006), but on the other hand limiting oneself to studying such events only brings back the problem of low frequency of events and issues related to small sample size.

Another aspect that effects the stability of π is the conditions for which it is estimated. It has been acknowledged that depending on the type of road users involved, manoeuvre they perform and speed regime π -value would be different (Hydén, 1987, Linderholm, 1981, Hydén, 1977). It was noted, however, that introduction of many conflict categories quickly results in very few conflicts and accidents per category thus making the estimates of π very uncertain.

Wu & Jovanis (2012) generalised this problem suggesting to use a discrete choice models (probit or logit) to estimate the probability for a surrogate event to develop into an accident. As the input for the model, a set of variables describing the surrogate event conditions are used. Wu & Jovanis (2012) successfully applied this approach on a case of road departures using the data from a naturalistic driving study in which accidents, near-misses and a long list of variables describing the situational conditions for those events were available.

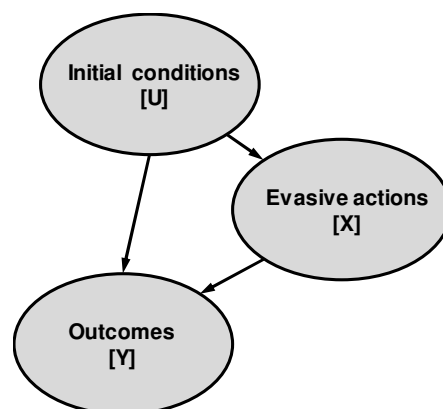


Figure 4. Causal model (adopted from Davis et al., 2011).

Davis et al. (2011) outlined a causal model that can be used to calculate the probability for any traffic event to develop into an accident (Figure 4). A set of initial conditions [U]

defines a set of evasive actions [X] with their probabilities. The final outcome [Y] is a probabilistic function of both the initial conditions and the evasive actions. A conflict is then defined as a set of initial conditions in which the final probability of a collision is above zero. The sum of all conflicts multiplied by their develop-to-accident-probabilities yields the expected accident number during the studied period. To demonstrate the approach, *Davis et al. (2011)* applied the model on a case of car-following interactions on a motorway using the distribution of driver reaction times and braking intensity from naturalistic data. It appears, however, that the task of defining the probabilistic functions that relate the model components for more complex traffic situations (e.g. urban traffic with mixed directions and types of road users) seem to be a much bigger challenge.

2.6. Validity

2.6.1. The concept of validity

Validity is a crucial aspect of any study or measurement. Validity, in general, relates to the approximate truth of an inference, and it is informed by both correspondence and coherence conceptions of truth, as well as pragmatism (*W. Shadish, T. Cook, D. Campbell/Shadish et al., 2002*). It is important to note that validity is not a matter of “yes or no”, but it is a matter of degree (*Carmines & Zeller, 1979*). Therefore, validity is a concept designating an ideal state. This implies that validity is a concept that has to be “pursued”, but that cannot be completely “attained” (*Brinberg & McGrath, 1985*). Whether a certain level of validity is considered “sufficient” is therefore usually rather a matter of argumentation, debate and consent than a measurable aspect that should exceed a certain threshold. Validity has to be assessed relative to purposes and circumstances (*Brinberg & McGrath, 1985*).

The validity of an indicator concerns the crucial relationship between concept and indicator. Validity relates to the use to which a specific measurement is put: does an indicator actually “measure” the property you want to measure (*Carmines & Zeller, 1979*). Therefore, one validates not the measuring instrument itself, but rather the measuring instrument in relation to the purpose for which it is being used (*Carmines & Zeller, 1979*). Validity is evidenced by the degree that a particular indicator measures what it is supposed to measure rather than reflecting some other phenomenon that is not intended to be measured (*Carmines & Zeller, 1979*).

The main goal of developing and using surrogate safety indicators is to measure traffic (un)safety. Therefore, validity of an indicator means to what extent it describes (un)safety which from Vision Zero-perspective equals to expected serious injuries in traffic.

There many ways to study validity of a surrogate safety measure – expert judgements, comparison with other “indirect” measures (e.g. micro-simulation results), comparison with observed/reported accidents, etc. The “strength” of validation vary depending on which approach is used. Throughout the report, we focus only on the validation studies that do compare the surrogate measures with the actual accident data.

2.6.2. Product validity

Product validity deals with how well a surrogate safety indicator is able to estimate expected number of accidents. According to *Hauer & Gårder (1986)* who looked into this issue on behalf of TCT “it should be clear that the performance of the TCT cannot

and should not be judged by its success in *predicting* future accidents. The number of accidents to occur in the future can no more be predicted than can the roll of a die. The proper question to ask is: how good is the TCT in estimating the number of *expected accidents?*”. In this sense the TCT or any other safety indicator, should be compared to other methods, e.g. accident data or exposure, and comparisons should be made between the variances of the estimates. *Hauer & Gårder (1986)* conclude, in their attempt to make a final definition of “validity”, that “a technique (method, device) for the estimation of safety is “valid” if it produces unbiased estimates, the variance of which is deemed to be satisfactory.” This further implies, in order to obtain a reliable and stable estimate of the expected number of accidents, the variance of this “accident-to-conflict ratio” should be sufficiently low.

2.6.3. Relative validity and process validity

In case a conflict observation study or indicator does not allow to reliably calculate the expected number of conflicts, but allows to reliably indicate the direction (and possibly order of magnitude) of change, one can speak of **relative validity**. If, as previously stated, validity is a matter of degree (*Carmines & Zeller, 1979*), then relative validity can be considered as a lower degree of validity than product validity.

Process validity indicates the extent to which safety indicators can be used for describing the process that leads to accidents (*Svensson, 1998*). In the process validation work of the Swedish Conflicts Technique, *Hydén (1987)* has compared the processes preceding injury accidents to those preceding conflicts. Analyses showed large similarities between accidents and conflicts when the comparison was based on TA values and conflicting speed. The analyses also showed that the type of evasive actions were very similar among accidents and conflicts. *van der Horst (2013)* conducted long-term video observations to collect data on the pre-crash phase of real accidents (what exactly happened just before the collision?). The video recordings of collisions were used to evaluate and validate the safety value of in-depth accident analyses, road scene analyses, and behavioural observations (including traffic conflicts). The conflicts that were scored clearly illustrated typical safety problems and resulted in several observations about the typical lay-out and functioning of the intersection at hand, well in line with the accident causation processes that resulted in the video-collected collisions.

2.7. Reliability

The concept of reliability refers to the accuracy and the consistency of measurements. In other words, the measured value very closely represents the “true” value and the measurement error should remain within the same limits regardless of measurement locations, time of the day, traffic situation, etc., thus ensuring that measured differences reflect the actual difference in the studied phenomenon and not in the measurement’s accuracy (*Laureshyn, 2010*).

From perspective of surrogate safety indicators, two main aspects should be considered:

- Accuracy of measurements for and individual traffic event (road users’ position, speed, etc.) and the detection errors related to that.

- The observation time necessary to collect sufficient number of individual events to be able to generalise their frequencies (e.g. estimate the “expected number of conflicts”).

The original traffic conflict techniques developed in 1970-80s relied heavily on human observer judgements both in detection and rating of the safety-critical traffic events (*Hydén, 1987, van der Horst & Kraay, 1986, Migletz et al., 1985, Baguley, 1982, Muhlrud, 1982*). Human cognitive capacity puts serious limitations on the complexity of the analysis that is feasible to perform in field conditions and in real time. As a consequence, the techniques operated with very few severity categories and often were based on verbal rather than quantitative classifiers. Even though, much critic was raised towards the reliability of the human observers as such.

When it comes to human estimates of objective measures, several validation studies showed that with proper training it is possible to get quite adequate accuracy. In general, humans are not very good at estimation of time-related measures (like time-to-impact, time-to-lane-crossing) and acceleration (*Kiefer et al., 2006, A. G. Galevan der Horst, 1991*). On the other hand, *Hydén (1987)* showed that it is possible to train observers to estimate speed and distance with sufficient accuracy (and based on those parameters even time-related measures can be calculated).

The automated data collection methods are objective per definition, but the performance of the technical tool might be influenced by the conditions in which it is used. In case of automated video analysis such factors, beside the choice of the video processing algorithm itself, are (*Morse et al., 2016*):

- Quality of the underlying calibration;
- Characteristics of the camera (e.g. resolution) and characteristics of the installation (height and angle);
- The complexity of the traffic scene;
- Environmental conditions (e.g. weather and lighting).

Performance measures of automated systems’ reliability can relate to a number of aspects of the data collection, for instance how many of the relevant situations are missed by the system, how accurate the measurements of behavioural and conflict indicators are, how well the algorithms can deal with changing weather and light conditions, etc. Little research has been done into measuring and comparing the performance of automated video systems for traffic-related applications in varied circumstances. Exceptions are recent studies by *Morse et al. (2016)* in which a number of different conditions are formally tested and *Laureshyn et al. (2013)* that looks into accuracy of estimated camera calibration parameters and their effect of the final measurements from video. *Saunier et al. (2014)* suggested a standard procedure to compare performance of different video analysis system based on the same input data and comparison of the output with manually produced ground truth.

Since the occurrence of safety-critical situations is also a random process (though with much higher frequency compared to accidents), it is important that the observations are done during a sufficiently long time so that stable estimates of their frequency can be obtained. The only systematic discussion on this subject we could find comes from *Hauer (1978)* who used data from several earlier traffic conflict studies in order to investigate the variation of observed conflicts over time. It was noted that the variation seemed to depend on the conflict definition. Conflicts with low frequency (less than 20 per day) were best described with the negative-binomial distribution while for high-

frequency conflicts (more than 20 per day) the normal distribution was sufficient. *Hauer (1978)* also illustrated how accuracy of the estimated “expected conflict rate” improves with the extension of the observation period which can be used for decisions on how long observation period “is long enough”. Even though the traffic conflict techniques have evolved a lot since the time this work was written and the suggested “rules of the thumb” are, most probably, no longer valid, the approach itself to the estimation of the observation period length appears to be very sound and still can be used for other surrogate safety indicators and estimation of frequency of events based on these indicators.

3. Methodology

We used a scoping review to retrieve and structure the information that is available on the topic of traffic conflict observations, and to identify the critical gaps in existing knowledge and the required steps forward in the field. In general, the aim of scoping reviews is to rapidly map the key concepts underpinning a research area and the main sources and types of evidence available (*Wilson et al., 2012, N. Fulop, P. Allen, A. Clarke, N. BlackMays et al., 2001*). A major advantage of scoping reviews is that they can provide a broad map of evidence that can be used by many, and for applications beyond the authors originally intended purpose (*Armstrong et al., 2011*). They are particularly useful in reducing duplication of efforts, and in guiding future research (*van Wee & Banister, 2016, Armstrong et al., 2011*), which are considered two critical points of attention in current surrogate safety research.

A systematic and transparent protocol was set up to find relevant studies and to extract knowledge from them. The main method for location of literature for this review was by searching the following databases available online: **ScienceDirect**, **TRID**, **Web of Science**, **Engineering Village** and **Scopus**. Applied search terms are shown in Table 1. Each database has been searched using the keywords presented in the table below. Additionally, the library of the Transport and Roads Department of Lund University was manually searched, since this library contains a lot of old reports and dissertations that are highly relevant to this paper, but that are usually not available in online repositories. To a limited degree, we have allowed for snowballing when some references found in papers turned out to be missing and seemed to be of high importance, and in case personal knowledge of the involved experts identified critical papers that were missed by this systematic procedure. For practical reasons, the studies have to be written in English, Swedish or Dutch. No time limit has been imposed on the search, all publications up until the end of 2015 are included.

Table 1. Search terms used for searching the online databases.

Keyword 1	Boolean operators	Keyword 2
Traffic conflict		
Conflict Technique		
Surrogate safety		
Safety critical event		
Indirect safety	AND	Traffic
Near-accident	AND	Traffic
Near-miss	AND	Traffic safety

All references were downloaded to EndNote reference management program, and duplicates were removed using the automatic procedure embedded in it. This resulted in 2445 hits. The screening of the papers to decide about inclusion or exclusion was performed by two researchers. In a first screening, the title and (if necessary) the abstract were used to judge whether a study should be included or not. In case this was insufficient to come to a conclusion, the full paper was checked.

The focus of the report is limited to road traffic conflicts that are observed from real traffic using on-site static cameras, sensors or observers. The following exclusion criteria were defined:

- Not about road traffic conflicts (e.g. air traffic conflicts);
- Not about on-site observations (e.g. naturalistic driving studies, microsimulation studies);
- Traffic conflicts are only supportive (e.g. traffic conflicts are used to validate validated driving simulator or microsimulation data, but are not analysed and reported on their own);
- Language (e.g. full paper in Chinese with English abstract);
- Duplicate (e.g. two publications about one study or other duplicates not automatically removed by EndNote).

Figure 5 shows the review decision process. It should be noted that a small number of methodological studies (4) were included despite them not focusing on on-site observation since they were relevant for this review.

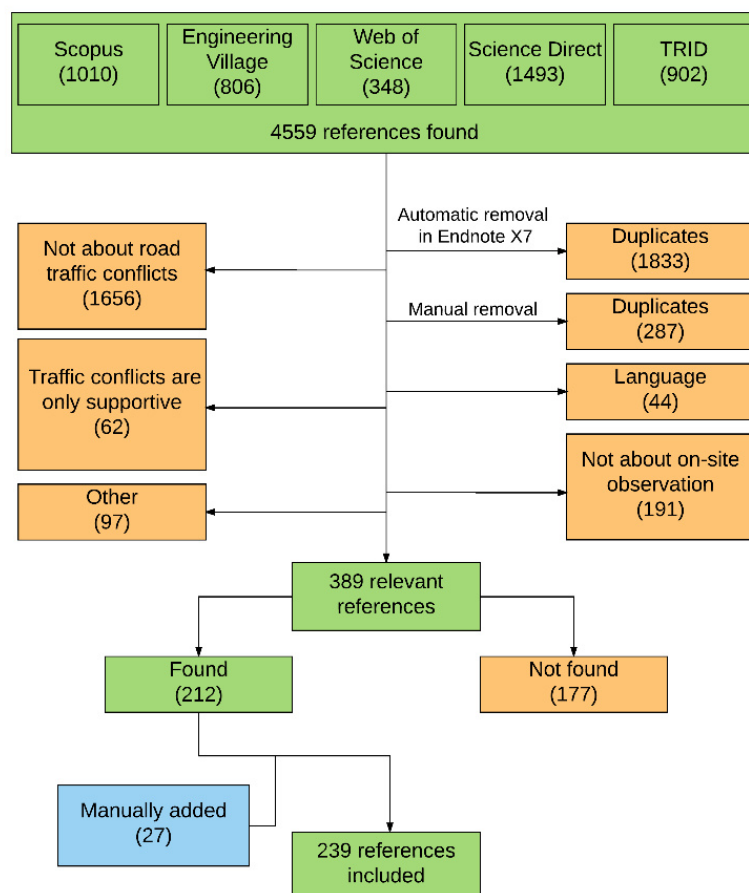


Figure 5. Review decision process flowchart.

Full papers were extracted automatically through the EndNote. Full papers that could not be automatically extracted were searched for manually through Google and Google Scholar search engines (e.g. full papers from ResearchGate, self-archiving databases of universities). Given enough time and resources, additional efforts to retrieve the remaining papers could have been made, such as contacting the authors or buying access to additional databases or individual papers, but this was not feasible within the frame of this review. It seems that the majority of studies that could not be retrieved are mostly small-scale studies, published as conference proceedings only or non-ISI ranked journals. While this is a limitation of the study, the added value of these papers to the review is considered limited. While this means that the report cannot provide a complete

representation of all available literature, we believe that the analysed dataset still provides an extensive and robust view of the field. Because of this limitation in available time and resources, a total of 177 references that, based on the title and abstract, might be relevant to the study could not be further analysed and included. It should be stressed that not all of these papers necessarily relate to the subject at hand, so the true number of appropriate papers that is missed because of this will be lower. A total of **239** references were included in the review.

Information about each publication was stored using a predefined codebook. The codebook included information about the aim of the study, the data collection, traffic conflict indicators and techniques that were used, threshold values that were used to distinguish between serious and non-serious events, data analysis techniques, possible links to other types of data and methodological aspects that were dealt with in the paper. The coding work was distributed among the same two researchers who decided on the in- or exclusion of papers. They were also the ones who designed the codebook (with additional input from the rest of the research team) and therefore have a very good understanding of the exact meaning of each element. Additionally, a guideline book was designed that accompanied the codebook that could serve as a lead in case of any doubts. The codebook was pre-tested on a small sample of the data and showed a good inter-coder agreement. Regular discussions and consistency checks between both researchers took place during the entire coding process.

Some papers were indicated as “must-read” papers. Those were read by the other authors involved in this report, and form the basis for the interpretive sections on the current status and future challenges. Reasons to indicate a paper as a must-read could be that they are major methodological contributions to the field (e.g. validation against accident data), exceptionally large size of the study, or presenting inspiring or promising theories about traffic conflicts and how to study them.

4. Results

This chapter aims to provide a comprehensive overview of the different indicators and conflict techniques that have been used or suggested in literature. To structure the covered material, the chapter is divided into five parts.

The overview provides a summary of the included literature focusing on the age of the publications, the split between indicators and techniques and the focus of the publications. The second section presents the different surrogate safety indicators used in the literature. All indicators are grouped after the main idea behind the indicator into several “families”. The most commonly used indicators in each family is given a broad presentation and other related indicators will be named shortly and finally, all found validation studies that use these indicators are presented. The third section focuses on the different traffic conflict techniques and their validation studies. The fourth section presents the indicators and techniques that have been used in studies that either focused on or included VRUs. The final section presents details regarding the surrogate safety study design, including observation length, number of sites, observation method, etc.

Describing the literature frequency, systematic distinction is made between the publications before and after 2005 which allows showing the share of the last-decade development. Out of the total 239 publications that are covered, 153 were published after 2005. A distinction is also made between applied and methodological studies. Applied studies use surrogate indicators or traffic conflict techniques to answer a particular question about road safety methodological ones study the indicators or conflict techniques themselves (e.g. their validity or theoretical grounds).

4.1. Overview of the literature

Figure 6 presents the distribution of the found publications over the time. The graph includes both the publications that are actually included in the study and those identified as potentially relevant, but for which no full text could be retrieved. The first number describes the publications that have actually been read and analysed while the second gives an idea about the “true” number of publications.

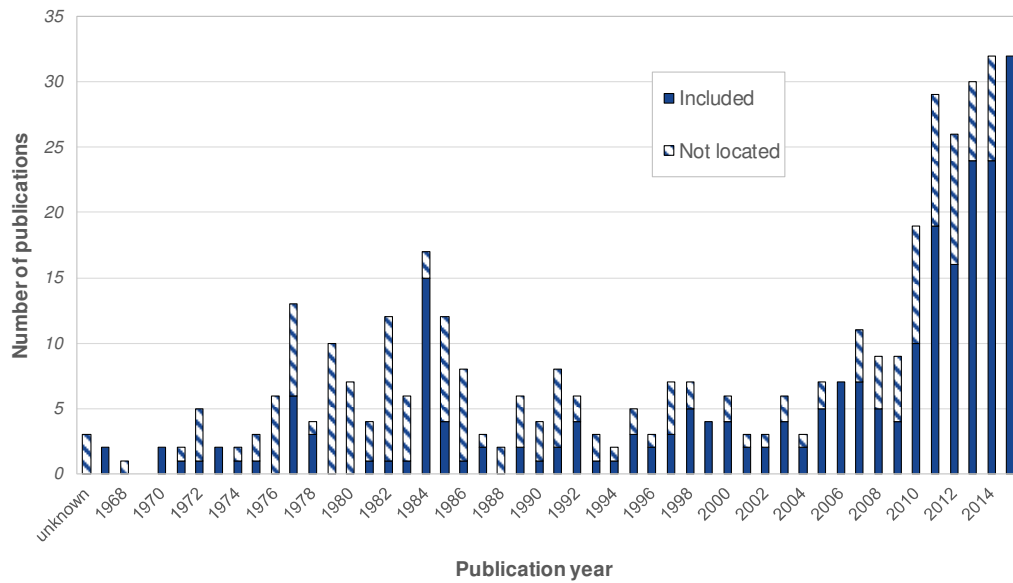


Figure 6. Distribution of publications over time.

Although the first documented application of traffic conflict observations was done by *Perkins & Harris (1967)*, the usefulness of traffic conflicts advocated by *Forbes (1957)* a decade before. The period of late 1970s and 1980s when many national traffic conflict techniques were introduced and tested is reflected by the peak in publication numbers. One can also observe a steep increase of publications started from 2010 which is, probably, a reflection of the recent development in the area of technical tools that allow to collect surrogate safety data in an efficient way.

There is a relatively high number of publications before the early 1990s that could not be found. This is due to the fact that many of the old publications are available in paper format only and thus hard to reach.

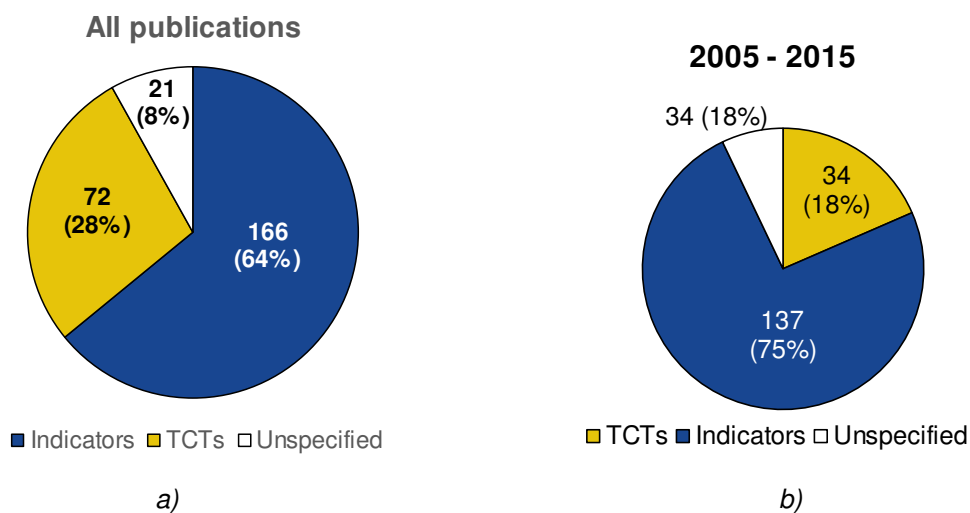


Figure 7. Split between indicators and traffic conflict techniques (TCTs) for all included (a) and last-decade (b) publications.

Figure 7 shows the split between use of individual surrogate safety indicators and traffic conflict techniques. The small “unspecified” category refers to publications in which the used definitions are unclear.

Figure 8 shows that most of the applied studies are published after 2005 while the methodological studies are more evenly spread over the time.

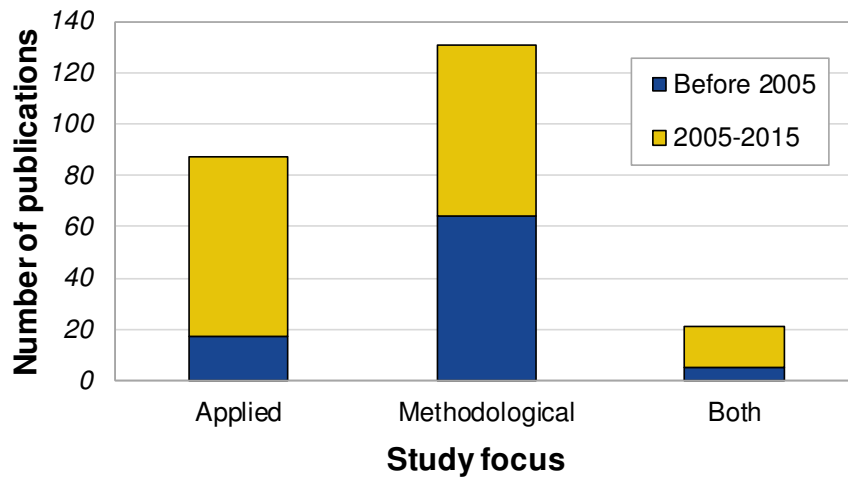


Figure 8. Distribution of study focus.

4.2. Surrogate safety indicators

4.2.1. Current practice of indicator use

Based on the basic idea that an indicator is based on, we grouped the indicators found in literature into several “families”. Each family is described closer in the following sections. The frequency of usage for the different indicator families in literature is presented in the Figure 9.

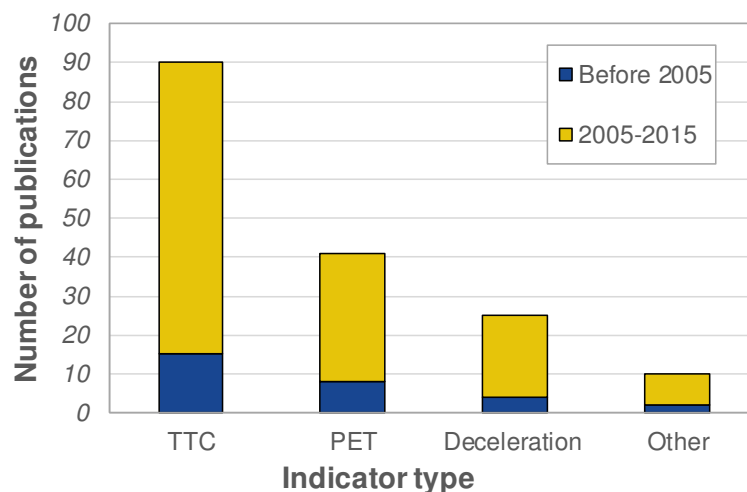


Figure 9. Usage of surrogate safety indicators.

The most common indicator family is TTC, followed by PET and deceleration. The most common single indicators are TTC_{min} (60 publications, 51 between 2005 and 2015) and PET (31 publications, 25 between 2005 and 2015). The category “unspecified” includes publications with unclear descriptions of the used indicator/technique. It should be noted that some publications apply more than one type of indicator, such publications are counted multiple times.

4.2.2. Collision course and motion prediction problem

Collision course is a key concept of many surrogate safety indicators and traffic conflict techniques. Collision course is a pre-condition for a collision – without being on a collision course at least at the very last phase, collision is not possible.

The basic idea of the collision course is that the two road users will collide if they continue “as is”. This requires, however, more precise instructions on how the future motion is to be predicted.

The earlier definitions involved travelling on “present course and at present rates” (*van der Horst, 1990, Hydén, 1987, Hayward, 1971*). However, continuing strictly with the same speed and direction is a quite unlikely scenario to actually occur. A more general, but also more realistic interpretation could be the planned travel given that the road users are unaware of each other (fail to detect the danger in time). This seems also to be the practical interpretation that has been used by the conflict observers in the field.

The generalized definition of a collision course is a situation in which there is at least one (physically) possible trajectory per road user that could lead to their collision at a future instant. In general, there are actually several such trajectories (with their respective probabilities).

Once the various trajectories that may lead two road users (or a road user and a static obstacle) are considered, it is a matter of finding a model that can predict road user trajectories from their past positions at any instant. After reviewing related work, in particular in the field of robotics, two families of motion prediction methods are proposed by *Mohamed & Saunier (2013)*.

Context-free kinematic methods. These methods do not take into account the situational context and include the traditional method of motion prediction at constant velocity (with only one predicted trajectory) and variants with constant angular velocity (assuming that the driver “freezes” with the steering wheel) or constant acceleration (driver’s foot “freezes” on a gas pedal). The projected trajectories with constant angular velocity often took strange shapes and went out of the road (*van der Horst, 1990*) while prediction at constant velocity seems to be appropriate for simple conditions like car-following (*Kiefer et al., 2005*) but not for more complex situations (or at least not any better than constant speed assumption - *van der Horst, 1990*). *Mohamed & Saunier (2013)* proposed probabilistic variants such as “normal adaptation” which adds some random noise to represent small, unconscious adjustments made by road users while they travel mostly straight.

Empirical methods. These methods are based on the observed movements of road users, as observed in general, for various road configurations or at a particular site. *Laureshyn et al. (2010)* suggested to use the actual trajectories during the interaction as approximations for the initially planned trajectories. This approach is questionable in case the road user adjusts the path in order to mitigate the dangerousness of the situation (e.g. taking a larger radius of a turn to win extra time). It was found to be particularly problematic in case of cyclists (and probably pedestrians, too) who often use “slalom technique” to avoid motor traffic (*Laureshyn et al., 2016*). Empirical methods have been experimented in particular in the form of motion patterns learnt from observed trajectories in (*St-Aubin et al., 2015b, Saunier et al., 2010, Saunier et al., 2007*). Motion patterns are representations or models of the road user trajectories at a

given road site. Two main models have been applied to estimate a collision course and compute indicators such as TTC:

- prototype trajectories where clusters of trajectories are represented by an observed trajectory that can be used to predict road user trajectories (*St-Aubin et al., 2015a, Saunier et al., 2010, Saunier & Sayed, 2008, Saunier et al., 2007*)
- discretized spatial distributions where the probabilities of each road user to reach a discretized region of the site from another, possibility based on other initial conditions such as initial speed, at a future time step are estimated from observations (*St-Aubin et al., 2015a*).

Using these motion prediction methods, especially the ones based on motion patterns, results in a more robust estimation of indicators that depend on motion prediction such as TTC as will be explained in the next section. Whether a collision course exists (i.e. TTC can be computed) is not affected anymore by small variations in initial speed or directions when predicting motion at constant velocity. Hence, TTC is measured for longer periods of time, which reflects the evolution of the interaction over time, along with other continuous indicators such as simple directly observed indicators characterizing the relationship between the two interacting road users such as distance, relative speed, etc. (*Saunier et al., 2010*).

If not specified in the remainder of this report, the collision course and indicators that depend on it are estimated based on the simplest and most common approach of motion prediction at constant velocity.

4.2.3. Time-to-Collision family

Definition of TTC

Time-to-Collision (TTC) is one of the most common indicators to describe the severity of a traffic event. *Hayward (1971)* defined TTC¹ as “the time until a collision between the vehicles would occur if they continued on their present course at their present rates”.

TTC possesses some distinct properties:

- TTC cannot be measured directly but is calculated based on future motion prediction.
- It can be calculated only as long as the road users are on a collision course.
- TTC is a continuous indicator and may be calculated for any time instance during the collision course.

The theoretical TTC curves are shown in Figure 10. According to *Hayward (1971)*, the curve starts when the road users get onto the collision course, reaches some minimum value (zero in case of a collision) and then “jumps” into infinity when the collision course seizes. *Laureshyn (2010)* showed that depending on the conditions of the evasive action, TTC can seize without the final “jump”.

¹ The original name proposed by *Hayward (1971)* was Time-Measured-to-Collision (TMTTC).

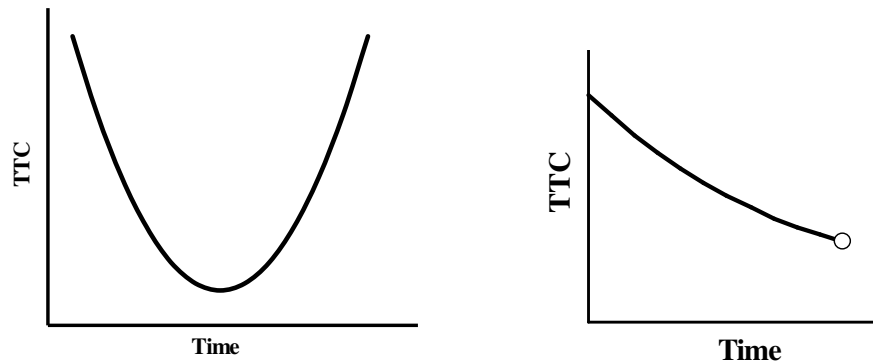


Figure 10. Theoretical TTC curve: a) original curve discussed by *Hayward (1971)*; b) a special case when the collision course seizes without TTC making a jump into infinity (*Laureshyn, 2010*).

Momentarily TTC values

Since an encounter between two road users is a continuous process, a question raises which time instance describes the severity of the event in a most representative way. At least two points on the TTC curve are associated with important moments in the interaction development:

- **TTC_{min}**, the lowest TTC value under the interaction (*van der Horst, 1990, Hayward, 1971*).
- **Time-to-Accident (TA)**, TTC value at the moment the first evasive action is taken by one of the road users (*Hydén, 1987*).

Figure 11 illustrates TA and TTC_{min} in situation of a car approaching a stationary object (*van der Horst, 1990*). TTC at the onset of braking (TTC_{br}, same as TA), represents the available manoeuvring space at the moment an evasive action starts (point A). Point B is the minimum TTC (TTC_{min}) reached during the approach process.

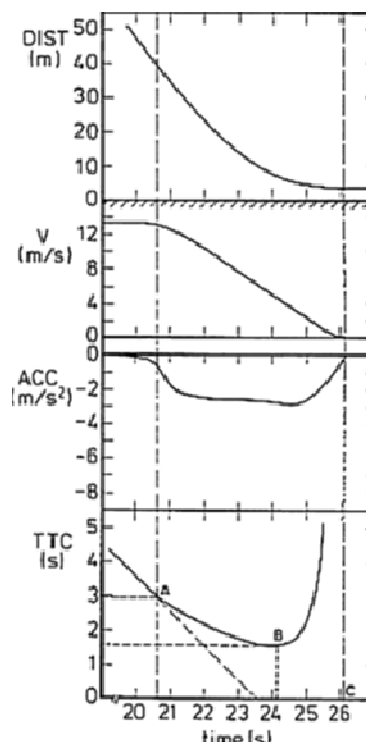


Figure 11. Time history of braking by a car approaching a stationary object: DIST = distance to object, V = velocity, ACC = acceleration and TTC = Time-To-Collision based on constant speed

and heading angle. Point A indicates the start of braking (TA) and point B indicates TTC_{min} (van der Horst, 1990).

Both indicators have pros and cons. TTC_{min} characterises how close in time the road users came to each other, which is a measure of the nearness to the actual collision and in a way how successful the evasive action was. TA, on the other hand, represents a critical moment when a hazard has been detected and the evasive action starts. The advantage of TA is that it can be successfully applied to distinguish severity of both collision and no-collision events, while TTC_{min} in case of a collision always equals zero. From the perspective of the theoretical definition of a traffic conflict always preceding a collision (see model in Figure 3a), TA is also more attractive. The severity measure suggested by TA is independent of the outcome of the evasive actions (collision or not), while non-zero TTC_{min} values will always indicate that the collision was avoided. On the other hand, the exact moment of the evasive action start is not always easy to detect, both when it comes to human observers and automated tools (Laureshyn, 2010).

Other related indicators

Some other modifications of TTC have been suggested.

Similar to TTC measuring the time of arrival to the collision point, it is possible to measure the arrival time to a certain position on a road. Várhelyi (1998) introduced **Time-to-Zebra** (T_{zeb}) which is a time remaining for a car to reach the pedestrian crossing. For driver warning system about lane departure, van Winsum et al. (2000) calculated **Time-to-Lane crossing** (TTL) which is a time remaining for a vehicle to reach the border of the current traffic lane (calculated based on the lateral movement of the vehicle within the lane).

Since increased severity of a traffic situation is associated with lower TTC values, Chin et al. (1992) suggested to use the **reciprocal of TTC**, i.e. $1/TTC$.

To take into consideration not only the momentarily TTC value, but to some degree also the duration of the safety-critical situation, Minderhoud & Bovy (2001), proposed two enhanced TTC-based indicators. The first one, **Time Exposed TTC** (TET) is the time during an encounter when the TTC is below a certain threshold value, TTC^* (Figure 12). TET reflects the duration of the most critical part of an encounter when TTC is low. The second indicator is called **Time Integrated TTC** (TIT) and calculated in the following way:

$$TIT = \int [TTC^* - TTC(t)] dt, \quad 0 \leq TTC(t) \leq TTC^*$$

TIT is the area between the threshold level TTC^* and the TTC curve when it goes below the threshold, thus reflecting both the lengths of the time with low TTC and the extent to which the TTC sinks below the threshold.

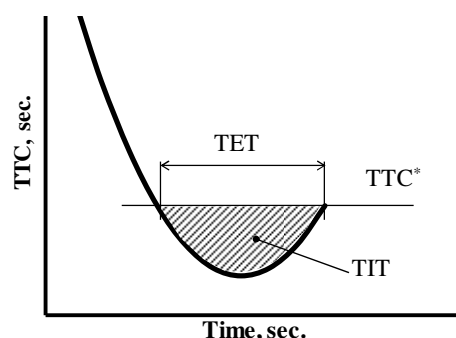


Figure 12. The definition of TET and TIT (adopted from Minderhoud & Bovy, 2001).

In many cases, calculations based on more accurate and frequent measurements of current speed and position (for example, extracted from video frame by frame using special software) reveal that very minor speed variations affect whether at a certain time instant two road users are on a collision course or not if using motion prediction at constant speed and direction. Strictly speaking, without a collision course TTC cannot be calculated, but this does not necessarily mean that the dangerousness of the situation changes dramatically – especially if in the next time instant, the new speed value makes it a collision course again. This is indirectly supported by *Svensson (1998)* who observed in situations with very small time margins road users behaving as if it was a collision course and taking evasive actions to mitigate the situation. To be able to classify such situations, too, *Svensson (1998)* “extended” the dimensions of the vehicles so that TTC calculations were possible.

Laureshyn et al. (2010) suggest to use the expected arrival time of the second (later) road user to the potential collision point as an extended version of the TTC-concept. This indicator (called T_2) equals to TTC in case of a collision course, and allows for smooth transition between collision course and non-collision course time instances providing a continuous curve over the entire interaction. Calculation of T_2 is no longer meaningful after the first road user left the collision point (collision is no longer possible). This natural limitation makes it quite convenient to use the minimal value of T_2 as a severity measure (T_2^{\min}): *i*) situations with large time margin will never reach low values and will not be selected as risky; *ii*) situations with small margins will have low values and considered risky which makes sense as a small speed deviation might have resulted in a collision course and ultimately in a collision; *iii*) in situations with collision course, T_2^{\min} equals TTC_{\min} . Another advantage of T_2 is that its latest value is very close¹ to PET indicator described in a later section making, again, a smooth connection between collision course and non-collision course indicators.

The lack of robustness of traditional collision course and TTC computation is caused by the restrictive and unrealistic assumption that road users continue their movement at constant speed and direction which has been discussed in the previous section. When the future motion of road users is represented in a form of a probabilistic model with multiple trajectories and corresponding probabilities (*Saunier et al., 2010*), several collision points with corresponding TTC values and their probabilities are possible (see Figure 13).

¹ Strictly speaking, T_2 at that moment equals to “expected PET” (predicted based on the current speed and path), but due to very little time left for any speed changes, the final PET value does not usually differ substantially.

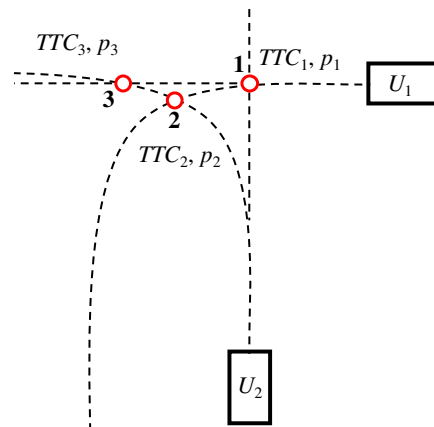


Figure 13. Multiple collision points calculated using a probabilistic model (based on *Saunier et al., 2010*).

In its generalised form, TTC then can be calculated at time instance t for two road users U_1 and U_2 as:

$$TTC(U_1, U_2, t) = \frac{\sum_{i=1}^n (p_i \cdot TTC_i)}{\sum_{i=1}^n p_i},$$

where p_i is the probability of the road users to collide at collision point i and TTC_i is the predicted time at which they will reach it.

TTC is defined only if there is at least one collision point, which corresponds to the traditional computation of TTC with only one predicted path per road user. In the general case, this yields the expected TTC over all possible collision points. Under the assumption that the road users are not aware of each other, the probability of reaching a collision point i is the product of the probabilities of each road user following (independently) the predicted trajectory that leads to it.

Instead of looking at single TTC values (momentarily or integral like TIT or TET), *Saunier & Mohamed (2014)* suggested examining the entire TTC curve and use the curve shape to build clusters and distinguish safety critical events. The initial tests of various algorithms for clustering time series data (*Saunier & Mohamed, 2014, Laureshyn et al., 2009*) show good performance, however the practical usability of this approach still requires further testing and validation on a larger data sets.

Calculation of TTC

Calculation of TTC for “ideal” conditions such as approach at a right angle or following situation is relatively straightforward. However, in a general case two vehicles can approach each other at any angle and, moreover, for the same angle different collision types are possible (Figure 14).

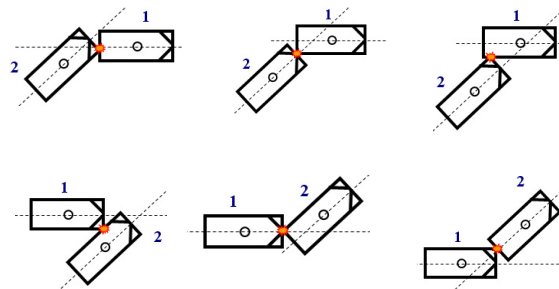


Figure 14. Possible collision types for the same approaching angle (adopted from *van der Horst, 1990*)

Assuming road users to have a rectangular form, *Laureshyn et al. (2010)* notes that the contact point will always involve a side of one rectangle and a corner of another and provides a calculation algorithm which considers all possible combinations of sides and corners. *Sobhani et al. (2012)* extend this algorithm for non-linear trajectories. Similar calculations are also provided in *Lua et al. (2012)*.

Ward et al. (2015) describe calculations of TTC based on the closest distance between two vehicles and the “closer rate” (speed at which this distance decreases). Since such calculations are possible even if the vehicles are not on a collision course (e.g. if they move towards each other but in separate lanes), it is suggested to use a test of “looming points” that ensures the presence of a collision course. The basic idea of the test is best illustrated by a hint used in the aircraft pilot training – an indication of a collision course with another approaching aircraft is that its relative position in the windscreen remains the same over time. This approach is particularly useful when the TTC measurements are done from inside a car and not in a co-ordinate system related to the road environment.

Validation of TTC – relation to accidents

Despite the very wide acceptance and use of TTC as a surrogate safety measure, the validation studies relating TTC to the actual accidents are extremely few. It is also important to note that it is only one study (*Sacchi et al., 2013*) that used fully automated measurements of TTC, in all other cases TTC was estimated by human observers and might be affected by their subjective perception of the situations (which is further discussed in Chapter 4.2.8).

Hydén (1977):

Conflict observations (an early version of the Swedish technique) were performed at 115 intersections in two large Swedish cities (Stockholm and Malmö). A severe conflict was defined as having Time-to-Accident (TA) less than 1.5 seconds. Conflicts were grouped in different “cells” based on the type of manoeuvre and road users involved (see Table 2).

Poisson distribution was assumed for both conflicts and accidents. The expected values (and 90% confidence intervals) for the ratio between number of accidents and conflicts per time unit for each cell are given in Figure 15.

Table 2. Conflict categories (cells) introduced in *Hydén (1977)*.

	car vs. car	car vs. cyclist/ pedestrian
Situations with low speeds/turning vehicles	Cell 1	Cell 3
Situations with high speeds/at least one straight going vehicle	Cell 2	Cell 4

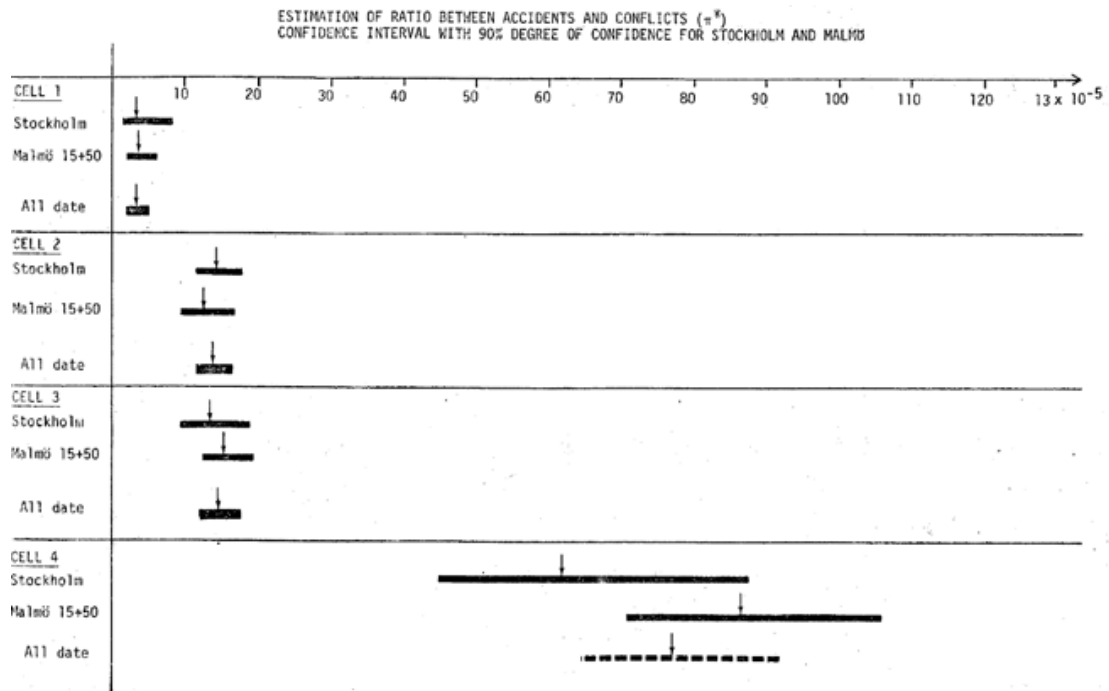


Figure 15. The ratio between number of accidents and conflicts (TA < 1.5 sec.) per time unit (Hydén, 1977).

Migletz et al. (1985):

Conflict studies (appears to be a modified version of the US conflict technique) were carried out at 46 signalised and non-signalised intersections in the greater Kansas City area. A severe conflict was defined as having Time-to-Accident (TA) less than 1.5 seconds. 12 basic conflict types were considered, a distinction of intersections by traffic volume (High, Medium and Low flow categories) was also done.

Two different validation approaches were investigated. Firstly, different correlation tests were used to calculate the relation between conflicts and accidents. Some weak correlations were found only. The second approach used the conflicts to estimate the expected number of accidents that was compared the expected number of accidents calculated based on accident data. Both of these values were then compared to the actual number of accidents for a given year. The value based on accident data was closer to the actual accident number but with a higher variance compared to the value based on conflicts. The report concludes that there seems to be no observable advantage of using accident data compared to conflict data. Potentially, combination of accident and conflict data could produce even better estimates (with lower variance) of the expected number of accidents compared to using only one data source.

Lord (1996):

The study tested several conflict definitions during observations of conflicts between left-turning traffic and crossing pedestrians at 8 signalized intersections in Ontario, Canada. One of the definitions included Time-to-Accident (TA) threshold of 1.5 seconds. The conflict data was compared to expected number of accidents obtained from a safety performance function for the same type of manoeuvre (Quaye et al., 1993). A number of statistical tests (Linear regression, spearman ranking, F-test) were used to investigate the relation between the number of observed conflicts and the expected number of accidents. The study concludes that conflicts with TA < 1.5 sec. had a significant correlation to the expected number of accidents.

Sacchi et al. (2013):

Autey et al. (2012) evaluated a new design for channelized right-turn lanes at 3 intersections in Penticton, British Columbia, using a fully automatic video analysis software that counted events with Time-to-Collision (TTC) below 3 seconds. Before-after design was used with about 16 hours of video processed for each period. The results showed a considerable reduction in the frequency and severity of traffic conflicts of different types.

Some years later *Sacchi et al. (2013)* made another before-after study for the same 3 intersections, this time based on accident data (Fully-Bayesian approach with 16 control sites). This study also showed a considerable improvement of safety levels after introduction of the new design of the right-turns. Very clear similarities between reductions in conflicts and collisions (in general and location-specific) were found.

El-Basyouny & Sayed (2013):

In this study, conflict and accident data from 51 signalised intersections in British Columbia were used to build a negative binomial regression model relating average hourly conflict rate with the number of accidents at a site. The conflicts were registered in an earlier study (*Sayed & Zein, 1999*) using the Canadian conflict technique. In the model, however, only conflicts with TTC below 1.5 seconds were used.

The model showed a very robust relation between conflicts and accidents. It performed better than a simple linear regression tested in *Sayed & Zein (1999)* (though the models are not directly comparable as different conflict selection criteria were used). This might, however, indicate that the assumption of the linear relation between conflicts and accidents is too simplistic and cannot be described by a single conversion factor.

Thresholds

It was only possible to make a meaningful summary of the used threshold to distinguish between severe and non-severe events for TTC_{min} (see Figure 16). The threshold values of 1.5s, 2s and 3s are the most common (11, 7 and 9 publications). 19 out of 55 publications did not make use of any thresholds at all but rather analysed all events where TTC_{min} calculation was possible. Majority of the publications that Figure 16 is based on have a methodological focus.

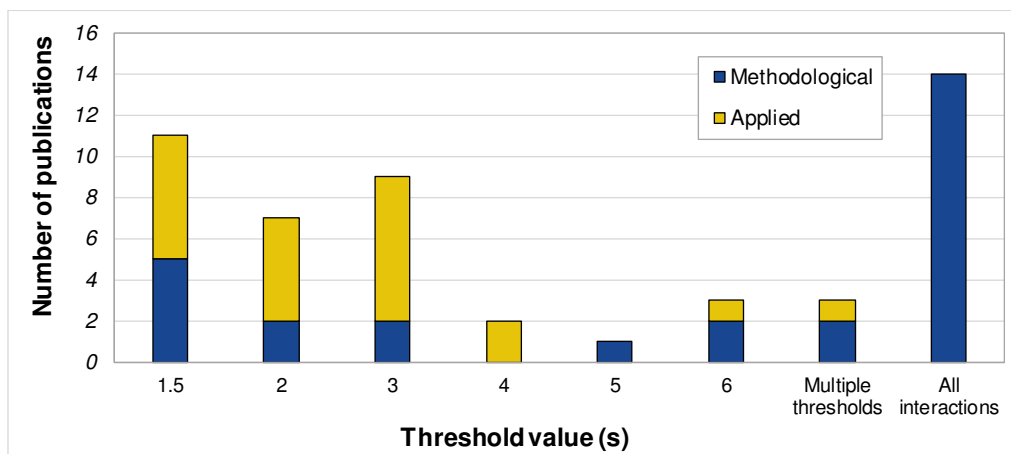


Figure 16. Threshold values for TTC_{min} .

4.2.4. Post-Encroachment Time family

Definition of PET

The Post Encroachment Time (PET), initially introduced by *Allen et al. (1978)*, is a measure that defines “near misses”, a situation in which a collision was avoided but with a relatively small margin. PET is calculated as the time between the moment that the first road-user leaves the path of the second and the moment that the second reaches the path of the first (see Figure 17), i.e. PET indicates the extent to which they missed each other.

Together with TTC, PET is the most frequently applied surrogate safety indicator. Whereas TTC requires a collision course, PET reflects the spatial and temporal proximity of road users. *Allen et al. (1978)* argued that the TTC and the TA are incomplete measures to define traffic conflicts since these measures become infinite in case there is no collision course, even if a collision is only avoided by a fraction of a second without an evasive action. Based on 25 observed collisions, *Allen et al. (1978)* stated that in part of these collisions an evasive action had not been present or could not be easily observed. Moreover, it was noted that a collision is most of the time a result of several sequential events, requiring more than one traffic conflict parameter to describe it adequately. Among these measurements PET was introduced - a measurement of how nearly a collision has been avoided.

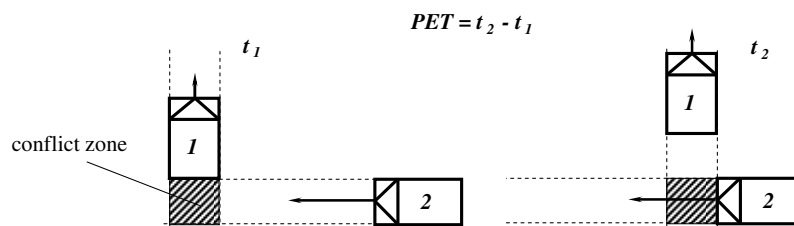


Figure 17. Definition of Post-Encroachment Time (PET).

Other related indicators

Gap Time (GT) is the time between the entries into the conflict spot of two vehicles, measured from the front bumper to the front bumper. It was originally described by *Allen et al. (1978)* in conjunction with PET and **Encroachment Time (ET)**. ET is the time that the first vehicle entering the conflict area infringes upon the predicted path of the second vehicle, measured from the rear bumper to the front bumper. GT and ET are both continuous indicators unlike PET, which can only be measured at one point in time and has one value. **Time Advantage (TAdv)**, originally proposed by *Hansson (1975)*, could be considered as an extension of the PET-concept (*Laureshyn et al., 2010*). TAdv is a predicted PET value provided that the road users continue with the same path and speed. According to *Laureshyn et al. (2010)*, TAdv (among other indicators) provides more insight in the continuous interaction between road-users over time and space compared to measuring a PET value, which can only be measured at one point in time.

Mohamed & Saunier (2013) proposed a variation of TAdv (called by the authors “predicted PET”, pPET) within the probabilistic framework described earlier (*Saunier et al., 2010*). The calculations are done considering the multiple trajectories and respective pPETs in a way similar to the generalised TTC calculation.

Time Headway is another indicator that may be used to estimate the severity of events in traffic, mainly for events involving vehicles moving in the same direction. Time

Headway is defined as the elapsed time between the front (resp. the rear) of the lead vehicle passing a point on the roadway and the front (resp. the rear) of the following vehicle passing the same point (*Vogel, 2003, Evans, 1991*). It was found that accident-involved drivers are more likely to follow with short headways (less than 1s) than accident-free drivers (*Evans & Wasielewski, 1983, 1982*). Alternatively, the measure can be expressed as a distance instead of a time period, and is labelled **Distance Headway**. In this case, the distance between the lead vehicle and the following vehicle at the same point in time is measured.

While the concept of Time Headway is quite well-known, its use seems to relate mostly to the field of traffic flow, while its use as an indicator for traffic conflict severity is more limited. *Vogel (2003)* compared the use of Time-to-Collision and Time Headway for car-following situations, and came to the conclusion that TTC is a more appropriate traffic conflict indicator. While low time headways indicate potentially dangerous situations, low TTC values indicate the actual occurrence of dangerous situations.

Validation of PET – relation to accidents

Cooper (1984):

Cooper (1984) describes two studies in which the relation between PET and the number of accidents were investigated. Cooper shows that PET is in general a better predictor of expected accidents than either past collision history or traffic volumes, even though the found relation was not very strong. It is argued that the most important improvement in the experimental design would be the inclusion of locations having higher overall accident frequencies.

Songchitruksa & Tarko (2006):

Songchitruksa & Tarko (2006) extracted PET values from video (8 hours per site) at 18 signalised intersections. Then the Extreme Value Theory methods were applied to estimate the frequency of the events in which PET reaches 0 (i.e. collisions). Results showed a relation between model estimates and accident data. However, due to the short observation period the variation in estimates was large. The authors estimated that 30-50 days of observations are necessary to obtain a reliable prediction.

Alhajyaseen (2015):

In this study the relationship between severe accidents and empirically estimated speeds and PET values (> 5 s) at conflict points (angle conflicts) were studied. Neither speed nor PET alone could provide a consistent safety assessment. Therefore, a Conflict Index (CI) was developed, that considered accident probability (based on PET measures) as well as severity (speed, weight and conflict angle). Results showed that based on CI a similar ranking of different intersections could be established as was done based on the number of severe angle collisions. The author noted, the index in its current form can only be used for comparative safety analysis, not to represent absolute accident frequencies at different sites.

Peesapati et al. (2013):

The aim of this study was to find an appropriate PET threshold value to select events that can be used for accident propensity assessment. The PET values for vehicles turning left and opposing through vehicles were collected from video at 18 intersections and related to accident statistics. The highest correlation between PET-selected events and reported accidents was found for the threshold of 1s ($R^2 = 0.63$). Even when the

selected events were presented as a proportion of the conflicting traffic volume, the best correlation with accidents was again for the threshold of 1s ($R^2 = 0.66$). Correlation of the conflicting traffic volume with accidents was found to be considerably lower ($R^2 = 0.46$) than for PET and PET combined with conflicting traffic volume.

Zheng et al. (2014a, 2014b):

These two studies explore application of the Extreme Value Theory methodology on PET-distributions, building on (and partly repeating) the work by *Tarko & Songchitruksa (2005)* and *Songchitruksa & Tarko (2006)*. The data was collected at 21 different sections of freeway in China (3 hours per site). Despite some improvements suggested, relatively high variation of the estimated accident number was still stated as a problem.

Thresholds

Figure 18 shows the frequency of the different PET thresholds used to distinguish serious and non-serious interactions. Even though the thresholds of 1s, 3s and 5 s have been used in some studies, the most common approach is to consider all interactions without any thresholds. In studies with multiple thresholds considered all threshold values were between 1 and 5 seconds.

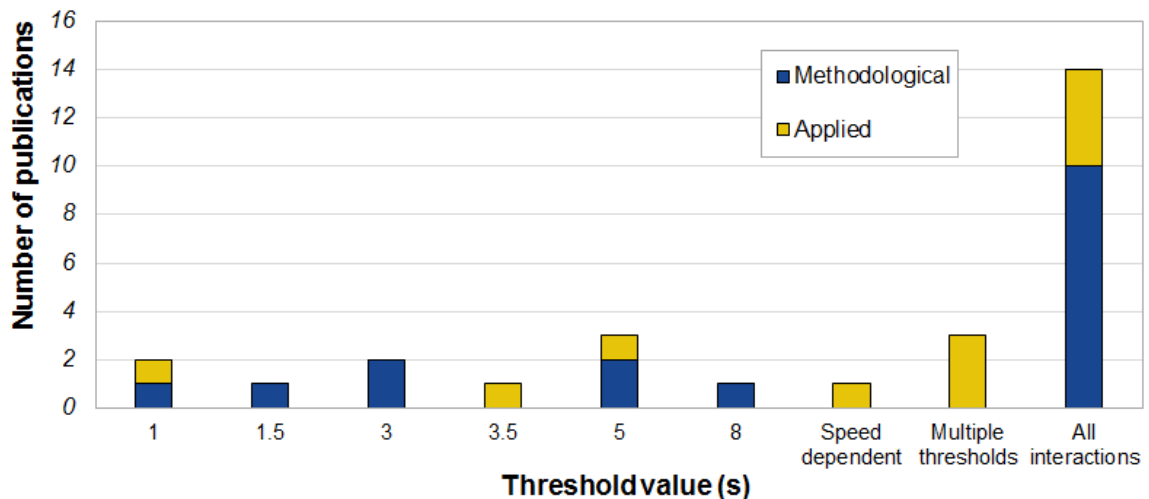


Figure 18. Threshold values for PET.

4.2.5. Deceleration-family

Deceleration is the most common evasive action taken by a vehicle to avoid a collision (*Hydén, 1987*). As a result, deceleration-related indicators can potentially be applied to a large variety of traffic situations, with or without a collision course.

One of the first researchers who used deceleration of vehicles in relation to safety was Edward Brill, who in 1972 developed a car-following model relating driver reaction time, temporal headway and deceleration response to rear-end collision frequency (*Brill, 1971*).

Several deceleration-based measures to describe the severity of a traffic situation have been suggested. **Deceleration Rate**, DR (or Initial Deceleration Rate) quantifies the magnitude of the deceleration action of a vehicle at the moment when it begins evasive braking manoeuvre. **Maximum Deceleration** (Max D) is the maximum deceleration rate of the vehicle observed during a conflict event (*Gettman & Head, 2003*). These

measures are based on observing the actual braking of a driver and might fail in case the driver misses the dangerous situation completely and does not react or applies an inappropriate braking force for the current situation (Zaki *et al.*, 2014). Experience from naturalistic driving studies suggests that there is a great variety in individual driving styles and what is a “critical” deceleration for one driver might be quite “normal” for another (Bagdadi & Várhelyi, 2011, af Wählberg, 2004, Nygård, 1999).

Hupfer (1997) suggested to describe the nearness to a collision through the minimal necessary deceleration for a driver to avoid the collision (to make a collision course-situation into PET-situation). This indicator was called **Deceleration-to-Safety Time** (DST) and it showed quite good performance applied on vehicle-pedestrian conflicts.

Similar to TTC, DST is a continuous indicator and can be calculated for each moment during a collision course. Hupfer (1997) also suggested the threshold values to grade the severity of the traffic conflict based on the DST (see Table 3). It is not explained, though, which DST value (e.g. at the start of the evasive action or the maximum value) is considered.

Table 3. Suggested definition of conflict levels based on DST (Hupfer, 1997).

DST _x ≤ 0 m/s ² :	Evasive action not necessary The predicted/reached safety time distance is bigger than x, an adaptation is not necessary.
DST ₀ < 1 m/s ² :	Adaptation necessary (interaction) The necessary evasive action is small, only a light adaptation has to be made.
DST ₀ < 2 m/s ² :	Reaction necessary (Conflict Level 1) The situation requires a noticeable deceleration of a road user. The situation is easy to control. There is time enough to consider other occurrences.
DST ₀ < 4 m/s ² :	Considerable reaction necessary (Conflict Level 2) The situation requires a considerable deceleration of one road user. The situation is controllable. Other occurrences can not be easy considered. (On a wet road this might be the maximum possible deceleration and is similar to an emergency braking: Conflict Level 4)
DST ₀ < 6 m/s ² :	heavy reaction necessary (Conflict Level 3) The situation requires a heavy reaction of at least one road user involved in the conflict. The reaction is hardly controllable. Other traffic occurrences can not be considered.
DST ₀ ≥ 6 m/s ² :	Emergency braking (Conflict Level 4) Uncontrollable reaction. Near miss accident.
Collision:	

Hakkert *et al.* (1977) suggested to use deceleration maps to identify deficient locations within the intersection area characterised by frequent and intensive decelerations.

Jerk is the derivative of acceleration with respect to time and reflects the “suddenness” of a braking. Studies of Nygård (1999) and Bagdadi & Várhelyi (2011) showed that safety critical events are easier to distinguish based on jerk rather than on acceleration profiles. A typical jerk profile of a sudden braking contains a negative and positive peaks (see Figure 19). In a study of Bagdadi & Várhelyi (2013), various indicators describing this shape (max negative, max positive, peak-to-peak values) were tested in a simulation of emergency braking by drivers in controlled conditions.

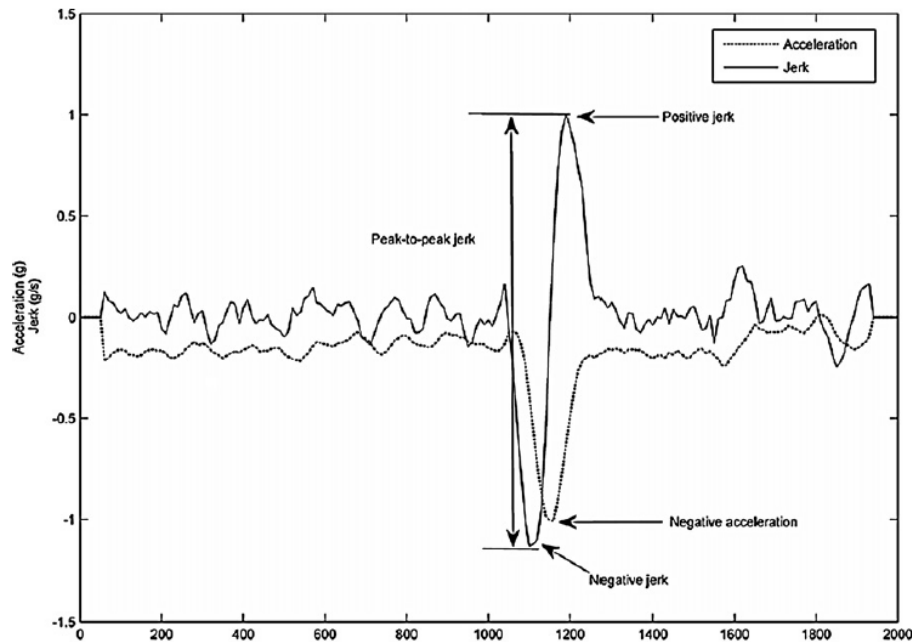


Figure 19. Jerk and deceleration profiles during a braking manoeuvre (*Bagdadi & Várhelyi, 2013*).

Other related indicators

Deceleration-to-Safety Time was primarily developed for right-angle collisions and uses PET-terminology in calculations. For car-following situations, **Deceleration Rate to Avoid Crash (DRAC)** was defined as the minimum deceleration rate required by the following vehicle to avoid a collision with the leading vehicle given unchanged speed of the leading vehicle (*Guido et al., 2013*). If a “wider” definition of PET for any approaching angle is used (*Laureshyn et al., 2010*), DST and DRAC are basically identical.

Kuang et al. (2015b) describes **Modified DRAC (MDRAC)** that also takes into account the driver’s perception and reaction time.

Oh et al. (2006) describes the concept of “safe stopping distance”, which is the minimum distance required for the following vehicle to safely reduce speed and avoid collision when the leading vehicle brakes with the maximum deceleration rate and stops. **Proportion of Stopping Distance (PSD)** is the ratio between the current distance and the minimal acceptable stopping distance (*Kuang et al., 2015b*).

Potential Collision Speed (PCS) is the expected impact speed in case of a sudden obstruction and limited time and distance to avoid it, for example in case of a chain collision in fog (*MacCarley et al., 2007*). The indicator takes into account both the reaction time and the braking during the time available before the collision.

Deceleration is often included as one of the components in complex indices to reflect collision risk on freeways (*Nasab et al., 2015, Tak et al., 2015, Xin-wei et al., 2013, Uno et al., 2002*). We omit detailed descriptions here as the conceptual forms of deceleration use in such calculations are covered by the indicators mentioned earlier.

Validation of deceleration-based indicators

While the list of possible deceleration-related indicators or complex indicators involving deceleration is quite long, the actual validation of these indicators in a view of relation to accidents is scarce. The discussion is much focused on whether deceleration can help

to distinguish critical events from non-critical. In this perspective, the deceleration itself seems to have quite limited value (*Zaki et al., 2014, Ismail et al., 2010a, af Wåhlberg, 2004, Nygård, 1999*), but it can be improved if combined with other indicators (*Ismail et al., 2010a*). Jerk seems to perform much better in distinguishing of safety critical events (*Bagdadi & Várhelyi, 2011, Nygård, 1999*).

4.2.6. Other indicators

This section will present other indicators that were found in the literature and do not fit into any of the previous categories.

Wei et al. (2014) defined a **critical conflict region** around a road user which changes depending on its speed, position and movement direction. Whenever two road users approach each other, the “pressure” increases and after a certain extent the road users are forced to take evasive action. The severity of a conflict was judged by a support vector machine that was trained using samples that had been judged severe or not severe by experts. A similar indicator, called **space occupancy index**, is proposed by *Ogawa (2007)*. This indicator changes depending on the speed of the road user and on the characteristics of the infrastructure. A conflict occurs whenever a road user enters space occupancy of another road user and an evasive action is taken to maintain the personal space.

Delta-V is a notation often used in physics to denote an object’s change of velocity because of an impact with another object. In the context of road accidents, Delta-V refers to the change of a velocity vector experienced by a road user during a collision. A large change in the magnitude and the direction of the speed of a road user over a short period of time implies extensive forces acting on the road user and has a strong effect on personal injuries. The finding that accidents involving higher Delta-V values tend to lead to higher injuries is supported by ample evidence (*Johnson & Gabler, 2012, Gabauer & Gabler, 2008, Evans, 1994*).

Shelby (2011) advocated for use of the expected Delta-V (what Delta-V would have been if the road users had collided at their current speeds) as an indicator for traffic conflict severity. *Laureshyn et al. (2017)* suggested **Extended Delta-V** indicator that takes into account the time remaining to take an evasive action and thus reduce the impact forces and also is possible to calculate for situations with or without a collision course. The authors suggest that such indicator better represents for the severity dimension as it combines both the nearness to a collision and the potential consequences. The indicator is particularly usable for situations with vulnerable road users involved since it is sensitive to the difference in mass of the two colliding objects which can be seen as a reflection of the vulnerability of the lightest one.

Kuang et al. (2015a) developed a new indicator called **Aggregated Crash Index (ACI)** based on the causal model presented by *Davis et al. (2011)*. The indicator is meant for car following scenarios and is made up of 4 conditions. The first condition determines the type of conflict which depends on the stopping time of the leading vehicle and the reaction time of the following vehicle. The second condition determines whether a collision occurs during the reaction time of the following vehicle. If no collision occurs directly, the third condition uses a new indicator called **Braking Rate to Accommodate a Disturbance (BRAD)** to calculate the severity of the evasive action and the fourth condition determines whether the evasive action is possible by comparing the BRAD to the Maximum Available Deceleration Rate. These conditions form a tree structure which

can lead to a successful evasive action or a collision. The ACI is then calculated as the accumulation of the collision probabilities of all possible outcomes.

Oh et al. (2009) based the definition of conflicts on signal violations in a signalized intersection. Events were divided into four severity levels depending on the consequences from the signal violation. *Oh et al. (2009)* uses what they call stopping distance to differentiate between the different levels.

4.2.7. Integration of different indicators to a single index

The general idea of safety indices is to integrate different indicators describing a traffic event into one single value. The rationale behind this approach is that many indicators are not sufficiently universal and cannot be applied to every event in traffic at any time. It is thus plausible that various conflict indicators represent partial images of the true severity of a traffic event (*Ismail et al., 2011*). It should be noted that many traffic conflict techniques can also be considered as a part of this category, for example the Swedish TCT based on TA and Conflict Speed, the DOCTOR technique based on TTC_{min} and PET and others that integrate objective indicators and subjective observer judgments. Because of their context and history, however, traffic conflict techniques are dealt with in more detail in section 4.3. Some other examples of this can be found in *Lu et al. (2012)* in which non-complete braking time and TTC are combined to calculate conflict severity and in *Wang & Stamatiadis (2014)* in which required braking rate, maximum available braking rate and TTC was used to create an Aggregate Crash Propensity Metric. A number of indicators that combine deceleration with other aspects such as radial acceleration (i.e. steering) or reaction time of the involved road users have also been suggested (*Nasab et al., 2015, Li et al., 2013, Oh et al., 2006, Uno et al., 2002, Balasha et al., 1979*). Further examples of this type of indicators that focus on VRUs are presented in section 4.4.

4.2.8. Subjective severity scores

A number of existing traffic conflict techniques rely to a large extent or at least complement the objective measures with some subjective score reflecting the severity of a conflict (*Kocárková, 2012, Brown, 1994, van der Horst & Kraay, 1986, Baguley, 1984*). Even though the reliability of subjective scores might be questioned, they seem to have some added value that otherwise is not reflected by the traditional objective indicators.

It appears that humans have some internal concept of “dangerousness” which is close to severity in its broad sense (i.e. nearness to severe consequences, see section 2.4 on severity definition). Several studies (*Kruysse & Wijlhuizen, 1992, Kruysse, 1991, Shinar, 1984, Lightburn & Howarth, 1979*) showed relatively good agreement in severity ranking of the video-recorded conflict situations between different observers. Individual observers seemed also to be very consistent in rating the same situation shown several times at different occasions. Even in real traffic conditions, the observers agree well at least in selecting the “severe” situations (*Hydén, 1987*). In a calibration study of several conflict techniques (*Grayson, 1984*) it was concluded that the main difference between the teams came in selection of situations, but, once selected, the severity scores agreed well.

This internal understanding of “dangerousness” seems to play important role in rating of traffic situations severity. *Shinar (1984)* instructed observers to base the severity rating

on objective criteria (time- and deceleration-related) and compared the scores with the actual measured values. Only a very weak relation between the subjective and objective scores was found while the observers agreed well with each other. In a study of *Kruysse & Wijlhuizen (1992)*, when observers (road safety professionals) were asked to explain which aspects (described in objective terms) were important for each conflict, again, a very weak relation between the presence of the named aspects and the final scores was found. Agreement among the observers was once again quite good.

To examine which clues the subjective “dangerousness” is related to, *Kruysse (1991)* used recorded conflict videos that were interrupted at different stages. The observers were asked to judge the severity based on the part of the sequence they had seen. It was found that the initial phase of the conflict development (start of the approaching and beginning of the evasive manoeuvre) have the most effect on the final severity, while seeing its final stage did not change the score. Also, if instead of judging “dangerousness” the observers were asked to judge the “suddenness”, the produced scores were quite similar (but not identical, which implies other aspects also playing role).

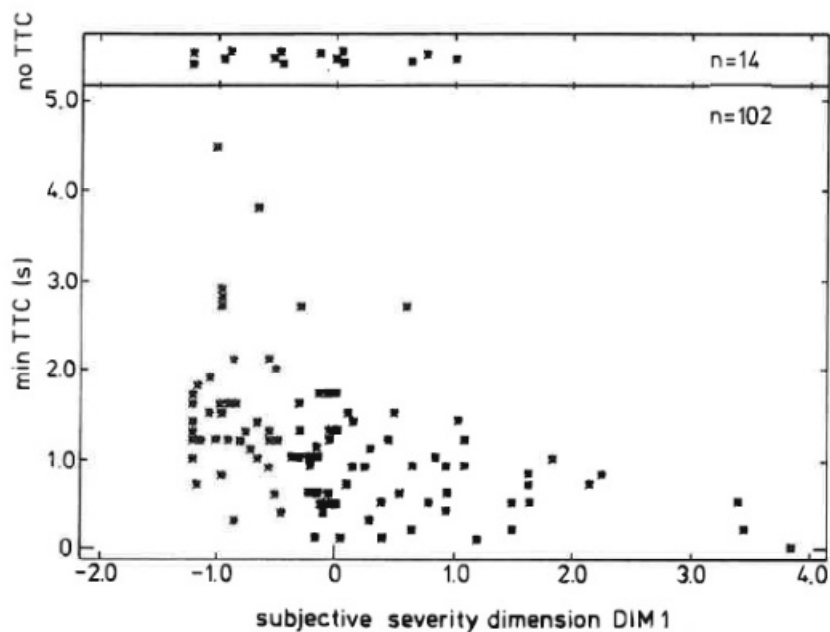


Figure 20. TTCmin values against a severity dimension. At the top the category of conflicts without a collision course (no TTC) is given (*van der Horst, 1984*).

A multivariate analysis of the traffic conflicts severity scores produced by the eight observer teams of the Malmö calibration study (*Grayson, 1984*) identified one common dimension that was interpreted as a “severity dimension”¹. When compared to a number of objective measures describing the conflicts (retrieved from video), TTC_{min} seemed to be the main contributor to this dimension, but not the only one (see Figure 20). Conflicts with a high severity score often had low TTC_{min} values, but not all conflicts with a low TTC_{min} were regarded as severe conflicts. Moreover, some conflicts with a relatively

¹ The scores were based on the definitions suggested by the different conflict techniques compared, but for many of them the definition included some subjective component.

high severity score did not have a collision course at all. For the greater part, these conflicts could be well described by the Post-Encroachment Time (*van der Horst, 1984*).

The relation of the subjective severity scores to the accident risk is hard to separate from the relation of the objective measures as in most available validation studies both were combined in rating of conflict severities. *Svensson (1992)* compared how well the expected number of accidents can be estimated using the serious conflicts classified strictly after the objective definition of Swedish TCT and based on subjective observers' judgments. The variation of the predicted values was lower for the conflicts selected based on subjective scores and the effect was even more clear for vulnerable road user accidents.

The enhancement of the traditional conflict techniques with objective and more accurate tools such as video analysis for measuring the safety-related indicators reveals that the subjective component played an important role in the earlier applications of the techniques (event if the formal definitions used were “objective”). For example, observers tend to set higher scores for situations involving vulnerable road users by “adjusting” speed and distance estimates (*Laureshyn et al., 2016*). Many of the situations experienced both by the road users and observers as having a collision course have in fact sufficient (though quite small) margins and would not result in a collision even if no evasive action was taken (*Laureshyn et al., 2016, Svensson, 1998*). The start of an evasive action (e.g. braking), clearly “seen” by an observer, is very difficult to precisely locate examining the speed profile (*Laureshyn, 2010*), which, again, raise a concern on how exactly the decision is taken by the observer.

4.3. Traffic conflict techniques

This section presents an overview of the different Traffic Conflict Techniques (TCTs). Various versions of TCTs were used in 75 of the 239 included publications and as seen in Figure 21 below, the US and the Swedish techniques appear to be used most frequently. The following section includes detailed descriptions of the Swedish TCT and the DOCTOR technique followed by slightly shorter descriptions of the US, British and Canadian techniques and ends by briefly mentioning the Finnish, French, Belgium, German and Czech techniques. The reason for providing more detailed descriptions of the Swedish TCT and DOCTOR is the extensive experience of the authors with these tools.

Most of the techniques have been developed over a long period of time which implies that not all publications regarding a specific TCT use exactly the same methodology. Publications that use modified versions of an original technique are presented together with the original technique. The following references may be used to find more detailed descriptions of the techniques:

- Swedish – *Hydén (1987)*
- Dutch (DOCTOR) – *Kraay et al. (2013)*
- US – *Parker & Zegeer (1989)*
- British – *Baguley (1984)*
- Canadian – *Sayed & Zein (1999)*
- Finnish – *Kulmala (1984)*
- French – *Muhlrad & Dupre (1984)*
- Belgian - *Mortelmans et al. (1986)*

- German – *Erke (1984)*
- Czech – *Kocárková (2012)*
- Austrian - *Risser & Schutzenhofer (1984)*

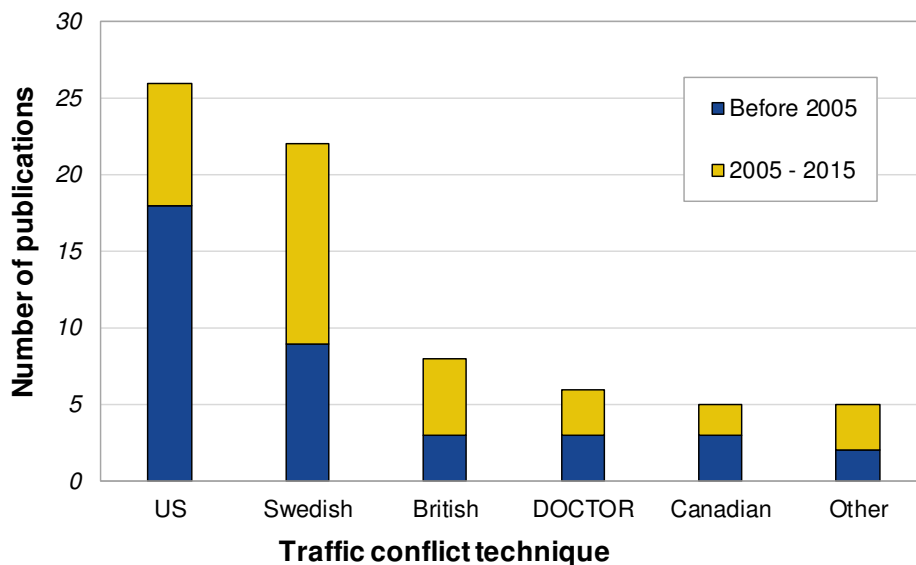


Figure 21. Usage of traffic conflict techniques.

4.3.1. The Swedish Traffic Conflict Technique

The Swedish TCT has been modified a few of times since it was first developed in 1976. The following basics have, however, remained the same:

- The presence of a collision course.
- The presence of evasive action in order to avoid a collision.
- The assumption that the interactions between road-users can be described as a continuum of safety related events. This implies that these events can be looked upon as different levels in a safety hierarchy, where the accidents are found at the very top and the "normal" passages at the bottom.

When developing the Swedish TCT it was found essential to distinguish the serious conflicts from the rest of the conflicts, as the serious conflicts were found to more strongly possess the quality of being an indicator of a breakdown in the interaction – a breakdown that could correspond to the breakdown in the interaction preceding an accident.

Below follow the descriptions of the different definitions of a serious conflict and different approaches to the relationship between serious conflicts and injury accidents, at different stages in the development of the Swedish TCT.

Hydén (1977)¹:

The **Time to Accident** (TA) is estimated from the moment one of the road users starts an evasive action and is the time that remains to an accident, if they had continued with

¹ Note that *Hydén (1977)* have already been described in section 4.2.3, however it is also included here due to its importance for the development of the Swedish TCT.

unchanged speeds and directions. The **Conflicting Speed** (CS) is the speed of the road user taking the evasive action at the moment just before the start of the evasive action. Based on watching video-taped conflicts, *Hydén (1977)* found indications that (almost) none of the involved road users voluntarily put themselves in a situation where the TA-value is below 1.5 seconds. Thus the threshold for a serious conflict was set to $TA \leq 1.5$ seconds (i.e. speed independent). When developing a model for the relationship between serious conflicts and injury accidents, data from 115 intersections was collected. Serious conflicts were registered during seven hours per intersection and injury accidents were collected for the last seven to eight years per intersection. Based on preliminary results from stepwise linear regression the relation between number of observed serious conflicts and number of injury accidents (per unit of time) should mainly depend on three variables for type of road user and four variables for vehicle speeds thus altogether 12 “cells”. For each of the 12 “cells” conversion factors (π) between serious conflicts and injury accidents were estimated. However, it turned out that some cells contained too few serious conflicts to produce statistically stable results. Therefore “cells” were merged based on similarities on probability of the road user obtaining injuries in a serious conflict and similarities in speeds. The merge reduced the matrix to four “cells”, see the previous Table 2 in section 4.2.3. Please note that in this version of the Swedish TCT conflicting speed was not used per se in the definition of a serious conflict, only to calculate the TA-value.

Linderholm (1981):

Linderholm (1981) used the same data set as *Hydén (1977)*, but had the aim of analysing conflict definitions and conversion factors based on likelihood of accident and likelihood of injury. The threshold for a serious conflict was the same i.e. $TA \leq 1.5$ seconds. But in order to categorize different conflict classes, $TA > 1.0s$ and $TA < 1.0s$ were introduced together with Conflicting Speed (CS) $> 35km/h$ and CS $< 35km/h$ (see Figure 22).

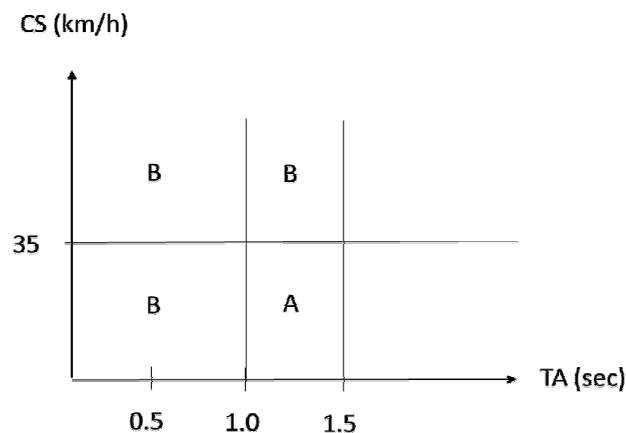


Figure 22. Definition of the two different conflict classes (*Linderholm, 1981*).

The hypotheses were the following: 1) The likelihood of a conflict ending up in an accident depends on the severity of the conflict which in turn depends on CS and TA value; 2) The likelihood of the accident resulting in injuries depends on: *i*) type of road users involved; *ii*) speed; and *iii*) angle of collision. Therefore, CS and TA value were used to distinguish two different conflict classes and car-car situations were split into parallel (rear-end and collision angle < 90 degrees) and perpendicular (meeting and collision angle ≥ 90 degrees). This resulted in six different “cells” with associated accident to conflict ratios π , see Table 4.

Table 4. Definition of the six cells and accident-to-conflict ratios $\pi \cdot 10^{-5}$ (Linderholm, 1981)

Situation → Conflict Class ↓	Car vs Car, Parallel*	Car vs Car, Perpendicular**	Car vs Cyclist/Pedestrian
A: CS < 35km/h and 1.0s ≤ TA ≤ 1.5s	0	2.4	9.6
B: All other conflicts with TA ≤ 1.5s	2.8	11.9	33.9

* = rear-end situations and situations with collision angle < 90 degrees

** = meeting situations and situations with collision angle ≥ 90 degrees

Gårder (1982):

Until this moment, conflict studies had only been conducted in urban areas. When performing conflict studies in rural areas, with higher speeds, it became evident that the urban definition of a serious conflict had to be modified. The speed independent $TA \leq 1.5s$ is replaced by a speed dependent curve formed by maximal braking on slightly moist asphalt ($TA = 1.5V/16.7e^{-0.148V}$, V in m/s). A safety margin of 0.5s was also added (Figure 23).

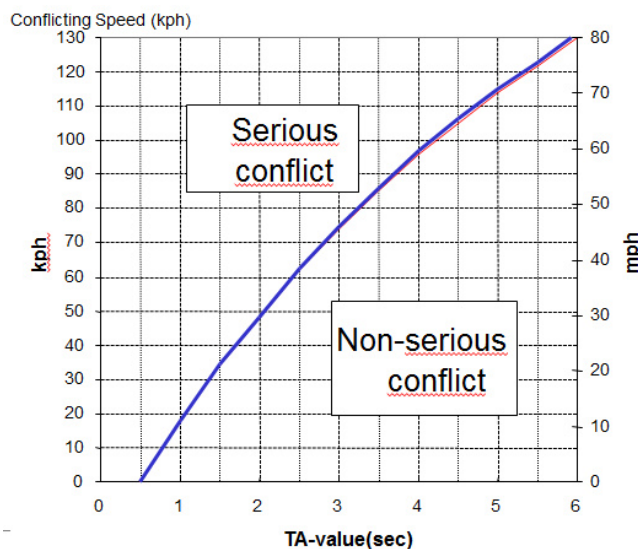


Figure 23. The speed dependent relation dividing conflicts into serious and non-serious conflicts (adopted from Gårder, 1982).

Hydén (1987):

In this doctoral thesis *Hydén (1987)* critically analysed the development of the Swedish TCT and the attempts so far to validate the technique. Now he elaborated on five different alternatives on how to define severity and severity zones (see Figure 24). For these alternatives he tested: *i)* severity of conflicts; *ii)* severity of accidents; *iii)* the proportion of accidents similar to conflicts; *iv)* the relevance of combining conflict and accident distributions. All in all, ALT.DEF 2 (Figure 24) turned out to perform best and *Hydén (1987)* suggested that serious conflicts should consist of situations within zone 2 and above in ALT.DEF. 2. This border is slightly different to *Gårder (1982)*.

Hydén (1987) also contributed to the process validation of the Swedish TCT (see chapter 2.6.3). The analyses showed that accidents and conflicts were similarly and continuously distributed in the TA/CS-graph with a tendency for the accidents being

located towards lower TA values and higher CS. This is very much in line with the hypothesis that accidents and conflicts are events in the same continuum.

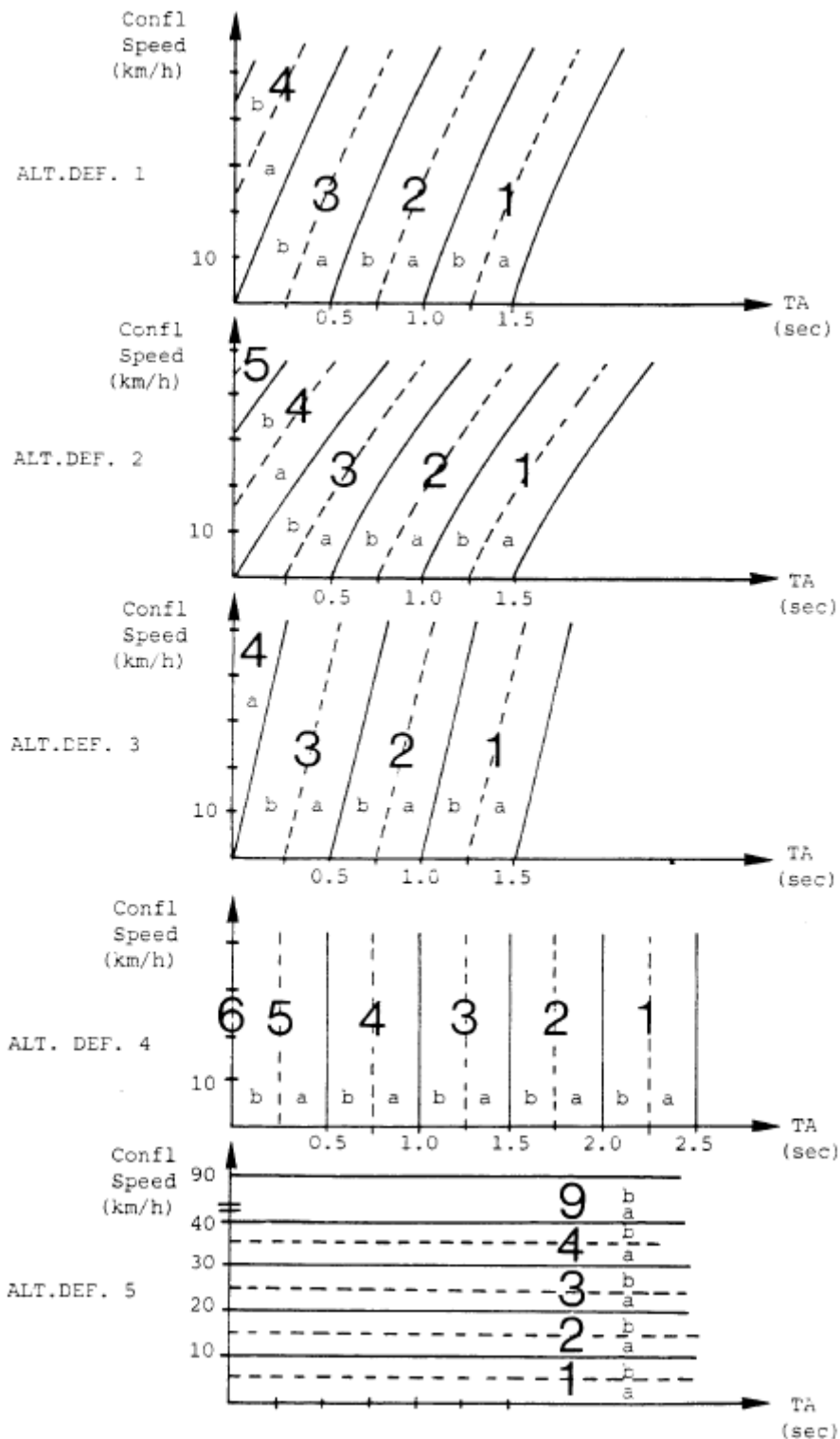


Figure 24. Definition of severity and severity zones for five different alternatives (Hydén, 1987).

Hydén (1987) also analysed conflict observers' reliability. In the *Malmö international calibration study*, there was an opportunity to check the subjective estimates with objective estimates in Grayson (1984). On average the Swedish observers' estimates of

TA were somewhat biased with a 0.05 second difference from the objective estimate. There was a tendency to underestimate high objectively measured TAs and overestimate low objectively measured TAs. When comparing estimates of CS, the analyses showed a small bias, observers' estimates of CS were on average 3 km/h lower than the objectively measured CS. The analysis also showed that the observers failed to score about 26% of the conflicts that should have been scored.

Svensson (1992):

In 1981 and 1982 additional data was collected in 115 intersections in the cities of Malmö and Lund. For each intersection information on design, traffic flow, accidents and conflicts were collected. Information on injury accidents for the past seven years were collected with restrictions to weekdays 9.00-18.00 and only during non-slippery conditions. Conflict studies were conducted during spring/summer period, during two weekdays and six hours per day. For each conflict the following was registered: type of road users involved, manoeuvres, CS, distance to collision point at the time of evasive action, type of evasive action and subjective severity rating on a six-point scale. The latter was a new component in Swedish TCT studies and aimed to reflect the human observer's estimate of the likelihood the situation could have resulted in injuries if the conflict had resulted in an accident. The aim was to analyse and compare these subjective estimates with objective estimates of a conflict's severity. A difference compared to previous conflict studies was also that the TA-value was calculated from estimates of CS and distance. Previously the TA-value was estimated directly. *Svensson (1992)* elaborated on different definitions of a serious conflict based on the different zones of ALT.DEF 2 in *Hydén (1987)* and subjective estimates and applied the product validation technique proposed by *Hauer & Gårder (1986)*. The main conclusion was that the subjective estimates and ALT.DEF 2 zone 2 and more severe produced the best estimates of the variance in expected number of accidents (λ), and that the subjective and objective estimates of severity followed each other nicely. This was very much in line with the results of the Malmö calibration Study (*Grayson, 1984*) that showed that observers also at "objective" registrations include an element of subjectivity which results in more similar registrations than perhaps the theoretical definitions could justify. The analyses further showed that at lower accident frequencies it was preferable to use one day of conflict studies instead of one year of accident data when estimating the expected number of accidents (λ). There were also indications (in terms of lower variance in λ) that it might be preferable, especially for situations involving vulnerable road users, to use observers' subjective rating compared to the strict "border" definition of a serious conflict.

Svensson (1998):

Svensson (1998) elaborated on the safety and severity hierarchy and how to understand these concepts. The traditional hierarchy is to be interpreted as a safety hierarchy and when applying the TA/Speed dimension, the severity hierarchy evolves. This study mainly contributed to the understanding that the shape of the severity hierarchy as such includes valuable safety information. She extended the TCT concept to also include "more normal" interactions but still with the requirements of collision course and evasive action. The interactions were positioned in the severity hierarchy (see Figure 25) with the estimated TA/Speed value obtained at the moment of evasive action.

The results suggested a border in the severity hierarchy at severity level 26 above which a high frequency of interactions is a sign of unsafety and just beneath which a

high occurrence rate of interactions is a sign of safety. It is, however, to be noted that not all interactions beneath the border are signs of safe locations. The earlier presented Figure 2, section 2.3, shows that a location with a high interaction frequency at low severity levels (signalised intersection, dotted line) instead seems to produce the conditions for occasional events with high injury accident potential. The convexity of these interactions with less severity was in this study widely spread over several severity levels. The opposite pattern (non-signalised intersection, solid line), a narrow convexity at reasonably high severities, thus just beneath the border, seems to be the insurance for preventing the most severe types of events from occurring. This is probably due to the learning process, i.e. the increased awareness of the road users brought about by involvement in interactions with reasonably high severity. It is, therefore, from a safety perspective not only interesting to analyse the part of the hierarchy with the most severe events, but also to take the convexity of the whole distribution into consideration.

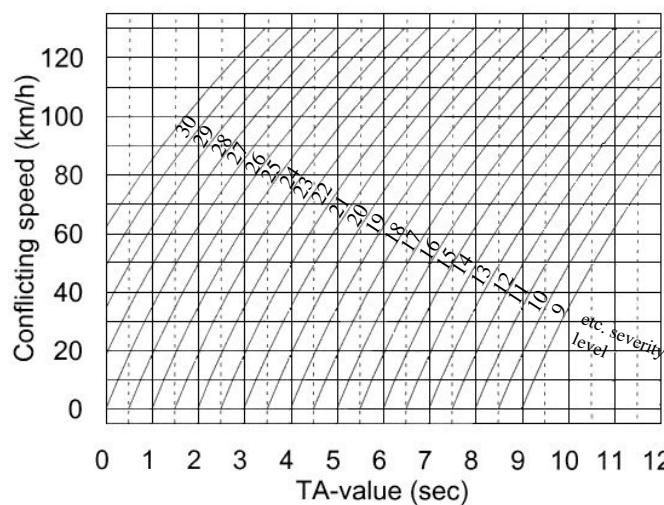


Figure 25. TA/Speed graph defining the different severity levels.

Shbeeb (2000):

Part of the aim of this study was to improve the Swedish TCT regarding vehicle-pedestrian conflicts based on the presumption that the current definition of a serious conflict produced less severe conflicts, than they should be, especially if the Relevant Road User (the road user who defines the severity of the conflict) was the pedestrian. For this aim 20 intersections were studied in Sweden, half of them signalised and half non-signalised. Conflicts were studied from video-recordings, six hours per day and for two days at the non-signalised intersections and for three days in signalised intersections. Traffic volumes were also obtained from the video recordings. Police reported injury accidents for 5 years (1993-1997) were collected for all 20 intersections. For the validation *Shbeeb (2000)* had a number of perspectives and analysis approaches:

Definition of Relevant Road User:

- Present: The road user who takes evasive action also defines the severity of the conflict
- High: The road user whose speed at the time of the evasive action and TA-value produces the more severe conflict regardless of who actually takes evasive action

- Low: The road user whose speed at the time of the evasive action and TA-value produces the least severe conflict regardless of who actually takes evasive action

Threshold Approach:

- GR5 = Present = The border between serious and non-serious conflicts intersects the TA-axis at 0.5 sec.
- GR6 = The border between serious and non-serious conflicts intersects the TA-axis at 0.25 sec.

For the validation the conflict to accident ratios (π) were calculated based on 7 of the 10 signalised and 7 of the 10 non-signalised intersections. Then the number of accidents were estimated at the remaining three signalised and non-signalised intersections. The results showed that the current definition of serious conflicts could be improved in the following ways:

- For signalised intersections: Adopt high definition of the Relevant Road User and threshold GR6;
- For non-signalised intersections: Present definition of Relevant Road User and adopt threshold GR6;
- For GR5 exclude conflicts where the speed of the Relevant Road User is less than 20km/h.

Current practice

The Swedish TCT conflict technique has also been used in recent years (*Fyhri et al., 2016, Svensson & Pauna-Gren, 2015, Sakshaug et al., 2010*). These examples most often use the division between serious and non-serious conflicts (Figure 23) suggested by *Gårder (1982)*. *Hydén (1987)* and *Svensson (1992)* suggested a small modification by including zone 2 in ALT DEF 2 (Figure 24) and *Shbeeb (2000)* also suggested some minor modifications. These modifications have, however, never gained support in the daily work with conflict studies. A possible reason might be the restricted capabilities connected to manual observers. It is easier to identify serious conflicts based on the currently applied thresholds than the suggested ones. New emerging technologies for collecting conflict data will, however, change the situation completely regarding what is possible to detect and could perhaps pave the way for trying out these other suggested thresholds.

4.3.2. DOCTOR - the Dutch Objective Conflict Technique for Operation and Research

Description of the technique

The DOCTOR method was developed in the Netherlands by the Institute of Road Safety Research (SWOV) and TNO Human Factors (*Kraay et al., 2013, van der Horst & Kraay, 1986*). According to this method, a critical situation is defined as a situation in which the available space for manoeuvre is less than is needed for normal reaction. If at least one of the parties involved needs to take action to avoid a collision, the situation is labelled as a conflict. In some cases, road users pass each other very narrowly without a noticeable evasive action. These situations can also be critical since a small disturbance in the approach process could easily have resulted in a collision. The severity of a conflict is scored on a scale from 1 (least severe) to 5 (collision), taking into

account: *i*) the probability of a collision and *ii*) the extent of the consequences if a collision had occurred.

The probability of a collision is determined by the following parameters:

- **Time-To-Collision (TTC).** The lower the TTC, the higher the collision probability will be. A minimum TTC during an interaction (TTC_{min}) of 1.5 sec or less is considered as critical.
- **Post-Encroachment Time (PET).** In urban areas PET values lower than 1 sec are considered as possibly critical.

The extent of the consequences is defined by the type of road users involved in the conflict, their speeds as well as the type of manoeuvres performed. For example, a conflict between a car and cyclists may entail much more serious consequences than a conflict between two cyclists, considered their relative vulnerability and speed. On the other hand, if a cyclists has a conflict with a car that drives < 20 km/h, consequences might again be relatively less serious. In Table 5 an overview is provided of what and how the different elements define the seriousness level of a conflict. Nevertheless, the technique contains a subjective component, since the observer always has also to take into account the behaviour of the road-users (do they for example undertake a controlled or uncontrolled evasive action) and extent of the consequences if a collision had taken place.

Table 5. Overall severity of a conflict according to the DOCTOR technique (Laureshyn et al., 2016).

Seriousness of injury	Probability of collision								
	no TTC	$TTC_{min}, sec.$					$PET, sec.$		
		>2	2-1.5	1.5-1	1.0-0.5	0.5-0	>1.0	1.0-0.5	0.5-0
<i>very small</i>	X	X	X	1	1	2	X	X	1
<i>small</i>	X	X	1	2	2/3	3	X	1	2
<i>fairly large</i>	X	1	2	2/3	3	4	1	2	3
<i>large</i>	1	2	2/3	3	4	5	2	3	4/5

The conflicts with a severity level of 2 or higher are usually considered in conflict studies based on DOCTOR.

Validation of DOCTOR

A comparison between video-taped conflicts and accidents (*van der Horst, 1984*) indicated that conflict severity scores, performed by individual observers of existing traffic conflict technique teams at that time (*Grayson, 1984*), were mainly correlated to Time-To-Collision (TTC) and type of accident.

P. C. Cacciabuevan der Horst (2007) conducted long-term video observations to collect data on the pre-crash phase of real accidents (what exactly happened just before the collision?) as a means to evaluate and validate the separate approaches of in-depth accident analyses, road scene analyses, and behavioural observations (including traffic conflicts). The rationale for this study was the notion that validation studies of traffic conflicts often hamper insufficient and incomplete accident data. This study included video observations for two years at four urban intersections (24h/day). In this period, in total 16 collisions could be identified. Apart from the selection of collisions, events with deviant road user behaviour or potential conflicts were collected from video on a more

or less random basis. Finally, for each intersection, one day had been selected to scan the video disks for potential conflicts. Collisions and conflicts have been analysed quantitatively using the VIDARTS (VIDEO Analysis of Road Traffic Scenes) approach and conflicts scored according to the criteria of the DOCTOR technique by human observers from video. The results of the collision- and conflict analyses have been compared with the results of the separately conducted road scene analyses (*van der Horst & Martens, 2007*) to evaluate the various approaches and to formulate recommendations for safety improvements at the four intersections. *P. C. Cacciabuevan der Horst (2007)* concluded that traffic conflicts and deviant behaviour, together with road scene analyses provide good insight into potential safety problems at specific locations which are both in line with a road user perspective and the results from the collision analyses. With respect to the latter, remarkably, in most cases, another (third) road user was (in)directly involved, either as a distracting or as a contributing element, for example by occluding the view of one of the road users involved. When looking at in-depth accident investigation methodology, this will be a critical point to collect reliable data on the influence of the presence of other road users in reconstructing the causal processes after the collision has occurred since these contributing circumstances will have disappeared shortly after the collision.

In a recent study (*van der Horst et al., 2016*) at three locations in Bangladesh, a combination of three research methods to monitor and evaluate the road safety interventions was applied in a Before-and-After study design. Because of the lack of reliable accident statistics in Bangladesh, firstly, a traffic accident recording system with trained local record keepers was developed. Secondly, laser-gun speed measurements of motorized traffic (both at intervention and control locations) were conducted, and, thirdly, DOCTOR was used to observe serious traffic conflicts at the intervention locations from video. All three evaluation measures pointed to a similar impact of the intervention program and unveil an improvement in road safety between 54% and 60%. The speed-reducing measures indeed considerably reduce the speed of motorised traffic, both the mean speed and 85th percentile values, both the number and severity of serious conflicts are reduced, and the actual number of accidents has decreased. Based upon the collection of the actual number of accidents occurring at the three intervention locations, the results indicate that both speed measurements and traffic conflict observations constitute a valid evaluation and have a good estimate of the expected safety effects, when no reliable accident data are available.

4.3.3. The American traffic conflict technique

Description of the technique

The American TCT was originally developed by *Perkins & Harris (1967)*. The conflicts were identified based on brake lights or weaving manoeuvres and five different types of conflicts were used:

- Left-turn conflicts
- Weave conflicts
- Cross-traffic conflicts
- Red-light violation conflicts
- Rear-end conflicts

Traffic violations were recorded as conflicts, regardless of the presence of other vehicles. In *Parker & Zegeer (1989)* a similar but slightly modified version of the

technique is presented. Their general definition of a conflict is “an event involving two or more road users, in which the action of one user causes the other user to make an evasive manoeuvre to avoid a collision” and they present the following type of conflicts:

- Same direction conflicts
- Opposing left turn conflicts
- Cross traffic conflicts
- Right-turn–on-red conflicts
- Pedestrian conflicts
- Secondary conflicts

The adjustment to different types of conflicts can also be observed in recent use of the technique. *Zhao et al. (2011)* used four different conflict types to study a left exit on a freeway and *Burnett & Sharma (2011)* used six different conflict types when studying the onset of yellow in signalized intersections.

Validation of the US technique

Baker (1972):

This study contains conflict data from 392 intersections (886 approaches). For each intersection two opposite approaches are studied. At each approach conflicts are counted for 5 hours by trained observers. In contrast to the original technique, red-light violation conflicts are not included. The analyses are mainly based on Pearson’s correlation coefficients between conflicts and accidents for different types of intersections and types of manoeuvres. The correlations seem to be very good, for many of the situations the results indicate statistical significance at the 95%-level. For 4-legged right angle signalised intersections, all but rear-end situations had statistically significant correlations. Less favourable results were found for three-legged (T) signalised intersections but this category was also so much smaller, 14 approaches compared to 122 of the type 4-legged right angle. For the non-signalised intersections, the results were even better, for both T-junctions and 4-legged right angle intersections, all situations but weaving situations at 4-legged right-angle intersections had significant correlation coefficients between conflicts and accidents.

Bearing in mind the difficulties connected to correlation coefficients as a verdict on whether conflicts are valid surrogates to accidents, argued by *Hauer & Gärder (1986)*, these results are quite remarkable. But despite the enormous size of this data set and very encouraging results, very few analyses are done and the conclusions are surprisingly weak – “The data compiled in this study tend to support the hypothesis that conflicts and accidents are associated” and “The traffic conflicts technique may be particularly valuable at low-volume rural intersections where accident reporting level is low”.

Paddock & Spence (1973):

This study studied at 611 intersection approaches and uses a regression model with flow and conflicts as input to generate a prediction equation that calculates the expected number of accidents. A division was made between signalized and non-signalized intersections. The resulting prediction errors between expected and observed accidents presented in Table 6.

Table 6. Prediction error estimates (*Paddock & Spence, 1973, cited in Glennon & Thorson, 1975*).

<u>Data Class</u>	<u>Number of Points</u>	<u>Prediction Error (Acc/Year)</u>		
		<u>50%</u>	<u>75%</u>	<u>95%</u>
All Data	611	± 1.2	± 2.0	± 4.2
Signalized	220	± 1.5	± 2.4	± 4.6
Unsignalized	391	± 1.1	± 1.8	± 3.8

This study has not been located during this literature review. The information regarding the study originates from *Glennon & Thorson (1975)*.

Pugh & Halpin (1974):

This study looked at 240 intersections in a similar way as *Baker (1972)*. A severity measure has been used where 1 is a routine conflict, 2 is a moderate hazard and 3 is a near-accident. A fair correlation between accidents and conflicts based on correlation coefficients was found. The authors concluded that the TCT seemed to be a valuable tool for assessing accident potential.

This study has not been located during this literature review. The information regarding the study originates from *Glennon & Thorson (1975)*.

Cooper (1973)

This study conducted two 14-hours day studies (counted in four 7-hour days) at 59 non-signalised intersections in four Canadian cities. Information on the extension of the accident data is not available. Some correlation (linear correlation coefficient of 0.453) was again found between the number of conflicts and accidents but a much stronger correlation was found between traffic flow and the number of accidents. Minor modification was made to the original technique as precautionary conflicts (i.e. braking for vehicle waiting to emerge, precautionary lane change, or anticipatory braking) were excluded. *Cooper (1973)* concludes that even though conflicts correlated significantly the correlation was not of a high order and even when combined with multiple regression analysis, only 50% of the accident sample variance was explained.

Migletz et al. (1985)

This study used a modified version of the original technique where a conflict was only regarded if it had a TA-value of less than 1.5 s. More details regarding this report have already been mentioned in section 4.2.3.

4.3.4. The British traffic conflict technique

The British technique (*Baguley, 1984*) is based on a 5 degree severity classification:

1. Controlled braking or lane change to avoid collision but with ample time for manoeuvre.
2. Braking or lane change to avoid collision with less time for manoeuvre than for a slight conflict or requiring complex or more severe action.
3. Rapid deceleration, lane change or stopping to avoid collision resulting in a near miss situation (no time for steady controlled manoeuvre).
4. Emergency braking or violent swerve to avoid collision resulting in a very near miss situation or minor collision.
5. Emergency action followed by collision.

According to *Baguley (1984)* this measure is influenced by: *i)* TA-value; *ii)* the severity of the evasive action; *iii)* the type of the evasive action; and *iv)* the distance between conflicting vehicles when evasive action is terminated. These four factors can be used to determine the severity (see Table 7).

Some recent studies have used this technique (*Agah & Sri, 2013*) or a modified version of it (*Kaparias et al., 2010*). *Lawalata & Agah (2011)* also describes several Indonesian studies that used the original British technique (published in Bahasa language, therefore we had to rely on the information provided by *Lawalata & Agah, 2011*).

Table 7. Conflict severity according to the British conflict technique (*Baguley, 1984*).

	Time	Long		Moderate			Short				
	Severity	Light	Medium	Light		Medium	Heavy	Medium		Heavy	Emergency
	Type	Simple/ complex	Simple/ complex	Simple	Complex	Simple/ complex	Simple/ complex	Simple	Complex	Simple/ complex	Simple/ complex
PROXIMITY	>2 car lengths	1	1	1	1	1	1	1	1	2	3
	1 to 2 car lengths	1	2	1	2	2	3	2	3	3	3
	1 car length or less	2	3	2	3	3	3	3	3	4	4
	Minor collision	3	4	3	3	4	4	4	4	4	4
	Major collision	5	5	5	5	5	5	5	5	5	5

Validation of the British technique - relation to accidents

Spicer (1973):

This study conducted conflict studies at 6 non-signalised intersections. Trained observers supported by video recordings observed each location for 10 hours. Only conflicts with a severity value of 2,3,4 or 5 were included. The analysis provided a correlation coefficient of 0.97 (statistically significant at 0.1%-level¹) between conflicts and number of accidents recorded in a 3-year period.

TRRL (1980):

This study used the same approach as *Spicer (1973)* but expanded the conflict study to 14 intersections. A high correlation was again found between accidents and conflicts as can be seen in Figure 26. Note that this study has not been located in this review and the information is based on *Baguley (1984)*.

¹ i.e. the risk of the correlation "by chance" is less than 1/1000.

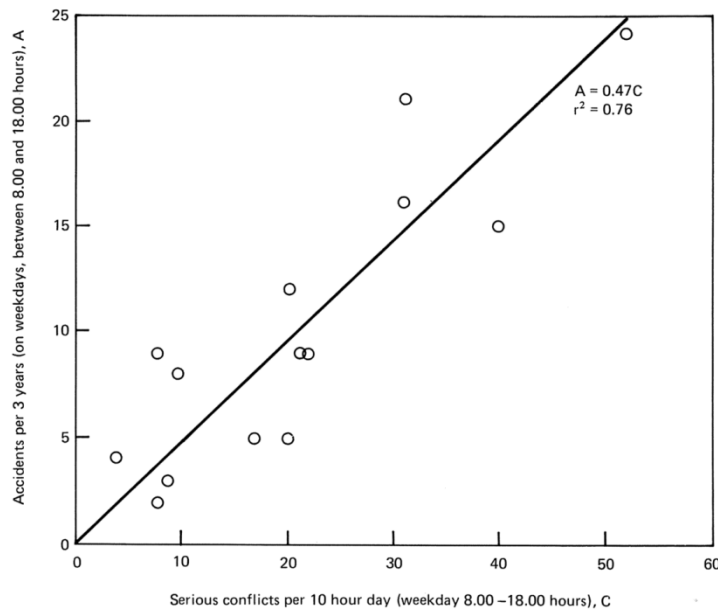


Figure 26. Correlation between conflicts and accidents (*TRRL, 1980, cit. Baguley, 1984*).

4.3.5. The Canadian traffic conflict technique

This technique was originally described by *Brown et al. (1984)*. Conflicts are classified based on the following indicators:

- TTC_{min} (3-2s, 2-1.5s, 1.5-1s, 1-0s);
- Estimated risk of collision, ROC (Small, Moderate, High, Very high);
- Type of evasive action (Braking, Swerving, Accelerating, Combination);
- Severity of evasive action (Light, Moderate, Heavy, Emergency);
- Proximity at end of evasive action (> 2 car lengths, 1-2 car lengths, < 1 car length, near-accident).

Observations are performed on-site by human observers. *Brown et al. (1984)* suggest to conduct at least two days of 8 hour observations at a specific location.

Validation of the Canadian traffic conflict technique

Brown et al. (1984):

This is a small case study of 4 intersections and 16 hours of observation per location. The conflicts were differentiated into 44 different manoeuvres. The case study found support for correlation between conflicts and accidents.

Brown (1994):

This study used a slightly different definition. It used “TTC from evasive action” (i.e. TA) and observer’ estimate of the Risk of Collision (ROC). These were divided into 3 instead of 4 groups (0-1s, 1-1.5s, 1.5–2s for TA and “slight”, “moderate” and “serious” for ROC). The two scales are given equal weight and are combined to give a composite five-point Likert-type scale on which the midpoint registers the critical conflict event, and corresponds to a TTC of 1.5 s or less with a “moderate” probability of collision. In the conflict studies and in the calculation of accident to conflict ratios, only conflicts within the threshold for a serious conflict (i.e. $TA < 1.5s$ seconds and ROC at least moderate) are included. 13 intersections were observed by trained observers for 16 hours at each location. Accident to conflict ratios (π) were calculated for the intersections both as a

whole and for specific manoeuvres. On a selection of 6 non-signalised intersections the π -values for the whole intersection had a lower variance than for the specific movements. *Brown, 1994* concludes that “using this definition of conflict severity indicated some association of conflicts with actual crashes, when disaggregated by movement, but in the application of the methods, dichotomous notions of conflicts and accidents were evident.”

Sayed & Zein (1999):

This is a summary of conflict studies at 94 intersections, 52 signalized and 42 un-signalized intersections. The conflict studies used the original indicators (TTC_{min} and ROC) but slightly different categories. Serious conflicts were defined as having TTC_{min} < 1.5s and a ROC of at least moderate risk. The results showed that all conflicts, as well as only serious conflicts, had strong correlation with accident records in signalized intersections but not in un-signalized intersections.

4.3.6. Other techniques

The Finnish traffic conflict technique

The Finnish technique (*Kulmala, 1984*) is similar to the Swedish technique but with some modifications. Interactions with a TA ≤ 1.5s are identified as conflicts. A conflict is considered serious if the evasive action is deemed uncontrolled. Other interactions of interest are called potential conflicts.

A threshold value of 3 seconds is used in rural areas instead of the 1.5 seconds and a threshold of 1 seconds is used if the conflict occurs between a pedestrian and bicyclists or between 2 bicyclists.

The French traffic conflict technique

The French TCT is described by *Muhlrad & Dupre (1984)*. The observation of conflicts is identified by evasive action taken by at least one of the involved road users. An evasive action is described as a discontinuity in the driving (cycling/walking) process that follows the occurrence of an unpredictable or surprising event. The differentiation between normal and emergency manoeuvres is not based on objective measures but rather “estimated” by a team of observers at the road side.

The conflicts are then differentiated by 5 different severity levels:

- Light conflict
- Moderate conflict
- Serious conflict
- Light collision
- Serious collision

The Belgium technique

The Belgium technique of near-accidents observation (*Mortelmans et al., 1986*) seems to be very similar to the French technique using the same 5 severity levels for conflict classification.

The German traffic conflict technique

The German TCT is described by *Erke (1984)*. Conflicts are indicated by a critical manoeuvre of at least one of the involved road users. The critical manoeuvres are:

- Braking
- Accelerating
- Swerving
- Stopping
- Running, jumping
- Combination of the above

The conflicts are located on a 3 level severity scale.

- Controlled braking and/or swerving.
- Strong braking and/or abrupt swerving, no time to perform a controlled manoeuvre.
- Emergency

Austrian technique

The Austrian TCT (*Risser et al., 1991, Risser & Schutzenhofer, 1984*) is a semantic technique thus it uses the human perception ability to evaluate the dangerousness of a situation without additional help of quantification. When training observers the aim is to achieve a high inter-correlation between trainer and trainees. In the Austrian TCT the observers should be able to distinguish between:

- controlled braking, swerving or acceleration, though necessary to avoid an imminent collision;
- strong or emergency actions of braking, swerving or accelerating, felt as last-second avoidance by the observers.

Discriminatory power is achieved by training that include discussions of events observed at the road-side and from video-recordings.

The most important applications in recent years are observation of conflicts involving motorcyclists in three European countries and observation of interaction processes between bicyclists and car drivers in Vienna, Austria.

Czech technique

The Czech TCT is described in *Kocárková (2012)* and uses five levels of subjective severity ratings. Level 0 is assigned to the controlled manoeuvre without any limitation or just with minor limitation. The difference between Level 1 and 2 is minor but two is a more severe, but still controlled, manoeuvre. Level 3 is assigned to situations when the road users are threatened and sharp manoeuvring or hard breaking is necessary to avoid an accident. Level 4 corresponds to situation that results in an accident.

4.4. Surrogate safety studies of VRU

4.4.1. Overview

This section will focus on surrogate safety measures used to study VRUs. Figure 27 shows the types of road users studied in the reviewed publications. The largest category includes motor vehicles only, but studies where VRU are included are not infrequent neither. Note that publications might include multiple types of road users and are then counted more than once.

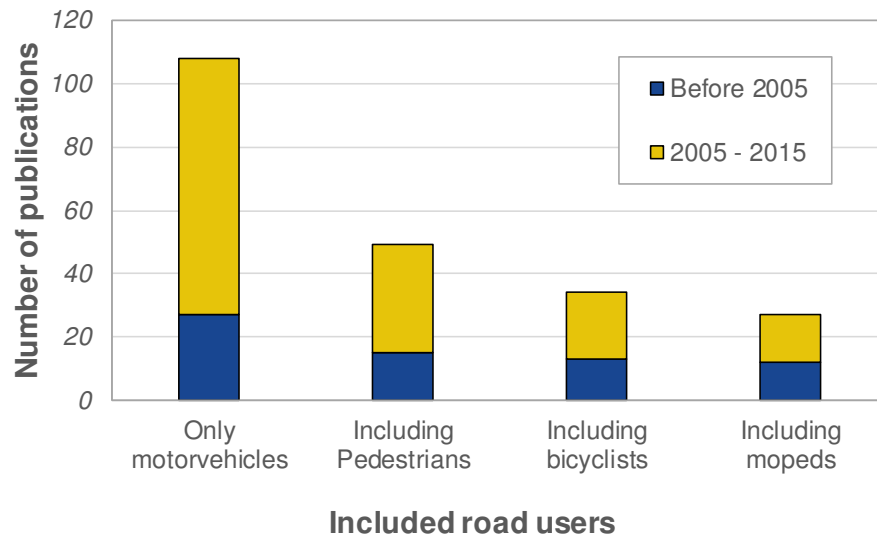


Figure 27. Type of road users observed.

Figure 28 shows the split between TCTs and individual indicators for used in all reviewed studies and the same split for publications that included VRUs. TCTs were used in approximately 28% of all publications but in 41% of all publications that included VRUs. Note that some publications use multiple TCTs or indicators and have been counted multiple times in the figures.

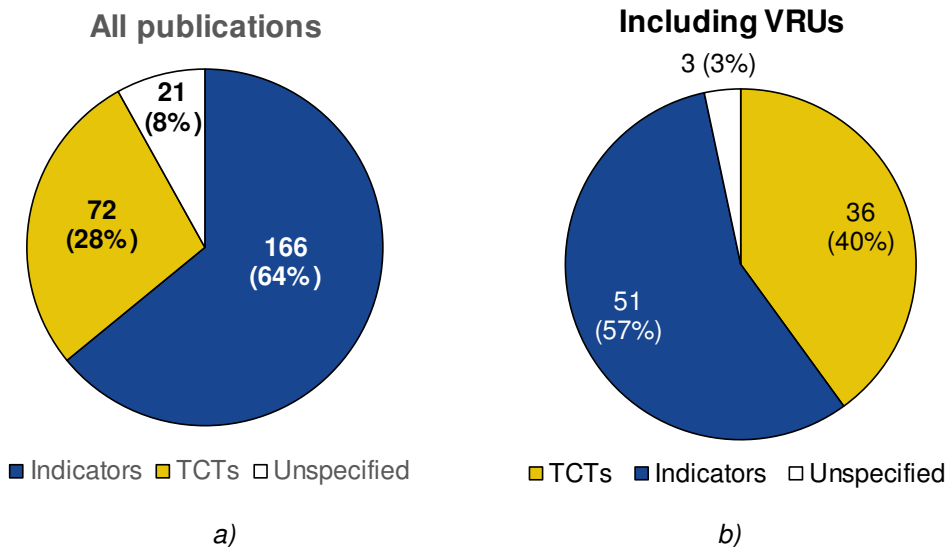


Figure 28. The split between the use of TCTs and individual indicators for all publications (a) and those including VRUs (b).

4.4.2. Indicator usage

Figure 29 shows the indicators used in the publications that studied VRUs. Time-to-Collision (TTC) and Post-Encroachment Time (PET) are the most commonly used indicators, followed by Deceleration-to-Safety Time (DST). It is also worth noting that all of the publications that use Gap time (GT) combines it with the use of DST, PET and TTC_{min} (e.g. Ni & Wang, 2015, Ismail et al., 2010b).

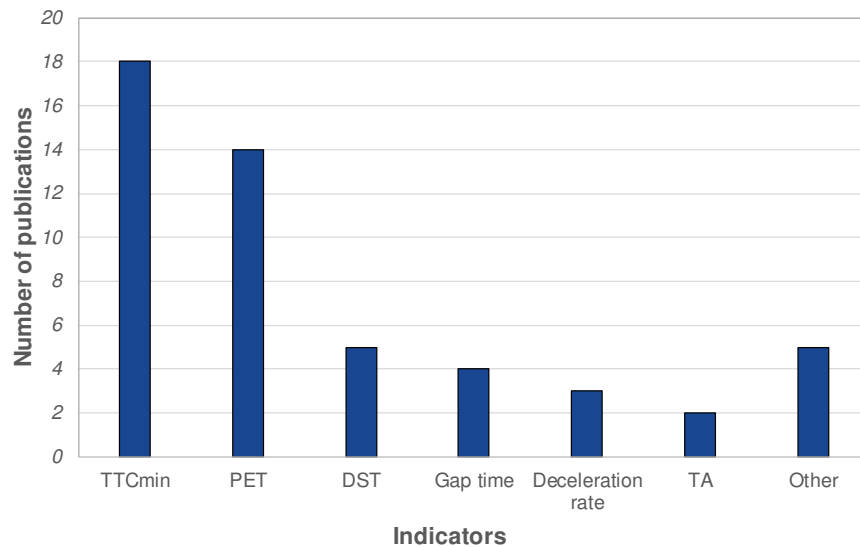


Figure 29. Indicators used to study VRUs.

The “Other” category includes indicators that have only been used once.

Tageldin et al. (2015) suggested examination of the jerk profile and the yaw rate to identify critical evasion for motorcyclists. Jerk profile is meant to estimate the intensity of the braking action, and the yaw rate is meant to quantify the swerving of the motorcyclist.

Cafiso et al. (2011) presented the pedestrian risk index (PRI). The PRI combines Time-to-Zebra crossing (TTZ) with assumptions regarding the driver’s reaction time, and the braking capabilities of the vehicle, to estimate the risk of collision as well as potential severity of a collision. The risk of collision is calculated as the time difference between TTZ and the estimated time needed to stop, and the severity is estimated as the potential impact speed squared. The final PRI is the risk of collision multiplied by the severity summed during the duration of a potential collision course.

Bagdadi (2013) suggested the “conflict severity” measure as a combination of the indicators Delta-V, TA, and the assumed maximum average deceleration. *Laureshyn et al. (2017)* extended this concept to a continuous indicator that can be applied for situations both with and without a collision course (described also in section 4.2.6).

Ogawa, 2007 discusses a space occupancy index based on personal space, which is an index that expresses the spatial sizes necessary to maintain safety on the road for bicycles, pedestrians, and vehicles. For each road user type, an area around the road user is defined based on the characteristics of the road user type. Figure 30 shows the personal space suggested for bicyclists. The number of conflicts is then estimated by the number of personal space incursions.

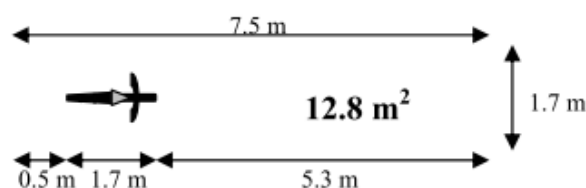


Figure 30. Area of personal space for a bicycle (Ogawa, 2007).

4.4.3. Traffic conflict techniques usage

Figure 31 shows the use of various TCTs in studies that included VRUs. The main difference compared to the general result is that the US technique is less prevalent, the majority of the publications used the Swedish technique.

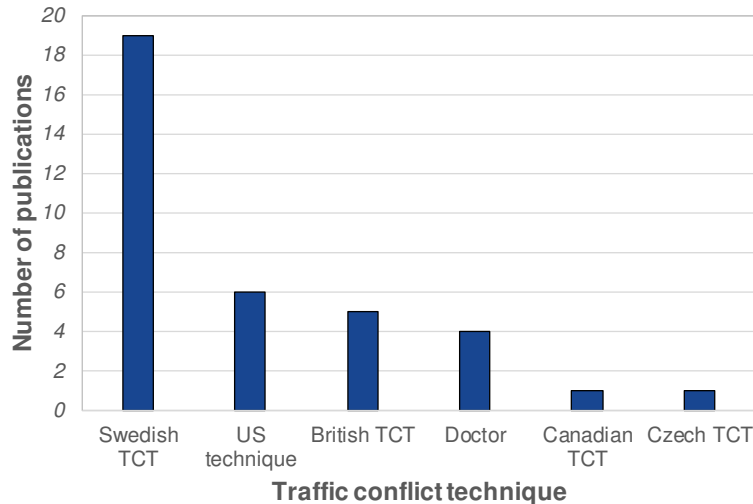


Figure 31. TCTs used in publications including VRUs.

Both the Swedish TCT and the DOCTOR included VRUs from the beginning of their development (*Hydén, 1987, van der Horst & Kraay, 1986*), and both have been used in recent studies focusing on pedestrians and bicyclists (*Fyhri et al., 2016, van der Horst et al., 2014, Sakshaug et al., 2010*). The British technique as well as the American technique have been modified so that they can be applied to VRUs, too (*Kaparias et al., 2010, Pietrantonio & Tourinho, 2006*).

4.5. Study design

This section presents the study designs and observational methods used in the reviewed publications. This is not the major focus of the report and therefore it will not be followed-up in the discussion section.

4.5.1. Number of locations and observation duration

Figure 32 shows the number of locations included in one study. The majority of studies are done using only one observation site. Approximately 30% of studies are using 5 and more observation sites.

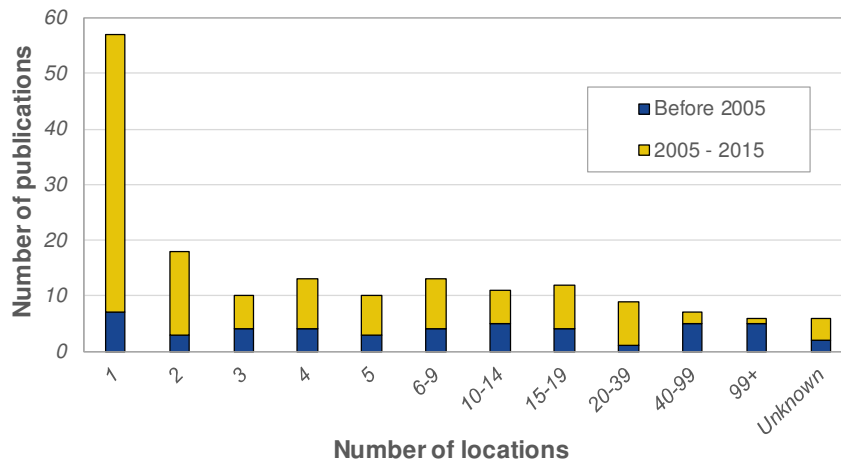


Figure 32. Number of observed locations.

Observations at only one site might be acceptable if the aim is to test performance of a traffic conflict technique or an indicator itself or to test some technical data collection tools. However, if the study is focused on applying observations for practical road safety study purposes, making observations at only one site can limit the possibilities to generalize results.

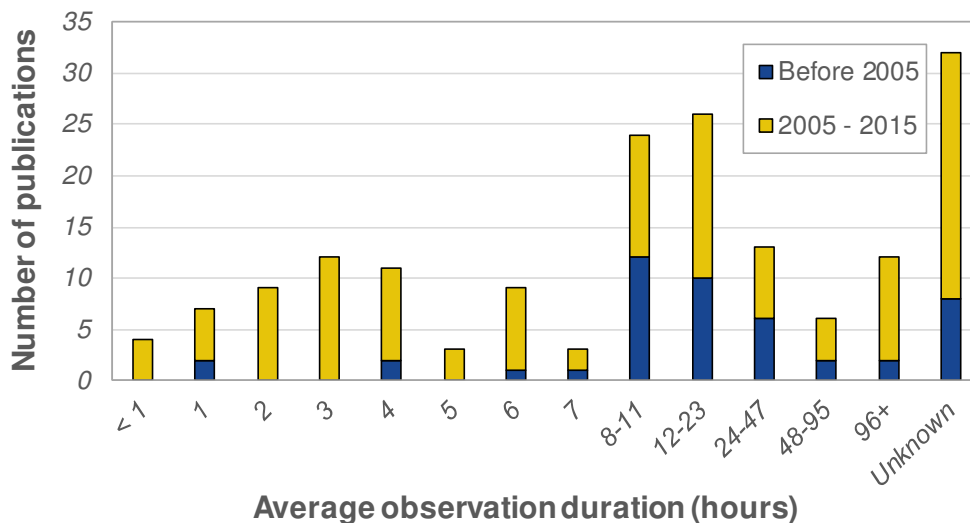


Figure 33. Average duration of observation per site.

Figure 33 shows the average duration of observations per location. Relatively short observation periods per observation site are quite common. 45% of all studies observed less than 8h per site (25% less than 4h), 22% of all studies observed more than 24h per observation site. Quite surprisingly, 23% of the included publications did not include any information about the duration of the observations at all.

Figure 34 zooms in at the studies with one observation site, and shows the duration of observations. 45% of studies had an observation period of less than 8h, and 28% had an observation period of less than 4h. 30% had an observation period of more than 24h, which is a slightly higher value than that of all studies.

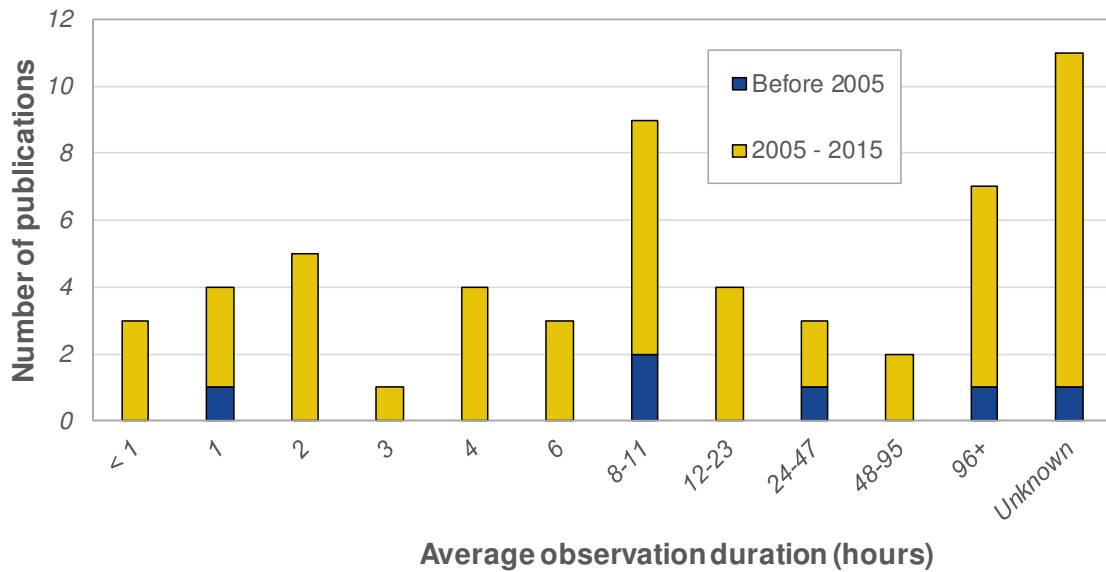


Figure 34. Average observation duration for studies with one site.

Figure 35 illustrates where the observation sites are located. It can be concluded that site-based observation studies mostly take place in urban areas. It can also be seen that this information is missing or unclear in a large number of publications, which can have important implications for the interpretation of study findings. The “other” category includes publications that specifically focuses on freeways, tunnels, work zones and toll stations.

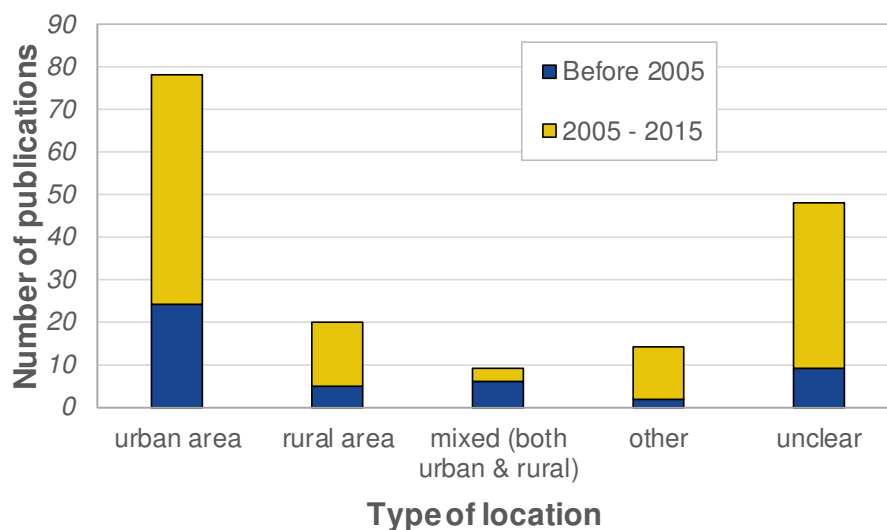


Figure 35. Type of location observed.

4.5.2. Data collection method

Figure 36 shows the data collection methods used in the studies. It can be seen that the different forms of manual observation have been most common over the years. However, the number of publications that apply video analysis software take up a significant share as well, especially over the last 10 years. Fully manual observations (i.e. human observers on-site without video support) are a relatively large category for all publication years taken together, but have rarely been applied during the last decade.

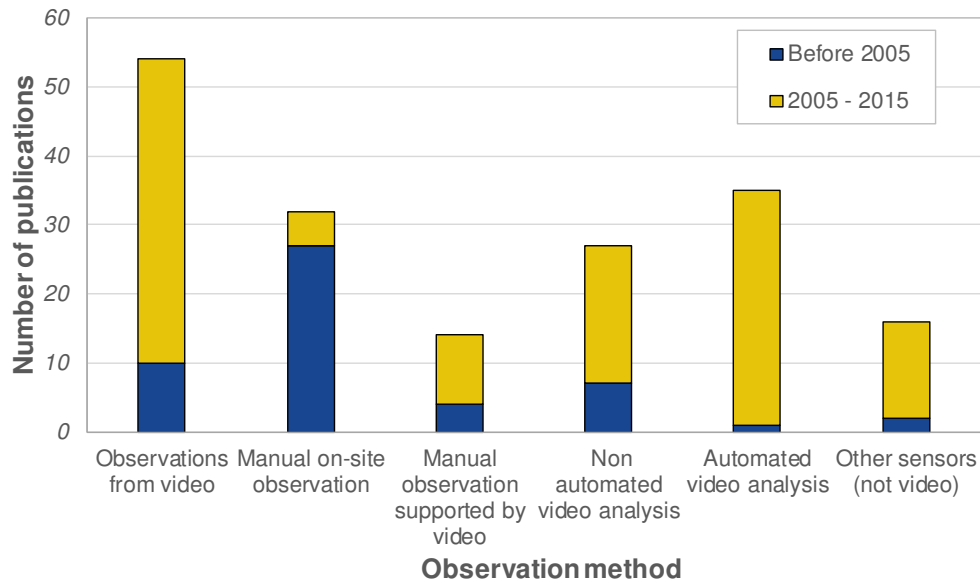


Figure 36. Data collection methods.

4.5.3. Data analysis

Figure 37 presents the additional data that is collected in conjunction with surrogate safety data. The most common additional data is the exposure (though defined in many different ways). Slightly less common are accident history, information about the infrastructure and systematic behavioural observations. The category “Other” is fairly large as well, including diverse types of data such as results from microsimulation or driving simulator studies, road user characteristics such as gender and age and survey or interview data.

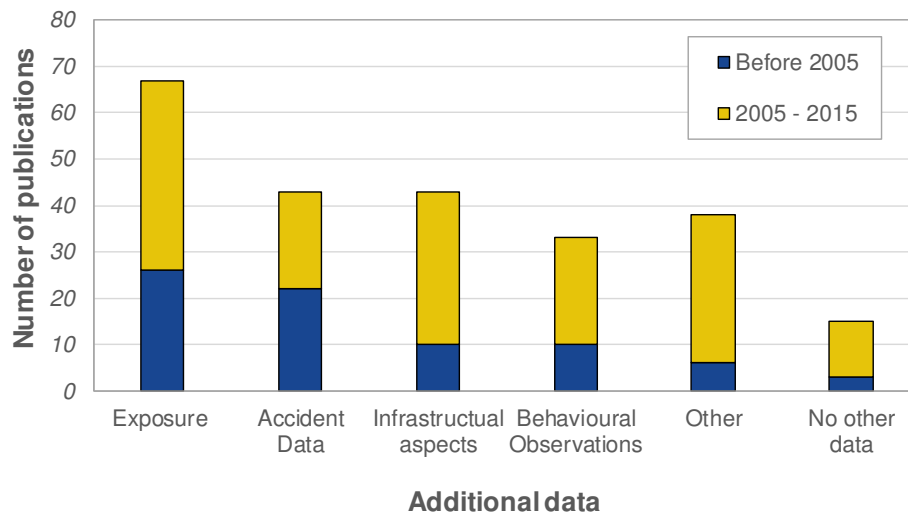


Figure 37. Additional data used.

Figure 38 presents the analysis methods applied on surrogate safety data.

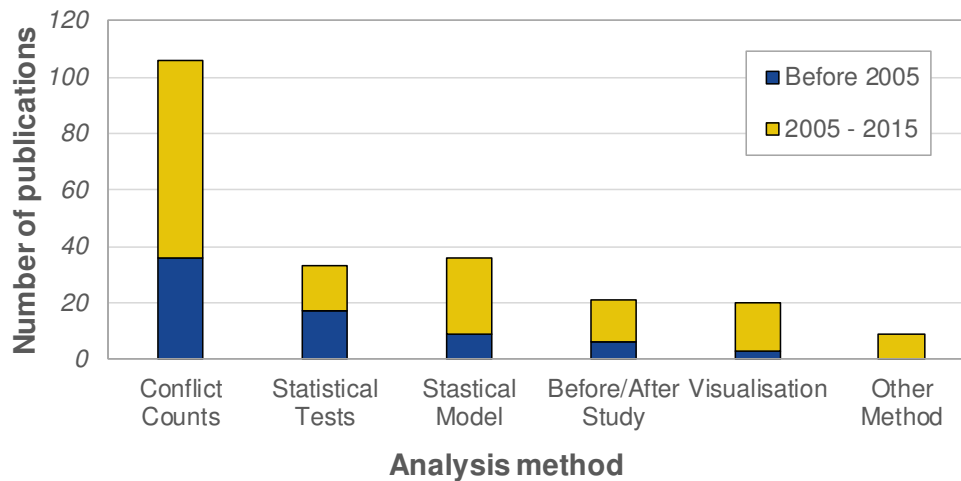


Figure 38. Analysis methods used.

Simple conflict counts are by far the most common type of analysis and are included in more than half of all studies that performed data collection and analysis (106 out of 169 for all studies, 70 out of 123 for studies between 2005 - 2015). In approximately half of these studies, it was the only form of data analysis. Statistical models and tests, before-and-after comparisons and visualization of the observed events on a map or aerial photo of the study sites are less common.

5. Discussion

5.1. Validity of applied techniques and indicators

As indicated at the beginning of this report, product (or “absolute”) validity can perhaps be considered as the most fundamental form, or the highest extent of validity. The usefulness of a surrogate measure of safety does, however, not (only) depend on the extent to which expected accident numbers can be correctly estimated (*Grayson, 1984*). The usefulness mainly depends on whether safety problems can be detected or not, and/or road safety counter-measures/treatments can be compared or evaluated (*Chin & Quek, 1997, J. A. Rothengatter, R. A. de Bruin Grayson & Hakkert, 1987, Hauer, 1978*). If we accept the premise that the ultimate goal of surrogate safety studies is not the estimation of expected accident numbers, we can see validity in different perspectives, i.e. the perspectives of relative validity and process validity.

Relative validity is easier to achieve than absolute validity, because it suffices to have sufficient evidence that a direction of effect on expected accidents can be inferred from a traffic conflicts study, instead of a way to convert the non-accident events to an expected number of accidents. It is the determination of a practical way to convert the non-accident events to the (expected) number of accidents ratio that seems to be the most problematic validity issue (*Hauer & Gårder, 1986*). This does however not mean that attaining absolute validity is not useful or worthwhile since an indicator with absolute validity still provides the most accurate and detailed information about safety performance. If we consider validity as a matter of degree, and not a “yes or no” concept, we can consider a high degree relative validity a lower (yet for some purposes acceptable) degree of validity than absolute validity.

Therefore, it is concluded that one should ultimately aim to attain absolute validity for a surrogate measure of safety, but that one should be aware that a measure that has a sufficient relative validity can be useful for specific study designs as well.

A parallel can be made here with other types of research, such as driving simulator research. The validity of driving simulator studies are fairly often questioned because one may doubt the extent to which behaviour in a simulated road environment corresponds to the participants’ actual driving behaviour in a real-life environment (*De Ceunynck et al., 2015, Fisher et al., 2011*). While absolute validity of the driving simulator as a research tool is not always attained, there is however plenty of research showing that driving simulators generally reach high relative validity (*Bella, 2009, Yan et al., 2008, Godley et al., 2002, Törnros, 1998*). This implies that the driving simulator is considered a valid tool for controlled experiments to compare safety aspects between different experimental conditions. As an illustration, suppose we are interested in comparing drivers’ speed behaviour while approaching two different types of road design. If we observe a significantly lower approach speed for road design A compared to road design B, relative validity implies that we can be confident that we would also observe a lower speed at road design A than at road design B in the real world. However, it is unsure whether the driving speeds in absolute terms would be exactly the same in the real world. The exact driving speeds in the real world (and the order of magnitude of the difference between both designs) might differ substantially. This is (the uncertainty about) absolute validity.

Process validity indicates the extent to which conflicts can be used for describing the process that leads to accidents (*Svensson, 1998*). Absolute (or relative) validity in itself may be enough to identify high-risk locations, to assess which road designs have a better safety performance than others, and which measures have a positive effect on the (expected) number of accidents that take place. In itself, however, this validity does not suffice to reveal causal chains underlying the accidents. In other words, such indicators cannot necessarily tell us *why* or *how* some locations perform better or worse than other locations. To be able to reveal such causal chains, the factors and processes that lead to conflicts should be similar to those that lead to accidents. When the factors are highly similar, studying the factors that lead to traffic conflicts can be considered a valid alternative for analysing the causal factors that lead to traffic accidents. Therefore, process validity is a highly relevant form of validity additional to product (respectively relative) validity.

Given the fact that validity as a concept has multiple dimensions and is a continuum, it is not an easy task to summarise the current status in the field. The literature shows mixed results. A number of publications indicate a poor relationship between the number of conflicts and accidents and have seriously questioned the usefulness of the traffic conflict technique (*Tiwari et al., 1998, Williams, 1981, Glennon et al., 1977*). Researchers analysing the reasons for the poor performance of the number of conflicts as a surrogate measure of safety however came to the conclusion that at least part of these issues can be attributed to unreliable and underreported accident data itself and operational and methodological issues to the studies themselves (such as ill-founded operations of the concept of “conflicts” and poor data collection methods) (*Peesapati et al., 2013, Chin & Quek, 1997, Muhlrad, 1982, Oppe, 1977*). On the other hand, there is a significant body of literature that has investigated the relationships between the number of traffic conflicts (operationalized in various ways) and accidents and came to favourable results (*Laureshyn et al., 2016, Zheng et al., 2014c, El-Basyouny & Sayed, 2013, Peesapati et al., 2013, Sacchi et al., 2013, Songchitruksa & Tarko, 2006, Lord, 1996, Brown, 1994, Hydén, 1987*).

Therefore, it seems fair to conclude that the concept of “conflicts” in the broader sense has shown a reasonable degree of relative and process validity as a surrogate measure of safety. Some indicators and techniques (e.g. the Swedish conflict technique) have been somewhat more elaborately validated than others, but no indicator or technique has been proven to outperform the others. Additionally, it must be added that many of the validation studies are relatively old studies, and a critical reassessment needs to show which findings still apply to current practices in surrogate safety studies (using video analysis software) and current traffic conditions (much busier traffic, traffic calming designs, safer vehicles, driver assistance technologies, etc.) Additionally, it seems that human observers can have a bias towards “adjusting” TTC, speed, etc. based on their subjective perception of the dangerousness. For example, it has been shown that many situations that a human observer would judge as an event with two road users on a collision course do not in fact have a collision course when it is measured precisely (*Laureshyn et al., 2016*). It therefore needs to be kept in mind that an automated tool replacing a human observer will likely not produce the same results even though formal definitions and thresholds are kept exactly the same. This can affect the transferability of results from (older) validity studies based on human observations to (newer) studies based on automated observations.

In conclusion, more research around the validity of traffic conflicts as a surrogate measure of safety is strongly recommended, but it should be kept in mind that high product validity does not seem to be a prerequisite to use traffic conflicts as a useful and valid tool for road safety studies. A sufficiently high level of process and/or relative validity allows for a wide range of road safety evaluation and diagnosis activities.

5.2. Safety continuum and continuous safety indicators

There seems to be a blind spot in traffic conflict studies and surrogate measures of safety about the interpretation of the continuous indicators like TTC, TAdv and speed-based indicators that result in time series for each conflict. In traditional TCTs, this data is reduced to a single value per conflict to identify and count serious conflicts, e.g. by applying a threshold on TA or TTC_{min} for example. Other approaches have been tested to derive a single indicator value from its time series, such as 15th centile (*St-Aubin et al., 2015a*). Although interpreting the number of serious conflicts (and less serious conflicts) over an observation period is the most common method to evaluate safety at a site, the complete distribution of indicators such as TTC_{min} can also be analysed, although the conclusions may be more difficult to draw (*St-Aubin et al., 2015a*). This is one of the ways of investigating empirically the safety hierarchy or continuum by ranking the severity of all observed traffic events on the same dimension. Research is still necessary to interpret these distributions and how different parts of the safety hierarchy may relate in different ways to safety (*Svensson, 1998, Svensson & Hydén, 2006, Saunier et al., 2011, Saunier et al., 2010*). Some evidence suggests that events further down the severity hierarchy (further than serious conflicts typically used in traditional TCTs) may actually indicate proper and safe interactive behaviour between the road users, especially at unsignalized intersections, including roundabouts, where interactive behaviour and road user awareness of each other is the intended mode of safe operation (*Svensson & Hydén, 2006*, see also the example of the two types of intersections and their distributions in Figure 2b). Further research is needed on the indicators used to define the safety hierarchy, i.e. how to rank the interactions in a safety hierarchy, and on the interpretation and comparison of these hierarchies, e.g. to identify indicator distributions that are typical for safe and unsafe situations.

There has been some research in the analysis of the continuous time series of road user interactions, instead of deriving only a single value. This has been implemented in the form of the clustering of interactions, in the case of a video dataset of collisions and conflicts, based on various time series such as distance, speed differential and TTC (*Saunier & Mohamed, 2014*). Appropriate similarity measures based on the longest common subsequence (LCSS) are used to compare time series, including their rate of change. The resulting clusters show that some conflicts appear to bear no similarity to observed collisions and should therefore not be used to draw conclusions about safety (e.g. counted as a serious conflict to compute the expected number of accidents). This is one of the only attempts to define empirically process validity, which can lead to better define which traffic events should be used for safety evaluations.

An important point is that such an approach is feasible only in an automated fashion and has been demonstrated over the last ten years through video data collection and computer vision.

5.3. Outcome severity in case of an accident

It was found that the vast majority of surrogate safety measures describe the severity of an event only in terms of nearness to an accident (i.e. the proximity of the involved road users in time and/or space or probability for a collision to occur). The potential outcome severity in case the event had led to an accident is rarely included. From a validity perspective, this can be considered as an important limitation as the holistic reflection of risk is not attained.

It was stated earlier that validity has to be assessed relative to the purposes and circumstances of the study (*Brinberg & McGrath, 1985*). Current road safety policies acknowledge that it is mostly the severe accidents that need to be avoided. Most road safety policy documents set ambitious targets for reducing the number of fatalities in the traffic system and quite a few set ambitious targets for reducing the number of severely injured victims as well. It seems, however, that few of them set explicit targets for reducing the number of slightly injured victims. Additionally, Vision Zero suggests that policymakers and road designers should strive towards a traffic system without fatalities or serious injuries, but acknowledges the fact that a traffic system without any accidents may be difficult to accomplish (*Johansson, 2009*). In that way, the vision “accepts” that property damage only or slight injury accidents may still happen.

Therefore, in order to be “valid” to support such policies, applied surrogate measures of safety need to better reflect the outcome severity. By using surrogate safety measures that reflect the occurrence of any accidents instead of the most severe ones, there is a risk that conclusions cannot meet the demands of road safety policy. A strong recommendation is therefore to further develop and apply surrogate safety measures that take the severity of the potential outcome into account as well.

5.4. Subjective component in safety studies

The subjective perception of traffic situations by road users and observers most probably includes both the probability of a collision and its possible consequences. It is hard to deny that, being a pedestrian, a meeting with a large vehicle is experienced as much riskier than a meeting with a cyclist approaching with the same speed and at the same distance. In this respect, the subjective perception comes closer to the “true” severity compared to many objective measures that most often reflect just one of the severity aspects.

There are some indications that such holistic approach to severity improves the accuracy of the safety estimates. If an objective indicator “fails” in some situations, an expert really “feels” that an event is serious and “adjusts” the measurements. However, such subjective corrections are to a high degree “a black box” and their reliability will always be questioned. To avoid such critic automated data collection methods are to be preferred.

Automation requires clearly defined objective measures that can describe a traffic situation in a holistic manner (similar to subjective perception of the “dangerousness”). However, the subjective “dangerousness” might be closer to the “true” severity, but still not equal to it. Therefore, to mimic subjective judgements with objective indicators might be a mid-target, but the ultimate goal is still to find (and validate) universal objective indicator(s) that produce accurate estimates of safety, regardless to types of manoeuvre and road users involved and other conditions.

Since the subjective component played very important role in the application of traffic conflict techniques until quite recent time (both due to use of subjective scores and use of observers to estimate the objective measures), the conclusion about the various objective indicators in the old validation studies should be taken with some criticism. It might be the case the measured with technical tools only (i.e. subjective component removed), their performance will change. New validation studies with the enhanced technical tools are thus necessary.

The current status of the automated tools for safety analysis still requires to some degree expert judgements. While further enhancement of the tools for taking measurements and calculation of safety indicators is still necessary, one can ask a question whether the ultimate goal in such development is a completely human-free decision system. Reality shows that conflicts and accidents often constitute “odd” or “uncommon” situations, which might not have been considered when a conflict indicator or technique was created (meaning the risk of a “biased” safety diagnosis). Observation of recorded safety-relevant traffic situations (especially in a “condensed” form video with less relevant “normal” traffic already been removed) usually results in much better understanding in the functioning of the traffic system and new insights in how the safety situation can be improved.

5.5. Surrogates safety measures for VRUs

Based on the results presented in section 4.4 it seems that most TCTs and indicators have been applied in VRU safety studies at least on some occasions. There are, however, some doubts whether there have been reflections on how suitable a TCT or indicator is for studying VRU safety. Some studies seem to apply a “standard” TCT or indicator without a second thought on the differences between car-car and car-VRU situations (*Hussein et al., 2015, Zaki et al., 2013*), while others consider the specific characteristics of the VRU in traffic (*Laureshyn et al., 2017, Bagdadi, 2013, Cafiso et al., 2011*). There are also a number of attempts of modifications of the “standard” TCTs and indicators to better reflect VRU properties (*Kaparias et al., 2010, Ismail et al., 2009, Malkhamah et al., 2005, Shbeeb, 2000*).

5.5.1. Collision risk

The difference in the potential for evasive action makes estimates of risk of collision in VRU-car situations quite different to car-car situations. Pedestrian speeds are relatively low and they can stop almost instantly. Bicycle and moped speeds are higher, but they can, on the other hand, more easily swerve to avoid a potential collision. How much these differences should be taken into account depends on how often it is the VRU who takes the evasive action. The studies of vehicle-pedestrian conflicts carried out by *Shbeeb (2000)* indicated that in 20% in Sweden and in 50% in Jordan it is the pedestrian who takes the evasive action. This is a relatively high share, and if in all these situations the underlying assumptions for risk estimates are wrong, quite misleading interpretations might be expected.

Two main types of indicators are used to estimate the risk of collision - the type of indicators that measure “nearness” between road users (such as TTC or PET) and the type of indicators that describe the “harshness” of the necessary evasive action to avoid a collision (such as DST).

For the first type, it is actually important who is about to collide into whom. In example in Figure 39, both situations have the same low value of PET. However, it is counter-intuitive to assume that the risk of a collision in both situation is the same – the situation A appears to be quite risky as, given high speed of the car and unavoidable braking distance travelled before the final stop, a delay by a fraction of a second by the pedestrian might lead to a collision. The situation B is quite “normal” as the pedestrian can stop at a very close distance from the car.

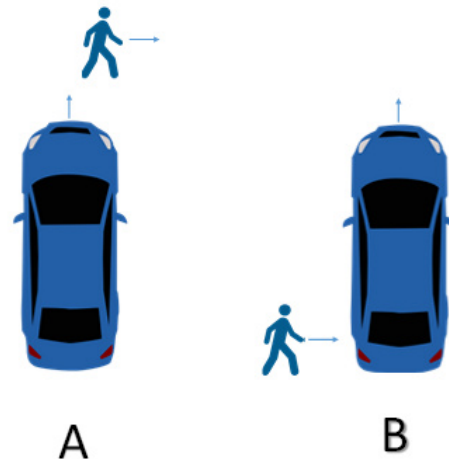


Figure 39. Two situations with the same PET value but different severity.

The indicators of the second type are also greatly affected since the types and intensity of the evasive actions for VRUs and motor vehicles are so different. This becomes particularly the case for deceleration-based indicators since these usually use assumptions regarding cars’ braking capabilities which of course are not relevant for VRUs. Closely related to that is the problem of the thresholds used to distinguish between the “severe” and “non-severe” events in traffic as they are also often based on the time/space required for a motor vehicle to stop.

5.5.2. Injury risk

The absence of a protective “shell” for a VRU results in much higher risk for injuries in case of a collision. Moreover, the vulnerability varies within the same type of VRUs as, for example, older persons have much higher risk for fractures or other serious injuries even in collisions at very low speeds.

Most of the indicators used to estimate the risk of collision fail to consider the risk of injury. Neither TTC nor PET, the most frequently used indicators, utilise information on the types of the road users, their speeds or vulnerability.

The aspect of “who collides into whom” is again important. In the example in Figure 39, the situation A appears to have much higher risk of a serious injury for the pedestrian since the car possesses high kinetic energy to dissipate during the collision, while in situation B the risk for injury is lower.

This vulnerability of the road users is partly considered in some TCTs that involve a subjective judgement of the potential consequences (e.g. *Kraay et al., 2013, Brown, 1994*). Some attempts to develop objective indicators that take into account the speed and mass differences have also been suggested (*Bagdadi, 2013, Shelby, 2011*). Using the weight of road users as a measure of their vulnerability still has some limitations.

For example, it does not allow to differentiate a young and old pedestrian, a bicyclist in a helmet or without which all could have different risks of injury.

5.5.3. Weighting together collision and injury risks

It is still an open question on the two components of the severity (risk of a collision and risk of an injury *in case* of a collision) should be weight together. Many conflict techniques use an objective indicator describing the first component and a subjective score to describe the second. To put them together and get a final conflict severity score usually some tables are used (see for example DOCTOR's definitions in Table 5), but the suggested score definitions seem be quite arbitrary and lack a clear motivation.

There are several attempts to define a complex objective severity measure that integrates different aspects of the severity, such as combinations of TTC_{min} , PET, DST and GT (*Ismail et al., 2010b*), Delta-V, TA and braking assumptions (*Bagdadi, 2013*), pedestrian risk index (*Cafiso et al., 2011*) and the extended Delta-V (*Laureshyn et al., 2017*). All these measures have some limitations and the issues named earlier when it comes to applications on VRU-situations are not fully addressed.

Another important aspect is the validation of such new measures which is nearly absent. It appears that the most validated indicator is the subjective score set by trained traffic conflict observers (*Svensson, 1992*). Human observers seem to have a holistic understanding of the severity and adjust the score properly when it comes to VRU-situations. On the other hand, subjective measures, as previously been discussed, have a number of other weaknesses, particularly when it comes to their transfer into an automated data collection toolbox.

5.6. Strengths and limitations (of this study)

The strength of this study is that it is a unique attempt to provide a systematic overview of the scientific literature around surrogate safety studies. It provides broad information about the field that can be used for a wide variety of goals (*Armstrong et al., 2011*). It can help to reduce duplication of research efforts and can guide future research in the field (*van Wee & Banister, 2016, Armstrong et al., 2011*).

A limitation of the study is the high number of publications for which no full text could be accessed. Because of this high number of missing publications, it cannot be claimed that this scoping review provides a complete overview of the existing literature. This is partly the consequence of the decision to include all types of publications found from the databases, including journal papers, research reports, conference papers, doctoral dissertations, book chapters, etc. A solution that could have partly avoided this problem would have been to limit the review to journal papers only. Limiting the included publications in a scoping review to journal papers only is not uncommon and leads to much higher percentages of found full texts because journal papers are more easily accessible than the other types of publications (*Pham et al., 2014*). However, we know from experience that many influential studies in this domain (especially older studies) have not been published as journal papers. It has therefore been decided not to limit the type of publication. While this has somewhat come at the cost of having less of a non-biased sample of publications than when we would have focused on journal publications only, we believe that it has strongly improved the content of the study.

A limitation of the wide scope of a scoping review, is that it does not allow to go into a very high level of detail into the different aspects. The use of a predefined code book to

systematically collect the content of the papers inevitably provides a strong simplification of reality. Future research could narrow the focus to specific subtopics of traffic conflict literature and deal with them into more detail. Additionally, it would be useful to conduct systematic reviews of related research fields (including naturalistic driving and microsimulation studies) with the specific focus of identifying aspects that can be of use in traffic conflict studies (such as new indicators, tests of threshold values, methods of data analysis, etc.).

6. Conclusions

In recent years, research and use of surrogate safety indicators has increased significantly. Often surrogates are used in combination with (semi-)automated video analysis methods. The applied indicators (such as TTC or PET) and traffic conflict techniques show an overwhelming variety and creativity. This alone reflects very distinctly the traffic safety research community's strong interest and need for surrogates of safety. Limitations in accident data, being the primary driving force for using surrogates in traffic safety analyses, are even more striking when it comes to VRU since accidents involving VRU are heavily underreported.

The relation between safety surrogates and accidents is an important aspect regarding the usability of the surrogates. The conducted validation studies are few, most of them are relatively old and use data collected by human observers. The transferability between these studies and studies using automated sensor techniques is uncertain. Therefore, more research around the validity of surrogates is strongly recommended. It should be kept in mind though that high product validity does not have to be a prerequisite to use surrogate measures as a useful and valid tool for road safety studies. A sufficiently high level of relative validity allows for a wide range of road safety evaluation and diagnosis activities while process validity allows for better understanding of the underlying factors leading to accidents.

It was also concluded that applied surrogate measures of safety need to reflect outcome severity in a better way. Bearing in mind VRUs higher chance of injuries if involved in an accident, this aspect cannot be simply ignored. Measuring the injury potential with the surrogate measures is very much in line with the philosophy of Vision Zero that sets highest priority on elimination of fatalities and severe injuries rather than prevention of any kind of accidents.

The subjective component of some traffic conflict techniques seems to improve the estimate of injury risk and may therefore be regarded as potentially very useful indicator for VRU safety studies. On the other hand, subjective measures have a number of other weaknesses, not the least when it comes to their transferability to automated technology.

7. Acknowledgements

The authors would like to thank prof. **Andrew Tarko** (Purdue University, USA) and prof. **Shalom Hakkert** (Technion, Israel) for their valuable comments on this manuscript.

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.

8. References

- af Wåhlberg, A. E. (2004) The stability of driver acceleration behavior, and a replication of its relation to bus accidents. *Accident Analysis & Prevention* 36 (1), pp. 83-92. [http://dx.doi.org/10.1016/S0001-4575\(02\)00130-6](http://dx.doi.org/10.1016/S0001-4575(02)00130-6).
- Agah, M. M., A. Sri (2013) Influence of red motorcycle box to the traffic conflict and traffic flow at the Ahmad Yani-Laswi signalized intersection. 16th Road Safety on Four Continents conference, Beijing, China, 15-17 May 2013.
- Alhajyaseen, W. K. M. (2015) The Integration of Conflict Probability and Severity for the Safety Assessment of Intersections. *Arabian Journal for Science and Engineering* 40 (2), pp. 421-430. <http://dx.doi.org/10.1007/s13369-014-1553-1>.
- Allen, B. L., B. T. Shin, P. J. Cooper (1978) Analysis of traffic conflicts and collisions. *Transportation Research Record* 667, pp. 67-74.
- Alsop, J., J. Langley (2001) Under-Reporting of Motor Vehicle Traffic Crash Victims in New Zealand. *Accident Analysis & Prevention* 33 (3), pp. 353-359. [http://dx.doi.org/10.1016/S0001-4575\(00\)00049-X](http://dx.doi.org/10.1016/S0001-4575(00)00049-X).
- Amoros, E., J.-L. Martin, B. Laumon (2006) Under-Reporting of Road Crash Casualties in France. *Accident Analysis & Prevention* 38 (4), pp. 627-635. <http://dx.doi.org/10.1016/j.aap.2005.11.006>.
- Armstrong, R., B. J. Hall, J. Doyle, E. Waters (2011) 'Scoping the scope' of a cochrane review. *Journal of Public Health* 33 (1), pp. 147-150. <http://dx.doi.org/10.1093/pubmed/fdr015>.
- Asmussen, E. (Ed.) (1984) Proceedings of the NATO advanced research workshop on international calibration study of traffic conflict techniques. 25-27 May 1983, Copenhagen, Denmark. NATO ASI Series F: Computer and System Science vol. 5.
- Autey, J., T. Sayed, M. H. Zaki (2012) Safety evaluation of right-turn smart channels using automated traffic conflict analysis. *Accident Analysis & Prevention* 45, pp. 120–130. <http://dx.doi.org/10.1016/j.aap.2011.11.015>.
- Bagdadi, O. (2013) Estimation of the severity of safety critical events. *Accident Analysis & Prevention* 50, pp. 167-174. <http://dx.doi.org/10.1016/j.aap.2012.04.007>.
- Bagdadi, O., A. Várhelyi (2011) Jerky driving – an indicator for accident proneness. *Accident Analysis & Prevention* 43, pp. 1359–1363. <http://dx.doi.org/10.1016/j.aap.2011.02.009>.
- Bagdadi, O., A. Várhelyi (2013) Development of a method for detecting jerks in safety critical events. *Accident Analysis & Prevention* 50, pp. 83-91. <http://dx.doi.org/10.1016/j.aap.2012.03.032>.
- Baguley, C. J. (1982) The British Traffic Conflicts Technique: State of the Art Report. Third International Workshop on Traffic Conflicts Techniques, Leidschendam, The Netherlands, April 1982.
- Baguley, C. J. (1984) The British traffic conflicts technique. NATO advanced research workshop on international calibration study of traffic conflict techniques, Copenhagen, Denmark, 25-27 May 1983.
- Baker, W. T. (1972) An evaluation of the traffic conflicts technique. 51st Annual Meeting of the Highway Research Board.
- Balasha, D., A. S. Hakkert, M. Livneh (1979) A quantitative definition of the near accident concept. Second international traffic conflicts technique workshop, Paris, France, 10-12 May 1979.

- Bella, F. (2009) Can Driving Simulators Contribute to Solving Critical Issues in Geometric Design? *Transportation Research Record* 2138, pp. 120-126. <http://dx.doi.org/10.3141/2138-16>.
- Brill, E. A. (1971) A car following model relating reaction time and temporal headways to accident frequency. Naval Postgraduate School.
- Brinberg, D., J. E. McGrath (1985) *Validity and the research process*. SAGE Publications, Inc, Beverly Hills.
- Brown, G., A. Chau, B. DeCastilho, F. Navin (1984) *Traffic conflict procedure and validation*. University of British Columbia. Department of Civil Engineering, Transport Safety Research Group.
- Brown, G. R. (1994) *Traffic conflicts for road user safety studies*. *Canadian Journal of Civil Engineering*.
- Burnett, N., A. Sharma (2011) Role of Information on Probability of Traffic Conflict on the Onset of Yellow. *Advances in Transportation Studies (RSS2011 Special Issue)*, pp. 29-40.
- Cafiso, S., A. G. García, R. Cavarra, M. A. R. Rojas (2011) *Crosswalk Safety Evaluation Using a Pedestrian Risk Index as Traffic Conflict Measure*. 3rd International Conference on Road Safety and Simulation, Indianapolis, USA, 14-16 September 2011.
- Campbell, K. L., H. C. Joksch, P. E. Green (1996) *A Bridging Analysis for Estimating the Benefits of Active Safety Technologies*. The University of Michigan, Transportation Research Institute. UMTRI-96-18.
- Carmines, E. G., R. A. Zeller (1979) *Reliability and validity assessment*. SAGE Publications, Inc, Beverly Hills.
- Chin, H. C., S. T. Quek (1997) Measurement of traffic conflicts. *Safety Science* 26 (3), pp. 169-185. [http://dx.doi.org/10.1016/S0925-7535\(97\)00041-6](http://dx.doi.org/10.1016/S0925-7535(97)00041-6).
- Chin, H. C., S. T. Quek, R. L. Cheu (1992) Quantitative examination of traffic conflicts. *Transportation Research Record* 1376, pp. 86-91.
- Cooper, P. J. (1973) *Predicting intersection accidents: the use of conflicts and other models based on traffic flow parameters to predict accident experience at non-signalised intersections*. Ministry of Transport, Ottawa, Canada.
- Cooper, P. J. (1984) *Experience with traffic conflicts in Canada with emphasis on "Post Encroachment Time" techniques*. NATO advanced research workshop on international calibration study of traffic conflict techniques, 25-27 May 1983.
- Davis, G. A., J. Hourdos, H. Xiong, I. Chatterjee (2011) Outline for a causal model of traffic conflicts and crashes. *Accident Analysis & Prevention* 43 (6), pp. 1907-1919. <http://dx.doi.org/10.1016/j.aap.2011.05.001>.
- De Ceunynck, T., C. Ariën, K. Brijs, T. Brijs, K. V. Vlierden, J. Kuppens, M. V. D. Linden, G. Wets (2015) Proactive evaluation of traffic signs using a traffic sign simulator. *European Journal of Transport & Infrastructure Research* 15 (2), pp. 184-204.
- El-Basyouny, K., T. Sayed (2013) Safety performance functions using traffic conflicts. *Safety Science* 51 (1), pp. 160-164. <http://dx.doi.org/10.1016/j.ssci.2012.04.015>.
- Elvik, R., A. Høye, T. Vaa, M. Sørensen (2009) *The handbook of road safety measures*. Emerald Group Publishing Limited, Bingley, UK.
- Erke, H. (1984) *The traffic conflict technique of the Federal Republic of Germany*. NATO advanced research workshop on international calibration study of traffic conflict techniques, Copenhagen, Denmark.
- Evans, L. (1991) *Traffic Safety and the Driver*. Van Nostrand Reinhold, New York.

- Evans, L. (1994) Driver injury and fatality risk in two-car crashes versus mass ratio inferred using Newtonian mechanics. *Accident Analysis & Prevention* 26 (5), pp. 609–616. [http://doi.org/10.1016/0001-4575\(94\)90022-1](http://doi.org/10.1016/0001-4575(94)90022-1).
- Evans, L. (2004) *Traffic Safety*. Science Serving Society, Michigan.
- Evans, L., P. Wasielewski (1982) Do accident-involved drivers exhibit riskier everyday driving behavior? *Accident Analysis & Prevention* 14 (1), pp. 57-64. [http://dx.doi.org/10.1016/0001-4575\(82\)90007-0](http://dx.doi.org/10.1016/0001-4575(82)90007-0).
- Evans, L., P. Wasielewski (1983) Risky driving related to driver and vehicle characteristics. *Accident Analysis & Prevention* 15 (2), pp. 121-136. [http://dx.doi.org/10.1016/0001-4575\(83\)90068-4](http://dx.doi.org/10.1016/0001-4575(83)90068-4).
- Fisher, D. L., M. Rizzo, J. Caird, J. D. Lee (2011) *Handbook of Driving Simulation for Engineering, Medicine, and Psychology*. CRC Press. ISBN 9781420061000.
- Forbes, T. W. (1957) Analysis of "Near Accident" Reports. *Highway Research Bulletin* 152.
- Fyhri, A., H. B. Sundfør, T. Bjørnskau, A. Laureshyn (2016) Safety in numbers for cyclists—conclusions from a multidisciplinary study of seasonal change in interplay and conflict. *Accident Analysis & Prevention* (in press). <http://dx.doi.org/10.1016/j.aap.2016.04.039>.
- Gabauer, D. J., H. C. Gabler (2008) Comparison of roadside crash injury metrics using event data recorders. *Accident Analysis & Prevention* 40 (2), pp. 548–558. <http://doi.org/10.1016/j.aap.2007.08.011>.
- Gettman, D., L. Head (2003) Surrogate safety measures from traffic simulation models. U. S. Department of Transportation, Federal Highway Administration. Report No. FHWA-RD-03-050.
- Glennon, J. C., W. D. Glauz, M. C. Sharp, B. A. Thorson (1977) Critique of the traffic-conflict technique. *Transportation Research Record* 630, pp. 32-38.
- Glennon, J. C., B. A. Thorson (1975) Evaluation of traffic conflict technique. Midwest Research Institute.
- Godley, S. T., T. J. Triggs, B. N. Fildes (2002) Driving simulator validation for speed research. *Accident Analysis & Prevention* 34 (5), pp. 589-600. [http://dx.doi.org/10.1016/S0001-4575\(01\)00056-2](http://dx.doi.org/10.1016/S0001-4575(01)00056-2).
- Grayson, G. B. (1984) The Malmö study: a calibration of traffic conflict techniques. SWOV. R-84-12.
- Grayson, G. B., A. S. Hakkert (1987) Accident analysis and conflict behaviour *in* J. A. Rothengatter, R. A. de Bruin (Eds.) *Road users and traffic safety* (pp. 27-59). Van Gorcum. ISBN 9023223160.
- Guido, G., A. Vitale, V. Gallelli, G. Figliomeni (2013) Level of safety on two-lane undivided rural highways. 2nd International Conference on Civil Engineering and Transportation, Guilin, China, 27-28 October 2012. <http://dx.doi.org/10.4028/www.scientific.net/AMM.253-255.1705>.
- Güttinger, V. A. (1982) From Accidents to Conflicts: Alternative Safety Measurement. Third International Workshop on Traffic Conflicts Techniques, Leidschendam, The Netherlands, April 1982.
- Gårder, P. (1982) Konfliktstudier i landsvägskorsningar (*in Swedish*) Conflict studies in rural intersections. Institutionen för Trafikteknik, LTH, Lund. Bulletin 42.
- Hakkert, A. S., D. Balasha, M. Livneh, J. Prashker (1977) Irregularities in Traffic Flow as an Estimate of Risk at Intersections. First Workshop on Traffic Conflicts, Oslo, Norway, 26-27 September 1977.

- Hansson, A. (1975) Studies in driver behaviour, with applications in traffic design and planning. Two examples (*in Swedish*). Lund Institute of Technology, Department of Traffic Planning and Engineering, University of Lund. Bulletin 11.
- Hauer, E. (1978) Traffic conflict surveys: some study design considerations. Transport and Road Research Laboratory. Supplementary report 352.
- Hauer, E. (1982) Traffic conflicts and exposure. *Accident Analysis & Prevention* 14 (5), pp. 359-364. [http://dx.doi.org/10.1016/0001-4575\(82\)90014-8](http://dx.doi.org/10.1016/0001-4575(82)90014-8).
- Hauer, E. (1997) Observational before-after studies in road safety. Pergamon. ISBN 9780080430539.
- Hauer, E., P. Gårder (1986) Research into the validity of the traffic conflict technique. *Accident Analysis & Prevention* 18 (6), pp. 471-481. [http://dx.doi.org/10.1016/0001-4575\(86\)90020-5](http://dx.doi.org/10.1016/0001-4575(86)90020-5).
- Hayward, J. C. (1971) Near misses as a measure of safety at urban intersections. Master thesis, The Pennsylvania State University, Department of Civil Engineering.
- Hupfer, C. (1997) Deceleration to Safety Time (DST) - a useful figure to evaluate traffic safety? 10th ICTCT workshop, Lund, Sweden, 5-7 November 1997.
- Hussein, M., T. Sayed, P. Reyad, L. Kim (2015) Automated Pedestrian Safety Diagnosis and Behavioural Study at a Signalized Intersection in New York City. 94th TRB Annual Meeting, Washington DC, USA, 11-15 January 2015.
- Hydén, C. (1977) A traffic conflict technique for determining risk. University of Lund, Lund Institute of Technology, Department of Traffic Planning and Engineering. LUTVDG/(TVTT-3005)/1-95/(1977) & Bulletin 15 B.
- Hydén, C. (1987) The development of a method for traffic safety evaluation: the Swedish traffic conflict technique. Doctoral thesis, Lund University, Department of Traffic Planning and Engineering.
- ICTCT (2016) International co-operation on Theories and Concepts in Traffic Safety. www.ictct.org/ictct/about-us Acc. 23 September 2016.
- Ismail, K., T. Sayed, N. Saunier (2010a) Automated analysis of pedestrian-vehicle conflicts: context for before-and-after studies. *Transportation Research Record* 2198, pp. 52-64. <http://doi.org/10.3141/2198-07>.
- Ismail, K., T. Sayed, N. Saunier (2010b) Automated safety analysis using video sensors: technology and case studies. Canadian Multidisciplinary Road Safety Conference, Ontario, Canada.
- Ismail, K., T. Sayed, N. Saunier (2011) Methodologies for aggregating indicators of traffic conflict. *Transportation Research Record*, pp. 10-19. <http://doi.org/10.3141/2237-02>.
- Ismail, K., T. Sayed, N. Saunier, C. Lim (2009) Automated analysis of pedestrian-vehicle conflicts using video data. 88th TRB Annual Meeting, Washington DC, USA, 11-15 January 2009.
- Johansson, R. (2009) Vision Zero - implementing a policy for traffic safety. *Safety Science* 47, pp. 826–831. <http://dx.doi.org/10.1016/j.ssci.2008.10.023>.
- Johnson, N. S., H. C. Gabler (2012) Accuracy of a damage-based reconstruction method in NHTSA side crash tests. *Traffic Injury Prevention* 13 (1), pp. 72-80. <http://doi.org/10.1080/15389588.2011.636592>.
- Kaparias, I., M. G. H. Bell, J. Greensted, S. Cheng, A. Miri, C. Taylor, B. Mount (2010) Development and implementation of a vehicle-pedestrian conflict analysis method: Adaptation of a vehicle-vehicle technique. *Transportation Research Record*, pp. 75-82. <http://dx.doi.org/10.3141/2198-09>.

- Kiefer, R. J., C. A. Flannagan, C. J. Jerome (2006) Time-to-collision judgments under realistic driving conditions. *Human Factors* 48 (2), pp. 334–345.
<http://dx.doi.org/10.1518/001872006777724499>.
- Kiefer, R. J., D. J. Le Blanc, C. A. Flannagan (2005) Developing an inverse time-to-collision crash alert timing approach based on drivers' last-second braking and steering judgments. *Accident Analysis and Prevention* 37, pp. 295-303.
<http://dx.doi.org/10.1016/j.aap.2004.09.003>.
- Kocárková, D. (2012) Traffic Conflict Techniques in Czech Republic. SIIV - 5th International Congress - Sustainability of Road Infrastructures, Rome, Italy, 29-31 October 2012. <http://dx.doi.org/10.1016/j.sbspro.2012.09.952>.
- Kraay, J. H., A. R. A. van der Horst, S. Oppe (2013) Manual conflict observation technique DOCTOR. Foundation Road safety for all, The Netherlands.
- Kruysse, H. W. (1991) The subjective evaluation of traffic conflicts based on an internal concept of dangerousness. *Accident Analysis & Prevention* 23 (1), pp. 53-65.
[http://dx.doi.org/10.1016/0001-4575\(91\)90035-4](http://dx.doi.org/10.1016/0001-4575(91)90035-4).
- Kruysse, H. W., G. J. Wijnhuizen (1992) Why are experts not better in judging the danger of filmed traffic conflicts? *Accident Analysis & Prevention* 26 (3), pp. 227-235. [http://dx.doi.org/10.1016/0001-4575\(92\)90002-Z](http://dx.doi.org/10.1016/0001-4575(92)90002-Z).
- Kuang, Y., X. Qu, S. Wang (2015a) A tree-structured crash surrogate measure for freeways. *Accident Analysis & Prevention* 77, pp. 137-148.
<http://dx.doi.org/10.1016/j.aap.2015.02.007>.
- Kuang, Y., X. Qu, J. Weng, A. Etemad-Shahidi (2015b) How Does the Driver's Perception Reaction Time Affect the Performances of Crash Surrogate Measures? *Plos One* 10 (9). <http://doi.org/10.1371/journal.pone.0138617>.
- Kulmala, R. (1984) The Finnish traffic conflict technique. NATO advanced research workshop on international calibration study of traffic conflict techniques, Copenhagen, Denmark.
- Laureshyn, A. (2010) Application of automated video analysis to road user behaviour. Doctoral thesis, Lund University, Faculty of Engineering, LTH, Department of Technology and Society, Transport & Roads.
- Laureshyn, A., T. De Ceunynck, C. Karlsson, Å. Svensson, S. Daniels (2017) In search of the severity dimension of traffic events: Extended Delta-V as a traffic conflict indicator. *Accident Analysis & Prevention* 98.
<http://dx.doi.org/10.1016/j.aap.2016.09.026>.
- Laureshyn, A., M. d. Goede, N. Saunier, A. Fyhri (2016) Cross-comparison of three surrogate safety methods to diagnose cyclist safety problems at intersections in Norway. *Accident Analysis & Prevention* (in press).
<http://dx.doi.org/10.1016/j.aap.2016.04.035>.
- Laureshyn, A., M. Nilsson, H. Ardö, Å. Svensson, A. Persson (2013) How accurately can we measure from video? 26th ICTCT workshop, Maribor, Slovenia.
- Laureshyn, A., Å. Svensson, C. Hydén (2010) Evaluation of traffic safety, based on micro-level behavioural data: theoretical framework and first implementation. *Accident Analysis & Prevention* 42, pp. 1637-1646.
<http://dx.doi.org/10.1016/j.aap.2010.03.021>.
- Laureshyn, A., K. Åström, K. Brundell-Freij (2009) From speed profile data to analysis of behaviour: classification by pattern recognition techniques. *IATSS Research* 33 (2), pp. 88-98. [http://dx.doi.org/10.1016/S0386-1112\(14\)60247-8](http://dx.doi.org/10.1016/S0386-1112(14)60247-8).
- Lawalata, G. M., H. R. Agah (2011) Traffic conflict analysis as a road safety diagnostic tool for urban road facilities. *International Journal of Technology* 2 (2), pp. 112-121.

- Li, X., J. Chen, H. Wang (2013) Study on Flow Direction Changing Method of Reversible Lanes on Urban Arterial Roadways in China. *Procedia - Social and Behavioral Sciences* 96, pp. 807-816.
<http://dx.doi.org/10.1016/j.sbspro.2013.08.092>.
- Lightburn, A., C. I. Howarth (1979) A study of observer variability and reliability in the detection and grading of traffic conflicts. Second international traffic conflicts technique workshop, Paris, France, 10-12 May 1979.
- Linderholm, L. (1981) Vidareutveckling av konflikttekniken för riskbestämning i trafiken (*in Swedish*) Further development of conflict technique for in traffic risk assessment. Lund University, Dept. of Traffic Planning and Engineering, Lund, Sweden. Rapport 7025.
- Lord, D. (1996) Analysis of Pedestrian Conflicts with Left-Turning Traffic. *Transportation Research Record* 1538, pp. 61-67. <http://dx.doi.org/10.3141/1538-08>.
- Lu, G., M. Liu, Y. Wang, G. Yu (2012) Quantifying the Severity of Traffic Conflict by Assuming Moving Elements as Rectangles at Intersection. *Procedia - Social and Behavioral Sciences* 43, pp. 255-264.
<http://dx.doi.org/10.1016/j.sbspro.2012.04.098>.
- Lua, G., M. Liua, Y. Wanga, G. Yua (2012) Quantifying the severity of traffic conflict by assuming moving elements as rectangles at intersection. 8th International Conference on Traffic and Transportation Studies, Changsha, China, August 1–3, 2012. <http://dx.doi.org/10.1016/j.sbspro.2012.04.098>.
- MacCarley, C. A., C. Ackles, T. Watts (2007) A Study of the Response of Highway Traffic to Dynamic Fog Warning and Speed Advisory Messages. 86th TRB Annual Meeting, Washington DC, USA, 21-25 January 2007.
- Malkhamah, S., M. Tight, F. Montgomery (2005) The development of an automatic method of safety monitoring at Pelican crossings. *Accident Analysis and Prevention* 37, pp. 938-946. <http://dx.doi.org/10.1016/j.aap.2005.04.012>.
- Mays, N., E. Roberts, J. Popay (2001) Synthesising research evidence *in* N. Fulop, P. Allen, A. Clarke, N. Black (Eds.) *Methods for studying the delivery and organisation of health services*. Routledge, London.
- Migletz, D. J., W. D. Glauz, K. M. Bauer (1985) Relationships between traffic conflicts and accidents. U.S. Department of Transportation, Federal Highway Administration. FHWA/RD-84/042.
- Minderhoud, M. M., P. H. L. Bovy (2001) Extended time-to-collision measures for road traffic safety assessment. *Accident Analysis & Prevention* 33 (1), pp. 89-97.
[http://dx.doi.org/10.1016/S0001-4575\(00\)00019-1](http://dx.doi.org/10.1016/S0001-4575(00)00019-1).
- Mohamed, M. G., N. Saunier (2013) Motion Prediction Methods for Surrogate Safety Analysis. *Transportation Research Record* 2386, pp. 168-178.
<http://dx.doi.org/10.3141/2386-19>.
- Morse, P., P. St-Aubin, L. Miranda-Moreno, N. Saunier (2016) Transferability Study of Video Tracking Optimization for Traffic Data Collection and Analysis. 95th TRB Annual Meeting, Washington DC, USA, 10-14 January 2016.
- Mortelmans, J., A. Vits, L. Venstermans, M. van Essche, M. Boogaerts, E. Vergisson (1986) Analyse van de verkeersveiligheid met behulp van de Bijna-Ongevallen methode (*in Dutch*) Analysis of road safety by means of the near-accident method. Katholieke Universiteit Leuven, Departement Bouwkunde.
- Muhlrad, N. (1982) The French Traffic Conflict Technique: a State of the Art Report. Third International Workshop on Traffic Conflicts Techniques, Leidschendam, The Netherlands, April 1982.

- Muhlrad, N., G. Dupre (1984) The French conflict technique. NATO advanced research workshop on international calibration study of traffic conflict techniques, Copenhagen, Denmark.
- Nasab, S. E. J., S. S. Naserlavi, N. Pourmoallem (2015) Investigating the influence of Vehicular Ad-hoc Networks on driver behaviours and traffic accidents. *Travel Behaviour and Society* 2 (3), pp. 156-165.
<http://dx.doi.org/10.1016/j.tbs.2015.01.001>.
- Ni, Y., M. Wang (2015) Behavior-based Method on Pedestrian-Vehicle Conflict Analysis at Intersections. 94th TRB Annual Meeting, Washington DC, USA, 11-15 January 2015.
- Nygård, M. (1999) A method for analysing traffic safety with help of speed profiles. Master thesis, Tampere University of Technology, Department of Civil Engineering.
- Ogawa, K. (2007) An Analysis of Traffic Conflict Phenomenon of Bicycles Using Space Occupancy Index. *Journal of the Eastern Asia Society for Transportation Studies* 7, pp. 1820-1827.
- Oh, C., S. Park, S. G. Ritchie (2006) A method for identifying rear-end collision risks using inductive loop detectors. *Accident Analysis & Prevention* 38 (2), pp. 295-301. <http://dx.doi.org/10.1016/j.aap.2005.09.009>.
- Oh, J., J.-y. Min, M. Kim, H. Cho (2009) Development of an Automatic Traffic Conflict Detection System Based on Image Tracking Technology. *Transportation Research Record* 2129, pp. 45-54. <http://dx.doi.org/10.3141/2129-06>.
- Oppe, S. (1977) What Task is a Traffic Conflicts Technique intended for? First Workshop on Traffic Conflicts, Oslo, Norway, 26-27 September 1977.
- Paddock, R. D., D. E. Spence (1973) The Conflicts Technique: An Accident Prediction Method. Ohio Department of Transportation, Division of Highways.
- Parker, M. R., C. V. Zegeer (1989) Traffic Conflict Techniques for Safety and Operation. US Department of Transportation, Federal Highway Administration. FHWA-IP-88-027.
- Peesapati, L. N., M. P. Hunter, M. O. Rodgers (2013) Evaluation of Postencroachment Time as Surrogate for Opposing Left-Turn Crashes. *Transportation Research Record* 2386, pp. 42–51. <http://dx.doi.org/10.3141/2386-06>.
- Perkins, S. R., J. I. Harris (1967) Criteria for Traffic Conflict Characteristics: Signalized Intersections. General Motors Corporation. Electro-Mechanical Department. Research Publication GMR-632.
- Pham, M. T., A. Rajić, J. D. Greig, J. M. Sargeant, A. Papadopoulos, S. A. McEwen (2014) A scoping review of scoping reviews: advancing the approach and enhancing the consistency. *Research Synthesis Methods* 5 (4), pp. 371-385.
<http://dx.doi.org/10.1002/jrsm.1123>.
- Pietrantonio, H., L. F. B. Tourinho (2006) A Decision-based Criterion for Selecting Parameters in the Evaluation of Pedestrian Safety Problems with the Traffic Conflict Analysis Technique. *Transportation Planning and Technology* 29 (3), pp. pp 183-216.
- Pugh, D. E., T. J. Halpin (1974) Traffic Conflicts in Washington State. Washington State Department of Highways.
- Quaye, K., L. Leden, E. Hauer (1993) Pedestrian accidents and left-turning traffic at signalized intersections. AAA Foundation for Traffic Safety.
- Risser, R., A. Schutzenhofer (1984) Application of traffic conflict technique in Austria. NATO advanced research workshop on international calibration study of traffic conflict techniques, Copenhagen, Denmark.

- Risser, R., W. D. Zuzan, W. Tamme, J. Steinbauer, A. Kaba (1991) Handbuch zur Erhebung von Verkehrskonflikten mit Anleitungen zur Beobachterschulung (*in German*) Handbook for the collection of traffic conflicts with instructions for observers' training. Literas Universitätsverlag.
- Sacchi, E., T. Sayed, P. de Leur (2013) A comparison of collision-based and conflict-based safety evaluations: the case of right-turn smart channels. *Accident Analysis & Prevention* 59, pp. 260-266. <http://doi.org/10.1016/j.aap.2013.06.002>.
- Sakshaug, L., A. Lareshyn, Å. Svensson, C. Hydén (2010) Cyclists in roundabouts – different design solutions. *Accident Analysis & Prevention* 42, pp. 1338-1351. <http://dx.doi.org/10.1016/j.aap.2010.02.015>.
- Saunier, N., H. Ardo, J.-P. Jodoin, A. Lareshyn, M. Nilsson, A. Svensson, L. F. Miranda-Moreno, G.-A. Bilodeau, K. Åstrom (2014) Public video data set for road transportation applications. 93rd TRB Annual Meeting, Washington DC, USA, 12-16 January 2016.
- Saunier, N., M. G. Mohamed (2014) Clustering Surrogate Safety Indicators to Understand Collision Processes. 93rd TRB Annual Meeting, Washington DC, USA, 12-16 January 2014.
- Saunier, N., N. Mourji, B. Agard (2011) Mining Microscopic Data of Vehicle Conflicts and Collisions to Investigate Collision Factors. *Transportation Research Record* 2237, pp. 41-50. <http://dx.doi.org/10.3141/2237-05>.
- Saunier, N., T. Sayed (2008) Probabilistic Framework for Automated Analysis of Exposure to Road Collisions. *Transportation Research Record* 2083, pp. 96-104. <http://dx.doi.org/10.3141/2083-11>.
- Saunier, N., T. Sayed, K. Ismail (2010) Large scale automated analysis of vehicle interactions and collisions. *Transportation Research Record* 2147, pp. 42-50. <http://dx.doi.org/10.3141/2147-06>.
- Saunier, N., T. Sayed, C. Lim (2007) Probabilistic collision prediction for vision-based automated road safety analysis. 10th International IEEE Conference on Intelligent Transportation Systems, Seattle, WA. <http://doi.org/10.1109/ITSC.2007.4357793>.
- Sayed, T., S. Zein (1999) Traffic conflict standards for intersections. *Transportation Planning & Technology* 22, pp. 309-323. <http://dx.doi.org/10.1016/j.ssci.2012.04.015>.
- Shadish, W., T. Cook, D. Campbell (2002) Statistical Conclusion Validity and Internal Validity *in* W. Shadish, T. Cook, D. Campbell (Eds.) *Experimental and Quasi-Experimental Designs for Generalized Causal Inference* (pp. 33-63). Houghton Mifflin, Boston.
- Shbeeb, L. (2000) Development of a traffic conflicts technique for different environments - a comparative study of pedestrian conflicts in Sweden and Jordan. Doctoral thesis, University of Lund, Lund Institute of Technology, Department of Technology and Society, Traffic Engineering.
- Shelby, S. G. (2011) Delta-V as a measure of traffic conflict severity. 90th TRB Annual Meeting, Washington DC, USA, 23-27 January 2011.
- Shinar, D. (1984) The traffic conflict technique: a subjective vs. objective approach. *Journal of Safety Research* 15, pp. 153-157. [http://dx.doi.org/10.1016/0022-4375\(84\)90046-X](http://dx.doi.org/10.1016/0022-4375(84)90046-X).
- Sobhani, A., W. Young, S. Bahrololoom, M. Sarvi (2012) Calculating Time-To-Collision for analysing right turning behaviour at signalised intersections. *Australasian Transport Research Forum* 2012, Perth, Australia, 26-28 September 2012.

- Songchitruksa, P., A. Tarko (2006) The extreme value theory approach to safety estimation. *Accident Analysis and Prevention* 38, pp. 811-822.
<http://dx.doi.org/10.1016/j.aap.2006.02.003>.
- Spicer, B. R. (1973) A study of traffic conflicts at six intersections. Transport and Road Research Laboratory (TRRL). Report LR551.
- St-Aubin, P., N. Saunier, L. F. Miranda-Moreno (2015a) Comparison of Various Time-to-Collision Prediction and Aggregation Methods for Surrogate Safety Analysis. 94th TRB Annual Meeting, Washington DC, USA, 11-15 January 2015.
- St-Aubin, P., N. Saunier, L. F. Miranda-Moreno (2015b) Large-scale automated proactive road safety analysis using video data. *Transportation Research Part C* 58 (B), pp. 363-379. <http://dx.doi.org/10.1016/j.trc.2015.04.007>.
- Svensson, Å. (1992) Vidareutveckling och validering av den svenska konflikttekniken (*in Swedish*) Further development and validation of the Swedish traffic conflict technique. Lund University, Institute of Technology, Dept. of Traffic Planning & Engineering. (unpublished).
- Svensson, Å. (1998) A method for analysing the traffic process in a safety perspective. Doctoral thesis, University of Lund, Lund Institute of Technology, Department of Traffic Planning and Engineering.
- Svensson, Å., C. Hydén (2006) Estimating the severity of safety related behaviour. *Accident Analysis & Prevention* 38, pp. 379-385.
<http://dx.doi.org/10.1016/j.aap.2005.10.009>.
- Svensson, Å., J. Pauna-Gren (2015) Safety at cycle crossings – the relationship between motor vehicle drivers' yielding behaviour and cyclists' traffic safety. 28th ICTCT Workshop, Ashdod, Israel, 29 – 30 October 2015.
- Tageldin, A., T. Sayed, X. Wang (2015) Can Time Proximity Measures Be Used as Safety Indicators in All Driving Cultures? A Case Study of Motorcycle Safety in China. 94th TRB Annual Meeting, Washington DC, US, 11-15 January 2015.
- Tak, S., S. Kim, H. Yeo (2015) Development of a Deceleration-Based Surrogate Safety Measure for Rear-End Collision Risk. *IEEE Transactions on Intelligent Transportation Systems* 16 (5), pp. 2435-2445.
<http://doi.org/10.1109/tits.2015.2409374>.
- Tarko, A., P. Songchitruksa (2005) Measuring road safety. Road Safety on Four Continents, Warsaw, Poland, 5-7 October 2005.
- Tarko, A. P. (2012) Use of crash surrogates and exceedance statistics to estimate road safety. *Accident Analysis & Prevention* 45, pp. 230-240.
<http://dx.doi.org/10.1016/j.aap.2011.07.008>.
- Tarko, A. P., M. A. Romero, V. K. Bandaru, K. B. Ariyur, C. G. Lizarazo (2017) Feasibility of Tracking Vehicles at Intersections with a Low-end LiDAR. 96th TRB Annual Meeting, 8-12 January 2017.
- Tiwari, G., D. Mohan, J. Fazio (1998) Conflict analysis for prediction of fatal crash locations in mixed traffic streams. *Accident Analysis & Prevention* 30 (2), pp. 207-215. [http://dx.doi.org/10.1016/S0001-4575\(97\)00082-1](http://dx.doi.org/10.1016/S0001-4575(97)00082-1).
- TRRL (1980) Traffic conflicts and accidents at road junctions. TRRL Leaflet LF918.
- Törnros, J. (1998) Driving behaviour in a real and a simulated road tunnel—a validation study. *Accident Analysis & Prevention* 30 (4), pp. 497-503.
[http://dx.doi.org/10.1016/S0001-4575\(97\)00099-7](http://dx.doi.org/10.1016/S0001-4575(97)00099-7).
- Uno, N., Y. Iida, S. Itsubo, S. Yasuhara (2002) A microscopic analysis of traffic conflict caused by lane-changing vehicle at weaving section. 13th Mini-EURO Conference: Handling Uncertainty in the Analysis of Traffic and Transportation Systems, Bari, Italy.

- van der Horst, A. R. A., M. de Goede, S. de Hair-Buijssen, R. Methorst (2014) Traffic conflicts on bicycle paths: A systematic observation of behaviour from video. *Accident Analysis & Prevention* 62, pp. 358-368. <http://dx.doi.org/10.1016/j.aap.2013.04.005>.
- van der Horst, A. R. A. (1984) The ICTCT calibration study at Malmö: a quantitative analysis of video recordings. Institute for Perception TNO. IZF 1984-37.
- van der Horst, A. R. A. (1990) A time-based analysis of road user behaviour in normal and critical encounters. Doctoral thesis, Delft University of Technology.
- van der Horst, A. R. A. (1991) Time-to-Collision as a cue for decision making in braking *in* A. G. Gale (Ed.) *Vision in vehicles - III* (pp. 19-26). Elsevier Science Publishers B.V., Amsterdam.
- van der Horst, A. R. A. (2007) Time-related measures for modelling risk in driver behaviour *in* P. C. Cacciabue (Ed.) *Modelling driving behaviour in automotive environments* (pp. 236-252). Springer.
- van der Horst, A. R. A. (2013) Video-recording of accidents, conflicts and road user behaviour: a step forward in traffic safety research. 8. Gemeinsames Symposium der DGVM und DGVP, Hamburg, Germany, 7-8 September 2012.
- van der Horst, A. R. A., R. J. Kraay (1986) The Dutch conflict technique – DOCTOR. Workshop on traffic conflicts and other intermediate measures, Budapest, Hungary, 8-10 September 1986.
- van der Horst, A. R. A., M. H. Martens (2007) Wegbeeldanalyses Integrale Aanpak Analyse Verkeersongevallen (*in Dutch*) Road scene analyses integral approach traffic accidents analysis. TNO Human Factors. TNO-DV 2007 D068.
- van der Horst, A. R. A., M. C. Thierry, J. M. Vet, A. F. Rahman (2016) An evaluation of speed management measures in Bangladesh based upon alternative accident recording, speed measurements, and DOCTOR traffic conflict observations. *Transportation Research Part F* (in press). <http://dx.doi.org/10.1016/j.trf.2016.05.006>.
- van Wee, B., D. Banister (2016) How to write a literature review paper? *Transport Reviews* 36 (2), pp. 278-288. <http://dx.doi.org/10.1080/01441647.2015.1065456>.
- van Winsum, W., K. A. Brookhuis, D. de Waard (2000) A comparison of different ways to approximate time-to-lane crossing (TLC) during car driving. *Accident Analysis and Prevention* 32, pp. 47-56. [http://dx.doi.org/10.1016/S0001-4575\(99\)00048-2](http://dx.doi.org/10.1016/S0001-4575(99)00048-2).
- Wang, C., N. Stamatiadis (2014) Evaluation of a simulation-based surrogate safety metric. *Accident Analysis & Prevention* 71, pp. 82-92. <http://dx.doi.org/10.1016/j.aap.2014.05.004>.
- Ward, J. R., G. Agamennoni, S. Worrall, A. Bender, E. Nebot (2015) Extending Time-to-Collision for probabilistic reasoning in general traffic scenarios. *Transportation Research Part C* 51, pp. 66-82. <http://dx.doi.org/10.1016/j.trc.2014.11.002>.
- Várhelyi, A. (1998) Driver's speed behaviour at a zebra crossing: a case study. *Accident Analysis & Prevention* 30 (6), pp. 731-743. [http://dx.doi.org/10.1016/S0001-4575\(98\)00026-8](http://dx.doi.org/10.1016/S0001-4575(98)00026-8).
- Wei, F., W. Guo, X. Liu, C. Liang, T. Feng (2014) Left-turning vehicle trajectory modeling and guide line setting at the intersection. *Discrete Dynamics in Nature and Society* 2014. <http://dx.doi.org/10.1155/2014/950219>.
- Williams, M. J. (1981) Validity of the traffic conflict technique. *Accident Analysis & Prevention* 13, pp. 133-145. [http://dx.doi.org/10.1016/0001-4575\(81\)90025-7](http://dx.doi.org/10.1016/0001-4575(81)90025-7).
- Wilson, M. G., J. N. Lavis, A. Guta (2012) Community-based organizations in the health sector: a scoping review. *Health Research Policy and Systems* 10 (36). <http://dx.doi.org/10.1186/1478-4505-10-36>.

- Vogel, K. (2003) A comparison of headway and time to collision as safety indicators. *Accident Analysis and Prevention* 38, pp. 427-433. [http://dx.doi.org/10.1016/S0001-4575\(02\)00022-2](http://dx.doi.org/10.1016/S0001-4575(02)00022-2).
- Wu, K.-F., P. P. Jovanis (2012) Crashes and crash-surrogate events: Exploratory modeling with naturalistic driving data. *Accident Analysis & Prevention* 45, pp. 507-516. <http://dx.doi.org/10.1016/j.aap.2011.09.002>.
- Xin-wei, L., W. Xiao-fei, F. Xin-sha (2013) Research on Driving Risk Model of Freeway Interchange Entrance Area for Accident Prevention. *Procedia - Social and Behavioral Sciences* 96, pp. 25-30. <http://dx.doi.org/10.1016/j.sbspro.2013.08.006>.
- Yan, X., M. Abdel-Aty, E. Radwan, X. Wang, P. Chilakapati (2008) Validating a driving simulator using surrogate safety measures. *Accident Analysis & Prevention* 40 (1), pp. 274-288. <http://dx.doi.org/10.1016/j.aap.2007.06.007>.
- Zaki, M., T. Sayed, A. Tageldin, M. Hussein (2013) Application of computer vision to diagnosis of pedestrian safety issues. *Transportation Research Record* 2393, pp. 75-84. <http://doi.org/10.3141/2393-09>.
- Zaki, M. H., T. Sayed, K. Shaaban (2014) Use of drivers' jerk profiles in computer vision-based traffic safety evaluations. *Transportation Research Record* 2434, pp. 103-112. <http://doi.org/10.3141/2434-13>.
- Zhao, J., H. Zhou, H. Chen, P. Hsu (2011) Safety Effects of Left Exit on Freeway: A Case Study in Florida, USA. *Advances in Transportation Studies* 25, pp. 75-90.
- Zheng, L., K. Ismail, X. Meng (2014a) Freeway safety estimation using extreme value theory approaches: a comparative study. *Accident Analysis & Prevention* 62, pp. 32-41. <http://dx.doi.org/10.1016/j.aap.2013.09.006>.
- Zheng, L., K. Ismail, X. Meng (2014b) Shifted Gamma-Generalized Pareto Distribution model to map the safety continuum and estimate crashes. *Safety Science* 64, pp. 155-162. <http://dx.doi.org/10.1016/j.ssci.2013.12.003>.
- Zheng, L., K. Ismail, X. Meng (2014c) Traffic conflict techniques for road safety analysis: open questions and some insights. *Canadian Journal of Civil Engineering* 41 (7), pp. 633-641. <http://doi.org/10.1139/cjce-2013-0558>.