Self-reports of traffic accidents

a critical evaluation of information validity, survey design and limitations of use

Møller, Katrine Meltofte

Publication date:
2019

Document Version
Publisher's PDF, also known as Version of record

Link to publication from Aalborg University

Citation for published version (APA):
Møller, K. M. (2019). *Self-reports of traffic accidents: a critical evaluation of information validity, survey design and limitations of use*. Aalborg Universitetsforlag. Ph.d.-serien for Det Ingeniør- og Naturvidenskabelige Fakultet, Aalborg Universitet

SELF-REPORTS OF TRAFFIC ACCIDENTS

A CRITICAL EVALUATION OF
INFORMATION VALIDITY, SURVEY DESIGN AND LIMITATIONS OF USE

BY
KATRINE MELTOFTE MØLLER

DISSERTATION SUBMITTED 2019

AALBORG UNIVERSITY
DENMARK

SELF-REPORTS OF TRAFFIC ACCIDENTS

A CRITICAL EVALUATION OF

INFORMATION VALIDITY, SURVEY DESIGN AND LIMITATIONS OF USE

by

Katrine Meltofte Møller

AALBORG UNIVERSITY
DENMARK

Dissertation submitted

.

Dissertation submitted: 2019

PhD supervisor: Associate Prof. Harry Lahrmann,
Aalborg University

PhD co-supervisor: Associate Prof. Anne Vingaard Olesen,
Aalborg University

PhD committee: Associate Professor Erik Kjems (chairman)
Aalborg University

PhD, Marketing Manager Michael Wøhlk Jæger Sørensen
Via Trafik, Aarhus

Senior Lecturer Aliaksei Laureshyn
Lund University

PhD Series: Faculty of Engineering and Science, Aalborg University

Department: Department of Civil Engineering

ISSN (online): 2446-1636
ISBN (online): 978-87-7210-497-3

Published by:
Aalborg University Press
Langagervej 2
DK – 9220 Aalborg Ø
Phone: +45 99407140
aauf@forlag.aau.dk
forlag.aau.dk

Printed in Denmark by Rosendahls, 2019

ACKNOWLEDGEMENTS

I might be the author of this thesis, but it would be amiss to claim that I have produced it singlehandedly. Many people have contributed to my thesis in one way or another. Colleagues have helped with advice, useful discussions and feedback, and some have suffered an extra workload while I was on two maternity leaves. Friends have shared my ups and downs, keeping me sane through the hardest periods of writing. Family have spent time with our girls providing me with extra working hours. And my husband Kim has always been there to cheer me on and to calm me down when the stress of balancing work with our growing family was getting to me. He has handled the strain I have put on him by undertaking a task such as this, with enormous understanding, forgiveness and love. I am forever grateful for having him in my life. Lastly, a vast sum of people, whom I have never met and never will, have participated in my surveys; without their contribution none of this research could be carried out.

My warmest thanks to all,
Katrine Meltofte Møller

CURRICULUM VITAE

Katrine Meltofte Møller (female) has been an employee of Aalborg University since 2014. Her primary focus is the application of self-reports to gain knowledge of traffic accidents and the study of the limitations, advantages and validity of self-reported traffic accident data. She holds a Master of Science degree in Transportation Engineering (Aalborg University 2012); her Master thesis explored the traffic-related components which cause the road to be viewed as a barrier for pedestrian crossings. 2012-2014 she was working at the Danish Road Directorate in Aalborg with signage and traffic safety at roadwork sites. She has also been teaching this subject at the two-day course Vejen som Arbejdsplads at VEJ-EU in 2014-2017. As part of her employment as a PhD-student she has been teaching and supervising the students at the Civil Engineering program at Aalborg University as well as writing a report on visibility of vulnerable road users. Her list of publications can be viewed at https://vbn.aau.dk/da/persons/119890.

ENGLISH SUMMARY

Self-reports are getting more frequently applied within traffic safety work and research, for instance due to underreporting in police records of traffic accidents, yet many of the methodological aspects related to self-reporting is unknown. The objective of this thesis is to examine the correctness of self-reported information and the effect of various sampling methods and survey designs as well as the possible application of the self-report methodology within the area of traffic safety.

This have been investigated through different self-report surveys. One survey have applied a novel approach to sampling in the emergency room, resulting in 411 self-reports of accidents. Analysis of the respondents' hospital and police records have been carried out on this dataset, combining all 411 records of accidents, as well as a preexisting one, combining 74 records of accidents. Another survey have been distributed in Denmark, Sweden and Belgium, totaling 7.689 respondents and 804 self-reported accidents. Lastly, eight semi-structured interviews have been conducted in a qualitative approach to the subject.

The survey results show that self-reports have moderate to perfect agreement with police recordings of the same accident ranging from 83% (κ=0,51) to 100% (κ=1) on factors such as weather conditions and transport modes. Self-reporters furthermore show consistency in their written description of the accident location and their pinpoint on a map (87% concordance). A large proportion of respondents (45-55%) state to have been in contact with the police, but no matching police records can be identified; the interviews indicate that wrongful police recording is a contributor to this. Social desirability is found to affect some questionnaire responses but the thesis exemplifies how this can be utilized as a strength of the self-report methodology to investigate beliefs and attitudes. E.g., it is discovered that the responsibility of an accident is not considered equal across transport modes; both car drivers and bicyclists tend to absolve bicyclists of fault in accidents.

Several sampling methods is evaluated. The highest response rate (15%) is found with invitation via Digital Post, yet this provides a composition of respondent with less young men and women and less elderly women than the population in general.

It is concluded that self-reports of the individual accident provides relatively trustworthy and correct information on numerous important accident parameters and as such could be applied within the site specific traffic safety work, yet it is not recommended for estimation of underreporting or in other aggregated estimations generalized to the population.

DANSK RESUME

Selvrapportering bliver mere og mere hyppigt anvendt i trafiksikkerhedsarbejdet, eksempelvis grundet problemer med underrapportering af politiregistrerede uheld, men mange af de metodiske aspekter, der knytter sig til selvrapportering, er ukendte. Formålet med denne afhandling er at undersøge korrektheden af den selvrapporterede information, effekten af forskellige sampling-metoder og undersøgelsesdesigns samt den mulige anvendelse af selvrapportering inden for trafiksikkerhedsarbejdet.

Dette er blevet undersøgt gennem forskellige selvrapporteringsundersøgelser. Én undersøgelse har anvendt en ny fremgangsmåde til sampling på skadestuen, resulterende i 411 selvrapporteringer af uheld. Analyse af respondenternes hospitals- og politirapporter er foretaget på dette dataset, hvor alle 411 rapporter kombineres, samt et allerede eksisterende dataset hvor 74 rapporter kombineres. En anden selvrapporteringsundersøgelse er foretaget i Danmark, Sverige og Belgien, resulterende i 7.689 respondenter og 804 selvrapporterede uheld. Endelig er der foretaget otte semistrukturerede interviews i en kvalitativ tilgang til emnet.

Undersøgelsesresultaterne viser, at selvrapporteringerne har moderat til perfekt overensstemmelse med politirapporterne på de samme uheld, varierende fra 83% (κ=0,51) til 100% (κ=1) på faktorer såsom vejrforhold og transportform. Respondenterne udviser desuden konsistens i deres beskrivelse af uheldsstedet og deres stedfæstelse på kort (87% overensstemmelse). En stor andel af respondenterne (45-55%) angiver at have været i kontakt med politiet, uden at en matchende politirapport dog kan findes; interviewene indikerer, at fejlagtige politiregistreringer er medvirkende til dette. Visse af spørgsmålssvarene er underlagt respondenternes ønske om at afgive socialt acceptable svar, men afhandlingen eksemplificerer, hvordan dette kan udnyttes som en styrke til at undersøge respondenternes holdninger. Eksempelvis er det fundet, at det opfattede ansvar for et uheld ikke deles ligeligt mellem forskellige transportformer; både bilister og cyklister fralægger til en vis grad cyklisterne for skyld i uheldene.

Adskillige sampling-metoder er blevet evalueret. Den højeste responsrate (15%) er fundet med undersøgelsesinvitation udsendt via Digital Post, men dette medfører en respondentkomposition med færre unge mænd og kvinder samt færre ældre kvinder end befolkningen i almindelighed.

Det konkluderes, at selvrapportering af det enkelte uheld giver relativt pålidelig og korrekt information på en lang række vigtige uheldsparametre, og som sådan kan metoden anvendes i det stedsbestemte trafiksikkerhedsarbejde, men den anbefales ikke til estimering af underrapportering eller i andre aggregerede estimater som generaliseres til den brede befolkning.

TABLE OF CONTENTS

1. INTRODUCTION

"[...] we may become so enamored of increasingly sophisticated statistical techniques and data processing capabilities that we may pay inadequate attention to the quality of the data obtained in our investigations – data which, in fact, become the raw material for these statistical and data processing methods. For no scientific discipline can be any better than the quality of its raw data".

Leon Gordis (1979) in American Journal of Epidemiology

Self-reports of traffic accidents are getting more and more frequently applied within the field of traffic safety (Kamaluddin *et al.* 2018). The reasons for this are numerous: Surveys can be designed so all needed information is gathered for a specific research purpose. Self-reports may also be applied to gather knowledge of accidents that are otherwise not recorded in official records (such as police or hospital records). They are relatively cheap to gather and there is usually very little time from survey start to data being ready for analysis. Their merits are many, yet much is unknown of their data quality. This PhD-thesis aims at fillling in some of the gaps of knowledge related to the self-reported data's quality as well as how this is affected by methodological considerations regarding survey and questionnaire design.

The thesis revolves around one main research question:

To what extent can self-reports be applied to minimize the impact of underreporting on accident prevention in Denmark?

Four sub-questions have been identified to support the research question:

- How is self-reports' agreement with other data sources?
- What is the correctness of the self-reported information?
- What is the representativity of the respondents who self-report?
- Which issues and accidents can be addressed by self-reports?

I will address these questions first by providing an overview of the issues related to the current practice of accident registration and the use of self-reports in traffic safety work in Chapter 2. I will then move on to discuss sampling of respondents and how different sampling and invitational modes affect the information gathered from self-reports in Chapter 3. Chapter 4 will address the different definitions of traffic accidents applied in the surveys and by the Danish police and Chapter 5 will focus on accident information obtained from self-reports. Finally, Chapter 6 and 7 will provide a discussion and a conclusion of results.

The thesis includes four papers in Appendix, which all addresses one or more of the abovementioned sub-questions. The papers will be referred to throughout the thesis as Paper I, II, III and IV:

Paper I	A case study on agreement between self-reported bicycle accidents and hospital and police records *Peer-reviewed in selected Proceedings from the Annual Transport Conference at Aalborg University 2015* *Møller, K. M., Madsen, T.K.O., Olesen, A.V. Lahrmann, H.*
Paper II	The quality of self-reports of accidents compared with hospital or police records. *Manuscript* *Møller, K., M., Olesen, A.V. Lahrmann, H.*
Paper III	Nonresponse and dropout in random samples' and volunteers' self-reports of traffic accidents *Manuscript* *Møller, K., M., Olesen, A.V. Lahrmann, H.*
Paper IV	Perceived fault and accident causation factors in self-reported traffic accidents *Manuscript* *Møller, K. M., Olesen, A.V., Martinussen, L., M., Lahrmann, H.*

Part of the research have been funded via the Horizon 2020 project InDeV (*In-depth Understanding of Accident Causation for Vulnerable Road Users*). I have written and co-authored more on self-reports as part of the project deliverables and a reader interested in this topic might find the publications worth reading; they are all available at the project homepage www.indev-project.eu, except for the literature review published in European Transport Research Review. The publications are as follows:

Møller, K. M., Andersen, C. S., Várhelyi, A., Schönebeck, S., Reumers, S., Hosta, P., & Szagała, P. (2017). *Accident Information from six European Countries Based on Self-reports*. Bruxelles: Publications Office of the European Union.

Møller, K. M.; Andersen, C. S. (2016). *Self-reported accidents. Review of current study methods for VRU safety: Part 1 – Main report.* red. Piotr Olszewski. Warsaw: Warsaw University of Technology, 2016. p. 79-91.

Kamaluddin, N. A, Andersen, C. S., Larsen, M. K., Meltofte, K. R. , & Várhelyi, A. (2018). Self-reporting traffic crashes – a systematic literature review. *European Transport Research Review*. 10:26. https://doi.org/10.1186/s12544-018-0301-0

Andersen, C. S., Kamaluddin, N. A., Várhelyi, A., Madsen, T. K. O., & Møller, K. M. (2016). *Review of current study methods for VRU safety: Appendix 7 – Systematic literature review: Self-reported accidents*. Warsaw: Warsaw University of Technology.

Andersen, C. S., Madsen, T. K. O., Agerholm, N., & Møller, K. M. (2018). Self-reporting of accidents and near-accidents. I E. Polders, & T. Brijs (red.), *How to analyse accident causation? A handbook with focus on vulnerable road users* (p. 71-92). Hasselt University.

2. THE USE OF SELF-REPORTS

This chapter will provide an introduction to the current level of knowledge of incompleteness of police accident records and exemplify some of the problems that arises from applying only police recorded accidents in traffic safety work. Secondly, the possibility of incorrectness of police records is presented. Hereafter an overview of the state-of-the-art regarding application of self-report studies within traffic safety is presented as well as the unresolved issues related to their validity. Paper II addresses the state-of-the-art further in the introduction. Lastly, the various self-report regimes, which have been conducted as part of this PhD-thesis, is introduced.

2.1 INCOMPLETE RECORDS OF TRAFFIC ACCIDENTS

Improvements of traffic safety often relies on knowledge of accidents; in order to improve current conditions and prevent future accidents, one requires knowledge of the accidents that already have taken place to ascertain the safety problems. To learn of the accidents, one normally examines the police recordings of accidents. But many accidents happen that are not recorded by the police, even though they should – this is termed *underreporting*. There are many reasons an accident ends up outside official records (see e.g. Derriks & Mak 2007), but the very first step towards accident registration is for someone with knowledge of the accident to contact the police. Hereafter the police have to register the accident for it to be part of official accident records. An overview of the processes that could lead to underreporting is presented in Figure 1. It is well known that police records of accidents are not complete measurements of the true number of accidents. Several studies have been made on the level of underreporting internationally (Elvik & Mysen (1999) provide an overview) and in Denmark (Janstrup *et al.* 2014). However, the level of underreporting varies greatly from country to country; this variation is only in some aspects explained by the severity of accidents (Elvik & Mysen 1999, Derriks & Mak 2007). On basis of a meta-analysis of studies on underreporting, Elvik & Mysen (1999) found that in 13 countries the average reporting level of fatal accidents was 95%, while the police only recorded 70 % of the severe accidents, 25% of the accidents with slighter injuries and 10% of accidents with very slight injuries.

The exact level of underreporting in Denmark is unknown, but a recent study by Janstrup *et al.* (2014) estimates the level by applying capture-recapture techniques on hospital and police data from the island of Funen. Table 1 provides a calculation of police reporting level based on their results.

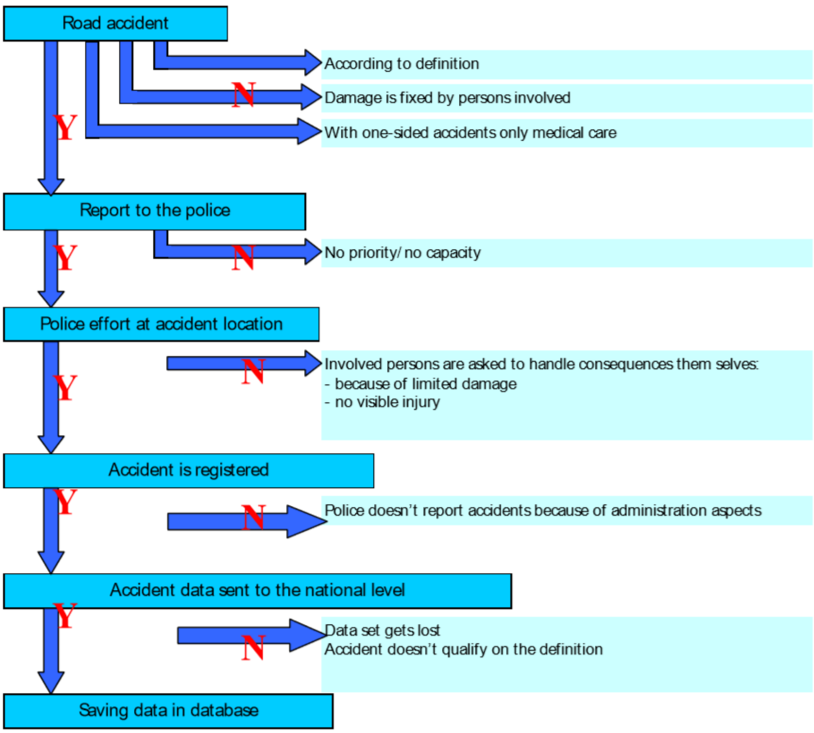


Figure 1: Illustration by Derriks & Mak (2007), which shows some of the many processes that could lead to an accident being exclude from the pool of official accident records.

	Fatal	**Severe**	**Slight**
Matched	111	2051	2172
Unmatched in police data	35	396	464
Unmatched in emergency room data	5	3373	11181
Level of reporting, Definition 1	0,96	0,38	0,16
Level of reporting, Definition 2	1,26	0,65	0,23
Level of reporting, Definition 3	0,97	0,42	0,19

Table 1: Results of the matching between hospital end police records by Janstrup et al. *(2014), but shown as totals from the period 2003-2007. The level of reporting is not part of the original work by Janstrup* et al. *(2014), but is calculated for use in this thesis. The definitions of reporting levels follows the work by Elvik & Mysen (1999).*

The results from Janstrup *et al.* (2014) in Table 1 show that the level of underreporting varies with injury severity; i.e. the more serious the accident the more likely it is that someone alerts the police and that they file a report. It can be noted that Table 1 provides different results of reporting level depending on which definition of reporting level is applied for calculation. This exemplifies the work of Elvik & Mysen (1999), where three definitions on how to calculate the reporting level were stated. Definitions can be seen in Figure 2; Table 1 shows the various results on reporting level that the application of the three definitions yield. Definition 3 is considered the most correct by Elvik & Mysen (1999), however it is seldom applied in litterature, as it requires a study of individual records. Yet here the definition can be applied with the detailed data from Janstrup *et al.* (2014).

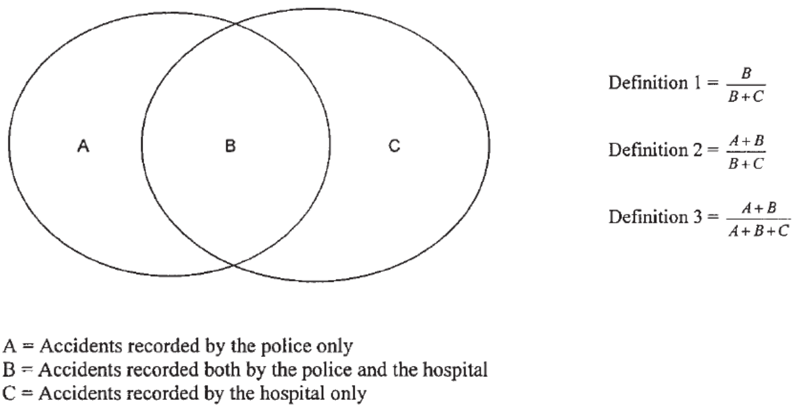


Figure 2: Illustration of three definitions of the level of accident reporting as presented by Elvik and Mysen (1999).

The level of underreporting in Denmark seems to be inconsistent over time. Figure 3 illustrates this by comparing the indexed sum of police records with the indexed sum of hospital records since 2001. The time trend shows that development is quite different in the two data sets; police records indicate that the number of injured in traffic accidents in Denmark have declined drastically over the last decade and a half, whereas the decline in hospital records of traffic accident victims have been much slower. This indicates the level of underreporting is not constant over time; it has increased since the beginning of the millennial. On the positive side, the development in hospital records and police records seem to be more aligned since 2010, and even though the two sets of records are not completely concordant, they do agree that the number of injured in traffic is declining.

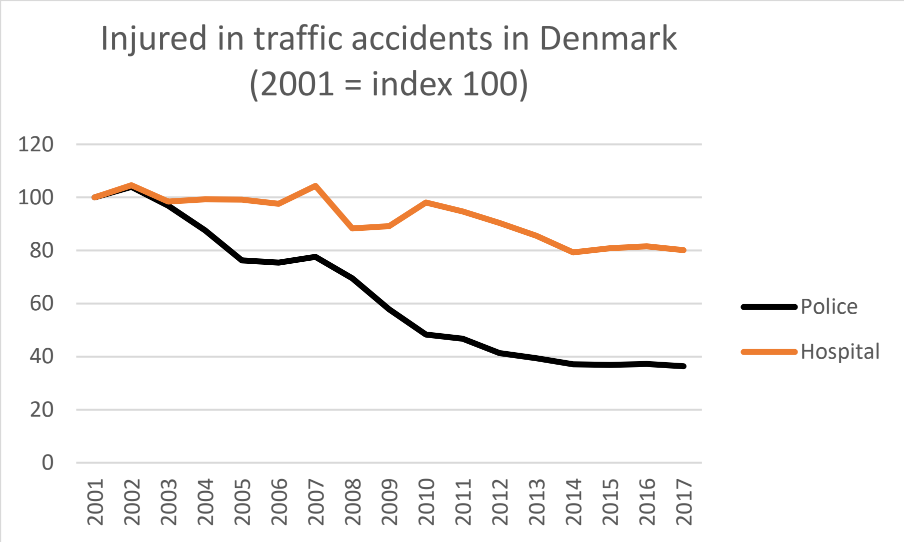


Figure 3: Time trend of accident reporting by police and hospital in Denmark 2001-2017. 2001 is chosen as index to emphasize the time trend. Numbers provided by Statistics Denmark, (Statistics Denmark 2019a).

It is also known that the level of underreporting in Denmark varies with mode of transport, (Janstrup *et al.* 2014). Janstrup *et al.* (2014) report their findings as results for the years 2003, 2004, 2005, 2006 and 2007, but Table 2 show their results as averages. It is clear that especially bicycle and motorcycle accidents are underreported with reporting levels as low as 6% for bicycle accidents with slight injury.

	Fatal	**Severe**	**Slight**
Pedestrian	100%	56%	24%
Cyclist	97%	14%	6%
Moped	100%	46%	23%
Motorcyclist	100%	38%	8%
Car	95%	68%	28%
Bus	-	13%	7%
Other*	100%	78%	45%

Table 3: Results from Janstrup et al. *(2014) on police catch rate depending on mode of transport and injury severity; original data on reporting levels are converted into averages from the period 2003-2007 for the use in this PhD-thesis. *Include users of vans, tractors and trucks.*

Besides severity level and transport mode, characteristics relate to the driver and the accident are associated with reporting rates. Abay (2015), who also applies Danish accident data in his studies, finds that accidents happening in peak hours have a lower reporting rate than accidents in working hours. Accidents with male drivers are found to have a higher reporting rate than those with females and accidents with drives older than 55 years have a higher reporting rate than those with younger drivers.

Resources and availability of police personnel might also influence the level of underreporting besides the aforementioned factors. For instance, the Danish Police Reform in 2007, in which the number of police districts as well as the number of stations with 24-hours service were drastically reduced, have been showed by Abay (2015) to have reduced reporting levels for all severity levels.

2.1.1. SOME IMPLICATIONS OF INCOMPLETE ACCIDENT REPORTING

The number of accidents recorded by the police is applied in varies aspects of traffic safety work, e.g. hazardous road location (*"black spots"*), effect studies of safety measures and development of campaigns. All of these are affected by the presence of underreporting. For instance, the study of effect of a traffic safety measure is based on the number of accidents before the measure was implemented and afterwards. If the level of underreporting was constant over time and independent of injury severity, underreporting would not affect the calculation of safety effects. But as previously stated, this is not the case. Moreover, calculation of effect of safety measures is complicated by the fact that traffic safety measures rarely have the same effect on all severity levels (Elvik *et al.* 2009, Jensen 2008). As reporting level varies over time, effect studies, which normally spans larger time periods, are inevitably affected by the reporting level's inconsistency and dependency on severity level.

In Denmark, it is common practice to identify hazardous locations with the help of the recorded accidents within a given time period as well as the amount of traffic on the location (Polders & Brijs 2018). That accident recording is affected by mode of transport will have consequences for which locations that are identified as hazardous. Thus segments with a high frequency of severe bicycle injuries might never be identified due to their low reporting rate of 14% (Janstrup *et al.* 2014), whereas segments with severe car accidents would have a higher chance of being located as their reporting rate is 68%, (Janstrup *et al.* 2014).

2.2 INCORRECT RECORDS OF TRAFFIC ACCIDENTS

Besides underreporting, another factor that impairs the outcome of traffic safety work is if information in traffic accident records is flawed. It is rarely only the number of accidents that is of relevance in traffic safety work; the nature of the accidents (transport mode, manoeuvres, surface conditions etc.) is investigated to help determine the appropriate safety measure(s). Thus, if accident information is not recorded correctly, the choice of safety measure might be less effective. The recording errors in police records are seldom investigated, but Ahmed *et al.* (2017) provide a meta-analysis of the existing studies and an average of recording errors in high-income countries is calculated. Ahmed *et al.* (2017) state that 39% of the information regarding the victim is erroneous. The same is true for 25% of information regarding the location, 15% regarding the environment and 12% regarding the vehicle. Their findings do not include Danish data, and it is not possible to ascertain how this would compare in a Danish context. Yet the paper by Ahmed *et al.* (2017) is of interest as *proof of concept* of erroneous police recordings also in Denmark. Only one study investigates erroneous recording practice in Denmark: Larsen *et al.* (2017) examines almost 80.000 records of extra-accidents. An extra-accident is an accident where the police find that there is so little injury or property damage that it does not constitute a reportable accident – hence only a limited amount of information is recorded by the police (Vejdirektoratet 2017). Larsen *et al.* (2017) find that 12% of the extra-accidents on rural roads were misclassified; the records contained written descriptions that showed that injury or large property damage were in fact present in the accidents. The misclassification on motorways was 10% and 1% on city streets. Paper II presents further findings on erroneous police records found as part of this thesis.

It is thus advocated to accept the police recordings of accidents with caution; not to expect recording errors among them would be amiss, even though few studies have quantified them. The aforementioned numerous factors that influence the propensity of an accident ending up in the police records also indicate that official accident records are both incomplete as well as biased. This does of course not entail that we should reject police records altogether, but serves as a reminder of the data quality of the accident records that current traffic safety work is based upon.

2.3 SELF-REPORTS AS RECORDS OF TRAFFIC ACCIDENTS

In an effort to minimize the effect of underreporting on traffic safety measures, other types of records could be used to supplement the recordings of the police. The scope of this PhD-thesis is to investigate self-reports, but of course, other possibilities might be explored such as ambulance data or extended emergency room records, (Møller, Clemmensen and Janstrup 2017).

Self-reports is a broad term and come in many varieties. Self-reports can be verbal; they can be carried out as either telephone interviews or face-to-face interviews. Written self-reports are often based on the respondents' completion of a questionnaire which can be either on paper or online. It is most common to apply a self-report regime based on questionnaires; a literature review finds that 65% of studies apply written self-reports and 23% verbal reports, (Kamaluddin *et al.* 2018). In this PhD-thesis I will apply the term *self-reports* to describe the fact that people provides information of their traffic accident themselves through an online questionnaire, i.e. with no interviewer or intermediary.

Road users own reports and descriptions applied in traffic research is quite common; the Driver Behaviour Questionnaire is a well-known example of this (Lajunen & Özkan 2011) as well as the Danish National Travel Survey, (Christiansen & Skougaard 2015). Yet when it comes to self-reports of accidents, self-reports are not necessarily as common practice. Silkeborg Municipality have tried to apply an app-based self-report regime yet with little success of obtaining knowledge of more than a few of the most minor accidents, (Møller, Clemmensen and Janstrup 2017). A recent PhD-thesis have, among other methods, applied self-reports of accidents to investigate vulnerable road users' road safety (Madsen 2018). But even though some self-report regimes are applied in Danish context, they have one thing in common: They are all time-limited projects, which are not an integrated part of the continuous road safety work of the road authorities, (Møller, Clemmensen and Janstrup 2017). There is no scientific literature that indicate that this is different in any other nation. Self-reports are thus limited to projects with certain focuses and a fixed duration of time. 144 studies have been located in a literature study of the self-report methodology applied to traffic accidents; the aim and focus of the self-report regimes can be many, (Andersen *et al.* 2016, Kamaluddin *et al.* 2018). Some self-report regimes seek out only to state if an accident have occurred or not, and others try to quantify underreporting. But most commonly, self-reports are applied to investigate either the effect of a specific safety measure or to investigate a specific crash causation factor (Andersen *et al.* 2016, Kamaluddin *et al.* 2018). An illustration of the most common motivations for traffic accident studies with self-reports is provided in Figure 4.

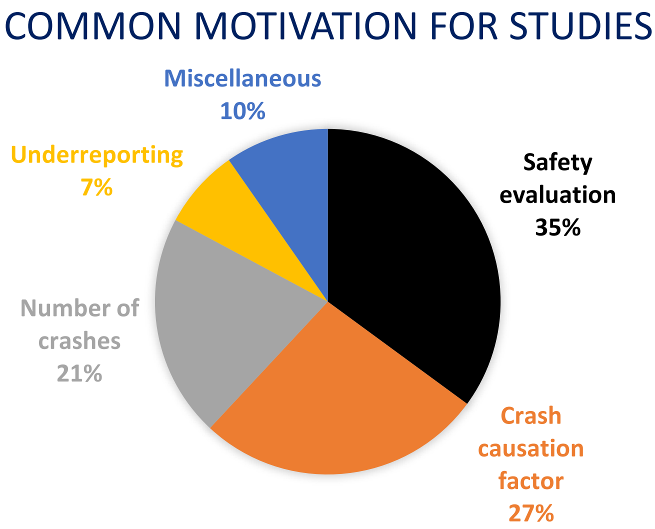


Figure 4: Motivation for studies with self-reports of traffic accidents as found by Kamaluddin et al. *(2018).*

Even though self-reports are getting more commonly applied within research, some researchers express distrust of the information gathered by self-reports, (e.g. Lajunen & Özkan 2011 and af Wåhlberg 2009). Two issues are often mentioned in research: Social desirability and memory effects. Social desirability is the inclination that respondents want to present themselves as better than they are, both intentionally and unintentionally, (see e.g. Paulhus & Reid 1991); Paper IV focuses more on this issue and its effect on some of the self-reported information. Memory effects are related to the recall period; if one asks a respondent how many accidents he/she have experienced in e.g. the previous thirty years, some accidents are expected to be forgotten. The deterioration of memory related to accidents is not that well researched, but some studies show a very high rate of forgetting, up to approximately 25% of accidents per year, (af Wåhlberg 2011). Sixty percent of the self-report regimes investigated by Kamaluddin *et al.* (2018) apply a recall period of 1-3 years, and only 11% have a recall period longer than five years. Paper II touches upon memory effects related to telescoping of accident dates, but memory effects in general are not suitable for investigation in the data sampled as part of this PhD-thesis as recall periods no longer than 3 months have been applied (see section 2.4).

One researcher in particular, Anders af Wåhlberg, has given voice to several concerns about self-reported information and sums up his apprehensions:

"…it would seem that the, somewhat limited, research on the validity of the self-reported collisions indicate:

1. *strong forgetting;*
2. *under-reporting by those with many crashes;*
3. *possible over-reporting by some sub-groups;*
4. *low agreement with other sources;*
5. *uncertainty about the correctness of reported details;*
6. *different predictive patterns as compared to other sources*."

af Wåhlberg (2009), page 37.

As discussed by e.g. Brookhuis (2011), not all agrees with af Wåhlbergs statements, but his concerns can at least be taken as an indication that there is not a substantial amount of methodological scientific literature to make final conclusions on the validity of self-report regimes within traffic research. This lack of knowledge is also proven by the literature review that finds that even though 134 publications of research utilizes self-reports, two thirds of these have a applied/practical focus with very little attention to the methodological considerations related to self-reports (Kamaluddin *et al.* 2018).

It is thus around this gap of knowledge in methodological issues that the PhD-thesis revolves. In recognition of the unknown validity of the self-reported information, both related to respondents ability and willingness to tell correctly about their accident as well as the representativity of respondents, several different approaches have been made within this PD-thesis to further develop the knowledge of self-report methodology; an overview of the gathered data is presented in the section below.

2.4 OBTAINMENT OF SELF-REPORTS USED IN THE PHD-THESIS

The thesis is based on three different datasets, hereafter denominated the Project Cykeljakke dataset, the Emergency Room dataset and the InDeV dataset. An overview of their characteristics can be seen in Figure 1.

DATASET: Project Cykeljakke	DATASET: Emergency Room	DATASET: InDeV

• Participants in another research project (effect of brightly coloured jacket on bicyclists)
• Project duration of one year, asked every month about accident occurence
• Only bicyclists
• Some linked to hospital records and police records
• 694 accident reports, herof 74 linked to other records
• Applied in Paper I
	• Patients in the emergency room in Region Nord who suffered a traffic accident as well as single pedestian falls on public road
	• Project duration of 13 months
	• All transport modes + single pedestrian accidents
	• All inked to hospital records and police records
	• 411 accident reports
	• Applied in Paper II
		• 3 subsamples in thesis
		  • Denmark: random sample of 40.000 invited through digital post
		  • Sweden: random sample of 40.000 invited through regular post
		  • Belgium: Call for survey voulounteers in numerous free medias
		• Project duration one year, asked every third month about accident occurence
		• All transport modes + single pedestrian accidents
		• All unliked to other datasources
		• 532 accident reports in Denmark , 100 in Sweden and 172 in Belgium
		• Applied in Paper III + IV

Figure 5: Main characteristics of the three datasets in the thesis

2.4.1. THE CYKELJAKKE DATASET

A research project on the safety effect of fluorescent bicycle jackets have previous to this PhD-thesis been carried out in the Traffic Safety Research Group at Aalborg University. As part of the project participants received a questionnaire every month for the duration of a year asking about any bicycle accidents they had experienced in the previous month. These self-reports of bicycle accidents are applied in Paper I, yet neither the data collection nor survey construction have been part of the thesis work. Numerous papers and reports describe the study and its results on safety effects (Lahrmann *et al.* 2018, Lahrmann *et al.* 2014, Lahrmann & Madsen 2014, Thedchannamoorthy *et al.* 2014, Hansen *et al.* 2014).

A total of 6793 bicyclists participated, reporting 694 accidents throughout the year. I have linked the self-reports of the 694 bicycle accidents with police and hospital records in a first attempt to examine their congruency with hospital records; Paper I (Møller *et al.* 2015) provides further insight.

2.4.2. THE INDEV DATASET

Some of the thesis is part of the Horizon 2020-project InDeV in which a survey was designed to gain information on accidents with vulnerable road users and their socio-economic costs. Even though the scope was to examine vulnerable road users, all transport modes were included in the survey. The survey was longitudinal; participants were asked to register for participation and were then sent a link to an online questionnaire every third month for the duration of one year. The questionnaire was compiled in six different languages. The project partners from each nation (Denmark, Sweden, Belgium, Germany, Poland and Spain) were each responsible for sampling of respondents within their country. The sampling was less successful in Germany, Poland and Sweden and the results from these nations are excluded from the PhD-thesis, but can be viewed in the project report (Møller *et al.* 2017). In Denmark and Sweden, a random sample of 40.000 inhabitants was obtained for invitation in the longitudinal study and in Belgium a call for participants was distributed through various free digital channels. In Denmark, the survey invitation was distributed via Digital Post; it was the first time on the Faculty of Engineering and Science at Aalborg University that Digital Post had been applied for survey distribution. Paper III goes into further detail on sampling differences and the consequences this have had for the composition of respondents and dropout.

A total of 804 accidents were self-reported in these three countries; Table 12 and 13 on page 50-51 present the characteristics of these accidents. Paper IV utilizes the data to examine the self-reported accident causation factors and their association with perceived fault of the accident.

2.4.3. THE EMERGENCY ROOM DATASET

TrygFonden, who supported a new approach to self-reporting at emergency rooms that was implemented in collaboration with Region Nord, has also partly funded the PhD-thesis. The goal was to provide patients, who sought the emergency room due to a traffic accident, with a questionnaire with questions similar to the information recorded by the police as well as numerous other questions on e.g. accident causation. It is not legal to contact patients who have been in contact with the hospital without their initial consent to contact. The survey design thus relied on the nurses, who staff the emergency room reception, to ask patients if they consented to receive a letter with information and a link to the questionnaire. Information flyers were provided to patients at the reception desk and several meetings were held with the nurses to ensure that they felt confident in providing patients with needed information. If consent was given, the letters containing the link to the questionnaire were subsequently sent to the patients by Digital Post. A telephone hotline was established to answer any questions by respondents or assist in the completion of the online questionnaire.

The three emergency rooms in Region Nord were visited by 3.809 patients due to a traffic accident, 930 patients gave permission to receive the questionnaire and 439 questionnaires were completed. The questionnaire asked patients if they would consent to have their questionnaire linked with other data sources, e.g. police records and hospital records. 27 answered *no* to this and these respondents have been removed from analysis, leaving 411 self-reports of accidents as a basis for investigation in this PhD-thesis. These self-reports of accidents are investigated in Paper II to examine agreement with police records and hospital records as well as the internal consistency of the questionnaire answers. Section 3.1 provides further information on the coverage and sampling errors in the dataset.

35 of the respondents were selected for invitation to a follow-up interview; 9 participated in the semi-structured interviews. The interviews were conducted to investigate the consistency between respondents written self-reports and their verbal description as well as to gain insight in the some of the methodological aspects of the self-report regime, e.g. reason for participation, ease of locating the accident etc. An interview guide is presented in Appendix E. The interviews are described in Paper II and section 5.3 provides further results.

2.5 ETHICAL CONSIDERATIONS RELATED TO THE USE OF SELF-REPORTS

A number of ethical considerations have been made in order to complete this thesis, both related to the obtainment of data and to the dissemination of data. First, the distribution of the self-report questionnaires in the InDeV dataset had to receive ethical approval in Sweden, though not in Denmark, Belgium, Germany, Poland or Spain, which encouraged numerous ethical reflections. Secondly, the confidentiality and anonymity promised to all participants in our numerous questionnaires have been carefully considered in dissemination of results.

Datatilsynet, to whom all research projects with personal data must be submitted for approval, sets guidelines for data management. These have been followed in order to ensure that data is stored and treated properly to avoid theft or hacking of the sensitive information. These standard guidelines on treatment of sensitive data are considered adequate for the research purposes in this thesis and compliance is considered a first step in ensuring the confidentiality of the research subjects.

The anonymity of the research subjects has been ensured when disseminating results from the studies, both within this thesis and throughout the many reports written on the surveys. It have been kept in mind that results can be considered sensitive, if a reader is able to guess the identity of a participant based on otherwise anonymous information, i.e. even though the participant's name is not stated. This gives cause for consideration when dealing with traffic accidents as a severe accident in a specific location is often a rather uncommon incident. Thus providing a reader with even a small bit of information that relates to a specific accident location or time of occurrence might make it possible for the reader to connect the information from the research project to information from other sources (e.g., information provided by the local newspaper). Respondents have provided descriptions regarding their explicit maneuvers and accident locations in their questionnaires; these have been altered in the disseminated quotations where deemed needed. A footnote indicates the quotations where modifications have been made as for example in Paper II.

Some of the statistical analysis in this thesis have been made possible by the use of linked personal data provided by Statistics Denmark. These data comes with the constriction that results, tables etc. cannot be published, if the frequency is lower than five. This consideration is of course also rooted in the need to preserve anonymity. For instance, many tables is presented in Paper II with ≤*5* as cell values.

Interviews have been conducted with nine respondents from the Emergency Room dataset, which have all been voice recorded with the permission of the interviewee. All have accepted transcriptions of their statements to be disseminated but without their names. Thus, the interviewees might recognize their quotes as belonging to themselves, but no information have been written that would enable a reader to make a plausible connection between the quotes and a specific accident.

Both completion of the self-report questionnaire as well as the interview prompt respondents to evaluate an event in their life, which, if the accident is considered a traumatic event, might take an emotional toll on the respondent. However, the participation in all surveys was stressed to be voluntary, which is expected to preclude the emotionally traumatized as *"emotionally fragile persons and those who don't feel they can talk about a problem usually don't volunteer to be interviewed."* (Corbin & Morse 2003, p. 338). But there is no guarantee that a respondent, even though he/she volunteered to participate in an interview, is not distressed by the questioning. Yet it can be argued that the negative aspects of being in distress to some extent can be counterbalanced by a beneficial effect of bringing *"[...] issues to a conscious level where they can be dealt with*" (Corbin & Morse, 2003, p. 338). The interviews were not focused on issues that were expected to be considered distressful for the interviewees, and the semi-structured form made it possible for interviewees to somewhat avoid topics they found too difficult to discuss. An interview guide is presented in Appendix E.

3. DIFFERENT SAMPLING METHODS

This chapter presents the different sampling methods that have been applied throughout the PhD-thesis. The methodological issues related to sampling such as representativity, sampling and coverage error and response rates are discussed. Paper III have provided a detailed examination of these issues related to the InDeV dataset, and though some findings regarding the InDeV dataset are addressed in the chapter, the focus is primarily on the Emergency Room dataset as Paper II does not cover this to a full extent. The chapter ends with a presentation of the results from all three datasets regarding future possibilities for sampling of self-reported accident information within other sample frames.

The three datasets are sampled in two very different approaches; Figure 6 provides an illustration of the main difference. The Project Cykeljakken dataset and the InDeV dataset are both obtained by asking a sample of respondents if they experienced any accidents – and if they acknowledged this, they were then given more questions about their accident (Method A). Both studies were longitudinal; the same respondents were asked every month/third month for the duration on one year. The Emergency Room dataset was obtained by sampling respondents who were present there due to a traffic accident (Method B). Thus 100% of the 411 respondents in the Emergency Room dataset experienced traffic accidents, whereas only 9,6% of the respondents in the InDeV dataset (Møller *et al.* 2017) and 10,2% of the respondents in the Project Cykeljakke dataset reported accidents (Lahrmann *et al.* 2014).

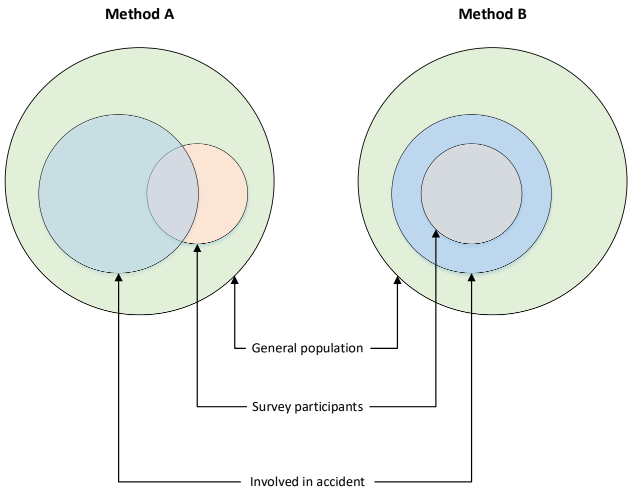


Figure 6: Illustration of sampling differences in the dataset. Method A: Sampling respondents within a sample frame of unknown accident involvement. Method B: Sampling of respondents in a sample frame where accident involvement is known.

When working with self-reports of traffic accidents, it is crucial to keep in mind that the self-reports are survey data and that many of the methodological considerations from the field of survey design applies to the self-report regimes as well. Thus self-reports are, as all other survey data, subjective to the four sources of error (coverage, sampling, non-response and measurement error) as described in detail by for instance Dillman *et al.* (2014). For the unfamiliar reader, a short description summarizing the definition of the errors is presented here as well as a demonstration of their influence in the surveys presented in the thesis:

Coverage error occurs, when the sample frame does not accurately represent the population on the characteristics, which one wishes to examine, (Dillman *et al.* 2014). The sample frame is the sum of people who *could* be contacted for survey participation. Hence coverage error is closely linked to how and where one sample ones' respondents from. For instance, one could find volunteers for survey participation as described in Paper I, where the recruitment of participants used many different channels – e.g. Facebook, the webpage of a number of papers, and Aalborg University's homepage to name a few (Hansen *et al.* 2014). Here it can be noted that all the communication channels are digital, which entails that the sample frame consists of people with access to the internet and the ability to browse these media on their electronic devices. Paper III also applies a dataset with similar coverage error, as the Belgian part of the InDeV dataset (Møller *et al.* 2017 and Paper III) were likewise invited through digital media. The Swedish part of the InDeV dataset utilizes postal addresses as invitation to survey participation, thus coverage error entails that people without any postal address is not part of the sample frame. In the Emergency Room dataset, there were 202 patients where nurses deemed consent to participation "*not possible*" (7%), see section 3.1. This contributes to coverage error, as the reason for this was that the patients had so severe injuries that they did not greet the nurses in the emergency room reception, but were moved directly to treatment instead. Hence, coverage error entails that injury severity among the sampled respondents is not completely comparable with that of the average patient recorded in emergency room records.

Sampling error is the error that occurs, when only some of the members of the sample frame are selected to the survey list, (Dillman *et al.* 2014). In a random sample, such as the Danish and Swedish part of the InDeV dataset, the error is – as the name suggests – random. I.e., one might draw a sample, which due to randomness, contains respondents with more single accidents than the general population. To minimize sampling error, larger samples is preferred in random samples. The sampling in the Emergency Room dataset have not been random; section 3.1 will address sampling error in the dataset in further detail.

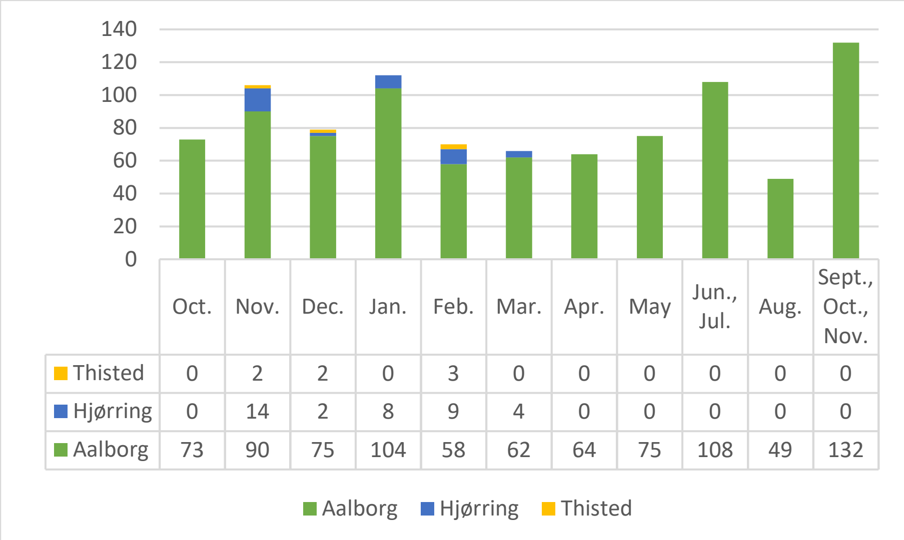
Non-response error occurs when those who chose to participate in a survey are somehow different to those who do not respond to the survey invitation on the parameters that one wishes to examine, (Dillman *et al.* 2014). Low response rates can

- though not necessarily - produce non-response error. Non-response is somewhat difficult to investigate as we very seldom have detailed knowledge of the non-responders; as they do not participate in the survey, we often have very limited information on them. In Paper II, a non-response characteristic is presented on the Emergency Room dataset, which shows that men in general lack participation in the survey as well as younger women and men. In section 3.2, a more in-depth discussion of the non-response errors in the InDeV dataset and in other related scientific studies will be presented. Dropout is often associated with non-response, with the underlying assumption that the respondents, who drop out of the study, share traits with the people who did not respond to invitation at all. Dropout analysis is conducted in both Paper II and III; the latter provides a more detailed insight into the questions that prompt respondents to drop out. Here we find that dropout is less associated with length of the questionnaire, but more commonly prompted by two specific items, which could be considered to have a private content.

Measurement error is a combination of many types of errors, which result in respondents providing inaccurate answers to the items in the questionnaire, (Dillman *et al.* 2014). For instance, measurement error could occur if the phrasing of an item is not understandable, if provided answer-options in a multiple-choice design do not reflect the experience of the respondents, or if respondents lack the knowledge of the answer to the asked question. Paper II investigates the respondents' inability to find a suitable answer among the stated multiple-choice options as well as the respondents' ability to locate their accident correctly on a map and chose an accident situation correctly. Another large contributor to measurement error is the effect of social desirability; that respondents answers are susceptible to self- and other-deception (Paulhus & Reid 1991, see Paper IV for more on this). Chapter 4 also discusses measurement error related to the respondents' and the professionals' definition of what constitutes as a traffic accident.

3.1 SAMPLING ERROR IN THE EMERGENCY ROOM DATASET

The project of sampling at the emergency room in Region Nord was most successful in the emergency room in Aalborg. The other emergency rooms in Region Nord either withdraw or were obtaining participation consent from so few of their patients that they were released from project participation in less than six months. Figure 7 show the distribution of respondents sampled in different emergency rooms throughout the project period; the small amount of participants from Hjørring and Thisted is evident. This exemplifies how important the hospital personnel are in such a research project; sampling is impossible without their dedication. That the interest and behavior of the hospital staff are so closely linked to sampling also entails that factors related to hospital personnel contribute to sampling and coverage error. As seen in Table 5, 2.817 patients were registered with a traffic accident in the emergency room in Aalborg in the survey period, but 1.443 of these were not asked if they wanted to participate in the self-report regime, even though they may have been eligible for sampling.



	Oct.	Nov.	Dec.	Jan.	Feb.	Mar.	Apr.	May	Jun., Jul.	Aug.	Sept., Oct., Nov.
Thisted	0	2	2	0	3	0	0	0	0	0	0
Hjørring	0	14	2	8	9	4	0	0	0	0	0
Aalborg	73	90	75	104	58	62	64	75	108	49	132

Figure 7: Respondents from the three participating emergency rooms in Region Nord.

This shows that even in the most successful of the emergency room setups, 51% of patients with traffic accidents were not asked to participate. This is expected to be primarily rooted in survey design: Nurses were not provided with any funds or extra working hours for asking of survey consent. Even though the project participation was agreed with their superior to be possible within their normal working procedure, some patients might have greeted the emergency room reception at an extraordinary busy period, which could have prevented nurses from asking about consent to the self-report regime. This is of course a source of sampling error. It is also possible that some nurses were better at remembering to ask patients of consent to survey participation than others, which would cause the sampling error to be somewhat related to which shift each nurse was working. Personal interpretations related to the patient were also an influential factor when the nurses had to decide whether to ask for survey participation or not. Some patients were intentionally not sampled for participation; 73 patients were deemed "*not relevant*" by the nurses and 29 had *"unknown relevancy"*; nurses explain that this was mainly patients who seemed not to speak Danish or seemed mentally less capable. As these sampling decisions had to be carried out as the nurses greeted the patient at the reception desk, their immediate perception of the patient was a contributory factor to sampling error. There is no data to explore this further or quantify the errors.

	All Region Nord emergency rooms	Aalborg emergency room
Consent to questionnaire	930	891
Consent to questionnaire, but were lacking information from administrative system to comply	34	31
No consent to questionnaire	168	149
Was not asked about consent	2.344	1.443
Consent not possible	202	201
Deemed *not relevant* by emergency room staff	95	73
Deemed *unknown relevancy* by emergency room staff	36	29
Total	**3.809**	**2.817**

Table 4: Sampling details from the emergency room in Aalborg as well as for the emergency rooms in Region Nord in total.

3.2 THE IMPORTANCE OF RESPONSE RATES AND RESPONSE PROPENSITIES

The InDeV dataset is examined in Paper III to investigate if the respondents are representative of the population in order to see if the accident frequency can be said to account for accident frequency in the population. In Paper III it was found that sampling of respondents through a call for volunteers in numerous free media were not successful in obtaining a representative sample, as there were more young and very few elderly among the fully compliant respondents. The random sample invited through regular post also had a significantly different age composition than the general population with many elderly and few younger people. From the investigations in Paper III, it is known that a random sample invited through Digital Post also yields a respondent age composition unrepresentative of the Danish population. The findings are not unexpected, as it is quite normal to stratify a survey on age and gender, as they are known to affect response propensities in various kinds of surveys also within the field of traffic research as in e.g. the Danish National Travel Survey, (Christiansen & Skougaard 2015). Yet the results from Paper III are of interest as they have shown how very different weights need to be applied within the different strata due to different invitational modes, and illustrates the benefits of applying a tailored design fitting the research question already in the stage of survey invitation. Besides the investigations on strata of age and gender carried out in Paper III, it have also been investigated if the size of the city, in which the Danish enrolled respondents live, is representative of the Danish population in general. This can be seen in Table 5.

	Inhabitants in the city where the Danish respondents live					
	< 5.000	5-20.000	20-100.000	>100.000	Do not know /other	Total
Number of respondents	1.344	1.011	1.468	1.476	233	5.532
Percentage of respondents	24,30%	18,28%	26,54%	26,68%	4,21%	100%
Percentage in population	33,98%	14,48%	19,24%	32,11%	0,19%	100%

Table 5: Self-reported number of inhabitants in the city where the Danish respondents live. Percentages on the distribution of the total population are calculated via Statistikbanken.dk (Statistics Denmark 2017). The numbers in the column "do not know/other" are not completely comparable, as the respondents who have answered "do not know" are not expected to be the same as e.g. homeless people that Statistics Denmark classifies as "other"; the category is thus excluded from statistical analysis.

A χ^2-test shows that there is significant (p<0,000) difference on the distribution of the general population and the distribution of respondents in Denmark. When conducting z-tests, they are all pairwise significantly different as well (p<0,000 for all); the inhabitants from small cities and rural areas are underrepresented in the study. Inhabitants from the largest cities in Denmark (more than 100.000 inhabitants) are also underrepresented in the study. There is no demographic information available from Statistics Sweden or StatBel to which one can compare the parts of the InDeV dataset that were obtained in Sweden or Belgium with. A study on Swedish traffic accident survey data finds that response rates decline the larger the city is, (Tivesten *et al*. 2012). The results from the Danish InDeV dataset are thus agreeing with the results from Tivesten *et al*. (2012) regarding the response rates from the largest cities, though they are in contrast to the finding from the rural areas where Tivesten *et al*. (2012) found response rates to be highest.

A main conclusion on self-reports' representativity is thus that none of the random samples can be used directly to account for accident frequency in the general population, as both age, gender and city size are skewed in the sample. A common solution is to stratify the sample on the variables one know to affect response propensities. Yet this is only possible on parameters, which researchers know to be skewed – i.e. gender, age and city size as established here. As found by Tivesten *et al*. (2012) and discussed in Paper III, there are numerous other factors that seem to influence the response propensities regarding traffic accidents. Among other things, Tivesten *et al*. (2012) find that vehicle age, type of ownership (private/company), driver age and gender and type of collision (single, animal or multiparty) significantly influence response propensities. It is not possible to apply the exact findings of weight functions from Tivesten *et al.* (2012) to other future studies, yet they prove that these variables do have an effect on response rates in their study and they would be expected to also affect response rates in other studies as well. The survey on attitudes towards accident registration of bicyclists by Kaplan, Janstrup and Prato (2016) also points out that numerous demographic variables, for instance gender, income level and having children affect how respondents view accident recording. As briefly presented in Paper IV, it is also very possible that circumstances related to the perceived fault of the accident influence the response propensity. In Paper IV, it was found that 67% believe the accident to be primarily their opponents fault – an explanation might simply be that someone not responsible for their accident is more likely to participate in a survey of the accident than someone who is at fault. This is also found by Tivesten *et al.* (2012); response rates of drivers in accidents where the other driver were at fault, were found to be higher than for drivers in single accidents. It would thus be amiss to simply perform stratification and weighting on the variables of age, gender and city size, even though it could be possible e.g. in the InDeV dataset in Paper III. It has to be concluded that currently there is no scientific literature that describes measures to weigh or stratify the randomly sampled data to a degree were the outcome of accident frequency is completely representative of the population and none of invitational modes applied in this thesis have obtained representative respondent compositions.

Fortunately, the self-reported accidents can be applied in other investigations than that of accident frequency – an example is provided in Paper IV of how perceived fault is associated with transport mode and how Odds Ratio can be applied to investigate attitudes on accident causation factors. Chapter 5 will apply a more direct approach to conclusions on the self-reported accident information. In these types of investigations it is important to obtain knowledge of as many accidents as possible. Here Method B, sampling respondents amongst a sample frame in which all have had a traffic accidents, show cost-effectiveness. If we sample in a sample frame in which accident involvement is *not* given, we would, if we invite 1.000 to a survey, catch 15 self-reports of traffic accidents - assuming that 10% of the population experience a traffic accident every year (as in e.g. Møller *et al.* 2017) as well as assuming that the response rate is 15%, which is the experience from the Danish numbers in the InDeV dataset (Møller *et al.* 2017). If we instead invite 1.000 respondents in a sample frame where accident involvement is given, we would yield 150 self-reports of accidents assuming the response rate is the same. However, the response rate have shown not to be the same. In the Emergency Room dataset, 1.132 patients were asked to participate and 439 completed questionnaires were returned. The effective response rate was thus 38%. If we apply this response rate in the calculation, we would yield with 380 self-reports of accidents for every 1.000 invited for participation. The large variation in response rates is to be expected; respondents have a higher response propensity if they deem the survey topic to be of relevance for them, (Dillman *et al.* 2014); thus, people who have just experienced an accident are more interested in participating in a traffic safety survey than those who have not recently had such an experience. When they were asked what their motivation was for participating in the self-report regime, the statements of interviewees exemplify this:

> *”I thought the accident was some of the worst that have ever happened to me, so if I in some way can help others so that it will not happen to them, then I would really like to help in that way. Normally I am someone who says NO to something like that, but because I had just experienced how much it had affected me, I chose to say yes.”*

> *“I said yes immediately, because I had this thought…that it was really nice than somebody was paying attention to what is going on - if it can contribute to something being changed in the long run about the conditions in that street, then….”*

> *“I said yes straightaway, I had no problems with that what so ever. If it can help others then…”Yes of course,” I thought.”*

> *“I think it’s just that I find it fascinating. I have recently finished my bachelor and back then I also needed a lot of research subjects – so why not. When you have tried doing such a project, you know how many people you have to reach. And then I am just interested in something happening with the traffic in Aalborg. Especially in that intersection.”*

The main explicit motivations for participation thus seem to be related to hopes of seeing future traffic safety changes, helping fellow road users and, among the younger interviewees who all had recent experience with administering surveys, the ability to identify with the role of research conductors. Future self-report regimes might try to stress some of these items in the survey invitation to explore if response rates might benefit.

3.4 WHAT OTHER SAMPLING POSSIBILITIES MIGHT BE EXPLORED?

It is not only in the emergency room that one could sample among those with known accident involvement. In Table 6, the self-reported information of contact with hospital, police and insurance company can be seen. There are large differences on the respondents’ frequency of contact to medical personnel, police and insurance across the three datasets. This is anticipated to be partly due to the various degrees of accident severity expected in the three datasets. Most important is the notion that roughly half of the respondents with an accident in the Danish InDeV dataset also contacted their insurance, whereas only one quarter had contact to medical personnel.

Self-reported contact to authorities		Project Cykeljakke dataset N=694	Emergency Room dataset N=411	InDeV (DK only) dataset N=532
Medical personnel	Yes - total	17,29% 120	100% 411	23,31% 124
	Own GP	3,03% 21	50,12% 206	N/A
	Emergence Doctor Service (*"Vagtlægen"*)	N/A	29,44% 121	N/A
	Hospital/emergency room	14,27% 99	100% 411	11,84% 63
Police personnel	Yes	3,31% 23	26,76% 110	6,58% 35
Insurance company	Yes	13,83% 96	47,93% 197	46,80% 249
Road authority	Yes	N/A	4,38% 18	N/A

Table 6: Self-reported information of whom the respondents had contact to due to their accident. Numbers in parenthesis, otherwise percentage of dataset.

This indicates a large potential for utilizing insurance companies as a facilitator of self-report questionnaires for obtainment of a large sample of accidents; a notion also recognized by Møller, Clemmensen and Janstrup (2017), though not quantified by them.

Very few respondents, less than five percent, in the Emergency Room dataset have contact with the road authority. An interviewee explains the following, when asked if he had any contact to the local road authority:

> *"As a rule I belong with those who think: Those authorities…who to ask and such? Then it's like: "Oh dear God…!". We don't know how the system works and basically we don't care. "*

That contact to road authorities is uncommon is not unexpected, keeping the aforementioned infrequent registration in Silkeborg Municipality's app in mind (Møller, Clemmensen and Janstrup 2017). The data thus consolidate insurance companies to be far preferable to road authorities if another sample frame with known accident involvement is to be recommended. Still, it is important to note that the Emergency Room dataset shows that only half of the respondents, who are registered with traffic accidents in the emergency room have also been in contact with their insurance company – a distribution quite similar is seen in the Euler-diagram presented in Paper I, (Møller *et al.* 2015). The diagram showed that 56% of respondents, who claim contact to the hospital, have also claimed contact to their insurance company, (Møller *et al.* 2015). A strategy of utilizing insurance companies as facilitators of self-report questionnaires would thus most likely yield both respondents with damage only accidents and some with injuries as well – though roughly half of the patients from the emergency rooms are expected not to be in the sample frame.

4. THE VARIOUS DEFINITIONS OF A TRAFFIC ACCIDENT

This chapter will clarify the different accident criteria applied throughout the thesis. First, the accident criteria for a reportable accident currently deployed by the Danish Police will be presented, secondly the definition applied in hospital records and lastly the three definitions applied in the self-report regimes. A discussion of the impossibility of applying the same accident definition in self-report regimes as in official records follows. As it have been found in Paper I and II that many respondents claim to have been in contact with the police even though they are not to be found in police records, a further investigation of the discrepancies is found in the final sections of this chapter.

4.1 WHAT IS AN ACCIDENT IN THE POLICE RECORDS?

For an incident to be considered a traffic accident by the Danish Police, two distinct criteria must be present, (Vejdirektoratet 2017):

1. The accident have happened on a road, square or area that is used for ordinary transport of one or more transport modes
2. At least one of the involved road users was driving

Hereafter, a traffic accident is deemed reportable by the Danish Police if one of the following circumstances is present, (Transport-, Bygnings- og Boligministeriet 2005):

- A person is injured
- Material damage of more than 50,000 DKK per vehicle or 5,000 DKK on other equipment
- A person with no fixed abode in Denmark is involved and a compensation claim is made against him/her
- The police is called to the scene of the accident and finds that a person employed by the police is involved in the accident
- There has been a violation of the Road Traffic Act that should lead to preliminary charges

It can be noted that while the requirements of material damage are rather specified in the legislation, the definition of personal injury is not provided. A help in interpretation is provided by the guidelines on accident reporting, which state that a person should be classified as uninjured ”*even if the person have received smaller abrasions, small cuts or small bruises”* (Vejdirektoratet, 2017, page 51).

Paper II presents findings of how the criteria of personal injury is not always adhered to by the police, for instance three of the eight interviewees are not to be found in police records even though they all sustained bone fractures in their accident.

4.2 WHAT IS A TRAFFIC ACCIDENT IN HOSPITAL RECORDS?

The hospital personnel applies no threshold for either material damage or injury for them to write a record of a person involved in a traffic accident. Instead, they operate with two classification variables: A) Reason for contact and B) Activity at the time of injury.

For a patient to be recorded as having had a traffic accident, the following must apply, (Sundhedsdatastyrelsen 2019):

- The reason for contact has to be classified as *contact reason 2 - Accident*
- The activity of the patient at the time of the accident has to be classified as either *paid transport work EUA0* or *transport in spare time EUA2*.

Thus, suicide attempts, which have to be registered with *contact reason 4 - suicide attempt*, are not recorded as traffic accidents. The fact that activity have to be classified as transport entails that other activities such as *sport and exercise (EUA5)* and other *vital activities (EUA7)* are not considered traffic accidents. Paper I illustrates how this can cause a wrongful estimation of the total number of bicycle accidents in Denmark, as 21 of the 74 respondents with hospital records were not recorded as traffic accidents, (Paper I, Møller *et. al* 2015).

4.3 WHAT IS AN ACCIDENT IN THE SELF-REPORT REGIMES?

The InDeV dataset and the Project Cykeljakke dataset have several similarities in their accident definitions. First, respondents were asked if they had experienced any traffic/bicycle accidents in the previous three/one month. If answering affirmative, respondents were then given several other questions regarding location of the accident, severity etc. that either ended the questionnaire or subsequently were used for cleanup of self-reports that did not meet the requirements for accident definition.

If only asking respondents if they "have experienced a traffic accident" the basis for measurement error arises, as it is very unlikely that all respondents are familiar with the rather intricate nature of the definition of a traffic accident according to the Danish Police. Respondents are not traffic safety experts and do not know any formal definition of what constitutes as a traffic accident. Furthermore, the formal definition (see section 4.1) is not easily understandable and consequently unfit as guideline for respondents. When asking respondents to report "if they have had a traffic accident", the respondents' answers thus very much relies on their own definition of whether something they have experienced is termed "a traffic accident". To minimize measurement error, a short example of traffic accidents was provided; in the InDeV dataset this was phrased: "*This also includes falling as a pedestrian or bicyclist, even* *though no-one else was involved in the accident, as well as all other traffic accidents that happened for instance while you were using motorized vehicles. Accidents that happened while you were a passenger should not be reported here."* (Møller *et al.* 2017, page 55). As the scope of the InDeV project was to gain knowledge of accidents with vulnerable road users, the stress of reporting single bicycle accidents as well as single pedestrian falls were highly relevant. The exclusion of passengers in the study was due to the fact that many of the items dealt with causation factors and accident specific details of which passenger recall have unknown validity. To minimize the potential for erroneous answers passengers were therefore excluded from the survey.

It was needed to introduce various items that could help to distinguish between near-accidents/conflicts and the accidents with little or no injury that respondents report among more serious accidents. The inclusion of conflicts/near-accidents in accident analysis will give results very different from official statistics; for instance, Madsen (2018), who investigated self-reported accidents and near-accidents with vulnerable road users, found significant difference on parameters such as road geometry, presence of a counterpart and counterpart's mode of transport when compared to official police-recorded statistics.

The development of items for distinguishing between conflicts, accidents with no/very little injury and accidents with injury have been very difficult. It is a balance between providing respondents with enough information to minimize measurement error without overloading them with too complex a description. In the InDeV dataset where accident involvement was unknown, two items were introduced to which answers could help determine if there was an accident or not. The items were as follows, (Møller *et al.* 2017, page 55):

"During your accident, were you or your vehicle in physical contact with another road user or another vehicle?"

"During your accident, did you fall or crash, were you hurt or were your belongings ruined?"

A reported incident was excluded from further analysis if the respondent did not answer acknowledging on at least one of these items. A similar definition have been applied elsewhere in Project InDeV, see Madsen (2018). In the InDeV dataset, 218 out of 1050 (21%) incidents were deemed near-accidents and removed from data (Møller *et al.* 2017). It must be recognized that there is a qualitative component of stating whether one has been hurt or not that makes an undisputable definition of a severity threshold impossible. For instance, a bruise, if large enough and placed unfortunate, can cause a person pain and to be unfit for work – in such a case it seems misguided to state that such an accident did not result in any personal injury at all. On the other hand, a bruise could also be completely insignificant for a person's wellbeing and would perhaps not cause the respondent to claim that he/she had been hurt.

The definition and subsequent cleanup of the Project Cykeljakke data have not been part of this PhD thesis, but is described in detail by Lahrmann *et al.* (2014). Here it is stated that the dataset consists of self-reports of bicycle accidents that have happened on public road in Denmark and in which the respondent either had physical contact with another road user, were injured, had material damage or crashed, (Lahrmann *et al.* 2014).

In the Emergency Room dataset, a slightly different approach to accident definition was applied, as it was self-evident that the respondents had physical contact with either another road user, the road or its equipment to sustain the injuries they were to be examined for at the emergency room. As described in Chapter 3 section 3.1, the nurses in the emergency room reception were the ones who decided if a patient were relevant for sampling or not. The nurses were instructed to include all patients who had a traffic accident that complied with the definition they apply in their normal registration practice (see section 4.2), as well as to take special care to include bicycle accidents, skateboard accidents etc. if they happened on public road. They were also asked to include pedestrians with falls on public roads in the sample.

No criteria for injury severity is applied in the Emergency Room dataset. As sampling have been carried out before respondents had their medical examination, there is no guarantee that respondents' injuries are severe – only that they have sought the emergency room to be examined. Instead of trying to define an undisputable yet practical applicable lower bound for injury severity, the specific sampling location is utilized as a threshold for injury severity. It is thus induced that persons who find it necessary to visit the emergency room, and not for instance their own general practitioner, are in so much pain and themselves believe their accident to be so severe, that their accident should also be relevant in a self-report regime.

The definition of a traffic accident is thus different in the three datasets; a summarization is provided in Figure 8.

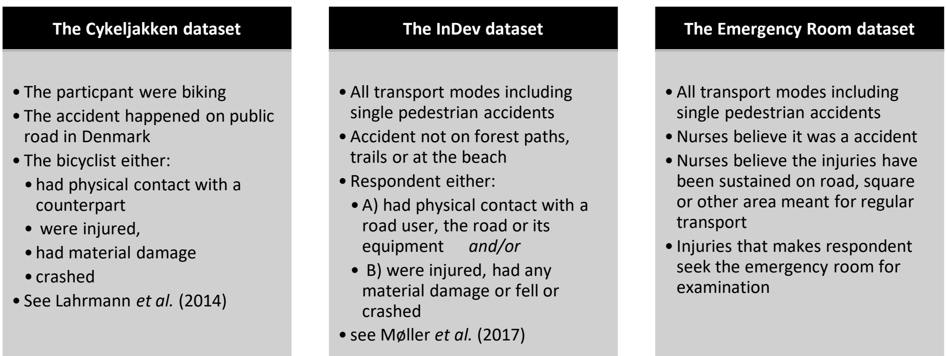


Figure 8: Overview of the accident definitions applied in the three datasets

4.3.1. WHY NOT APPLY THE SAME DEFINITION AS THE POLICE?

A deliberation have been made of trying to apply the same accident definition in the self-report questionnaires as the one applied by the police presented in section 4.1. Of course the respondents do not necessarily have information on his/her counterpart's injuries or amount of material damage, but it have been considered if the respondent could determine if his/her own experience of the accident lives up to the definition applied by the police. Two key issues present themselves. The first is related to the location of the accident, which is defined in this way in the official police definition: "*The accident have happened on a road, square or area that is used for ordinary transport of one or more transport modes."* (Vejdirektoratet 2017). This is considered somewhat difficult to understand for the respondents; measurement error could e.g. arise from the following questions: What is a road? Does it have to be paved? Is a separate bicycle path a road? Is the sidewalk a road? What about a gravel path in the forest? And what is covered by the term ordinary transport?

The second issue is related to the official definition's threshold for injury severity: "*Injured: Persons with injuries for which medical treatment or hospitalization (even if only for observation) is normally required. Persons with lesser wounds, minor cuts and bruises are not recorded as injured."* (Hemdorff, 2004, page 1). This definition is difficult to apply in self-report regimes for several reasons. First of all, there are injuries where no medical treatment is possible, but which are still so severe that is seems counterintuitive that a respondent is not injured – for instance a broken rib, which cannot be bandaged or put in a cast, yet still causes pain and moderate bedrest. Respondents with such injuries will assumable feel that they have been injured in the accident; it can be difficult for them to understand and accept that their injury is not considered relevant. It could also be argued that such injuries should in fact be of interest in self-report regimes as well as in official records. Secondly, the definition of a "lesser wound" or a "minor cut" is somewhat open to interpretation: Does it refer to the length of the wound? Does physical placement matter or is the depth of the wound of interest? Likewise with "bruises": A black eye can cause blurred vision and difficulties with opening the eye, something that would cause the patient great discomfort and to be temporarily unfit for work, yet it could also merely be of cosmetically inconvenience. Therefore, some bruises might seem more appropriately classified as an injury, albeit a slight one. If applying the phrasing on definition of reportable injuries, as provided above by Hemdorff (2004), in a self-report questionnaire, measurement error would thus arise, as the respondents are left to interpret their physical injuries through quite vague definitions.

To sum up; the direct application of the definition of reportable injuries used by the police is not conceivable within a self-report regime – it is simply not possible for respondents to know if their contact to the police have or should have resulted in a police report being written. Firstly, there is no notification provided to citizens if they appear in a police report; respondents do not have the knowledge of a police report being filed or not. Secondly, in multiparty accidents respondents have no way of knowing if their accident lives up to the reporting criteria (for instance the injuries of their counterpart or his/her possible employment in the police). Thirdly, both the official definition of a traffic accident and the injury criteria are highly susceptible to measurement error.

4.4 THE DISCREPANCY BETWEEN CLAIMED CONTACT TO THE POLICE AND THE NUMBER OF POLICE REPORTS

As presented in Chapter 3 section 3.3, the respondents were asked if they have had any contact to the police due to their accident; the frequency varies from 3% to 27% depending on the dataset. Yet is important to keep in mind that just because a respondent have been in contact with the police, the police does not necessarily have to file a report; the criteria mentioned in section 4.1 have to be present for an accident to be considered reportable by the police.

Keeping the difference between definitions of traffic accidents in police records versus the self-report regimes in mind (section 4.1 and 4.2), it would be wrong to conclude that all self-reported accidents with police contact should also be reported by the police. However, a comparison of the reporting levels in the datasets, as well as with the national average reporting level, provide interesting information on the nature of the self-reported accidents. In the Project Cykeljakke and Emergency Room dataset (Paper I and II) respondents have been matched to their police records by CPR-number. The results can be seen in Figure 9. In the dataset from Project Cykeljakken (Paper I, Møller *et. al* 2015) 55% of the number of respondents that have stated to have been in contact with the police equals a police report. The same is true for 45% of respondents sampled in the emergency room (Paper II). Due to privacy issues, respondents from the InDeV dataset could not be matched to police records.

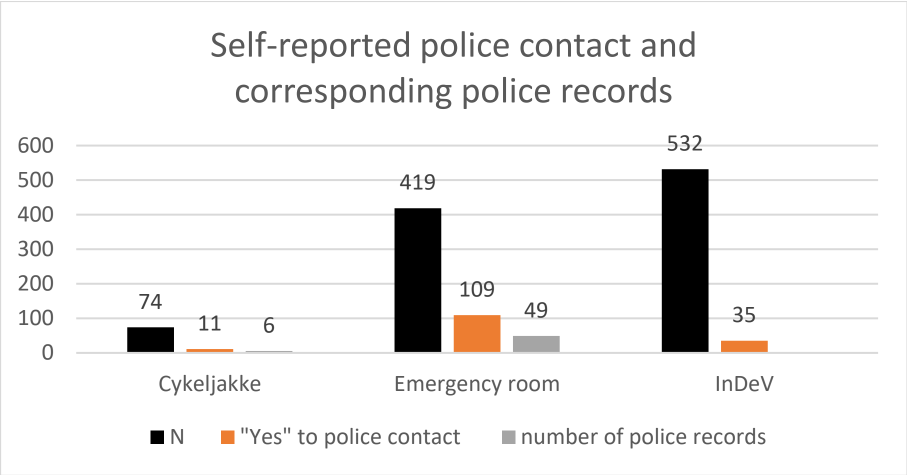


Figure 9: Number of respondents and their claimed contact to police. The number of matched police records in the InDeV dataset is unknown.

The reporting levels can be compared with the findings of Janstrup *et al.* (2014), who found a police catch rate on emergency room recorded bicycle crashes on 14-15% for severe injuries and 6-7% on slight injuries. In that perspective, the 11% reporting rate found on bicycle accidents in the Project Cykeljakken dataset (Table 7) indicates that the overall severity of the self-reports matched to hospital records are somewhere in between slight and severe. In the Emergency Room dataset (Table 8), the police catch rate is 15%. As these records also include other transport modes than bicycles, it is unexpected that the percentage is so low as Janstrup *et al.* (2014) found police catch rate to be the lowest for bicyclists (e.g. severe car accidents 66-73%, slight car accidents 25-31%). Hence, the implementation of other transport modes in the sample was expected to drive police catch rate up. This speaks to the previously discussed notion that the self-reported accidents overall have a relatively slight degree of severity compared with hospital records in general, as the nurses could not ask the most severely injured to participate in the survey (see section 3.1).

The Project Cykeljakken dataset shows that when 53 respondents are found to be registered in the hospital records with a traffic accident, 6 are to be found in the police database. The Emergency Room dataset shows that 49 of the 411 registered traffic accidents can be found in police records. To see if the ratio of police to hospital records in the two studies could be said to representative of the national mean, a comparison with official statistics is carried out. A comparison between the reporting level in the Project Cykeljakken dataset and the reporting level of all bicycle accidents in Denmark in average in 2012 and 2013 can be seen in Table 7. A two-tailed z-test is applied, which shows that the difference in reporting levels is significant (z= 2,1349, p=0,033).

	Bicycle accidents, mean 2012-2013	**Cykeljakken dataset**
Police records	811	6
Hospital records	16.331	53
Level of reporting by police	4,97 %	11,32%

Table 7: The number of accidents with bicyclists (solo accidents and multiple parties) in police records and hospital records. Data from 2012 and 2013 (Statistics Denmark 2019). The level of reporting by police is calculated as the number of police records divided by number of hospital records multiplied by 100.

Table 8 presents the reporting level in the Emergency Room dataset as well as the national average from 2016-2017. The two-tailed z-test proves significant difference on the reporting levels from the study and the national average (z=2,592, p=0,001).

	All accidents, mean 2016-2017	Emergency Room dataset
Police records	3.274	49
Emergency room records	31.114	330
Level of reporting by police	10,52%	14,85%

Table 8: The number of all accidents in police records and hospital records. Data from 2016 and 2017, (Statistics Denmark 2019). The level of reporting by police is calculated as the number of police records divided by number of hospital records multiplied by 100. Even though the data originally contained 411 records, 81 of these were stated to be single pedestrian falls – I have removed these to improve clarity of the comparison with real life registration.

Both datasets have a significantly higher proportion of police to hospital records than the national mean. As previously argued a larger overall severity is not believed to be the reason for this in the Emergency Room Dataset, rather a lesser severity is expected. That reporting rates are significantly higher in the studies could quite possibly be attributed to respondent characteristics. Abay (2015) presents variables associated with higher reporting rate to the police, both crash and driver related. For instance, he finds that people above 55 years are more likely to have a police report as well as an emergency room report as compared to a baseline of 46-55. Men are also more likely than women to have a police record as well as an emergency room record. Crash characteristics also influence reporting levels; Abay (2015) finds e.g. that accidents with speeds above 110km/h are more likely to appear in both police and hospital records as well as accidents with vehicles turning in intersections. Part of the explanation of his findings derives from injury severity, and personal characteristics and attitudes towards police and reporting in general might also contribute. Kaplan, Janstrup and Prato (2016) investigated intentions to report bicycle crashes and found several associations between demographic characteristics and intention to report bicycle accidents. For instance, respondents finding reporting to be useless were associated with being women, having a high income and residing in the Copenhagen Capital Region, and distrust in the police was found associated with being male, having a low or medium income and residing in the Copenhagen Capital Region. In addition, aversion to medical consultation were found to be associated with being male, having a low income and residing in rural areas (Kaplan, Janstrup and Prato 2016). These beliefs are expected to affect the response propensities of different demographic subgroups in contact to the hospital and the police respectively. Likewise, it can expected that similar beliefs affect the participation in self-report regimes as well, though not explicitly examined by Kaplan, Janstrup and Prato (2016). It can be hypothesized that persons, who are positive towards contacting the police to have their accident recorded, are also positive towards participating in a self-report survey on their traffic accident. It is not possible to test this hypothesis with the present datasets, but the significant difference in reporting rates in self-report regimes and the national average indicates that response propensities are somehow different.

5. ACCIDENT INFORMATION

The amount of information gathered from respondents as part of this thesis is vast. It is not possible – and also outside the scope of the research question – to communicate the entire sum of knowledge of accidents that have been accumulated, but it is relevant for answering part of the research question to veer from the methodological focus to address some of the survey results. Paper II and IV addresses some of the survey results as well; Paper II in an investigation of the coherence with police records and Paper IV in an example of how the items most susceptible to social desirable responding might be applied within the field of traffic safety research. The following subsections provide information gathered from the questionnaires to explore which types of accidents the self-report regimes provides information on, but also to show some of the differences in gathered information from the Emergency Room dataset and the Danish subset of the InDeV dataset.

5.1 PEDESTRIAN FALLS

Even though pedestrians single accidents are not considered traffic accidents by the Danish Police (see section 4.1), it has been one of the focus points of the InDeV-project that pedestrian single accidents could be include in traffic accident analysis just as single bicycle or single car accidents (Olszewski *et al*. 2017). The arguments for this are many, but most prominently: Pedestrians are road users as well as the road users on wheels and should be considered on equal terms in accident analysis. It is also argued that pedestrian single accidents can lead to quite severe injuries, especially for the elderly, and that the condition of the sidewalk is the responsibility of the road authority, just as the condition of the road itself or of a bicycle lane.

In support of this recommendation, the pedestrian single accidents have been recorded in both the Emergency Room dataset as well as the InDeV dataset. In the InDeV dataset 54 out of 59 (92%) recorded pedestrian accidents in Denmark was single accidents, thus 54 out of the 532 (10%) self-reports are single pedestrian accidents. The Emergency Room dataset have 87 pedestrian accidents, hereof 81 single pedestrian accidents (93%): pedestrian single accidents thus comprises 20% of the 411 self-reports.

In the Emergency Room dataset, pedestrians with single accidents have been asked of the reason for their accident. The format was multiple-choice, keeping in line with the other items in the questionnaire. All of the possibilities were external factors related to the road and its condition, which are not as susceptible to the influence of social desirability as the accident causation factors evaluated in Paper IV. The results are shown in Table 9.

Self-reported reason for pedestrian accident	N	%
I fell/twisted my ankle due to an uneven surface	32	34,78%
I fell because the surface was slippery due to snow/ice	18	19,57%
I fell/twisted my ankle due to a curb	17	18,48%
I did not fall, but had another sort of accident *(e.g. walked into something, an object fell on me etc.)*	5	5,43%
I fell because the surface was slippery due to wet leaves, mud etc.	1	1,09%
Other	19	20,65%
Total	**92**	**100%**

Table 9: Self-reported causes for pedestrian falls from the Emergency Room dataset.

The most common external factor related to single pedestrian falls is uneven surfaces (35%) and second most common is falls due to snow/ice (20%). The self-report regimes are thus a valuable tool in gaining knowledge of these accidents as they would otherwise be unknown and thus impossible to take precautionary measures against by the road authority.

5.2 ACCIDENT CAUSATION

In the literature review by Andersen *et al.* (2016) and Kamaluddin *et al.* (2018) it was discovered that 27% of the reviewed projects applied self-reports to gain knowledge of one or more crash causation factors. It is thus fairly common to rely on the ability of the self-report regimes to provide knowledge of pre-crash circumstances, which might be, at least partly, a causation factor for the accident. Paper IV goes into detail in an analysis of the self-reported accident causation factors from the InDeV dataset. Here one of the conclusions is that there is little consistency between respondents statements of accident causation factors (e.g. inattention, mechanical error etc.) and the frequency of these factors found in either other self-report studies (e.g. Meunier, Dupont, Mersch & Van den Berghe (2018)) or Danish accident in-depth investigations (e.g. AIB (2014)). As suggested in Paper IV, it seems more appropriate to view the respondents' choice of accident causation factors in another scope, for instance their association with perception of fault in the accident, instead of taking them at face value. The results presented in Paper IV show that both respondents' own transport mode as well as their counterpart's have significant influence on perceived fault; bicyclists are less likely to be held at fault for the accident than other road users. Odds Ratios were calculated for the respondent finding himself at fault when acknowledging the various causation factors. It was discovered that factors related to lack of attention and being distracted had the highest Odds Ratios, indicating that these factors are believed by the respondents to be more associated with accident causation than e.g. fatigue. Thus the findings from Paper IV illustrate how results from self-reports might be applied in the non-site specific traffic safety work as information on which road users or which traffic behavior might need addressing in campaigns etc.

Another finding from Paper IV is that there are many (65%) of the respondents who do not acknowledge any of the accident causation factors presented in multiple-choice format. Table 10 presents the findings from Paper IV as well as the causation factors stated in the Emergency Room dataset; here the frequencies of no chosen causation factor are down to 41% and 37% in multiparty and single accidents respectively. Chi^2-test shows that there is no significant difference within the Emergency Room dataset on the 41% and 37% from multiparty and single vehicle accidents (p=0,436). However, when comparing responses on multiparty accidents in the InDeV dataset and in the Emergency Room dataset, the difference on proportions of respondents that chose no causation factor is significant (Chi^2-test yields p<0,001). There are thus significantly more in the Emergency Room dataset who find one of the accident causation factors applicable.

Besides the significant difference in the proportion of respondents who chose a causation factor, there is also difference in the proportion of respondents perceiving their counterpart to be primarily at fault in the accident. There are 68% in the InDeV-sample who find the accident to be primarily their opponents fault (see Paper IV); in the Emergency Room dataset this is significantly (p=0,049) lowered to 58%. This demonstrates that the sampling strategy applied in obtaining the Emergency Room dataset is better at recruiting respondents who find themselves at fault in their accidents, thus contributing to a larger proportion of accident causation factors present in the dataset.

	InDeV dataset		Emergency Room dataset			
	Multiparty		Multiparty		Single	
	N =300	*%*	*N=165*	*%*	*N= 234*	*%*
Fatigue total	**13**	**4,33%**	**6**	**3,64%**	**14**	**5,98%**
- I was tired	13	4,33%	6	3,64%	14	5,98%
Attention total	**110**	**36,67%**	**29**	**17,58%**	**8**	**3,42%**
- I was immersed in thoughts.	19	6,33%	11	6,67%	8	3,42%
- I was not aware of the other road user	94	31,33%	24	14,55%	NA	NA
Interpretation and assessment total	**123**	**41,00%**	**50**	**30,30%**	**NA**	**NA**
- I thought the other road user was aware of my presence	123	41,00%	50	30,30%	NA	NA
Inappropriate driving behavior total	**12**	**4,00%**	**3**	**1,82%**	**10**	**4,27%**
- My speed exceeded the speed limit.	1	0,33%	0	0,00%	1	0,43%
- I disregarded another road user's priority.	11	3,67%	NA	NA	NA	NA
- I was affected by alcohol etc.	1	0,33%	3	1,82%	9	3,85%
Orientation and distraction total	**36**	**12,00%**	**15**	**9,09%**	**10**	**4,27%**
- I was distracted by something	30	10,00%	12	7,27%	10	4,27%
- I was looking at my radio or navigation system	1	0,33%	1	0,61%	0	0,00%
- I was looking at my cell phone	5	1,67%	2	1,21%	0	0,00%
Vehicle errors total	**3**	**1,00%**	**0**	**0,00%**	**6**	**2,56%**
- There was a mechanical error/technical fault	3	1,00%	0	0,00%	6	2,56%
Stress total	**39**	**13,00%**	**8**	**4,85%**	**15**	**6,41%**
- I was in a hurry	39	13,00%	8	4,85%	15	6,41%
Auditory tasks total	**30**	**10,00%**	**17**	**10,30%**	**16**	**6,84%**
- I was listening to music	14	4,67%	13	7,88%	12	5,13%
- I was talking to someone	18	6,00%	4	2,42%	4	1,71%
Illness total	**4**	**1,33%**	**0**	**0,00%**	**3**	**1,28%**
- I was ill/unwell	4	1,33%	0	0,00%	3	1,28%
No causation factor	**196**	**65,33%**	**67**	**40,61%**	**86**	**36,75%**

Table 10: The causation factors from the InDeV dataset as presented in Paper IV compared with causation factors from the Emergency Room dataset.

5.3 INTERVIEW INFORMATION

The concordance between self-reported accident data from the questionnaires in the Emergency Room dataset and their matching police records have been thoroughly investigated in Paper II. Here it was found that mode of transport (both own and counterpart's) had the highest agreement, which is also the case for interview data compared with questionnaire data. The agreement between accident information provided in the questionnaire and in the interviews can be seen in Table 11.

Concordance between interview data and questionnaire data

	Int. 1	Int. 2	Int. 3	Int. 4	Int. 5	Int. 6	Int. 7	Int.8
Transport mode	yes	yes	yes	yes	yes	yes	yes	yes
Location	yes	no	yes	yes	yes	yes	yes	yes
Accident situation	NA*	yes	yes	yes	yes	NA*	yes	yes
Time	yes	yes	no	yes	yes	yes	yes	yes
Daylight conditions	yes	yes	yes	yes	yes	yes	yes	yes
Weather conditions	yes	yes	NA**	yes	yes	yes	no	yes
Surface conditions	yes	yes	NA**	yes	no	no	no	yes
Sustained injuries	yes	yes	NA*	NA*	yes	yes	yes	yes

**) Unspecified in questionnaire*
***) Cannot remember in interview*

Table 11: Concordance between accident information obtained through interview and through completed questionnaire from the eight interviewees (int.) in the Emergency Room dataset.

Accident date was found in Paper II to have a high agreement (91%, ICC=0,88) for police records compared to questionnaire data; as seen in Table 11 all interviewees except one state the time of their accident in concordance with the information they provided in the questionnaire. Weather conditions were found to have the lowest agreement between police reports and self-reports with κ=0,51 and 83% agreement (Paper II). There are one respondent who cannot remember the weather conditions when asked in the interview, and one who states the weather to be different than she previously stated in her questionnaire. The largest discrepancy between questionnaire information and interview information stems from the description of surface conditions which one interviewee cannot remember and to which three interviewees make statements disagreeing with their questionnaire choices. Paper II presents κ=0,68 and 89% agreement of self-reports and police records on road surface conditions, which was the second poorest agreement of the factors in Paper II. This could indicate that factors with high agreement between police records and respondents in general are also the factors which respondents show high consistency in throughout questionnaire and interview. There have not been conducted enough interviews for statistical analysis of this.

The semi-structured nature of the interviews allowed for them to unfold in different directions depending on the interest of the interviewee, see the interview guide in Appendix E. Many of the interviewees speak of their counterpart, and many speculate of the factors that they believe have been causing the accident, e.g.:

"I think she had the sun in her eyes or else she simply did not notice that I was there"

"He must also have been in a hurry. He might have been on his way for something. Perhaps he was in his own world."

Approximation of age and gender of their counterpart is stated by all interviewees as they all talked with their counterpart immediately after their accident. Several of them have had contact afterwards; two have had telephonic contact and two have met with their counterpart afterwards. The interviewees, who have met their counterpart in person, both tell that their counterpart felt great remorse and was very emotionally affected at their meeting. The subsequent contact seems unrelated to the injury severity of the accident; for instance one interviewee, who only had a material damage only accident, were visited by her counterpart afterwards. Many of the interviewees express that they believe their counterpart shares their negative experience of the accident, i.e.:

"I am certain she is as sorry as I am."
"It was something horrible for him as well."
"…they were more chocked than me I think."
"Of course she was sad as well."

The concern for their counterpart is not due to the interviewees finding themselves to be at fault in the accident – in fact none of the interviewees state in their questionnaire that they find the accident to be primarily their own fault. Six state that they believe the accident to be primarily the counterpart's fault and two state that they were equally at fault in the accident.

5.4 ACCIDENT CHARACTERISTICS

Many of the accident characteristics is presented in Paper II for the Emergency Room dataset or already published in the report by Møller *et al.* (2017) for the InDeV dataset. Table 12 and Table 13 present some of the characteristics which have not already been published or presented in Paper II, as well as a few of the findings which are also described elsewhere but where the result of viewing the characteristics in comparison with national records is novel. The tables also present characteristics from the Danish police records of 2017 provided by Vejdirektoratet (2018), Statistics Denmark (2019b) or the accident database *vejman.dk* by Vejdirektoratet (2019).

Many of the self-reported accidents have happened on sidewalks or bicycle facilities. In the Emergency Room dataset, sidewalks comprises 13% of the accident locations and the Danish subset of the InDeV dataset have 9% of accidents located on sidewalks (see Table 12); the frequencies are significantly different ($p=0,04$). When comparing the self-reported information with police records, the frequency of accidents on sidewalks is significantly higher than the 1% in police records ($p<0,001$ for both datasets). There is no significant difference of the proportion of accidents happening on bicycle facilities in the Emergency Room dataset (22%) and the Danish subset of the InDeV dataset (19%), but the frequency is significantly higher than the 1% in police records, ($p<0,001$ for both datasets). This displays one of the strengths of the various self-report regimes, namely the many vulnerable road users who have reported their accidents. Thus bicycle accidents comprises 32% and 43% in the two Danish datasets respectively compared with 24% in official Danish police recorded statistics, see Table 13. The InDeV dataset shows 15% of accidents happened in squares or parking lots, which is significantly higher ($p<0,000$) than the 2% in official Danish statistics. The Emergency Room dataset have 5% of accidents taking place at squares or parking lots, which is also significantly higher ($p<0,000$) than the 2% recorded by the police.

χ^2-tests have been made on the Danish InDeV dataset, the Emergency Room dataset and the official Danish accident statistics where possible in Table 13. The only investigated distribution that does not show to be significantly different in the three datasets is daylight conditions ($p=0,842$). This is consistent with another recent PhD-thesis which have found significant differences on many accident characteristics of self-reported accidents and police reported accidents, though no significant difference is found on time of day for bicycle accidents, (Madsen 2018).

	InDeV dataset Regular post (S)		Mass information (B)		Digital post (DK)		Emergency room dataset		Danish police records 2017 (accidents with injury or death)	
Respondents	970		1183		5536		411		5.748.769	
Accidents (N)	100		172		532		411		2789	
Percent involved in accident	10%		14%		10%		100%		0,05%	
Which type of facility?										
Busstop	2	2%	3	2%	5	1%	8	2%	11	0%
Sidewalk	5	5%	14	8%	48	9%	54	13%	26	1%
zebracrossing	3	3%	2	1%	11	2%	16	4%	117	4%
Square or parkinglot	22	22%	28	16%	80	15%	20	5%	64	2%
Bicycle facility	22	22%	28	16%	101	19%	89	22%	45	1%
Other	11	11%	14	8%	37	7%	17	4%	N/A	-
Road	35	35%	84	49%	250	47%	207	50%	N/A	-
Property damage										
No	37	37%	51	30%	173	33%	N/A	-	N/A	-
Yes	63	63%	120	70%	357	67%	N/A	-	8884*	-
Absence from work										
No	91	91%	151	88%	480	90%	91	22%	N/A	-
Yes	9	9%	21	12%	52	10%	307	75%	N/A	-

**) Property damage only, not counting accidents with both property damage and injury*

Table 12: Accident information from the two self-report datasets, InDeV and the Emergency Room, as well as police reported accident information from all accidents in Denmark 2017 with injury or death.

	InDeV dataset						**Emergency room dataset**		**Danish police records 2017** (accidents with injury or death)		χ^2 (Danish self-rep. vs. police)
	Regular post (S)		Mass information (B)		**Digital post (DK)**						
Means of transport											
car	30	30%	91	53%	281	53%	117	28%	1460	52%	
bicycle	33	33%	59	34%	169	32%	175	43%	793	28%	
on foot	34	34%	12	7%	60	11%	87	21%	416	15%	p<0,001
moped	1	1%	0	0%	6	1%	2	0%	301	11%	
motorcycle	0	0%	4	2%	5	1%	6	1%	223	8%	
other	2	2%	6	3%	11	2%	24	6%	125	4%	
Counterpart											
Single accident	58	58%	78	45%	300	56%	213	52%	514	18%	p<0,001
Multiparty accident	42	42%	94	55%	232	44%	157	38%	2275	82%	
Lighting											
Daylight	69	69%	117	68%	367	70%	256	62%	1926	69%	p=0,842
Darkness	22	22%	29	17%	117	22%	76	18%	635	23%	
Twilight	9	9%	23	13%	43	8%	35	9%	228	8%	
Road geometry											
Straight road	33	33%	62	36%	206	39%	217	53%	1437	43%	
Road access	18	18%	32	19%	67	12%	22	5%	79	3%	
Intersection	20	20%	42	24%	131	25%	78	19%	1301	39%	p<0,001
Roundabout	5	5%	7	4%	26	5%	10	2%	89	3%	
Other	24	24%	29	17%	99	19%	84	20%	412	12%	
Intersections											
Signalized	10	10%	13	7%	78	15%	26	6%	N/A	-	
Unsignalized	10	10%	29	17%	53	10%	51	12%	N/A	-	

Table 13: Accident information from the two self-report datasets, InDeV and the Emergency Room, as well as police reported accident information from all accidents in Denmark 2017 with injury or death. The p-values are for χ^2 tests of the distributions in the Emergency Room dataset, the police records and the Danish subset of the InDeV dataset.

Paper II provides further accident information on the accidents from the Emergency Room dataset in its appendix, e.g. the specific accident situations and counterpart's mode of transport. The main findings here are that situation 040 (*Accident on roadway, e.g. fall from two-wheeled vehicle*) and 011 (*Accident on right-hand side of road, while driving on straight road/at intersections*) are the most frequent accident situations for single vehicle accidents with 31 (26%) and 18 (15%) accidents respectively among the 118 reported one-vehicle accidents – see Figure 10 for an illustration. The multiparty accidents are most frequently from main category 1 *(Vehicles on same road going in same direction without turning from road*) with 43 out of 123 multiparty accidents (35%) belonging in this main category. 28 of 43 (65%) stems from situation 140 *(Accident with vehicle coming straight from behind)* making it the largest contributor to accidents across all main categories.

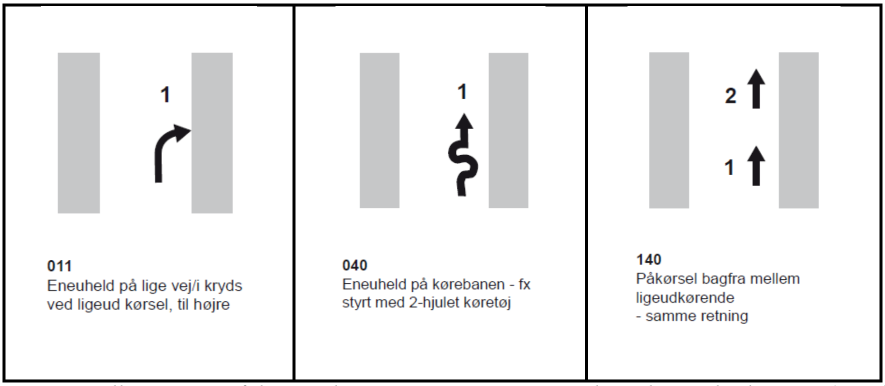


Figure 10: Illustrations of the accident situations 011, 040 and 140 by Vejdirektoratet (2017).

The distribution on main accident category in the Emergency Room dataset compared with that of Danish police records is shown in Table 14. The distribution in the two datasets show significant difference when conducting χ^2-test (p< 0,001).

There are thus significant difference on numerous of the accident characteristics when comparing the self-reported accidents to the police recorded accidents. Both the distribution of main accident categories, the mode of transport and the distribution of single accidents versus multiparty accident are different in the self-report regimes compared with the official police records. This chapter have are also presented several differences on accident characteristics, causation factors and perceived fault within the two datasets from self-report regimes in Denmark. This illuminates both how self-report regimes can provide knowledge of accidents very different from those in official records, but also that the sampling and survey design is instrumental in the outcome of which types of accidents that the survey gains information about.

	Main accident category	**Emergency room dataset** *(total 232)*	**Police records** *(total 3318)*
0:	Single accidents	109 (47%)	595 (18%)
1:	Same road with same direction	43 (19%)	438 (13%)
2:	Same road with opposite direction	7 (3%)	350 (11%)
3:	Same road with same direction and turning	8 (3%)	382 (12%)
4:	Same road with opposite direction and turning	8 (3%)	262 (8%)
5:	Different roads, no turning	4 (2%)	336 (10%)
6:	Different roads and turning	9 (4%)	374 (11%)
7:	Parked vehicle	4 (2%)	117 (4%)
8:	Pedestrian involved	12 (5%)	423 (13%)
9:	Animals, objects etc.	0 (-)	41 (1%)
	Unspecified	28 (12%)	0 (-)

Table 14: Accident situation distributed on main situations – a comparison between the questionnaire data from the Emergency Room dataset and police records. Data from Vejdirektoratet (2018) is applied for police records.

6. DISCUSSION

This chapter will first present a discussion of the findings of this PhD-thesis related to validity – this is ordered as replies to the concerns of validity of the self-reported information stated by af Wåhlberg (2009) *and cited in section 2.3. Afterwards recommendations for future studies with self-reported accident data will be presented in a discussion of possible methodological improvements and lastly a summarization of the practical experiences from this thesis work is disclosed.*

6.1 THE VALIDITY OF SELF-REPORTS OF ACCIDENTS?

In section 2.3, the concerns of Anders af Wåhlberg (2009) were cited as he is among the most prominent voices of distrust of self-reported accident information. His concerns are utilized as a framework for discussion as they are believed to cover a comprehensive view of the state-of-the-art methodological deliberations of validity related to self-reports of traffic accidents. In the following, his concerns will be applied as headlines (see section 2.3 for the contiguous quote) and each section will present the thesis findings in relation hereto.

Strong forgetting

af Wåhlbergs concern of strong forgetting (af Wåhlberg, 2009) has not been explicitly investigated as part of the analysis in this PhD-thesis; instead the focus have been on designing a self-report regime that minimized the possible memory effects. Thus none of the surveys apply a period of recall longer than three months, which was chosen to minimize the possibility of forgetting over time. Likewise, sampling in a frame with known accident occurrence have negated the possibility of respondents not recalling the accident happening – though of course accident specific details could have been forgotten. Yet the highest proportion of respondents answering "I do not recall" was 14%, which was in the item that asked respondents if streetlights were on or off at the time of their accident – after all a rather insignificant detail. But af Wahlberg's concern of forgetting is of course relevant if one applies a longer period of recall than e.g. three months – both the number of accidents as well as specific accident information will most likely be forgotten over time, though an investigation of how much and at what rate are completely outside the scope of this thesis. Nevertheless, it seems as if the concern can be overcome with a careful survey design; either utilizing a sample frame with known accident occurrence or minimizing the period of recall can negate at least some of the possibility of forgetting.

Under-reporting by those with many crashes

None of the findings show explicitly that those with many crashes does not either participate in the surveys or report fewer accidents than they have experienced, but the thesis supports that samples are skewed on various parameters and for instance the accident-prone younger males are often lacking from compliance. The elderly women are likewise underrepresented in the Danish random sample; in that age group women contribute to more accidents than men do. Sampling strategy and invitational mode are found to play an important role in the final composition of respondents; for instance the sample invited through mass-information have an overrepresentation of younger men and women in contrast to the sample invited through digital post. This PhD-thesis have found response propensities to be affected by invitational mode, age, gender and size of the city in which the respondent live; more factors influencing response propensities are suggested in the scientific literature (e.g. Tivesten *et al*. 2012 or Kaplan, Janstrup and Prato 2016). As we thus know that the non-responders or dropout-respondents are (in some samples) significantly different from the general population on demographic parameters, it is likely, though not demonstrated, that they would also differ on accident-proneness. As both response propensities as well as accident propensities are intricate and not yet fully known, it is impossible to take means to counteract this possible under-reporting by adapting sampling strategy or the subsequent statistical analysis. The worry of under-reporting in self-report regimes thus seems a valid concern no matter which sampling strategy or survey design is applied, and the application of aggregated self-reported information in generalization to the entire population should not be undertaken without severe reservations.

Possible over-reporting by some sub-groups

The concern of over-reporting – in other words reporting an accident that have not happened – is closely linked to the methodological concerns of measurement error and definitions of reportable accidents. Two possibilities exist regarding over-reporting: Either respondents make up stories of accidents or they report incidents that were not accidents as accidents. That respondents should be outright lying about accidents is not supported by any findings in this PhD-thesis or by any logical consideration. There is a rather high internal consistency in respondents' answers regarding the description of location and accident situation and outright lying would entail that the respondents would have conjured up an accident in many details and not strayed from the description of the make-believe situation throughout the questionnaire. Furthermore, the completion time of the questionnaires were estimated at 20 minutes; the incentive for respondents to voluntarily invest this much time in describing a make-believe accident seems non-existing. But that respondents report incidents as traffic accidents, even though they are in fact not accidents by a researcher's definition, is concurrent with the finding in the thesis – 21% of the reported accidents in the InDeV dataset were found to be near-accidents (Møller *et al.* 2017). This is by itself not a problem of the self-reported data, as items can be inserted that makes it feasible to discard the near-accidents. As long as the self-reported data is subjected to a cleanup before data is applied for analysis, the possible inclusion of near-accidents can be minimized. The true problem is related to measurement error; it is simply not possible for respondents to apply the same definition of reportable traffic accidents as the one applied by the police, see Chapter 4. Sampling within a frame with known accident occurrence can help to minimize the potential for over-reporting – if sampling could be conducted with either patients with known injuries or insurance policyholders with known amount of material damage, the self-reported accidents reportability by the police would be assured. The concern of possible over-reporting is thus something that to some extend can be handled through implementations of items to catch near-accidents being reported as well as accidents and through choice of sampling strategy. However, if sampling cannot be handled within a sample frame consisting of accidents that concur with the police's applied definition of accidents, the concern of over-reporting as well as under-reporting due to different accident definitions is valid.

Low agreement with other sources
This concern is disproven by the analysis in this PhD-thesis on many items. Moderate to almost perfect agreement with police records is found on various items of accident information and moderate agreement on counterpart is found with hospital records. The agreement is thus not low when it comes to the information provided in the self-report regimes. But af Wåhlberg is correct in contemplating that agreement is low if we only focus on accident frequency. All of the conducted self-report surveys tell of a higher accident frequency than what is found in police and/or hospital records. This can partly be due to the variations in accident definitions as argued in Chapter 4. But another reason also presents itself on basis of the investigations in the PhD-thesis, and that is the presence of misclassification and underreporting in police and hospital records to which the self-reported accidents are compared. Hospital records were found to grossly underestimate the number of bicycle accidents and the conducted interviews revealed several misclassifications by the police as well as one case of underreporting by the police. Thus, the low agreement with other sources of information regarding accident frequency is not solely due to errors in the self-reported data but is also reflecting the shortcomings of police and hospital records.

Uncertainty about the correctness of reported details
The uncertainty on correctness of reported accident location or accident situation have been resolved. More than 87% of respondents report the correct location of their accident and almost 85% of the self-reported accident situations are correct. The almost perfect agreement with police records on numerous items of accident specific information (transport modes, date and accident situation) also suggest that these items are often correctly reported by respondents. The environmental factors such as daylight, road surface and weather conditions have the lowest agreement with police records, and road surface conditions is also found as being least consistently described in interviews versus questionnaires. As especially accident location, transport mode and accident situation are relevant for accident analysis it is encouraging that these details seem quite correctly reported.

Different predictive patterns as compared to other sources
The self-reported accidents are in fact different from both police and hospital records. From the respondents motivation to report, the respondents demographic attributes to the nature of the self-reported accidents. But in itself, this difference should not be a concern. Not unless one believes the other sources of information to be a gold

standard, which would entail that the difference would be a sign of the self-reports' inferiority. An important note from the work with this PhD-thesis is that both police and hospital records are found to be flawed and should not be considered a gold standard. In many of the conducted interviews, the matching police records were found to be inconsistent with the self-reports. Moreover, in cases where police recorded information did not match the self-reported information, the interviews lead to the conclusion that it was the police reports, which were erroneous, due to the high consistency within the questionnaire and interview as well as inconsistency within the police records. It is therefore quite expected that the predictive patterns are different from other sources, and this concern of af Wåhlberg's is thus challenged – though not that his statement is incorrect, but that it should be a concern – as long as the correctness of police reports and hospital reports can be disputed.

6.2 FUTURE WORK AND RECOMMENDATIONS

6.2.1. IMPROVING THE AMOUNT OF "KNOWN" ACCIDENTS

As stated in Chapter 3, sampling within a population with known accident involvement is to be preferred, either via the emergency rooms or via the insurance companies, if the scope is to get information of as many traffic accidents as possible.

A possibility of improvement of survey design could be to expand on the reminders for survey participation. For instance, 930 patients gave permission in the emergency rooms to receive the questionnaire and 439 questionnaires were completed. Even though a reminder were sent after one month, this could possibly be handled by telephone instead of digital post; such a personal reminder might drive response rates further upwards, (Dillman *et al.* 2014). Another step towards improving the sampling in the emergency room would be to increase the number of patients that was invited to the survey; 51% of the patients in the emergency room was not asked by the emergency room staff to give consent to the questionnaire. Additionally, 7% were so severely injured that it was not possible to ask for consent in the emergency room. Two approaches might be possible for improvements on survey design:

1. The hospital could receive sufficient funding to prioritize the task of informing all patients at the emergency room reception as well as doing follow-up on severely injured patients, who are not greeted in the emergency room reception (i.e. making it mandatory in patient journal completion).
2. Legal changes could be introduced to make it possible to contact patients after their visit to the hospital without the need for their initial consent at the hospital.

The first solution will increase the sampled proportion of respondents, but the disadvantage is, as with e.g. extended emergency room recording, that currently the cost often falls to the road authorities who have to find means for this within their budget, thus presenting them with political, economic and ethical issues in prioritizing tasks. It must also be considered that as long as the task of recording or sampling falls to hospital personnel who are trained as medical staff, the need for them to fulfill their duties in a medical capacity would seem to overshadow the need for them to sample respondents in times with extraordinary workload. The latter approach would minimize the monetary cost of sampling respondents, yet the difficulties of changing the rules and regulations for conducting medical research are vast.

The insurance companies have also been addressed as a possible sample frame where accident occurrence is known. It would seem quite feasible for the companies to include a survey invitation in any response they send their policyholders who report a traffic accident. Yet their incentive for such an action is small, unless a direct safety effect is the outcome of the gathered self-reports; economic motivation could consist in the fact that fewer traffic accidents would entail less disbursed insurance money. A larger interaction between recording bodies might also enhance the number of accidents being recorded as part of the official statistics simply by changing the propensity of respondents calling the police when an accident happen. A practical example of this would be if a filed police report was needed for policyholders to obtain their insurance money; a setup much similar to the practice in Germany.

If ones goal is to increase the available data for traffic accident analysis, a first step towards gaining more knowledge of accidents might be to strengthen recording procedures at police departments. Especially as Paper II presents examples of interviewees, whose accidents have been wrongfully filed as extra-accidents (no injury, no damage) and are hence excluded from official accident statistics. One of the interviewees even explains that he called the police after the accident but was told that nothing could be done as no criminal actions had happened; his accident was not recorded at all. Even though the sample of interviewees is too small to be applied for statistical argumentations, it can be deduced that the police are most likely informed of more accidents than what is recorded and in addition misfile some accidents as extra-accidents despite of injuries being severe. If the police were to record civil registration numbers on material-damage-only accidents this misclassification could be subjected to further investigations.

There are thus numerous possibilities for implementation and improvement of future self-report regimes, yet none of them are easy, and free solutions does not exist for the obtainment of self-reports or improvement of official records. And the question remains: Whose responsibility and whose expenditure is it to make accident data available for road safety work? Is it the police's, the hospitals', the local road authorities', the Regions' or the national government's?

6.2.2. IMPROVING THE INFORMATION GATHERED IN SELF-REPORT REGIMES

The surveys have shown to have several items which could be improved if one wishes to improve on survey design (i.e. lower dropout, less "do not know/none of the above"-answers, shorten questionnaire time etc.). A first step of improvement would be to change the functionality of respondents' choice of accident situations. As 85% of respondents were correct in their choice, it seems promising to have respondents self-report their accident situation. Yet the survey design can definitely improve as the question about maneuvers right before the accident (which was applied to select which of the accident illustrations that were presented to the respondent) could not be answered by 18% of respondents, and as 43 of the 232 respondents presented with accident situations could not choose among the presented illustrations. There were thus only 189 choices of accident situations in the sample of 411 self-reports. Another approach might consist of having respondents draw their accident situation, which our interviewees all did easily and understandably. However, if applying an electronic survey design this might not be a desirable solution as drawing with mouse or touchpad is not necessarily a skill all respondents possess. A solution of drag-and-drop arrows, road users and geometric features might also be developed and tested; it would be interesting to see if further development of the item with accident situations could increase the proportion of respondents with a self-reported illustration.

Another possible improvement could be on the item of self-reported accident location. Almost all respondents (92%) provided GPS coordinates to the scene of their accident, but 12% of respondents did not provide a coordinate that matched their accident description. A change in approach might help respondents provide pinpoints that are more correct; an idea could be to implement a second item related to location. For instance, a *street view*-photo of their chosen location could be presented to respondents with a question of whether or not this was the correct choice of accident situation.

Several possibilities of either combing the self-reported data with other data sources or simplifying the questionnaire also comes to mind. Weather conditions might be gathered from historic weather databases based on respondents' choice of accident location and time. Lighting conditions could likewise be imported based on the respondents' choice of time of accident and location. Injuries could, with the right permissions given, be correlated with the respondents hospital records if they have been examined there. Likewise, the amount of material damage could also, to some extent, be gathered from the respondent's insurance company. As 94% of the respondents in the Emergency Room dataset consented to have their data combined with other data sources, such combinations should be considered relevant for future self-report regimes.

6.2.3. PRACTICAL ADVICE FOR DESIGNING SELF-REPORT REGIMES

There is no universally perfect design of a self-report regime – which design is preferable is closely linked to what information one wishes to obtain. None the less, some of the experiences made during the work of this thesis are relevant in close to all self-report regimes, and therefore some down-to-earth, practical advice for conducting self-report surveys are gathered here:

Careful consideration of the sample frame is important. If the scope of the self-report regime is not to account for a generalized accident frequency, it would be highly preferably to sample within a frame with known accident occurrence. Such a sampling method is found to heighten response rates, it eliminates some of the methodological concerns of accident definitions and is furthermore likely to be more cost-efficient with regards to how many accident reports are obtained through the survey. As stated in Chapter 3, emergency rooms or insurance companies might be considered for such sampling frames.

The survey invitation affects the outcome. The three different invitational modes (regular post, digital post and mass-information) were shown to provide very different outcomes regarding respondent characteristics – this is presented in Paper III and again in Figure 11 for easy an easy overview here.

	Advantages	Disadvantages
Random sample with Digital Post	+ Highest response rate in enrollment + High percentage of full-compliance respondents + Highest number of respondents and completed questionnaires	÷ Not free ÷ High dropout in enrollment ÷ Highest dropout in accident questionnaires ÷ Elderly females underrepresented
Random sample with regular post	+ Low dropout in enrollment + Lowest dropout in accident questionnaires	÷ Not free ÷ Low number of enrolled participants ÷ Poor at gaining knowledge of young people
Volunteers with mass information	+ Can be carried out free of costs + High percentage of full-compliance respondents + Very good at gathering data from young men and women	÷ High dropout rate in enrollment due to the question about providing email address ÷ Very poor at gaining knowledge of elderly people

Figure 11: An overview of the advantages and disadvantages of the different invitational modes. The figure is also presented in Paper III.

Motivation of any staff involved in sampling, e.g. at the emergency room or in any other place, is important to minimizing coverage error. It is highly advised to add resources for the extra working hours needed and try to increase sampling staff's dedication and motivation by keeping frequent contact and sharing the reporting progress continuously.

Decide on an accident definition. It is important to realize that respondents' accidents are not the same accidents as those reportable by the police. Starting a self-report regime should begin by a discussion of which accident definition is preferable – is it needed to adhere closely to the police's definition of a traffic accident or could another definition serve the purpose better. Even if one wishes to come as close to the criteria for police recording as possible, caution should be taken as it is impossible to apply completely congruent accident definitions; respondents do not know the injuries or material damage of their counterpart, any violations of the road traffic act or their counterpart's employment by the police.

Distinguish between accidents and near-accidents. The high frequency of reported conflicts in the InDeV dataset emphasizes the need for introduction of items to help remove conflicts from self-report regimes if a sample with unknown accident occurrence is applied. Items should be tailored to eliminate reported incidents that do not live up to the applied definition of accidents.

Check the calendar. Distribution of surveys and reminders in the summer holidays have been found to drastically lower response rates. If doing a longitudinal study, making a plan for survey distribution throughout the period and cross-examining it with national holidays is important to keep response rates up.

Be aware of social desirable responding. That respondents wish to present themselves as better than they are, is found to affect questions of fault and accident causation factors differently depending on which sampling strategy is applied. Answers highly susceptible to social desirable responding should not be taken at face value and analysis of the answers should be undertaken with this in mind.

7. CONCLUSION

A thesis statement with four research questions was presented in the introduction in Chapter 1. This conclusion will highlight the most prominent answers provided within the thesis and lastly a summarization of the results for the overall thesis statement is offered.

How is self-reports' agreement with other data sources?

Paper I and II have investigated agreement with both emergency room registrations and police recordings in two different datasets (the Project Cykeljakken dataset and the Emergency Room dataset). The two datasets show different agreement with hospital records. As shown in Paper I, there is moderate agreement (κ =0,63) between the self-reported information of counterpart and the counterpart registered in hospital records. The accidents investigated here are all bicycle accidents; a study on all modes of transport is presented in Paper II. Here agreement have risen to κ= 0,77, which indicates substantial agreement. But the difference of the two results is not expected to be due to the difference in the respondents' own mode of transport, but more likely due to the fact that the hospital personnel responsible for recording were aware of the research project in Paper II and therefore might have improved their registration practice. The moderate agreement of κ=0,63 between self-reported information and hospital records on counterpart is thus considered most accurate.

Paper II presents agreement between police records and self-reported accident information. The highest agreement (κ=0,93-1) show in classification of transport mode for both respondents themselves and their counterpart. There is also almost perfect agreement in choice of accident situations (κ=0,82). Substantial agreement is found in classification of road surface conditions (κ=0,68) and daylight conditions (κ=0,74). Weather conditions is found to have the poorest agreement of the investigated items with only moderate agreement (κ=0,51).

What is the correctness of the self-reported information?

The agreement with police records in itself is not a measure of how correct the self-reported information is; other studies (Ahmed *et al.* 2017, Chung & Chang 2015 and Larsen *et al.* 2017) have shown that police records have erroneous recordings of information. Examples of erroneous police recordings were also found when the information from the conducted interviews was compared with their matching police record, see Paper II. The correctness of the self-reported information is instead deduced from consistency in respondents' description of the accident throughout the many items in the questionnaire as well as the congruency of information between interview and questionnaire data.

In Chapter 5, a schematic overview of the congruency of questionnaire data and interview data was presented; transport modes and daylight conditions were perfectly congruent. The time and date of the accident and the injuries sustained from the accident were also quite consistent. The environmental conditions such as weather and road surface are the factors that show the poorest concordance.

The analysis conducted in Paper II finds that 87% of respondents are consistent in their written description of the accident location and their pinpoint of it on Google Maps. 85% of respondents have consistent accident descriptions and choice of accident situation.

What is the representativity of the respondents who self-report?
The self-report regime among patients with known accident occurrence (the Emergency Room dataset) have been described in Paper II. Here we found that men, who comprised 41% of the respondents, were underrepresented compared to the national average of male hospital recorded traffic accident patients. Young (below 18) and elderly (above 65) were also underrepresented among survey respondents. We also found that the non-responders deviated significantly from the responders for both men and women as the age group of 25-34 were very prone to non-response. Furthermore, as described in Chapter 3, a large proportion of patients (51% in the most successful emergency room) was not asked to participate in the survey, even if they might have been eligible for sampling, which entails that the potential for coverage error is large.

How representativity concerning age and gender of the respondents is influenced by survey invitation mode have been thoroughly investigated in Paper III. The different invitational modes are shown to yield very different compositions of respondents and the representativity of the self-report respondents are thus very much affected of the invitational mode. The invitation by digital post yields a final composition of full-compliance respondents with fewer young men and women and fewer elderly women than in the general population (Paper III). The invitation through mass media showed to yield more young and middle-aged respondents than the general population, but the elderly respondents were very underrepresented. Invitation through regular post results in a sample of respondents that do not show any significant difference from that of the overall population regarding neither age nor gender. This should not in itself be seen as an advocacy of using regular post, as it also must be considered that the response rate to regular post was very low (2%) compared to that of digital post (15%).

It is argued that even if random samples are utilized for self-report surveys and stratified on age and gender, such a measure is not believed to singlehandedly negate the unrepresentativeness of the obtained respondents. In Chapter 3 the distribution of the Danish respondents were shown to be unrepresentative of the population regarding the size of the city in which the respondents lives. Thus, inhabitants from smaller cities (less than 5000 inhabitants) and the largest cities (more than 100.000 inhabitants) are underrepresented among the respondents. Besides variation in response propensities related to demographic variables, such as age, gender and city size, other factors are also expected to influence response propensities. As found in Chapter 5, it might also be that respondents who believe their opponent to be at fault in the accident are more likely to participate in a self-report survey at least in random samples. Literature suggests that attitudes towards accident reporting in general depend on various demographic variables; if such attitudes and their variation across different demographic subgroups also affects participation in self-report regimes of traffic accidents, representativity is additionally hard to achieve.

The representativity of the respondents who self-report is thus highly intertwined with the sampling procedure and the invitational mode. None of the various designs applied for sampling within this thesis yield samples representative of the sample population and no easy solution of stratification or weighting might address the problem thoroughly. Therefore, careful deliberation and further work are advised if self-report survey data is considered to be applied in studies where results are needed to reflect the population.

Which issues and accidents can be addressed by self-reports?
Chapter 5 presents that especially accidents with vulnerable road users, single accidents and accidents that happen on bicycle facilities, sidewalks and parking lots or squares are more frequent among the self-reports than in police records. The most frequent accident situation is rear-end collision 140 *(accident with vehicle coming straight from behind)*. Knowledge of these accidents can thus greatly improve by utilizing self-reports.

In two separate studies we have combined self-reports of accidents with their matching police and or hospital records, see Paper I and II. Paper I presents that 21 of the 74 hospital records, which matched the self-reported bicycle accidents, were hospital records not registered as traffic accidents, (Møller *et al.* 2015). That 28% of the bicycle accidents are in fact not registered as traffic accidents in hospital records indicates that statistics applying hospital records greatly underestimate the number of bicycle accidents in Denmark. We have also found that 55% of respondents from the Project Cykeljakke dataset (Paper I, Møller *et al.* 2015) and 45% respondents from the Emergency Room dataset, who state to have been in contact with the police, have a matching police report (Paper II). The conducted interviews indicate that at least part of the many accidents unreported by the police are in fact reportable by them, but the accidents are either misclassified as extra-accidents or a report is not filed at all. This knowledge of misclassification and underreporting in both police and hospital records is only made possible when employing self-report regimes.

Paper IV shows that even answers to items highly susceptible to social desirable responding can be utilized in traffic safety work. Here an example was made of calculating Odds Ratio on respondents' answers on perceived fault in the accident and the admission of a variety of accident causation factors. This showed that being in a hurry or being occupied with auditory tasks were the two causation factors that the respondents believed to be least associated with being at fault in the accident. This indicates e.g. the need for campaigns to address the dangers of hurrying while driving. Thus, social desirability in answers is not necessarily a hindrance for the utilization of self-reported information.

Comparing the survey answers from the Emergency Room dataset to the Danish subset of the InDeV dataset, there is found significant difference in both perception of fault, choice of causation factors and accident specific information such as mode of transport and place of accident. This illustrates to what extent sampling methods affect the outcome of the self-report surveys. It is an important reminder that comparison of results is very difficult across various self-report regimes and demonstrates why careful consideration should be undertaken at the early stages of survey design to ensure that the chosen design is fitting for obtainment of the needed information.

To what extent can self-reports be applied to minimize the impact of underreporting on accident prevention in Denmark?

The thesis finds limitations in the application of self-reports in the accident prevention work in Denmark. Due to variations in response propensities, self-reports of traffic accidents seem unfit for estimation of accident frequency in the general population or any other calculations where representativity of the population is needed. Another important shortcoming is the difference in accident definition between police records and what can be applied within a self-report regime, which makes it erroneous to compare the number of self-reported accidents with the police records to estimate underreporting. Self-reports are therefore not recommended for estimation of underreporting or in other aggregated generalized estimations.

Yet when representativity is not as important, the self-reports of traffic accidents are found to be highly applicable. The self-report regimes show quite successful in obtaining information from respondents; among respondents with accidents, the response rate is high and dropout rate is low. Whereas the problem of representativity arises when wanting to aggregate the self-reported data, only minor issues with the validity of the self-reported information on the individual level is found. Thus, the correctness of self-reported accident information is reasonable – there is a fair correctness of reported information and consistency within answers. Many respondents are able to locate their accident on a map or chose a correct accident situation. Much of the self-reported questionnaire data is consistent both within the questionnaire but also when compared with respondent interviews. Disagreement with other data sources have to some extent been found to be attributable to erroneous recordings in the other data sources. Carefulness should still be applied when interpreting some of the answers from self-report regimes as social desirability is expected to affects some items, especially on accident causation factors. Nevertheless, even the answers affected most by social desirability might have an application in traffic safety work. As such, self-reports seem rightfully applied in the site-specific safety work, where each report could contribute as supplement to police recordings.

REFERENCES

Abay, K. A. (2015). Investigating the nature and impact of reporting bias in road crash data. *Transportation Research Part A: Policy and Practice,* Volume 71, Pages 31-45. https://doi.org/10.1016/j.tra.2014.11.002

Ahmed, A., Sadullah, A., Yahya, A. (2017). Errors in accident data, its types, causes and methods of rectification - analysis of the literature. *Accident Analysis & Prevention*, 2017 in Press. https://doi.org/10.1016/j.aap.2017.07.018

AIB – The Danish Road Traffic Accident Investigation Board (2014). *Why do road traffic accidents happen?* København: Vejdirektoratet. Available: http://www.hvu.dk/SiteCollectionDocuments/HVUdec14_UK_HvorforSkerUlykkerne.pdf

Andersen, C. S., Kamaluddin, N. A., Várhelyi, A., Madsen, T. K. O., & Møller, K. M. (2016). *Review of current study methods for VRU safety: Appendix 7 – Systematic literature review: Self-reported accidents.* Warsaw: Warsaw University of Technology. Available: https://vbn.aau.dk/ws/portalfiles/portal/263610045/Review_of_current_study_methods_for_VRU_safety_Appendix_7_Systematic_literature_review_Self_reported_accidents.pdf

Brookhuis, K. (2011). Anders af Wåhlberg: Driver Behaviour and Accident Research Methodology; Unresolved Problems. *European Journal of Transport and Infrastructure Research.* Volume 11, issue 1, pp. 96-97. Available: https://d1rkab7tlqy5f1.cloudfront.net/TBM/Over%20faculteit/Afdelingen/Engineering%20Systems%20and%20Services/EJTIR/Back%20issues/11.1/2011_01_05%20Book%20review.pdf

Christiansen, H. & Skougaard, B. Z. (2015). *Documentation of the Danish National Travel Survey. Report 10.* Published by: DTU Transport, Data- and Modelcenter, Kgs. Lyngby, Denmark. ISSN: 1601-9458 (Electronic version). Available: https://www.cta.man.dtu.dk/-/media/Centre/Modelcenter/tu/TU_english-version-2015.ashx?la=da&hash=D6574DA0B9DC551907D7CBB75D58E-6EB297CE41A

Chung, Y., Chang, I., (2015). How accurate is accident data in road safety research? An application of vehicle black box data regarding pedestrian-to-taxi accidents in Korea. *Accident Analysis & Prevention,* Volume 84, Pages 1-8. https://doi.org/10.1016/j.aap.2015.08.001

Corbin, J., & Morse, J. M. (2003). The Unstructured Interactive Interview: Issues of Reciprocity and Risks when Dealing with Sensitive Topics. *Qualitative Inquiry*, *9*(3), 335–354. https://doi.org/10.1177/1077800403009003001

Derriks, H. M., & Mak, P. M. (2007). *IRTAD special report – Underreporting of road traffic casualties*. The Netherlands: Ministry of Transport, Public Works and Water management.
Available:
https://www.who.int/roadsafety/publications/irtad_underreporting.pdf

Dillman, D. A., Christian, L. M., & Smyth, J. D. (2014). *Internet, phone, mail, and mixed-mode surveys - the tailored design method* (4th ed.) John Wiley & Sons, Incorporated. ISBN: 9781118921302 (e-book).

Elvik, R. and Mysen, A.B. (1999). Incomplete Accident Reporting - Meta-Analysis of Studies Made in 13 Countries. *Transportation Research Record*, (1665), pp. 133-140.

Elvik, R., Hoye, A, Vaa, T. Soerensen, M. (2009). *The Handbook of Road Safety Measures* (2nd ed.) Emerald Group Publishing Limited. ISBN: 9781848552500

Gordis, L. (1979). Assuring quality of questionnaire data in epidemiologic research. *American Journal of Epidemiology*. Volume 109, issue 4, pp.21–24. https://doi.org/10.1093/oxfordjournals.aje.a112654

Hansen, S., Thedchanamoorthy, S., & Bloch, A-G. (2014). *NOTAT - Projekt Cykeljakken: Rekruttering og gennemførelse.* Aalborg: Department of Civil Engineering, Aalborg University. DCE Technical reports, Nr. 169.
Available:
https://vbn.aau.dk/ws/portalfiles/portal/197572441/NOTAT_Projekt_Cykeljakken_Rekruttering_og_gennemf_relse.pdf

Hemdorff, S. (2004). *Parameters and values in the Danish road accident registration 2004. Translation Danish – English.* Road Directorate, Copenhagen, Denmark.
Available:

http://docplayer.dk/4146762-Parameters-and-values-in-the-danish-road-accident-registration-translation-danish-english.html

Janstrup, K. H., Hels, T., Kaplan, S., Sommer, H. M., & Lauritsen, J. (2014). Understanding traffic crash underreporting: linking police and medical records to individual and crash characteristics. *In Transport Research Arena 2014, Paris, France.*
Available:
http://orbit.dtu.dk/files/112762602/TRA_2014_Janstrup_Hels_Kaplan_Sommer_Lauritsen.pdf

Jensen, S. U. (2008). *Effektkatalog – Viden til bedre trafiksikkerhed*. Trafitec. Lyngby, Denmark.
Available:
https://www.trafitec.dk/sites/default/files/publications/effektkatalog_viden%20til%20bedre%20trafiksikkerhed.pdf

Kamaluddin, N. A, Andersen, C. S., Larsen, M. K., Meltofte, K. R. , & Várhelyi, A. (2018). Self-reporting traffic crashes – a systematic literature review. *European Transport Research Review*. 10:26.
https://doi.org/10.1186/s12544-018-0301-0

Kaplan, S., Janstrup, K. H., Prato, C.G. (2016). Investigating the reasons behind the intention to report cycling crashes to the police and hospitals in Denmark. *Transportation Research Part F: Traffic Psychology and Behaviour.* Volume 44, January 2017, Pages 159-167.
https://doi.org/10.1016/j.trf.2016.11.009

Lahrmann, H., Madsen, T. K. O., Olesen, A. V., Madsen, J. C. O., & Hels, T. (2018). The effect of a yellow bicycle jacket on cyclist accidents. *Safety Science,* 108, 209-217.
https://doi.org/10.1016/j.ssci.2017.08.001

Lahrmann, H., Madsen, J.C.O., Madsen, T.K.O., Olesen, A.V., Hansen, S., Thedchanamoorthy, S., Bloch, A., (2014). *Projekt Cykeljakken: Den sikkerhedsmæssige effekt af en gul cykeljakke.* DCE Technical Report nr. 167. Aalborg: Aalborg University, Department of Civil Engineering.
Available:
https://vbn.aau.dk/ws/portalfiles/portal/197570851/Projekt_cykeljakken_den_sikkerhedsm_ssige_effekt_af_en_gul_cykeljakke.pdf

Lahrmann, H., & Madsen, J. C. O. (2014). *NOTAT - Projekt Cykeljakken: to webbasserede spørgeskemaer*. Aalborg: Department of Civil Engineering, Aalborg University. DCE Technical reports, Nr. 171.

Available: https://vbn.aau.dk/ws/portalfiles/portal/197571935/NOTAT_Projekt_Cykeljakken_to_webbaserede_sp_rgeskemaer.pdf

Lajunen, T. & Özkan, T (2011). *Chapter 4 - Self-Report Instruments and Methods*. In: Handbook of traffic psychology. Academic Press, San Diego, pp 43–59. https://doi.org/10.1016/C2009-0-01975-8

Larsen, M.K., Jensen A.K.S., Andersen, C.S. (2017). Potentiale for flere uheldsdata på baggrund af fejlklassificeringer af ekstrauheld. *Proceedings from the Annual Transport Conference at Aalborg University*. Aalborg, Denmark. Available: http://www.trafikdage.dk/td/papers/papers17/474_AnneKriegbaum.pdf

Madsen, T. K. O. (2018). *Video Analysis and Mapping of Vulnerable Road Users' Safety*. (Ph.D. thesis). Aalborg Universitetsforlag. Ph.d.-serien for Det Ingeniør- og Naturvidenskabelige Fakultet, Aalborg Universitet. https://doi.org/10.5278/vbn.phd.eng.00057

Meunier, J.-C., Dupont, E., Mersch, J. & Van den Berghe, W. (2018). *My Life After the Crash - MyLAC: An international study on medical, psycho-social and economic consequences of road injuries.* Brussels, Belgium: Vias institute. Available: https://fevr.org/wp-content/uploads/2018/05/DEF_MyLACReport.pdf

Møller, K. M., Madsen, T.K.O., Olesen, A.V. Lahrmann, H. (2015) *A case study on agreement between self-reported bicycle accidents and hospital and police records.* Selected Proceedings from the Annual Transport Conference at Aalborg University 2015. Available: http://www.trafikdage.dk/td/papers/papers15/324_KatrineMeltofte_fagfaellebedoemt.pdf

Møller, K. M., Andersen, C. S., Várhelyi, A., Schönebeck, S., Reumers, S., Hosta, P., & Szagała, P. (2017). *Accident Information from six European Countries Based on Self-reports*. Bruxelles: Publications Office of the European Union. Available: http://vbn.aau.dk/files/262153504/Accident_Information_from_six_European_Countries_Based_on_Self_reports.pdf

Møller, M., Clemmensen, M. B., & Janstrup, K. H. (2017). *Bedre trafikuheldsdata: Danske erfaringer med brug af forskellige typer trafikuheldsdata.* DTU Management. Lyngby, Denmark.

Available:
http://orbit.dtu.dk/files/145213300/Rapport_Bedre_trafikuheldsdata.pdf

Olszewski, P., Osińska, B., Skoczyński, P., Zielińska, A. (2017). *Review of current study methods for VRU safety, Part 1 – Main report . Chapter 4.* Bruxelles: Publications Office of the European Union.
Available:
https://www.indev-project.eu/InDeV/EN/Documents/pdf/2-1-1.pdf?__blob=publicationFile&v=2

Paulhus, D. & Reid D. B. (1991). Enhancement and denial in Socially Desirable Responding. *Journal of Personality and Social Psychology*. Vol 60, No 2, 307–317.
doi:10.1037/0022-3514.60.2.307

Polders, E. & Brijs, T. (ed.) (2018), *How to analyse accident causation?: A handbook with focus on vulnerable road users*. Hasselt University.
Available:
https://www.indev-project.eu/InDeV/EN/Handbook/Handbook-InDeV.pdf?__blob=publicationFile&v=1

Statistics Denmark (2019a). Statistikbanken: *Tabel Moerke*.
Available:
https://www.statistikbanken.dk/statbank5a/SelectVarVal/Define.asp?MainTable=MOERKE&PLanguage=0&PXSId=0&wsid=cftree

Statistics Denmark (2019b). Statistikbanken: *Tabel UHELD4.*
Available:
https://www.statistikbanken.dk/statbank5a/SelectVarVal/Define.asp?MainTable=UHELD4&PLanguage=0&PXSId=0&wsid=cftree

Statistics Denmark (2017). Statistikbanken: *Tabel BY2.*
Available:
https://www.statistikbanken.dk/statbank5a/SelectVarVal/Define.asp?MainTable=BY2&PLanguage=0&PXSId=0&wsid=cftree

Sundhedsdatastyrelsen (2019). *Fællesindhold for Basisregistrering af Sygehuspatienter. Vejledningsdel.* 28. udgave af rapport nr. 6. Sundhedsdatastyrelsen, Copenhagen, Denmark.
Available:
https://sundhedsdatastyrelsen.dk/da/rammer-og-retningslinjer/om-patientregistrering/patientregistrering-feallesindhold

Thedchannamoorthy, S., Madsen, T. K. O., Araghi, B. N., & Lahrmann, H. (2014). *NOTAT - Projekt Cykeljakken: En spørgeskemaundersøgelse*. Aalborg: Department of Civil Engineering, Aalborg University. DCE Technical reports, Nr. 170.
Available:
https://vbn.aau.dk/ws/portalfiles/portal/220287420/NOTAT_Projekt_Cykeljakken_En_sp_rgeskemaunders_gelse.pdf

Tivesten, E., Jonsson, S., Jakobsson, L., & Norin, H. (2012). Nonresponse analysis and adjustment in a mail survey on car accidents. *Accident Analysis and Prevention,* 48, 401-145.
https://doi.org/10.1016/j.aap.2012.02.017

Transport-, Bygnings- og Boligministeriet (2005). *CIR nr. 57 af 11/07/2005.* European legislation identifier: /eli/mt/2005/57.
Available:
https://www.retsinformation.dk/Forms/R0710.aspx?id=2157

Vejdirektoratet (2017). *Indberetning af færdselsuheld - Rapport 580.*Vejdirektoratet, København, Denmark.
Available:
http://www.vejdirektoratet.dk/DA/viden_og_data/publikationer/Lists/Publikationer/Attachments/957/Indberetrning%20af%20færdselsuheld_web.pdf

Vejdirektoratet (2018). *Trafikulykker for året 2017.* København, Denmark.
Available:
https://www.vejdirektoratet.dk/api/drupal/sites/default/files/publications/trafikulykker_for_ret_2017.pdf

Vejdirektoratet (2019). *Vejman, version 2.4.* Last update march 13th 2019.
Program Available:
http://vejman.vd.dk/update/mui/index.jsp

af Wåhlberg, A. (2009). *Driver Behaviour and Accident Research Methodology: Unresolved Problems*, CRC Press LLC, 2009. ProQuest Ebook Central.
Retrieved from:
https://ebookcentral.proquest.com/lib/aalborguniv-ebooks/detail.action?docID=476442.

af Wåhlberg, A. E. (2011). Memory effects in self-reports of road traffic crashes. In *International conference on driver behaviour and training, 5th, Paris, France* (pp. 283–288). Farnham: Ashgate.

APPENDICES

TABLE OF CONTENTS

APPENDIX A PAPER I

A Case Study on Agreement between Self-Reported Bicycle Accidents and Hospital and Police Records

Katrine Meltofte Møller

Tanja Kidholm Osmann Madsen

Anne Vingaard Olesen

Harry Lahrmann

Selected Proceedings from the Annual Transport Conference at Aalborg University 2015

Denne artikel er publiceret i det elektroniske tidsskrift
Udvalgte artikler fra Trafikdage på Aalborg Universitet
(Selected Proceedings from the Annual Transport Conference at Aalborg University)
ISSN 1903-1092
www.trafikdage.dk/artikelarkiv
Modtaget: 01.10.2015
Godkendt: 01.11.2018

A case study on agreement between self-reported bicycle accidents and hospital and police records

PhD student Katrine Meltofte Møller (krm@civil.aau.dk), Aalborg University
PhD student Tanja K. O. Madsen (tkom@civil.aau.dk), Aalborg University
Assistant Professor Anne Vingaard Olesen (avo@civil.aau.dk), Aalborg University
Associate Professor Harry Lahrmann (hsl@civil.aau.dk), Aalborg University

Abstract

A self-report questionnaire on bicycle accidents was distributed to 6,793 respondents every month for one year (2012/2013). This paper evaluates the agreement between hospital data, police data, and self-reports concerning the number of accidents, as well as the information recorded regarding each accident.

Six hundred and ninety-four accidents were self-reported by 6,793 respondents. In 99 cases, the respondents reported getting medical care at a hospital. Of these self-reported contacts, 74 could be matched to a hospital record, but only 53 of the 74 were registered in the hospital records as traffic accidents. Information on the primary opposite party in the 53 accidents was compared, and moderate agreement was found between self-reports and hospital records (κ=0.63).

In 23 of the self-reported accidents, it was stated that a police report had been recorded, but only 6 accidents could be located in the police records. The sample size was too small to calculate statistical agreement between police records and self-reports.

These findings are relevant to discuss of the level of underreporting of bicycle accidents in Denmark because they indicate that there could be substantially more accidents than noted in official records. Furthermore, this paper provides insight into methodological issues associated with self-reporting of traffic accidents.

1. Introduction

"Underreporting" of road accidents is defined as incompleteness in road accident records. If the level of underreporting is biased (e.g., affected by injury severity, mode of transportation, or age of the road user involved) or not constant over time (i.e., time of year, over the years), the number of accidents reported will be systematically skewed and thus incorrect. The traffic safety initiatives based on these numbers will then be misdirected and inefficient.

Many studies have been conducted in the field of underreporting of traffic accidents (e.g., Elvik, Mysen, 1999; Derriks, Mak, 2007). In order to estimate underreporting, one must compare two or more data sources. But no matter which data sources are compared, we are faced with the same challenge: The rather hypothetical "true" number of accidents is unknown, and so is the completeness and correctness of any given data source. Data from hospital records are usually compared or matched to data from police records to estimate either the degree of underreporting in police records or the total number of accidents (Bull, Roberts, 1973; Broughton, Keigan et al., 2010; Boufous, Ivers et al., 2010; Janstrup, Kaplan et al., 2016). Fewer studies compare police records with other data sources, such as insurance records or self-reports of accidents (Isaksson-Hellman, 2012; Finestone, Guo et al., 2011).

In Denmark, it is estimated that only about 14% of severe bicycle accidents registered at hospitals are also registered by the police. Moreover, the police records have been shown to be systematically skewed (, Janstrup, Kaplan et al., 2016). On the contrary, the completeness and correctness of two other data sources—hospital records and self-reports of accidents—have not yet been evaluated in the scientific literature.

The aim of this research was to identify the discrepancies between three data sources: self-reported accidents, hospital records, and police records. Two questions were asked:

1) What is the probability of the existence of a police/hospital record given that a respondent claims its existence?
2) If the record exits, how consistent is the data provided by the respondent with the data in the police/hospital record?

The main aim of this research was to reveal the differences in information given on the same accident among different data sources, thus contributing to the knowledge of possible conclusional flaws that can arise when basing decisions and calculations on data from these sources.

2. Method

2.1 Participants and questionnaire design

A study based on self-reports of accidents was conducted as part of a project on the traffic safety effects of wearing a brightly-colored jacket while bicycling (Lahrmann, Madsen et al., 2018), in which 6,793 participants were divided into a test group (N=3,402) and a control group (N=3,391) and asked to answer a questionnaire each month for the duration of a year, beginning in November 2012 and ending in October 2013. In the online questionnaire, the respondents were asked whether they had had a bicycle accident in the previous month. If they answered in the affirmative, they were then given more questions about the nature of their accident.

The data used in this paper stem from the 6,793 participants' monthly questionnaires. More information about the project and the questionnaire design can be found in existing publications (Thedchanamoorthy, Madsen et al., 2014; Hansen, Thedchanamoorthy et al., 2014; Lahrmann, Madsen, 2014; Lahrmann, Madsen et al., 2018). It is important to note that the participants were all volunteers; therefore, this sample is not necessarily representative of neither the Danish population in general or the part of the population that cycle regularly. Men made up a small majority of this group (57%), and 76% of the participants were between 35 and 65 years of age. The Danish population consists of 50% men, and 41% of the population is between the ages of 35 and 65 years (Statistics Denmark, 2017). The bicycle is most frequently used by people aged 21-29 and the amount of kilometers per day is quite evenly distributed between men and women

(Transportministeriet 2013). It is unknown whether the participants' views on relevant themes (e.g., attitude towards accident registration and faith in authority) deviated from the general population norm. However, the fact that the participants answered 97.5% of the monthly questionnaires about accidents (Lahrmann, Madsen et al., 2014) indicates that they were extraordinarily dedicated. In comparison, a study from Sweden on accidents with a questionnaire comparable to that used in this study had a return rate of 35% (Tivesten, Jonsson et al., 2012).

2.2 Data sources and definitions

A traffic accident is reportable by the Danish police if one of the following circumstances is present (Hemdorff, Lund et al., 2003):

- A person was injured (bruises and abrasions do not count as injuries)
- Material damage of more than 50,000 DKK per vehicle or 5,000 DKK per other equipment occurred
- Foreigners were involved, and a compensation claim is made against them
- A person employed by the police was involved in the accident
- There was a violation of the Road Traffic Act that should lead to preliminary charges.

In contrast to this, the hospital personnel keep records of all people who seek medical care, no matter the severity of their injuries or the extent of material damages.

In the self-report questionnaire, respondents could list all accidents that they had encountered, thus also accounting for minor accidents with little or no personal injury or material damage.

Thus, there are differences among the three types of data sources in terms of the severity of accidents one could expect to find recorded. Hospital data cover accidents resulting in personal injuries only, police data cover a combination of accidents resulting in personal injuries and accidents resulting in material damage, and finally, self-reported accidents are expected to cover all accidents regardless of severity.

Different definitions of what constitutes an accident are also applied among the three data sources.
The police define a traffic accident as an accident occurring on a public road/square/area in which at least one of the involved parties was driving (Hemdorff, Lund et al., 2003). The hospital defines a traffic accident as an accident involving a means of transport or an animal being used to transport persons or goods from one place to another (Sundhedsstyrelsen, 2009). The self-report survey defines an accident as an event occurring on a public road in Denmark in which the self-reporter was traveling by bike and at least one of the following instances occurred (Lahrmann, Madsen et al., 2014):

- The bicyclist made physical contact with an opposite party
- The bicyclist fell over, was hurt, or experienced material damage because of an opposite party's behavior; this also includes events in which no physical contact occurred
- The bicyclist fell over or was hurt while biking, even though there were no other parties involved

The information recorded in the three data sources varies as well. Table 1 illustrates a simplified form of the information requested via the questionnaire, as well as information recorded by the police and the hospital.

Police records contain various categories of accident information, both as multiple choice data and descriptive fields, and police personnel are required to fill out all fields (cf. Table 1). Hospital records contain, as a standard, very little mandatory accident-related information (cf. Table 1). The self-reported records contain information on the same themes as police records, with some additions (cf. Table 1).

Table 1: The design of the monthly questionnaire in comparison with police and hospital records. Only data from police and hospital records comparable with the self-reports are shown in the table; police and hospital records contain more information than shown below (e.g., alcohol consumption in police records). [M]=Multiple choice, [D]=Descriptive, [Opt]=Optional, not mandatory to record, [A]=Only recorded if activity is filed as a traffic accident. The emergency room in Odense on the Danish island of Funen carries out extended recording (Ulykkes Analyse Gruppen, 2014), which is not shown here.

Topic	Self-reported information	Police records	Hospital records
Time and date of accident	Date[M] Month[M] Year[M] Time (hourly interval)[M]	Date[M] Month[M] Year[M] Time (exact)[M]	
Place	Location – Coordinates from maps.google.dk Place[D]	Location – Distance from a selected point	Place [M]
Circumstances	Road surface[M] Visibility[M] Day/night[M] Weather[M] Street light[M]	Road surface[M] Visibility[M] Day/night[M] Weather[M] Street light[M]	Road surface[Opt] Activity [M]
Safety measures	Use of lights on bike[M] Use of fluorescent bike jacket[M] Use of helmet[M]	Use of lights on bike[M] Use of helmet[M]	Use of helmet[Opt]
Accident	Other parties or solo accident[M] How many parties[M] Primary opposite party [M] Accident description[D]	Other parties or solo accident[M] How many parties[M] Primary opposite party [M] Accident description[M]	Other parties or solo accident[M, A] Primary opposite party [M, A]
Severity	Most severe injury[M] Injuries[D] Medical treatment[M] Absenteeism[M] Police report[M] Insurance report[M]	Most severe injury[M] Medical treatment[M]	Injuries[M] Medical treatment[M]

2.3 Matching records

All participants had to provide their CPR number (*Central Person Register*; a Danish civil registration system that provides each citizen with a unique identifier) in their participant registration. A match was based on matching CPR numbers across different datasets. Time constraints must also be taken into account; a respondent could have had several different accidents during the one-year study period, with no guarantee that every experienced accident was either self-reported or recorded by the police or the hospital. We considered the records to match if the date of the contact to the hospital was between one month prior to the self-reported accident date and two months after. Because respondents were asked to self-report every month, it would be highly unlikely for them to forget their accident date and state the accident date as more than one month off. To ensure that respondents who did not seek medical attention immediately after their accident were not excluded from the match, we considered accidents that were recorded by the hospital up to two months after the self-reported accident date to be a match.

If it was clear from the data in the LPR (*Landspatientregisteret;* the National Patient Registry, which is the source of the hospital data) that an entry could not describe a traffic accident, we excluded the match from the study. When we could not exclude the possibility that the injuries were sustained from a traffic accident, the records were considered a match.

Whereas police records are based on accidents, hospital records are based on injuries and contact with the hospital. Thus, one person can have multiple records in the hospital database due to one accident. To avoid counting duplicates of accidents in hospital records, we manually handled all the matching hospital records and discarded duplicates.

The same time constraints were applied to the matching of self-reports and police records, along with the criteria that the respondent was riding a bicycle during the accident.

A schematic overview of the matching procedure is shown in Figure 1.

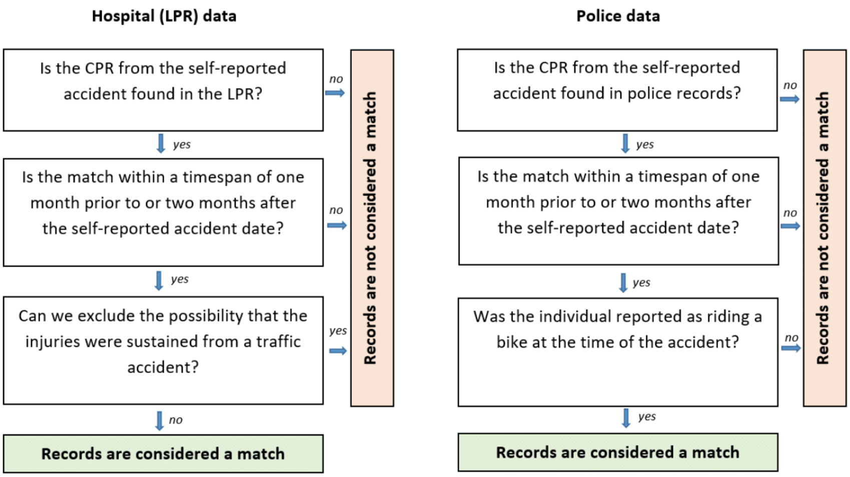


Figure 1: Schematic overview of the matching procedure of self-reported accidents with LPR data and police data.

3. Results

The participants reported 833 accidents during the project period, but some were discarded by the project team (139) because they did not meet the definition of an accident set within the study. Hence, the study is based on 694 self-reported bicycle accidents.

3.1 Self-reported contacts to hospital, police, and insurance companies

As can be seen from Table 2, the hospital was the most frequently contacted, closely followed by insurance companies. Police contact and visits to a general practitioner (GP) were much more uncommon. The overlap of records can be seen in Figure 2 in Appendix 1.

Table 2: Entities with which self-reporters claim to have made contact due to their accident; N=694. (*)The self-reporters did not specifically answer "no" to this because the question about seeking medical attention had a different phrasing.

	Yes	No	Do not know
Police	23	667	4
Insurance	96	585	13
Hospital / Emergency room	99	595*	0
Only GP and not hospital	21	673*	0

The rules for matching, as described in the *Method* section, yielded 74 matches between respondents' self-reported accidents and accidents registered in the LPR.

Twenty-one out of the 74 matching records were not registered in the LPR as traffic accidents (cf. Table 3 in Appendix 1) but with other activities leading to the accident.

Eleven of the 74 respondents who were matched to a record in the hospital database stated that they did not receive medical care (cf. Table 4 in Appendix 1).

The rules for matching, as described in the *Method* section, yielded 6 recorded accidents in the police records.

3.2 Consistency between LPR data and self-reports

The 53 matches in the LPR registered as traffic accidents provide information about the primary opposite party that can be compared with the self-reported information. Table 5 shows the primary opposite parties listed in the hospital traffic accident records and the primary opposite parties listed in the self-reported questionnaires. More solo accidents and fewer bicyclists as opposite parties were registered in the hospital records.

Table 5: The consistency between primary opposite parties listed in self-reports and opposite parties entered in the hospital database, N=53. [*]=The opposite party could not be registered by EUM-code (EUM-codes are part of the classification system used in the LPR; for a thorough introduction to the coding system, the reader is referred to Sundhedsstyrelsen (2009)).

Primary opposite party	LPR [number of cases]	Self-registered [number of cases]
Solo accident – no opposite party (EUM0 or EUU0)	32	24
Pedestrian (EUM1)	3	2
Bicyclist (EUM2)	4	10
Moped (EUM3)	1	1
Motorcycle (EUM4)	2	1
Car (EUM5)	10	13
Not specified (EUM9), or other (Truck/bus, animal)[*]	1	2

Cohen's Kappa was calculated to obtain the degree of agreement between the primary opposite party registered in LPR and that registered in self-reports, resulting in κ=0.63 ($p<0.001$), 95% CI (0.7946, 0.4654), which is in the lower region of the moderate interpretation (weak: κ=0.40–0.59, moderate: κ=0.60–0.79), (McHugh, 2012).

3.3 Consistency of information between police records and self-reports

The rules for matching only yielded 6 recorded accidents in the police records. The small sample size means that calculation of Cohen's kappa is discouraged; normally, a sample size of at least 30 comparisons is necessary (McHugh, 2012). Although no statistical testing could be conducted, the comparison of information present in the two datasets is given in Table 6; it indicates that most of the information matches.

Table 6: A comparison of information given in self-report questionnaires and registered in police records of the same accident. The number indicates how many of the 6 records contain the given information.

	Self-reports	Police records
Weather	6: No rain	6: No rain
Road surface	6: Dry	5: Dry 1: Wet
Daylight	4: Daylight 1: Night 1: Does not remember	5: Daylight 1: Night
Lighting	1: No street lights present	1: No street lights present

	1: Street lights lit 3: Street lights unlit 1: Does not remember	1: Street lights lit 4: Street lights unlit
Visibility	5: Good visibility 1: Does not remember	6: Good visibility
Helmet	5: Using helmet 1: Not using helmet	5: Using helmet 1: Not using helmet
Injuries	1: Femoral fracture 1: Concussion 1: Bruises and similar 3: Other	2: Fracture or ligament injury in hip/leg/foot 2: Slight injuries or admitted for observation (concussion) 1: Lesion on spine/neck/pelvis 1: Fracture or ligament injury in shoulder/arm/hand

4. Discussion

The rather low probability of a self-reported contact with the hospital resulting in a matching record could be affected by several factors. Measurement error (i.e., participants providing inaccurate answers to the questions asked [Dillman, Christian et al., 2014]) due to the questionnaire design seems highly relevant in a discussion of this because the Danish medical service enables people to call a doctor outside office hours[1] and to come in for a physical examination. This emergency service doctor is often, although not always, located at the hospital. Thus, it is quite possible that some respondents contacted the emergency service doctor and regarded this as having contacted the hospital due to the location of the doctor's office at the hospital. However, the emergency service doctor is in fact not a ward in the hospital and does not report his or her patients to the LPR; thus, the emergency doctor service records are not included in this study.

The low probability of the existence of a police record when a respondent claims to have been in contact with the police could be influenced by sampling error (i.e., error that occurs when only some of the members of a sample frame are selected to be surveyed [Dillman, Christian et al., 2014]). Because the sample list consists of volunteers, the individuals surveyed might be more interested in traffic safety than cyclists or the public in general and very keen to contribute to information on bicycle accidents. This could result in the respondents being more likely to contact the police than the average bicyclist, even with accidents that do not meet the police criteria for classification of traffic accidents, thus biasing the results. That self-reports can be subject to statistical bias due to possible overreporting by some subgroups has previously been argued (Tivesten, Wiberg, 2013; Wåhlberg 2009). But even if the sample is viewed as an overreporting subgroup, this does not explain the 12 respondents who claimed they had been in contact with the police while also claiming injuries worse than bruises and abrasions. It is not possible to determine whether the discrepancies can be attributed to measurement errors or erroneous recording practices by the police; this would require further study.

The overall trend of respondents reporting more accidents than can be found in the official datasets coincides with the results of other research comparing self-reports with police-recorded accidents (Arthur, Tubre et

[1] In Danish terms: "Vagtlægen"

al., 2001), (Boufous, Ivers et al., 2010). However, there are no studies comparing self-reported medical attention and official hospital records with which to compare these findings.

The accidents recorded as traffic accidents in the LPR are usually used to calculate the level of underreporting of accidents Denmark (Statistics Denmark, 2014). Our findings indicate that the level of underreporting by the police could actually be much higher than previously thought due to the mislabeled reports in the hospital database, as well as the high proportion of injured individuals seeking medical care at their GP instead of the hospital. There is a possibility that the number of injuries not labeled in the LPR as traffic injuries could be counterbalanced to some degree by people seeking medical care for injuries sustained due to something other than a traffic accident and being mislabeled as victims of traffic accidents. This is mentioned as a source of error by Roberts, Vingilis et al. (2008), but because it is much more time consuming for hospital personnel to register someone as injured in traffic than as otherwise injured (due to the fact that more information must be included in records of traffic accidents than in any other type of accident), it seems more likely that the bias in a busy emergency room would lean towards registering fewer traffic accidents. This speculation is consistent with the results shown in Table 3, in which 10 of the 21 mislabels are somewhat vague (EUA79, EUA8, EUA9). Table 3 also shows that 8 of the mislabels are found in the category EUA5: *Sport and exercise*. This indicates that further studies of the records in that category could yield knowledge about accidents that are mislabeled as traffic accidents. These findings are limited to bicycle accidents. It is not known how the rate or distribution of mislabels is affected by the mode of transportation.

5. Conclusion

What is the probability of the existence of a police/hospital record given that a respondent claims its existence?

Of the respondents, 99 reported receiving medical care at the hospital or the emergency room due to their accidents, 74 could be matched by their CPR number with an accident that happened between one month prior to the self-reported accident date and two months after, and of these 74 entries, 21 were not labeled as traffic accidents in the hospital records. Thus, the reporting level of bicycle accidents in hospital records was 53/99 * 100 = 53.54%.

Twenty-four respondents reported contact with the police due to their accidents. Only 6 respondents could be matched to police records. Thus, there was a police record for 25% of the respondents who claimed to have been in contact with the police. Of the 24 respondents, 12 stated that their most serious injury was bruises, abrasions (7), or no physical injury at all; thus, half of the accidents reported to the police did not meet the injury criteria for police reportable accidents in Denmark (cf. section 2.3, *Data sources and definitions*). The self-reported injuries of the remaining 12 respondents who contacted the police were so severe that they met the criteria for accidents reportable by the police.

If the record exists, how consistent is the data provided by the respondent with the data in the police or hospital record?

Cohen's kappa was calculated to obtain the degree of agreement between information in hospital records and in self-reports of accidents, yielding κ=0.63, which is a moderate degree of agreement between the two data sources. The main contributory factor to the discrepancies seems to be that the hospital wrongfully recorded bicycle accidents with another bicyclist as counterpart as single accidents.

The 6 matching police records and the self-reported accident information seemed fairly consistent with each other, but due to the low number of matching records, it was not possible to calculate Cohen's kappa.

Self-reports have not previously been applied to explore underreporting and consistency in hospital data, but the fact that hospital data suffers from underreporting is in line with the findings of Janstrup, Kaplan et al. (2016). The finding that 28% of the matches in hospital data are not recorded as traffic accidents and that there is only moderate agreement between the information provided by self-reporters and the information in the matching hospital records calls for awareness. Thus, researchers should take care when utilizing hospital data for traffic research purposes if steps are not taken to heighten the focus of hospital personnel to improve their recording practices.

References

ARTHUR, W.J., TUBRE, T., DAY, E.A., SHEEHAN, M.K., SANCHEZ-KU, M.L., PAUL, D., PAULUS, L. and ARCHULETA, K., 2001. Motor Vehicle Crash Involvement and Moving Violations: Convergence of Self-Report and Archival Data. *Human Factors: The Journal of the Human Factors and Ergonomics Society*, 43(1), pp. 1-11.

AUSTIN, K.P., 1994. *The Identification of Mistakes in Road Accident Records – Part Two. Working Paper* 407. University of Leeds, Institute of Transport: White Rose University Consortium.

AUSTIN, K.P., 1993.*The Identification of Mistakes in Road Accident Records – Part One: The Use of Geographic Information Systems.* Working Paper 406. University of Leeds, Institute of Transport: White Rose University Consortium.

BOUFOUS, S., IVERS, R., SENSERRICK, T., STEVENSON, M., NORTON, R. and WILLIAMSON, A., 2010. Accuracy of Self-report of On-road Crashes and Traffic Offences in a Cohort of Young Drivers: The DRIVE Study. *Injury Prevention*, (16), pp. 275-277.

BROUGHTON, J., KEIGAN, M., YANNIS, G., EVGENIKOS, P., CHAZIRIS, A., PAPADIMITRIOU, E., BOS, N.M., HOEGLINGER, S., PÉREZ, K., AMOROS, E., HOLLÓ, P. and TECL, J., 2010. Estimation of the Real Number of Road Casualties in Europe. *Safety Science*, 48, pp. 365-371.

BULL, J.P. and ROBERTS, B.J., 1973. Road Accident Statistics - A Comparison of Police and Hospital Information. *Accident Analysis and Prevention*, 5, pp. 45-53.

DERRIKS, H.M. and MAK, P.M., 2007. *IRTAD SPECIAL REPORT - UNDERREPORTING OF ROAD TRAFFIC CASUALTIES.* The Netherlands: Ministry of Transport, Public Works and Water Management.

DILLMAN, D.A., CHRISTIAN, L.M. and SMYTH, J.D., 2014. *Internet, Phone, Mail, and Mixed-Mode Surveys - The Tailored Design Method.* 4 edn. John Wiley & Sons, Incorporated.

ELVIK, R. and MYSEN, A.B., 1999. Incomplete Accident Reporting - Meta-Analysis of Studies Made in 13 Countries. *Transportation Research Record*, (1665), pp. 133-140.

HANSEN, S., THEDCHANAMOORTHY, S. and BLOCH, A., 2014. *NOTAT - Projekt Cykeljakken: Rekruttering og gennemførelse.* Aalborg: Department of Civil Engineering, Aalborg University.

HEMDORFF, S., LUND, H.V. and TAUL, B., 2003. *Indberetning af færdselsuheld: Vejledning.* Raport 277. Copenhagen, Denmark: Vejdirektoratet.

JANSTRUP, K.H., KAPLAN, S., HELS, T., LAURITSEN, J. and PRATO, C.G., 2016. Understanding Traffic Crash Under-reporting. Linking Police and Medical Records to Individual and Crash Characteristics. *Traffic Injury Prevention*, 17(6), pp. 580-584.

LAHRMANN, H., MADSEN, J.C.O., MADSEN, T.K.O., OLESEN, A.V., HANSEN, S., THEDCHANAMOORTHY, S. and BLOCH, A., 2014. *Projekt Cykeljakken: Den sikkerhedsmæssige effekt af en gul cykeljakke.* DCE Technical Report nr. 167. Aalborg: Aalborg University, Department of Civil Engineering.

LAHRMANN, H., MADSEN, T.K.O., OLESEN, A.V., MADSEN, J.C.O. and HELS, T., 2018. The Effect of a Yellow Bicycle Jacket on Cyclist Accidents. *Safety Science*, 108 pp. 209-217.

MCHUGH, M.L., 2012. Interrater Reliability: The Kappa Statistic. *Biochemia Media (Zagreb)*, 22(3), pp. 276-282.

ROBERTS, S.E., VINGILIS, E., WILK, P. and SEELAY, J., 2008. A Comparison of Self-reported Motor Vehicle Collision Injuries Compared with Official Collision Data: An Analysis of Age and Sex Trends Using the Canadian National Population Health Survey and Transport Canada Data. *Accident Analysis and Prevention*, 40, pp. 559-566.

STATISTICS DENMARK, 2017-last update, Befolkningstal, FOLK22017.
Available: http://www.statistikbanken.dk/FOLK2 [2017]

STATISTICS DENMARK, 2014-last update, Tabel MOERKE.
Available: http://www.statistikbanken.dk/MOERKE [2015].

SUNDHEDSSTYRELSEN, 2009. *Fællesindhold for basisregistrering af sygehuspatienter 2010.* www.sst.dk: Sundhedsstyrelsen.

THEDCHANAMOORTHY, S., MADSEN, T.K.O., ARAGNI, B.N. and LAHRMANN, H., 2014. *NOTAT - Projekt Cykeljakken: En spørgeskemaundersøgelse.* Aalborg: Department of Civil Engineering, Aalborg University.

TIVESTEN, E., JONSSON, S., JAKOBSSON, L. and NORIN, H., 2012. Nonresponse Analysis and Adjustment in a Mail Survey on Car Accidents. *Accident Analysis and Prevention*, 48, pp. 401-145.

TIVESTEN, E. and WIBERG, H., 2013. What Can the Drivers' Own Description from Combined Sources Provide in an Analysis of Driver Distraction and Low Vigilance in Accident Situations? *Accident Analysis and Prevention*, (52), pp. 51-63.

TRANSPORTMINISTERIET 2013. *Arbejdspapir 1: Cykling i Danmark - hvor står vi? Den nationale cykelstrategi 2013*. København, Denmark. Available: https://www.trm.dk/da/temaer/ny-cykelstrategi-2013/den-nationale-cykelstrategi [30. oktober 2018]

ULYKKES ANALYSE GRUPPEN**,** 2014-last update, Ulykkes Analyse Gruppen.
Available: http://www.ouh.dk/wm140123 [28. maj, 2015].

WÅHLBERG, A.A., 2009. Driver Behaviour and Accident Research Methodology: Unresolved Problems. *Human Factors in Road and Rail Transport.* Farnham, England: Ashgate.

Appendix 1

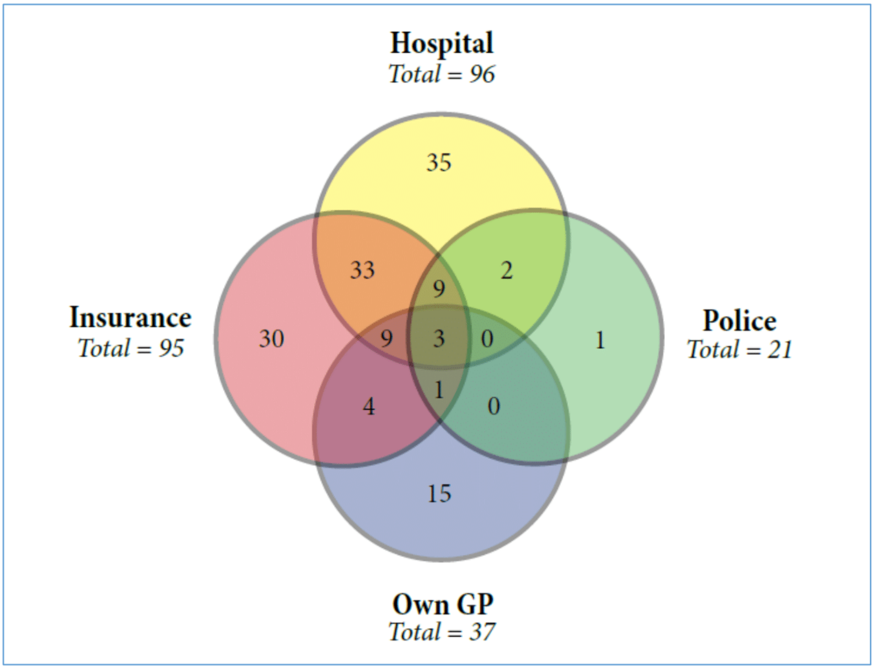


Figure 2: An Euler diagram of the self-reported data. Overlap between contact with police and contact with insurance company only=6, overlap between contact with own GP and contact with hospital only=5. Answers of "do not know" were excluded from the diagram; this is the reason for the inconsistency between the total numbers and the total numbers shown in the Euler diagram.

Table 3: The recorded activities leading to injuries registered in the LPR for all respondents for which there was a match between hospital records and self-reported accidents.

Activity	Code	Number of accidents	Considered a traffic accident in the LPR?
Paid transportation work	EUA0	1	Yes
Transport between workplaces	EUA02	2	
Paid work	EUA1	1	No
Paid work, unspecified	EUA19	1	
Transportation in spare time	EUA2	39	Yes
Transport to/from work	EUA21	1	
Other unspecified transportation	EUA28	1	
Transportation, unspecified	EUA29	9	
House activities and unpaid work	EUA3	1	No
Sports during school/education	EUA42	2	No
Sport and exercise	EUA5	8	No
Vital activity, unspecified	EUA79	1	No
Other activity	EUA8	5	No
Activity not specified	EUA9	4	No
Total		74	*Yes= 53* *No=21*

Table 4: The self-reported information regarding receipt of medical care.

		All matched [N=74]	Matched and registered in LPR with a traffic accident [N=53]
Did you receive medical treatment?	Yes	63	49
	No	11	4

APPENDIX B PAPER II

The Quality of Self-Reports of Accidents Compared with Hospital and Police Records

Katrine Meltofte Møller

Anne Vingaard Olesen

Harry Lahrmann

Manuscript

THE QUALITY OF SELF-REPORTS OF ACCIDENTS COMPARED WITH HOSPITAL AND POLICE RECORDS

Katrine Meltofte Møller — *Ph.D. fellow, Civil Engineering Department, Aalborg University, Denmark*

Anne Vingaard Olesen — *Associate professor, Civil Engineering Department, Aalborg University, Denmark*

Harry Lahrmann — *Associate professor, Civil Engineering Department, Aalborg University, Denmark*

Abstract

Self-reports of traffic accidents are becoming more commonly used in traffic safety work, but little is known of their agreement with official records. This paper evaluates a self-report regime in which patients in the emergency room were sampled. The information in self-reports of 411 accidents was compared with hospital records of the same accidents. High agreement (kappa = 0.77–0.94) regarding transport modes was found. Forty-nine self-reports of accidents were compared with their corresponding police records, and moderate to high agreement was found on a number of accident parameters, including transport mode (kappa = 0.93–1), accident situation (kappa = 0.82), and environmental conditions (kappa = 0.51–0.74). The quality of the self-report regime was investigated in two ways: A) By comparing respondents' choice of accident location and accident situation to their written descriptions; the results showed that 85–87% of the information was coherent. B) By investigating the proportion of respondents unable to provide answers to the multiple-choice questions in the questionnaire; results here showed that questions regarding maneuvers before the accident took place as well as choosing an accident situation proved difficult for 17–23% of the respondents. Interviews were conducted with eight respondents; their answers provided insight into the possible reasons for inconsistency between records. Overall, it is concluded that the self-reports showed a high quality in relation to the information's agreement with other official records, and we advocate that self-reported accident information could be considered trustworthy and applicable in analysis for traffic safety purposes. The most notable issues regarding the conducted self-report regime were the respondents' ability to answer questions regarding accident situations and maneuvers, indicating that, if future works could improve questionnaire design, the number of self-reports with complete, trustworthy data would be more numerous.

1. INTRODUCTION

Problems with underreporting of traffic accidents in police records have been investigated for several decades (Elvik & Mysen, 1999; Derriks & Mak, 2007). For instance, one study found that 66–73% of people severely injured in car accidents and recorded at the emergency room (ER) were also recorded by the police, whereas only 6–7% of slightly injured bicyclists from ER records could be matched to a police record (Janstrup, Kaplan, Hels, Lauritsen, & Prato, 2016). To compensate for the lack of complete accident data, many different approaches have been tried in Denmark. Most common are projects that utilize or expand on existing data sources, such as ambulance or hospital records, or a self-report regime providing accident-involved persons with an app/website/questionnaire, etc., through which they can report their accident; an overview is provided by Møller, Clemmensen, and Janstrup (2017). This paper focuses on self-report schemes as a supplement to, or surrogate for, official accident records.

Most studies on the validity, accuracy, and agreement of self-reports and official records investigate the extent to which the number of accidents in the two data sources agrees. The inability of self-reports to reproduce the same accident information as found in official records can be used as an argument for their inaccuracy (e.g., McGwin, Owsley, and Ball (1998), Finestone et al. (2011), and af Wåhlberg (2011)). The implied issues are often that respondents either forget accidents, do not wish to self-report their accident due to social desirability effects, or misplace them in time (i.e., *telescoping*: when people believe an incident is closer to the present time than in reality (Janssen, Chessa, & Murre (2006)). For instance, a study by Porter (2018) that investigated elderly drivers found that, even though the drivers did not self-report some of their accidents in police or insurance records, they self-reported more accidents than what was reported in official records. That self-reports provide more numerous accidents than police records was also found by e.g., Marottoli, Cooney, and Tinetti (1997), Anstey, Wood, Caldwell, Kerr, and Lord (2009), Boufous et al. (2010), and Finestone et al. (2011). This is the opposite to the findings of McGwin et al. (1998), who found the largest discrepancy in the police records of accidents that were not self-reported. A slightly more in-depth analysis by Tin Tin, Woodward, and Ameratunga (2013) reported that agreement between self-reported bicycle accidents and official records was influenced by respondent characteristics: higher levels of agreement were associated with e.g., being male and having a higher level of education. This trend is backed by Singletary et al. (2017), although only one characteristic significantly affected the kappa value for agreement: respondents with impaired contrast sensitivity (*night-blindness*) produced higher agreement with official records. Overall, Singletary et al. (2017) found substantial agreement between self-reports and police records of involvement in motor vehicle collisions and found self-reports to be a valid method of identifying crash involvement.

While the agreement of accident records in self-reports and official data sets has been investigated before, literature on the information validity of accident details (i.e., whether the information that respondents provide regarding their accident is congruent with the information in official records on the same accident) is contrastingly sparse, hence the topic of this research. A previous study by the authors investigated information concordance in a sample of cyclists, although only on the congruency of hospital records (*Landspatientregisteret* hereafter LPR) and self-reported accidents (Møller, Madsen, Olesen, & Lahrmann, 2015). That study showed that there was only moderate agreement between information on the transport mode of the counterpart registered in LPR and the mode registered in self-report questionnaires. Furthermore, that study found that 28% of the self-reported bicycle accidents were not registered in LPR as traffic accidents. This paper therefore examines if the findings from Møller et al. (2015) can be extrapolated to transport modes other than cycling and investigates accident information concordance in self-reports and police records.

To the authors' knowledge, only two other studies have examined information concordance in self-reports and other records, i.e., Begg, Langley, and Williams (1999) and Versteegh (2004). Begg et al. (1999) compared self-reported crash information with police records in a small sample of young adults (n = 25) in New Zealand and found the self-reported crash information to be a valid source of information, with agreement ranging from 100% to 76%. However, the study by Begg et al. (1999) has several limitations. They investigated only six pieces of information regarding the accident (year, day of the week, time of day, speed limit, road user status, and vehicle type), their respondents were all 21 years old, and the recall period was three years. Versteegh (2004) compared accuracy of information on Australian accidents gathered from driver interviews with in-depth accident investigation data (n = 73) and found agreement ranging from 99% to 66% on factors such as lighting conditions and speed zones. Versteegh's study comprehensively listed the examined factors and the methodology, but it did not illustrate usual police recording practice in comparison with the drivers' own descriptions, as all the accidents were so severe that in-depth investigations were carried out and compared.

This paper will add to the knowledge produced by Versteegh (2004), Begg et al. (1999), and Møller (2015) by investigating the accuracy of information provided on the same accident in self-reports, police records, and hospital records. The approach is novel, as sampling for the self-report regime was carried out among patients in the ER, representing different age groups, transport modes, and severity levels.

In contrast to Versteegh (2004) and Begg et al. (1999), we applied kappa values as well as percentage agreement for a better understanding of the self-reported information's correctness. The underlying assumption of our chosen statistical analysis was that we do not regard either police records or hospital records as a gold standard to which self-reports must agree for them to be regarded true. Research has shown that police reports are not flawless. Ahmed, Sadullah, and Yahya (2017) presented a meta-analysis that showed that in high-income countries the errors in police records averaged 25% regarding accident location, 39% on victim's information, 12% for vehicle information, and 15% for environmental factors. Chung and Chang (2015) used vehicle black boxes to investigate accuracy of police records in Korea and also found flawed information in police recording of time, place, and vehicle speeds. How much of this inaccuracy is present in Danish police reports cannot immediately be ascertained, but a study by Larsen, Jensen, and Andersen (2017) showed that 11.5% of police reports of so-called extra-accidents (no injury and no substantial property damage) on rural roads were actually misclassified and contained either injury or property damage. We therefore found it more appropriate to report on the level of agreement between the different reporting modes and apply Cohen's Kappa (see section *2.5 Statistics*) as an indication of how well the different reporting modes agreed. The issues described above related to the flawed nature of police records are also discussed further in this paper. We carried out semi-structured interviews with eight respondents, which exemplified some of the police misrecordings.

The aim of this paper is thus to examine the quality of accident information provided by self-reporters in order to ascertain if the applied method of sampling self-reports in the ER is a viable solution for obtaining accident information. This gives cause for two research questions:

- How well does the self-reported accident information agree with other recorded information on the same accident and what could be the cause of disagreement?

- Can we consider the information provided in the multiple-choice self-report regime trustworthy, i.e., does it show internal consistency?

2. METHODS

2.1 EMERGENCY ROOM SAMPLING

For the thirteen months (October 2016 to November 2017, from here on *the survey period*), patients who had been in contact with the ER in Region Nord in Denmark due to a traffic accident were asked by the nurse on duty in the ER reception if they wanted to participate in a survey regarding their accident. If respondents agreed to take part, an informational letter was sent out to the patient via digital post (see i.e. LifeinDenmark.com 2019) containing a link to the self-report questionnaire. A reminder was sent to non-responders after one month.

Pedestrian falls are not recognized as reportable traffic accidents by the Danish police (Vejdirektoratet 2017). However, hospital personnel were instructed to include patients who experienced a single pedestrian accident (falls on a public road) in the survey. This was the reason for the many pedestrian falls included in the ER registration; while the registration procedure is not representative of normal registration procedure in LPR, it was chosen as pedestrian falls were relevant for other aspects of the research project. We see no reason not to include the answers of pedestrians with falls in the analysis of the quality of information obtained through self-reports, as we have no reason to believe that pedestrians with falls would give poorer or better questionnaire answers than pedestrians with a counterpart in their accident.

2.2 SURVEY DESIGN

The questionnaire contained 35 main items and had an estimated completion time of 20 minutes. Some answers led to a question that would not have been presented if the respondent answered differently—for instance, respondents who indicated that their accident took place at an intersection were concurrently asked if the intersection had signal lights or not. A telephone hotline and email service were available if respondents had difficulties completing the questionnaire.

Most items provided answers in a multiple-choice form, with a variation of *other/none of the above/cannot say* present among the possible choices. Only three items asked the respondent to write text themselves, although it was not mandatory to provide written descriptions. An overview of the item themes is provided below:

Item 1: Time and date of the accident.

Item 2-6: Location of the accidents and the road geometry. The respondent was asked to locate the scene of their accident on a map (maps.google.com). Respondents could zoom, maneuver, and set a pin where the accident happened. The questionnaire software then logged the coordinates of the pinpointed location. Another question asked the respondent to describe in a text field, as accurately as possible, where the accident took place.

Item 7-14: Respondents’ own mode of transport, the use of any safety equipment (such as a helmet if using a bicycle or a seatbelt if driving a car), and how old they were when they obtained their license (if applicable for their chosen transport mode). If responding that they had an accident while in a vehicle, they were also asked to state if they were passengers or drivers.

Item 15-18: Lighting, road surface, and weather conditions.

Item 19: Purpose of the trip, e.g., shopping, business trip, exercise, etc.

Item 20-21: Involvement of any other road users in the accident and, if so, what their means of transport was.

Item 22-24: Based on the respondents' choice of road geometry and whether any other road users were involved in the accident (*counterpart*), they were then asked a question about their movements just before the accident. For instance, if a respondent stated that they had a multiparty accident on a straight road the item would be:

Describe the movements of you and your counterpart just before the accident happened.

- □ *The other road user moved in **the same direction** as me*
- □ *The other road user moved in **the opposite direction** as me*
- □ *One of us made a **U-turn***
- □ *One of us **reversed** when the accident happened*
- □ *I hit a **parked vehicle***
- □ *Other*

Respondents were then asked to choose a simplified accident illustration that described their accident. The questionnaire was tailored so that each respondent was only presented with illustrations that could be possible on the basis of his/her previous answers on geometry, counterpart, and movements.

The illustrations included geometry and the movements of the respondents and their counterpart (if any) immediately prior to the accident— examples of such illustrations can be seen in Figure 1.

The illustrations are similar in overall design to those applied by the Danish police (Vejdirektoratet 2017), with a few additions primarily featuring roundabouts. Illustrations in similar style have previously been applied in a self-report regime in the Horizon 2020 project InDeV (Madsen et al. 2018 and Madsen 2018). Passengers and respondents who stated that their accident took place off public roads were not asked to choose an illustration, and pedestrians who stated they had a fall were not asked either. An item with a text field asked respondents to describe how their accident happened in as much detail as possible.

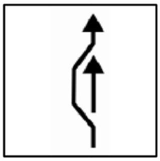
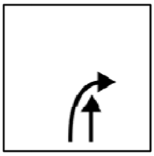
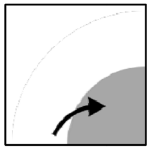
		
I was overtaking another road user on his/her left side or he/she overtook me from my left side	One of us drove straight ahead while the other turned to the right in front of the straight-going road user.	I had an accident on the right-hand side of a right bend.
Situation no. 111	**Situation no. 312**	**Situation no. 23**

Figure 1: Examples of the accident illustrations from the questionnaire. The situation numbers correspond to the equivalent situations applied in police records (Vejdirektoratet 2017).

Item 25: The next item asked who, if anyone, the respondent had been in contact with due to their accident. Multiple-choice options included the police, insurance, road authority, general practitioner, and the Emergency Service Doctor ("vagtlægen" in Danish).

Item 26-28: Physical injuries sustained and their severity in terms of days of absence from work or school.

Item 29-30: Possible accident causation factors and whom the respondent found to be at fault in the accident (if they had a counterpart).

Item 31-36. A text field gave the respondent the opportunity to tell more about their accident if they had anything else they would like to state. Lastly, four items with demographic themes were presented to respondents.

2.3 MATCHING

The matching between respondents and the LPR data was carried out via the respondents' civil registration number (*det Centrale Personregister*, hereafter CPR number). As all respondents were sampled through their contact with the ER, they were all found in LPR. The items investigated in the self-report questionnaire (transport mode, counterpart, and trip purpose) were also all registered in the initial contact with the ER staff.

Matching between police records and respondents was also by CPR number and carried out by Statistics Denmark. We considered all records of the respondents in the police accident database within the survey period to be a match.

The respondents who were interviewed were manually matched to police records in the Danish accident database (vejman.dk) via a search matching the accident descriptions they provided in their interview (e.g., accident date and location, their age, gender, and transport mode), as the recorded CPR number is not searchable or viewable in the database.

2.4 INTERVIEWS

Thirty-five respondents were selected for invitation to a follow-up interview. The selection criteria were:

- The respondents had to have given permission for combining their self-reports with other databases;
- The respondents had to have stated that they had been in contact with the police due to their accident;
- The respondents had to live within postal codes 9000-9270 (within a 15 minute drive from Aalborg University to minimize transport)

A postal invitation to participate in an interview was sent; respondents could respond either by telephone or email. Nine respondents (26%) committed to interviews, which were held at either the University or the respondent's home, as the respondents preferred. One voice recording was lost due to technical issues, leaving eight conducted interviews as a basis for analysis.

The semi-structured interviews had two purposes. Firstly, the interviews were another way of gaining knowledge of the accident. We prompted the respondent to talk about the aspects also handled in the self-report questionnaire, but in a semi-structured form. The information on accident-specific details was then investigated to gauge the congruence between the self-report questionnaires and the interviews. As part of the interview, the respondents were also asked to pinpoint the accident location on a map, printed by the interviewer beforehand, and to draw an illustration of their accident on a blank piece of paper. Secondly, the interview had some methodological themes: for instance, we asked respondents about their reasons for participating in the study and their experience with using the online map to locate their accident.

2.5 STATISTICS

Cohen's Kappa, κ (Cohen, 1960), was applied as a measure of agreement on how respondents answered in their self-report questionnaire compared with the information they provided to ER personnel or to the police. Cohen's Kappa describes how large the agreement is between data obtained through the two different methods (or between two judges) and takes into account that some of the agreement between datasets might be due to chance. A similar approach of applying Cohen's Kappa to ascertain agreement between self-reports and police records was applied in the study by Møller (2015).

The Kappa value was calculated (Cohen 1960):

$$\kappa = \frac{f_o - f_c}{N - f_c}$$

where
f_o is the number of units on which the raters agree
f_c is the number of units for which agreement is expected by chance
N is the total number of units rated
Our qualitative interpretation of the kappa values followed Landis and Koch (1997): 0–0.20 = slight agreement, 0.21–0.40 = fair agreement, 0.41–0.60 = moderate agreement, 0.61–0.80 = substantial agreement and 0.81–1.00 = almost perfect agreement.

To calculate Cohen's Kappa, data must be categorical. Thus, it was not possible to calculate kappa values for accident dates. Instead, the intraclass correlation (ICC) for average measures was calculated in SPSS with a two-way mixed model for absolute agreement, chosen by the guidelines provided by Koo and Li (2016).

Data were analyzed with descriptive statistics. The significance level was $\alpha = 0.05$, and 95% confidence intervals (CIs) are shown where percentages were calculated. Comparisons of the age composition of respondents and non-respondents as well as comparison with the overall composition of ER-recorded traffic accident victims in Denmark was carried out by a χ^2-test, and comparisons of gender compositions were made by a z-test.

3. RESULTS

In ERs in Region Nord 3,809 patients had contact due to a traffic accident. Of these, 1,132 patients were asked to participate in the survey, and 930 patients agreed to receive a questionnaire. Figure 2 shows their distribution throughout the survey period, and Table 1 shows their ages and gender.

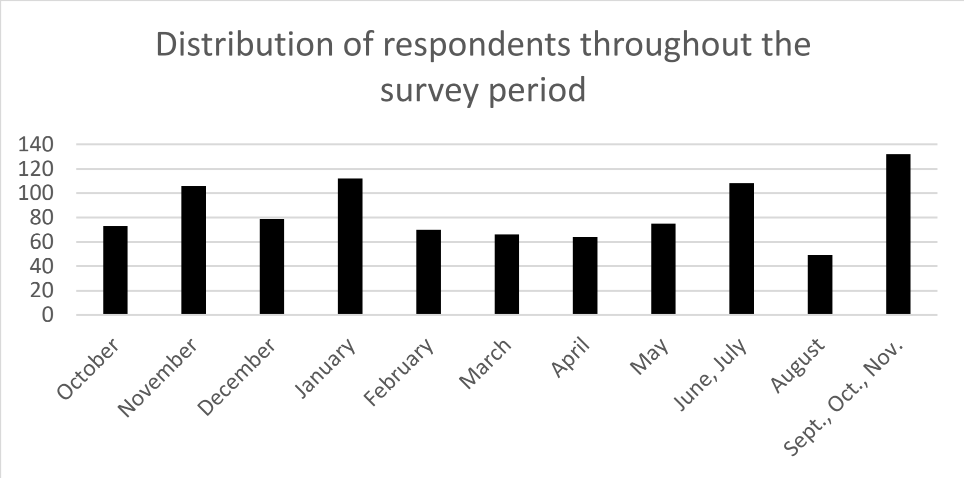


Figure 2: Number of patients consenting to receive a self-report questionnaire distributed over the research period. The month is when respondents first had contact with the ER. Numbers from the last three months are compiled as these respondents all had the questionnaire distributed simultaneously.

Not all the respondents completed the questionnaire: 474 respondents began answering the questionnaire, but 35 (7.38%) did not complete the questionnaire, leaving 439 completed questionnaires. The overall response rate on the distributed questionnaires is thus 47%. Of these, 27 (6.15%) stated that they did not give permission for the researchers to combine their questionnaires with data from other registers, and one questionnaire suffered from data corruption and could not be matched. The final sample thus consisted of 411 self-reports of accidents.

Table 15: Age of people in contact with the emergency rooms in Region Nord who consented to receiving a questionnaire

	Received a self-report questionnaire					
	Men		**Women**		**Total**	
	n	%	*n*	%	*n*	%
Below 18	36	3.87	30	3.22	66	7.09
18–24	87	9.35	84	9.03	171	18.38
25–34	81	8.71	93	10.00	174	18.70
35–44	47	5.05	64	6.88	111	11.93
45–54	67	7.20	68	7.31	135	14.51
55–64	51	5.48	76	8.17	127	13.65
65 and above	52	5.59	94	10.10	146	15.69
Total	421	45.27	509	54.70	930	100.00

When testing the responders' demography (Table 2) against the non-responders (Table A2, Appendix A) we found significant differences for both men (χ^2 = 37.4506, df = 6, p < 0.000) and women (χ^2= 53,3714, df = 6, p < 0.000). This is mainly due to the fact that both men and women aged 25–34 are prone to non-response whereas women aged 45–54 frequently chose to respond.

There were significantly more women responders than men (58.50% compared with 41.26%, z = -5.10148151, p < 0.001), cf. Table 2. When examining the gender distribution among all those involved in a traffic accident recorded by the ERs in Denmark (Table A1, Appendix A), we found that men comprised 54.60% of the national records. This is significantly higher than the 41.26% among our responders (z= -5.489942209, p < 0.001). Women are thus overrepresented in the current study.

On age, when we compared the age distribution of our responders against the distribution of traffic accident victims recorded at ERs on a national level (cf. Table A1, Appendix A), we found that for both men and women the sample was significantly different (women's χ^2 = 25.6863955, df = 3, p < 0.001; men's χ^2= 25.42261991 df = 3, p < 0.001), as we had a very small proportion of respondents below the age of 18.

Table 2: Demographic information of respondents

	Responded to self-report questionnaire					
	Men		**Women**		**Total**	
	n	*%*	*n*	*%*	*n*	*%*
Below 18	8	1.95	15	3.65	23	5.60
18-24	37	9.00	45	10.95	82	19.95
25-34*	19	4.62	29	7.06	49	11.92
35-44	20	4.87	27	6.57	47	11.44
45-54	29	7.06	46	11.19	75	18.25
55-64	29	7.06	36	8.76	65	15.82
65 and above	26	6.33	42	10.22	68	16.55
Will not tell	1	0.24	1	0.24	2	0.49
Total	169	41.12	241	58.64	411	100.00

** One respondent did not state their gender*

Most respondents (47%) were using a bicycle at the time of the accident, although car drivers (28%) and pedestrians (21%) were also represented well in the sample. Pedestrian falls were also registered in this research project, and 81 of the 87 pedestrian respondents experienced single accidents. For most purposes in this paper we included them in the analysis, even though they are not technically classified as traffic accidents (Vejdirektoratet 2017). Most participants experienced single accidents (52%). However, if we remove the pedestrian falls from the sample, it is more common to have experienced an accident with a counterpart than a single accident. In most cases, the counterpart was a car driver, and 58% of respondents with multiparty accidents found the accident to be primarily the other party's fault. Seventy-five percent had one or more days of absence due to their accident, with 9.79 as the average number of days and 55 as the maximum.

3.1 CONGRUENCY OF INFORMATION IN HOSPITAL RECORDS AND SELF-REPORTS

All 411 respondents were registered in LPR as a traffic accident, and thus all LPR records contained fields with possible information on the transport mode and counterpart. However, some were labeled "unspecified", and some were not answered in the questionnaire, leaving 399 respondents for whom we can ascertain if the record of their own means of transport in LPR was congruent with the means of transport in their self-report. The congruency can be seen in Table 3**.**

Of the 399 respondents (3.76%) presented different information. Seven of these inconsistencies stemmed from hospital personnel registering more as on foot than were self-reported, as seen in Table A4 in Appendix A. This gives a Cohen's Kappa value of 0.94, indicating almost perfect agreement.

Table 3: Congruency between LPR and self-reports

	Total used to classify	**Agreement**			
		n	*%*	*Cohen's Kappa*	*95% CI Kappa*
Own mode of transport	399	384	96.24	0.94	0.92–0.97
Counterpart mode of transport	366	320	87.43	0.77	0.69–0.82

Forty-six out of 366 (12.57%) did not present congruent information on the counterpart; 23 of these ertr due to the fact that the hospital recorded more respondents as having no counterpart than were self-reported, as seen in Table A5 in Appendix A. This equals a Cohen's Kappa of 0.77, indicating substantial agreement.

3.2 CONGRUENCY OF INFORMATION IN POLICE RECORDS AND SELF-REPORTS

One hundred and nine respondents stated that they had been in contact with the police due to their accident. Of these, 68 (62.39%) could not be matched to a police report. Eight respondents who stated that they had not been in contact with the police could be matched to a police record (2.65% of respondents stating no police contact). This gives a total of 49 matches between questionnaires and police reports (11.92%).

It is possible that some of the 68 respondents who could not be matched to a police record were instead recorded in police records as material damage-only accidents. As

CPR number or other personal information is not recorded in material damage only accidents, this renders matching impossible. The eight interviews could illustrate these discrepancies between claimed contact with the police and the inability to match respondents to a police record. Three of the interviewees were matched to a police record with their CPR number. Four of the eight interviewees could not be matched with their CPR number, as they were recorded as material damage-only accidents, even though all their injuries were quite extensive. One interviewee fractured her hand and injured ligaments in her hip, keeping her on strict bedrest for one month and crutches for half a year. Another bruised her ribs and afterwards received more than 20 physiotherapy treatments. One fractured his chest bone (and the passenger in the car fractured two fingers), and another fractured her hip. The inconsistency between the police-recorded severity and self-reported severity goes both ways; one of the police reports stated that the interviewee was seriously injured, yet in fact she stated that she received only small bruises from the accident and was back at school later the same day. One interviewee who could not be matched to any records stated:*" I called the police the day after, but they said that there wasn't really anything to do when nothing kind of criminal had happened."* It is thus also a possibility that some of the 68 respondents with no matching police record had in fact been in contact with the police, but the police had simply not written a report, despite the accident being of a severity that merited a report.

Police registered 39 respondents as having received medical attention at the hospital and ten were registered as having had no contact with the ER. Yet, as all respondents were sampled through the ER, we know that all had in fact been in contact with the ER. Table 4 compares the police-reported severity with respondents' statements about having been absent from work due to their accident. We found that there were both some who were classified as unharmed, who stated they took one or more days of absence from work due to their injury, as well as some among those who were classified as severely injured, who stated that they did not need any days of absence due to their injury.

Table 4: Police-reported level of severity compared with self-reported absence from work due to the accident. Data from Statistics Denmark cannot be shown when frequency is less than five.

	Unharmed	Slightly injured	Severely injured
Absent	≤ 5	26	6
Not absent	≤ 5	≤ 5	≤ 5
Do not know	0	≤ 5	0

3.3 ACCIDENT DATE

Five entries in the self-report questionnaires could not be used for comparison; a software error meant that respondents who did not choose a year for their accident were wrongfully coded. This left 44 respondents with a comparable accident date. Forty were in full compliance with police records. The other four self-reported accidents happened on dates prior to the date in the police report—the discrepancies can be seen in Figure 3.

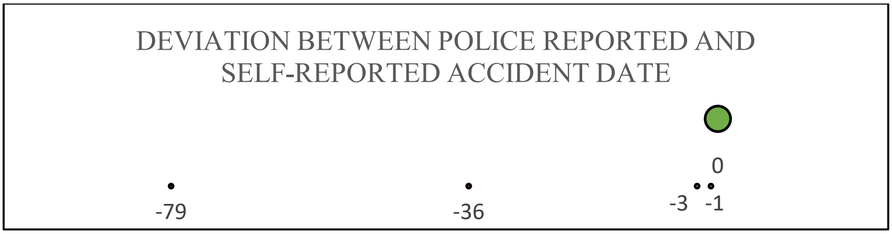


Figure 3: Number of days the self-reported accident date deviated from the police-recorded accident date. The area of the bubbles represents the number of observations and the offset on the X-axis corresponds to the number of days. The offset on the Y-axis is trivial and made to improve viewability of data.

Table 5 shows the kappa values for police reports and self-reports; the data are provided in Table A6-A10 in Appendix A.

Table 5: Congruency of police reports and self-reports

	Total used to classify	**Agreement**			
		n	*%*	*Cohen's Kappa*	*95% CI Kappa*
Counterparts' mode of transport	37	37	100%	1	-**
Own mode of transport	49	49	95.92%	0.93	0.83–1.03
Date of accident	44	40	90.91%	0.88*	0.79–0.94
Accident situation	25	21	84.00%	0.82	0.66–0.98
Daylight conditions	39	34	87.18%	0.74	0.55–0.93
Road surface conditions	36	32	88.89%	0.68	0.41–0.96
Weather conditions	41	34	82.93%	0.51	0.24–0.79

* ICC value, not Kappa, due to continuous data.

** Cannot be calculated due to perfect agreement

Transport modes showed the highest kappa values. There was perfect agreement between police and respondents' classification of the counterpart, resulting in a kappa value of 1. Agreement between the police and respondents regarding their own mode of transport was also almost perfect (kappa = 0.93); the only discrepancies were two respondents labeled as car drivers by police that were labeled a cyclist and a van driver, respectively, by respondents.

The ICC for accident dates was also high (0.88). There was almost perfect agreement on the classification of accident situations, as indicated by the kappa value of 0.82. But the 95% confidence interval is quite large, partly because only 25 of the respondents chose an accident situation among the illustrations in the questionnaire, thus providing a smaller sample size for calculation.

In the conducted interviews we also found inconsistency between the accident situations recorded by police and what was described by respondents. All the interviewees managed to make a drawing on a piece of paper that made it possible to allocate a specific accident situation. The results were as follows:

- Two interviewees were in full concordance with their chosen accident situation in the questionnaire, the interview, and the police record.
- Two interviewees could not choose an accident situation in their questionnaire, but their interview and the police choice of accident situation were consistent.
- Two interviewees showed consistent description in both their questionnaire and in their interview as well as in the drawing they made, but this was not in concordance with what was recorded by the police. In one of these police reports, it was also clear that the police's choice of accident situation did not match the written description by the police of how the accident happened.
- One interviewee could not choose the accident situation in the questionnaire but was consistent in her oral description and her drawing; yet this did not concur with the police choice of accident situation.

The interviews thus clearly showed that the accident situation recorded by the police should not be considered a gold standard, as only four interviews showed consistency with the police's choice of accident situation, and three did not.

Daylight, road surface and weather conditions yielded the lowest kappa values. The deviation on classification of daylight conditions primarily stemmed from inconsistency in dusk and dark ratings between respondents and police (see Table A6, Appendix A). For road surface conditions, there was no pattern in the misclassifications (see Table A7, Appendix A), leading to the relatively low kappa value of 0.72, indicating substantial agreement. Weather conditions had the lowest kappa value of 0.51, indicating moderate agreement; as seen in Table A8 in Appendix A this mainly stemmed from respondents choosing to describe weather as either dry or wet, whereas police personnel applied other descriptions.

The inconsistency between environmental factors in the police records and the questionnaires was not present in the interviews. Notably, many of the respondents seemed to not directly remember the environmental factors, but deduced them based on other facts when asked, e.g.:

"It cannot have been raining, because then I would have been wearing my raincoat."

"*Fine weather. No, I cannot remember that. It is quite possible that it was raining; it has been raining almost the entire spring."*

One respondent clearly stated about the environmental factors:
"*It is not something that have had a very strong presence* [in my thoughts]."

3.4 INTERNAL RELIABILITY OF QUESTIONNAIRES

3.4.1 INABILITY TO FIND A SUITABLE MULTIPLE-CHOICE OPTION

If many respondents were not able to choose a suitable answer among the provided multiple choice options, it might indicate problems with the question, either in terms of poor understanding, an option missing, or asking something that respondents were not able to answer. In Table 6, an overview is provided of the percentage of respondents unable to find an applicable option among the provided answers in multiple-choice format, and thus answering *do not know*, *none of the above* or *other*.

Respondents had the most difficulty with choice of movements before the accidents as well as choice of specific accident situations. The reason for pedestrian single accidents was also among the questions that had a high proportion of *none of the above*. Whether or not the streetlights were on or off (if dusk or dark) was fairly often forgotten (14%), and 12% of respondents provided with the question would not say who they found to be at fault in the accident.

Table 6: frequency of *do not know/none of the above/other* in multiple choice answers

Item theme	Frequency of unsuitable answer options	out of	%
Accident situation (multiparty)	28	123	22.76
Movements before accident	26	143	18.18
Accident situation (single)	19	109	17.43
Streetlight on/off	16	115	13.91
Responsibility for accident	24	195	12.31
Reason for pedestrian accident	9	81	11.11
Trip purpose	34	411	8.27
Locate accident on map	31	403	7.69
Weather conditions	23	371	6.20
Road surface conditions	22	370	5.95
Unfit for work	13	411	3.16
Type of facility (plaza, road, bicycle lane, etc.)	9	411	2.19
Counterpart's mode of transport	3	157	1.91
Geometry	6	383	1.57
Own postal code	5	411	1.22
Daylight conditions	4	371	1.08
Own age	3	411	0.73
Own mode of transport	1	408	0.25
Gender	1	411	0.24

3.4.2 LOCATION OF THE ACCIDENT

Seventy-five of the 411 respondents used the items with a descriptive field to write a description of where the accident happened, although this was not mandatory. This descriptive location was then compared to the Global Positioning System (GPS) coordinates of the accident location that respondents had pinpointed on the map.

A match of coordinates and text description was accepted if the description matched a location within 100 meters of the GPS coordinates. Descriptions varied in detail, e.g.,[1]:

"I am driving westwards at Sohngaardsholmsvej – right before Vonsyldsgade…"

"…after turning from Hadsundvej onto Filstedvej…"

"On the bicycle path in front of Fitnessworld Vesterbro."

[1] Street and shop names have been altered to preserve anonymity.

The results of the comparison between descriptive fields and coordinates is in Table 7; 87% showed congruency between the GPS coordinates and description of the accident location.

Table 7: Matching of information from coordinates and descriptive fields

	Congruency of respondents' pinpoint on the map and their written description		
	n	*%*	*95% CI*
Location as both GPS and description	71		
Matching information	62	87.32	79.58–95.06
GPS-coordinates not a match	9	12.68	4.94–20.42

The interviews showed that there was consistency (within a 15-meter radius) between the location pinpointed by the respondents and by the police in six of the seven accidents. The accident with an incongruent location had locations placed approximately 100 meters apart. When interviewees were asked if they found it difficult to locate their accident on a map, many instantly stated that their accident took place in an area they were familiar with as a reason for the ease of location. For instance:

"Well, I bicycle there every day!"

"I did not find it difficult in that sense. I did it as well as I could. I am also raised up here…well I know that area."

"I actually found it quite easy. Also because I am quite well-versed in that neighborhood; I used to live there."

3.4.3 ACCIDENT SITUATION

Table A3 and A11 in Appendix A show that 80% of respondents presented with the accident illustrations (185 out of 232) were able to find an applicable accident illustration among the provided sketches. When the z-test was conducted, there was no significant difference in the proportion of respondents able to apply an illustration, depending on whether there was a counterpart or not involved (83% compared with 77%, $z = 1.01$, $p = 0.31$).

Some respondents provided descriptions of their accident and their maneuvers immediately before the accident in the descriptive field. Examples of descriptions that could be used to find a fitting illustration could be[2]:

[2] Directions and transport modes have been altered to keep anonymity.

“She drove into me from the left while I was driving straight ahead.”

“Oncoming moped drove on the left side of a friend with whom he was talking and thus wasn’t paying attention to where he was driving. He ended up colliding with me head on.”

”The oncoming cyclist took a left turn in front of me.”

Accident classification was possible for the 117 respondents who both found an applicable accident illustration and provided enough information in the descriptive field. Table 8 shows the congruency of illustrations chosen by respondents to their accident description provided in the text fields. Almost 85% of the accident illustrations applied by respondents were congruent with the description provided in text.

Table 8: Congruency of applied illustrations with textual description

	Congruency of respondents' illustration choice and their written description		
	n	*%*	*95% CI*
Accident situation in both descriptive and multiple choice	117	100	-
Matching situation	99	84.62	81.28–87.95
Situation not a match	18	15.38	12.05 – 18.72

4. DISCUSSION & CONCLUSION

The sampling method of asking accident-involved patients in the ER to participate in a self-report regime proved highly efficient, with a high response rate (47%) and low dropout (7%). We found that respondents’ ability to locate their accident was quite trustworthy; almost all respondents (92%) managed to pinpoint their accident, and 87% showed consistency in their pinpoint and written description. The conducted interviews revealed that a partial explanation might be that respondents found it easier to pinpoint locations along a familiar route or area.

Eighty-five percent showed consistency between the chosen accident illustration and the text description of their accident situation. Comparing accident situations with police records, we found almost perfect agreement (kappa = 0.82), as 84% of the records agreed. Our interviews revealed that some discrepancies between police records and self-reports could be rooted in erroneous recordings by the police. When examining the questions to which respondents had issues finding a suitable answer, it is clear that the accident situation and the description of their own maneuvers as well as their counterparts were items which many failed to answer. Approximately 18% of respondents could not choose among the multiple-choice options describing their maneuvers before the accident happened, and 20% could not choose among the provided illustrations of accident situations. The interviewees who could not find a suitable answer in their completion of the questionnaire all managed to draw illustrations on paper and give detailed oral descriptions of their accident. This indicates that inability to answer these items is more related to the survey methodology than the respondents' lack of knowledge or recall.

Begg et al. (1999) found the largest disagreement (8–24%) in factors describing the time of accident (i.e., day of the week, hour of the day and year of the crash). This is also in trend with our findings, in which 9% of the police and self-reports showed divergent accident dates. The use of a three-year recall period by Begg et al. might explain why some of their time components had lower agreement than we found. The four divergent accident dates in our study were all to be backwards telescoped; the respondents placed them further back in time than the police recorded. This is congruent with the findings of Janssen, Chessa, and Murre (2006), who observed that recent events are more likely to be backwards telescoped than remote events. It is of course also possible that the discrepancies can be attributed to error in either matching or self-reporting. But one might keep in mind a practical aspect regarding the date of the accident: how often is an exact date necessary when doing accident prevention work? Two of the four displacements were dated within five days of the police report and the other two were dated within the three months. It could very well be argued that this displacement in time is trivial when conducting e.g., *black spot* analysis.

Our findings that weather conditions had the lowest agreement (83%) among the environmental factors is similar to the findings of Versteegh (2004), in which 81% of the records agreed on weather conditions. He found 99% agreement regarding lighting conditions and 95% on road surface conditions, higher than our findings of 87% and 89%, respectively. Our interviews revealed that respondents, at least in hindsight, seemed to deduce the environmental factors based on other facts, yet we do not know if this reasoning was also present at the time of questionnaire completion, which was commonly 1–3 months after their accident. A few respondents (6%) stated that they could not answer the question on how the weather conditions were at the time of the accident. A reason for the discrepancies could thus very well be the subjective nature of some weather situations. For instance, how much should it rain before one would classify it as raining?

That the highest levels of agreement came from items regarding transport mode is in concordance with the study by Begg et al. (1999), who found 100% agreement between self-reported and police-reported vehicle type. Somewhat surprisingly, we also found very high agreement on transport mode when comparing self-reports with hospital records. Descriptions of the respondents' own mode of transport in LPR was congruent in 96% of the records when compared with their own statements in the questionnaire, an almost perfect agreement (kappa = 0.94). We found the same high agreement regarding the counterpart mode of transport: 87% of hospital records and self-reports had congruent information. This equals a kappa value of 0.77, indicating substantial agreement. This is higher than the previous study by Møller et al. (2015), where a kappa value of 0.63 was found on counterparts' mode of transport. An obvious explanation for this is that ER personnel, who are responsible for registration of traffic accidents, transport mode, counterpart, etc. in the LPR, are the same personnel who were responsible for asking the patients if they wanted to participate in the survey. Thus, the ER personnel were highly aware of the added focus on registration practice of traffic accidents; this could have led to more thorough questioning and registration of the patients at the ER. This suspected change in registration praxis can also have been due to the fact that all 411 accidents were recorded as traffic accidents in the hospital records. This is in contrast with a previous study, where 21 out of 74 bicycle accidents were found not to be recorded as traffic accidents in LPR (Møller al. 2015). Sampling in that study was done prior to the accidents, and the cyclists were matched to their hospital records after completing a self-report questionnaire to explore information coherence. In the current study, ER personnel were instructed to ask patients about survey participation even if they had doubts whether the accident could be classified as a traffic accident or not. This could have prompted personnel to adopt a more lenient registration practice in the LPR as well. The results on congruency between self-reports and LPR data must thus be considered a best-case scenario; without the extra focus by an ongoing research project it is quite possible that concordance between records might be lower.

Overall, we see potential for improvements in the questionnaire design for movements and accidents illustrations, which might help to minimize answers of *do not know*. Yet we can also conclude that the respondents who did manage to choose an accident illustration (194 respondents) were often correct in their choice. We have applied a new sampling method in the ERs, which seems quite effective in terms of response rates, yet with a lack of younger males and females among the responders and a larger proportion of women aged 45–54 as participants. We conclude that the applied method for self-reporting seems a viable solution for obtaining accident information, as both internal consistency and agreement with other data sources was quite high.

ACKNOWLEDGEMENTS

The study was financially supported by the Danish foundation TrygFonden. The authors would like to thank TrygFonden for their funding.

REFERENCES

Ahmed, A., Sadullah, A., & Yahya, A. (2017). Errors in accident data, its types, causes and methods of rectification - analysis of the literature. *Accident Analysis & Prevention*. Advance online publication. https://doi.org/10.1016/j.aap.2017.07.018

Anstey, K. J., Wood, J., Caldwell, H., Kerr, G., & Lord, S. R. (2009) Comparison of self-reported crashes, state crash records and an on-road driving assessment in a population-based sample of drivers aged 69-95 years. *Traffic Injury Prevention*, *10*(1), 84–90. https://doi.org/10.1080/15389580802486399

Begg, D. J., Langley, J. D., & Williams, S. M (1999). Validity of self-reported crashes and injuries in a longitudinal study of young adults. *Injury Prevention*, *5*, 142–144. https://doi.org/10.1136/ip.5.2.142

Boufous, S., Ivers, R., Senserrick, T., Stevenson, M., Norton, R., & Williamson, A. (2010). Accuracy of self-report of on-road crashes and traffic offences in a cohort of young drivers: The DRIVE study. *Injury Prevention,* 2010; *16*, 275–277. http://dx.doi.org/10.1136/ip.2009.024877

Cohen, J. (1960). A coefficient of agreement for nominal scales. *Educational and Psychological Measurement*, *20*(1), 37–46. https://doi.org/10.1177/001316446002000104

Chung, Y., Chang, I. (2015). How accurate is accident data in road safety research? An application of vehicle black box data regarding pedestrian-to-taxi accidents in Korea. *Accident Analysis & Prevention, 84*, 1–8. https://doi.org/10.1016/j.aap.2015.08.001

Derriks, H.M., & Mak, P.M. (2007). *IRTAD special report - underreporting of road traffic casualties.* The Netherlands: Ministry of Transport, Public Works and Water Management.
Retrieved from https://www.who.int/roadsafety/publications/irtad_underreporting.pdf

Elvik, R., & Mysen, A.B. (1999). Incomplete accident reporting - meta-analysis of studies made in 13 countries. *Transportation Research Record*, *1665*, 133–140.

Finestone, H., Guo, M., O'Hara, P., Greene-Finestone, L., Marshall, S., Hunt, L., . . . Biggs, J. (2011) Department of Transportation vs Self-reported Data on Motor Vehicle Collisions and Driving Convictions for Stroke Survivors: Do They Agree? *Traffic Injury Prevention*, 12:4, 327–332, DOI: 10.1080/15389588.2011.581716

Janstrup, K.H., Kaplan, S., Hels, T., Lauritsen, J., & Prato, C.G., (2016). Understanding traffic crash under-reporting: Linking police and medical records to individual and crash characteristics. *Traffic Injury Prevention*, *17*(6), 580–584. https://doi.org/10.1080/15389588.2015.1128533

Janssen, S., Chessa, A., Murre, J. (2006). Memory for time: How people date events. *Memory & Cognition*, 2006, *34*(1), 138–147. https://doi.org/10.3758/BF03193393

Koo, T. K., & Li, M. Y. (2016). A guideline of selecting and reporting intraclass correlation coefficients for reliability research. *Journal of Chiropractic Medicine*, *15*(2), 155–163. https://doi.org/10.1016/j.jcm.2016.02.012

Landis, J., & Koch, G. (1977). The measurement of observer agreement for categorical data. *Biometrics*, *33*(1), 159–174. https://doi.org/10.2307/2529310

Larsen, M.K., Jensen A.K.S., & Andersen, C.S. (2017). *Potentiale for flere uheldsdata på baggrund af fejlklassificeringer af ekstrauheld*. Proceedings from the annual transport conference at Aalborg University. Aalborg, Denmark. Retrieved from http://www.trafikdage.dk/td/papers/papers17/474_AnneKriegbaum.pdf

LifeinDenmark.com (2019). *Living in Denmark – Digital post.* Webpage visited May 1st 2019 at https://lifeindenmark.borger.dk/living-in-denmark/digital-post

Madsen, T. K. O. (2018). *Video analysis and mapping of vulnerable road users' safety*. (Doctoral dissertation). Aalborg Universitetsforlag. Ph.d.-serien for Det Ingeniør- og Naturvidenskabelige Fakultet, Aalborg Universitet. Retrieved from https://doi.org/10.5278/vbn.phd.eng.00057

Madsen, T. K. O., Várhelyi, A., Polders, E., Reumers, S., Hosta, P., Bibiloni, D. J., ... Lahrmann, H. S. (2018). *Assessment of safety of VRUs based on self-reporting of accidents and near-accidents.* Brussels: Publications Office of the European Union.

Retrieved from https://www.indev-project.eu/InDeV/EN/Documents/pdf/3-2.pdf?__blob=publicationFile&v=2

Marottoli, R., Cooney, L., & Tinetti, M. (1997). Self-report versus state records for identifying crashes among older drivers. *Journals of Gerontology: Series A*, *52A*(3), M184–M187. https://doi.org/10.1093/gerona/52A.3.M184

McGwin, G., Owsley, C., & Ball, K. (1998). Identifying crash involvement among older drivers: Agreement between self-report and state records. *Accident Analysis & Prevention*, *30*(6), 781–791.

Møller, M., Clemmensen, M. B., & Janstrup, K. H. (2017). *Bedre trafikuheldsdata: Danske erfaringer med brug af forskellige typer trafikuheldsdata.* DTU Management. Retrieved from http://orbit.dtu.dk/files/145213300/Rapport_Bedre_trafikuheldsdata.pdf

Møller, K. M., Madsen, T.K.O., Olesen, A.V., & Lahrmann, H. (2015) *A case study on agreement between self-reported bicycle accidents and hospital and police records.* Selected proceedings from the annual transport conference at Aalborg University 2015. Retrieved from http://www.trafikdage.dk/td/papers/papers15/324_KatrineMeltofte_fagfaellebedoemt.pdf

Porter, M. M. (2018). An examination of the concordance between self-reported collisions, driver records, and insurance claims in older drivers, *Journal of Safety Research,* 67, 211–215. https://doi.org/10.1016/j.jsr.2018.07.007

Statistikbanken (2019). *Tabel Moerke*. Retrieved from http://www.statistikbanken.dk/statbank5a/default.asp?w=1455

Singletary, B.A., Do, A.N., Donnelly, J.P., Huisingh, C., Mefford, M.T., Modi, R., . . . McGwin, G. (2017). Self-reported vs state-recorded motor vehicle collisions among older community dwelling individuals. *Accident Analysis & Prevention*, 101, 22–27. https://doi.org/10.1016/j.aap.2017.01.021

Tin Tin, S., Woodward, A., & Ameratunga, S. (2013). Completeness and accuracy of crash outcome data in a cohort of cyclists: A validation study. *BMC Public Health, 13*, 420. http://dx.doi.org/10.1186/1471-2458-13-420

Vejdirektoratet (2017). *Indberetning af færdselsuheld. Rapport 580*. Retrieved from http://www.vejdirektoratet.dk/DA/viden_og_data/publikationer/Lists/Publikationer/Attachments/957/Indberetrning%20af%20færdselsuheld_web.pdf

Versteegh, S. (2004). The accuracy of driver accounts of vehicle accidents. *2004 road safety research, policing and education conference. Session: Crash and injury analysis.* Perth, Australia. Retrieved from https://core.ac.uk/download/pdf/12769779.pdf

af Wåhlberg, A. E. (2011). Memory effects in self-reports of road traffic crashes. In *International conference on driver behaviour and training, 5th, Paris, France* (pp. 283–288). Farnham: Ashgate.

APPENDIX A

Table A1: Age and gender of traffic accident victims registered in emergency rooms in Denmark in 2016 and 2017 (Statistikbanken 2019)

	Men			Women			Total		
	n	*%*	*95% CI*	*n*	*%*	*95% CI*	*n*	*%*	*95% CI*
Below 18	6469	19.04	18.83–19.25	4359	15.43	15.22–15.65	10828	17.40	17.25–17.55
18–24	6150	18.10	17.89–18.31	5336	18.89	18.66–19.12	11486	18.46	18.30–18.61
25–64	18292	53.84	53.57–54.11	15581	55.15	54.86–55.45	33873	54.43	54.23–54.63
Above 65	3066	9.02	8.87–9.18	2974	10.53	10.34–10.71	6040	9.71	9.59–9.83
Total	33977			28250	100		62227		

Table A16: Age and gender of non–responders among the invited patients

	Men		Women		Total	
	n	*%*	*n*	*%*	*n*	*%*
Below 18	28	5.39	15	2.89	66	12.72
18–24	50	9.63	39	7.51	171	32.95
25–34	62	11.95	64	12.33	174	33.53
35–44	27	5.20	37	7.13	111	21.39
45–54	38	7.32	22	4.24	135	26.01
55–64	22	4.24	40	7.71	127	24.47
65 and above	26	5.01	52	10.02	146	28.13
Total	252	48.55	267	51.45	519	100.00

Table A3: Respondents' capability of choosing an accident illustration fitting their accident

		Found applicable illustration	No applicable illustration	Total
Single accidents	*n*	90	19	109
	%	82.57	17.43	
	95% CI	75.45–89.69	10.31–24.55	
Multiparty accidents	*n*	95	28	123
	%	77.24	22.76	
	95% CI	69.83–84.65	15.35–30.17	
Total	*n*	185	47	232
	%	79.74	20.26	
	95% CI	74.7–84.91	15.09 –25.43	

Table A4: Comparison of self-reported means of transport and the LPR-recorded means of transport

	Self-recorded means of transport							
LPR-recorded means of transport	Car	Bicycle	On foot	Moped	Motor-cycle	Truck	Van	Total
Car*	114	0	0	0	0	0	≤5	117
Bicycle*	≤5	169	≤5	0	0	0	≤5	172
On foot*	≤5	≤5	86	0	0	≤5	0	94
Moped*	0	0	0	9	≤5	0	0	10
Motorcycle *	0	0	0	0	≤5	0	0	≤5
Truck or bus*	0	0	0	0	0	≤5	0	≤5
Van	0	0	0	0	0	0	0	0
Total*	117	174	87	9	6	≤5	≤5	399

*Data from Statistics Denmark cannot be shown when frequency is less than five.

Table A5: Comparison of self-reported counterpart and his/her means of transport and the LPR registration

LPR-recorded counterpart	Self-recorded counterpart: No counterpart	Car	Bi-cycle	On foot	Moped	Motor-cycle	Truck	Van	Tot.
No counterpart*	200	9	6	≤5	≤5	0	≤5	0	223
Car*	9	92	0	≤5	0	0	0	8	110
Bicycle*	≤5	≤5	16	0	0	0	0	0	18
On foot*	≤5	0	0	≤5	0	0	0	0	6
Moped*	0	≤5	0	0	≤5	0	0	0	≤5
Motorcycle*	0	0	0	0	0	0	0	≤5	≤5
Truck or bus*	≤5	≤5	0	0	0	0	≤5	≤5	≤5
Van	0	0	0	0	0	0	0	0	0
Total*	212	105	22	10	≤5	0	≤5	10	366

*Data from Statistics Denmark cannot be shown when frequency is less than five.

Table A6: Lighting conditions as stated by police and respondents

Police-recorded	Self-recorded lighting conditions: Daylight	Dusk	Dark	Empty or do not recall
Daylight*	26	0	0	≤5
Dusk*	≤5	≤5	≤5	0
Dark*	0	≤5	≤5	≤5
No information*	≤5	0	0	0

* Data from Statistics Denmark cannot be shown when frequency is less than five.

Table A7: Road surface as stated by police and respondents

Police recorded	Self-recorded road surface conditions: Dry	Wet	Slippery due to snow or ice	No recall/none of the above	Empty
Dry*	27	≤5	0	≤5	≤5
Wet*	≤5	≤5	0	≤5	≤5
Slippery snow/ice*	≤5	≤5	≤5	0	≤5
Slippery due to other*	≤5	0	0	0	0
No information*	≤5	≤5	0	0	0

* Data from Statistics Denmark cannot be shown when frequency is less than five.

Table A8: Weather conditions as stated by police and respondents

Police-recorded	Self-recorded weather conditions: Dry	Rain	other	Empty
Dry*	30	0	≤5	8
Rain*	≤5	≤5	0	0
Other*	≤5	≤5	≤5	0

* Data from Statistics Denmark cannot be shown when frequency is less than five.

Table A9: Own mode of transport as stated by police and respondents

Police-recorded	Self-recorded mode of transport: Car	Motorcycle	Bicycle	Pedestrian	Van
Car	26	0	≤5	0	≤5
Motorcycle	0	≤5	0	0	0
Bicycle	0	0	18	0	0
Pedestrian	0	0	0	≤5	0
Van	0	0	0	0	0

*Data from Statistics Denmark cannot be shown when frequency is less than five.

Table A10: Counterpart's mode as stated by police and respondents.

Police recorded	Self-recorded counterpart's mode of transport: Car	van	Moped	Bicycle	More than one and I cannot tell	Empty
Car*	33	0	0	0	≤5	6
Van*	0	≤5	0	0	0	≤5
Moped*	0	0	≤5	0	0	0
Bicycle*	0	0	0	≤5	0	0
Road equipment*	0	0	0	0	0	≤5

*Data from Statistics Denmark cannot be shown when frequency is less than five.

Table A11: Distribution of chosen accident illustrations in self-reports. The situation numbers (sit. no.) correspond to the numbers applied by Danish police in official accident registration.

Multi-party accidents		No applicable option	**Total 28**	*n*	28								
				Sit.no	999								
	Misc.	8: Pedestrian involved	**Total 12**	*n*	2	2	2	1	2	1	2		
				Sit.no	811	812	820	841	871	878	880		
		7: Parked vehicle	**Total 4**	*n*	2	1	1						
				Sit.no	710	720	740						
	Inter-sections	6: Different roads and turning	**Total 9**	*n*	4	1	1	2	1				
				Sit.no	610	643	650	660	670				
		5: Different road, no turning	**Total 4**	*n*	1	3							
				Sit.no	510	520							
		4: same road opposite direction and turning	**Total 8**	*n*	8								
				Sit.no	410								
		3: Same road, same direction and turning	**Total 8**	*n*	1	4	2	1					
				Sit.no	311	312	321	322					
	Straight road	2: Same road opposite direction	**Total 7**	*n*	1	4	1	1					
				Sit.no	211	241	242	270					
		1: Same road with same direction	**Total 43**	*n*	5	28	3	5	2				
				Sit.no	111	140	152	160	170				
Single accidents		No applicable option	**Total 19**	*n*	19								
				Sit.no	999								
		0: Single accidents	**Total 90**	*n*	18	9	5	4	5	4	12	31	2
				Sit.no	011	012	021	022	023	024	032	040	050

APPENDIX C PAPER III

Nonresponse and Dropout in Random Samples' and Volunteers' Self-Reports Of Traffic Accidents

Katrine Meltofte Møller

Anne Vingaard Olesen

Harry Lahrmann

Manuscript

NONRESPONSE AND DROPOUT IN RANDOM SAMPLES' AND VOLUNTEERS' SELF-REPORTS OF TRAFFIC ACCIDENTS

Katrine Meltofte Møller	*PhD.-fellow, Civil Engineering Department, Aalborg University, Denmark*
Anne Vingaard Olesen	*Associate professor, Civil Engineering Department, Aalborg University, Denmark*
Harry Lahrmann	*Associate professor, Civil Engineering Department, Aalborg University, Denmark*

Abstract

This paper investigates how three different survey invitation modes result in different compositions of respondents with regard to age and gender as well as different response and completion rates on traffic accident questionnaires. More than 75,000 people were invited through different modes to join a longitudinal study of traffic accidents. The three evaluated invitational modes were: A) A postal invitation to a random sample of 40,000 Swedes, B) an internet invitation sent via Digital Post (a secure mailbox provided to Danish citizens through which communication with authorities is received) to a random sample of 36,000 Danes, and C) a mass-distributed email and call for survey volunteers in numerous newsletters and homepages in Belgium with as random a target as possible. The participants all had the same invitational text to enroll in a longitudinal study of accident involvement and received the same four questionnaires (though in different native languages). The invitation through Digital Post led to the highest response rate to the longitudinal study (15%) but also yielded the highest proportion of respondent dropout (23%). Despite this, the digital invitational mode is recommended instead of invitation by regular post, where very low response rates were found (2%). The Digital Post invitation resulted in a final composition of full-compliant respondents, where young men and women as well as elderly women were underrepresented. The invitation by mass information in numerous free digital medias and at-hand email lists yielded high questionnaire return rates in the longitudinal study (83%) and a final composition of full-compliant respondents with more young and middle-aged respondents than the population in general, while lacking elderly participants.

1. INTRODUCTION

Self-reports, which are becoming increasingly frequent in traffic research studies (Lajunen, Özkan 2011), have many advantages, such as the cost-effectiveness of obtaining large sets of data. As the registered number of traffic accidents is decreasing for most countries in Europe (ERSO 2017)), self-reports can be used to complement official statistics to minimize the effect of underreporting in official data sets (Kamaluddin 2018, Lajunen & Özkan 2011, Møller et al. 2015). This is also the final scope for the data sets at hand; the study is part of a larger European cross-national longitudinal survey in the Horizon 2020 project InDeV (In-Depth understanding of accident causation for Vulnerable road users), with the aim of investigating the underreporting of traffic accidents and accident costs utilizing a self-report scheme. The paper at hand will not focus on the results regarding underreporting and accident causation but will instead present a methodological focus on the obtainment of self-reported data and its representativeness. This focus is sparsely covered in existing publications—a systematic literature review of self-reported traffic accidents finds that only 5 of 134 papers on the self-reporting of traffic accidents are considered to have a strong methodological focus (Kamaluddin et al. 2018). This paper will evaluate how different modes of invitation to survey participation influence the response rate and composition of respondents with regard to age and gender. While much is known about the effect of various aspects of the initial contact with respondents via paper formats, e.g., the salutation in a cover letter or the title of the signatory, a study shows that these findings are not transferable to email-based invitations (Porter, Whitcomb 2003). There are contradictory findings on the effect on response rates given the mode of invitation for survey participation. Sometimes, postal invitations yield higher response rates than emails; sometimes the opposite (Tai et al. 2018, Kaplowitz 2012). This paper will build to the knowledge of how invitational mode affects response rates, here, investigated through postal invitation, Digital Post (a secure mailbox provided to Danish citizens handling communication from authorities), and email/newsletter mass distribution.

The focus on response rates and sampling is uncommon, as response rates are often missing from literature on self-reported traffic accidents; the systematic literature review finds that 42% of the reviewed papers (134 in total) do not present response rates or the data to calculate them (Kamaluddin 2018). A study of invitational modes' effect on response rates in traffic safety studies has not been found. The aim of this paper is thus to address the gap in knowledge regarding the role that invitational mode plays in questionnaires in the context of self-reports of traffic accidents. This knowledge is important, as it is used to structure and design better self-report regimes with higher response rates. Problems with declining response rates have been an increasing issue in the last decades (Sax et al. 2003). As traffic accidents are rather uncommon events, response rates are of high relevance in traffic safety research studies; a large sample is needed to capture a sufficient amount of traffic accidents for statistical analysis. But most importantly, response rates are closely linked to nonresponse bias—the error that arises when not all invited to participate in the survey choose to respond (Dillman et al. 2014). This is problematic when the nonresponders differ in the characteristics relevant for the study, e.g., gender and age or being more or less likely to experience traffic accidents or having different types of traffic accidents than the responders. A solution to problems with a nonresponse bias, besides trying to increase response rates in tailoring the survey design, might also consist of using statistical weights to counterbalance the missing answers. Weighting rests on an assumption that the answers obtained from the survey respondents can be extrapolated to account for the answers nonresponders might have provided had they chosen to take part in the survey. For this to be the case, one must have knowledge of response propensities from different respondents (i.e., age, gender, etc.) and how they interact with the survey variables (accident data such as cause of accident, mode of transport, etc.). Little is known of this within the scope of accident analysis, thus using weighted answers is relatively uncommon. To the authors' knowledge, only a single study, carried out by Tivesten et al. (2012), goes into detail to examine response propensities with regard to questionnaires on traffic accident data; the study provides insight into which auxiliary variables one might use in further studies of response propensities and weighting but does not derive exact weight functions. As such, the work with weighted answers within questionnaires on traffic accidents is still relatively rough and unfinished. To develop and apply a weight function, one must also have knowledge of the nonresponders' characteristics—an impossibility when inviting volunteers via mass information, as in this study. Even with random samples, it might be difficult to obtain the nonresponders' characteristics, as this requires the sample database to have the necessary auxiliary variables recorded, e.g., income level, educational level, employment status, etc. Further, the accident itself is seldom as relevant as its specific nature, expressed in a variety of factors (e.g., transport mode, counterpart, severity, age, gender), the use of statistical weights cannot necessarily be applied due to sparse data. Weighting is therefore not the focus applied here; instead, we investigate how issues with nonresponse bias might be handled through careful consideration of the invitational mode.

The paper addresses the following questions within the field of surveys on self-reported traffic accidents:

- What do different invitational modes entail for response rates and dropout?
- How do respondents who have answered all four questionnaires in the longitudinal study differ from the overall population depending on invitational mode?
- What does this entail for the conclusions of surveys with self-reports of accidents?

We will address these questions in the following manner:

First, we will comment on response rates and dropout related to enrollment in the longitudinal study and investigate the representativeness of the participants with regard to gender and age.

Second, we will address dropout in the accident questionnaires; what are the age and gender of respondents who start to answer the accident questionnaire but do not finish it? Are there any common traits in the questions that prompt them to drop out?

Finally, we will address the final composition of full-compliance respondents with regard to age and gender depending on how the respondents were recruited.

2. METHOD

2.1 SAMPLE SELECTION AND SIZE

The different invitational modes were chosen to encompass the most commonly used methods in traffic accident self-reporting: volunteers and random samples (Kamaluddin 2018). The three different sampling methods were applied in three countries, respectively.

A) Digital Post to random sample: In Denmark, the sample was selected through the statistical bureau of Denmark who provided a random sample of 40,000 CPR numbers (a civil registration system that allocates a unique number for each citizen) of Danes above the age of 18. The Danish sample received the information letter with a link to the enrollment questionnaire via Digital Post (an online, secure mailbox used for contact from government authorities to Danish citizens.) Since 10 % of the Danish population is exempted from Digital Post, for instance, due to very high age or mental disabilities, the Danish sample consists of circa 36,000 Danes above the age of 18 with access to the Digital Post.

B) Regular post to random sample: In Sweden, the sample was selected by contacting PostNord, the Swedish postal company. They then provided the addresses of 40,000 Swedish inhabitants above the age of 18. The Swedish sample received the information letter with a link to the enrollment questionnaire via PostNord— that is, by physical, paper postal service. Thus, the Swedish sample consists of 40,000 adult Swedes with a valid postal address.

C) Mass information to volunteers: In Belgium, a mass information strategy was applied to get as many inhabitants as possible to volunteer as survey participants. The only constriction was that the call for volunteers should be free of charge for the researchers; thus, no paid internet panels or advertisement could be used. This was supposed to illustrate the circumstances under which researchers might operate without funding for sampling. The information letter with the link to the enrollment questionnaire was included in emails to students and staff at Hasselt University, participants from previous traffic research projects, and user groups of the Policy Research Centre. It was featured in newsletters and on homepages of the city of Antwerp, the Flemish Foundation of Traffic Knowledge (VSV), and the Flemish Association for Prevention and Protection (Prebes) as well as the Cyclist Association and the Pedestrian Association. Furthermore, the researchers' own personal contacts were used via email, Facebook, and LinkedIn. Due to the variety of media through which the call for participants was spread, it is impossible to determine how many have been made aware of the call for volunteers, and nonresponse cannot be calculated.

2.2 SURVEY DESIGN

The survey consists of two different questionnaires: an online enrollment questionnaire (EQ) and four online accident questionnaires (AQ).

The people who wanted to sign up for the study were asked to open the online EQ, which asked for the participant's email address (to which the AQs were sent) along with information regarding gender, age, and the size of the city in which the respondent lived.

The enrolled participants then took part in a longitudinal study of traffic accidents. An email with a hyperlink to the questionnaire regarding any involvement in traffic accidents (AQ) was sent to the respondent once every third month for a total of four times over one year. A reminder was sent after two weeks. The AQ consisted of one primary question: *"Have you, in the previous three months, been involved in one or more traffic accidents?"* If the respondents answered *"no*" the questionnaire ended, if they answered *"yes"* they were then provided with numerous questions regarding the accident. The questionnaire had a total of 35 items of which only the last did not elicit a forced answer. Otherwise, item nonresponse was not possible without exiting the questionnaire. All questions, except the last, were multiple choice with fixed response alternatives usually supplemented with an option of *do not know* or *none of the above*. The first 16 questions covered accident details, e.g., mode of transport and weather conditions at the time of the accident. The next 16 questions covered issues regarding the outcome of the accident, such as the need for medical care, rehabilitation, and material damage. Finally, two demographic questions regarding employment and income as well as a free text field were provided. The completion time was estimated at 20 minutes. For more information, the reader is referred to Møller et al. (2017), where the invitational letter and questionnaire could be found.

2.3 NONRESPONSE AND DROPOUT

Figure 1 provides an overview of the different terminologies regarding dropout and nonresponse applied throughout the paper.

Nonresponse to the items in the EQ de facto makes respondents unable to participate, as information, e.g., regarding the respondent's email were necessary for the longitudinal design. Item nonresponse in the AQ prompted the questionnaire to end, i.e., item nonresponse equals dropout

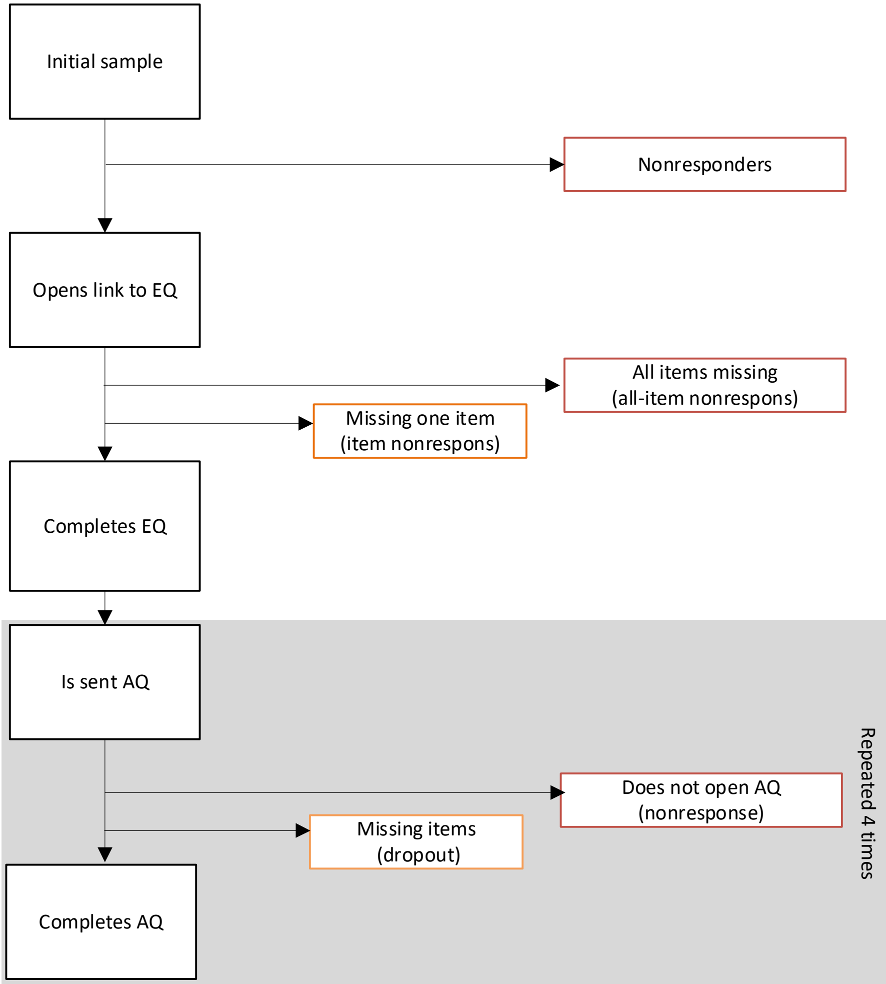


Figure 12: The different stages of nonresponse and dropout in the study. EQ: Enrollment questionnaire, AQ: Accident questionnaire. AQs were distributed four times to all respondents even if they did not answer the previous AQ.

2.4 STATISTICAL ANALYSIS

Data is analyzed through descriptive statistics, and 95% confidence intervals (CI) are calculated where percentages are shown. Comparisons of two percentages are carried out by z-test, i.e., evaluation of the test in the standard normal distribution (Kirkwood

& Sterne 2003). Otherwise, χ^2-tests are applied where appropriate. Significance level is $\alpha=0.05$.

The population statistics used to compare the participants' demography is found through the statistical bureau of each country for the year 2017 (Danmarks Statistik [2017], Statistiska Centralbyrån [2017], and StatBel [2017]).

3. RESULTS

3.1 EARLY DROPOUT AND RESPONSE RATES IN THE EQ

The response to the EQs is different depending on the sampling strategy; Table 1 provides an overview.

The response rate is significantly higher with Digital Post (15.4%) than with regular post (2.4%), (z-test=63.1; p<0.0001). On the other hand, the dropout rate is significantly lower with regular post (8.8%) than with Digital Post (15.9%) (z-test=7.2; p<0.0001). The dropout reason also varies with the invitational mode. While the respondents invited through Digital or regular post have more total nonresponses than single-item nonresponses, i.e., they have opened the questionnaire but not answered any of the questions, this is not the case for the respondents invited through mass information. The volunteers invited through mass information show more single-item nonresponses, most of these due to missing email addresses.

	Digital Post [DK]	**Regular post** [S]	**Mass information** [B]
Invited	36.000	40.000	**
Opened link to EQ	6.576	1.064	1.517
Nonresponse to all items	1.039	94	62
Nonresponse to one item	5	0	272*
Total dropout	1.044	94	334
Completed EQ	5.532	970	1.183
Dropout rate	15.88%	8.83%	22.02%
Response rate	15.37%	2.35%	**

**) 272 people filled out some but not all information (commonly, email is lacking).*
***) It is not possible to calculate the response rate for the volunteers from Belgium, as it is impossible to determine how many people have been made aware of the invitation for enrollment.*

Table 17: Response and dropout in the enrollment questionnaire

The enrolled participants' demographics derived from the three invitational methods can be seen in Table A1 in Appendix A. None of the invitational modes result in gender compositions that are significantly different from those in each country's overall population, but all invitational modes result in age compositions for each gender that are significantly different from that of each country's population (see test statistics in Appendix A). The mass distribution of invitations yielded many younger people and almost no elderly, whereas both Digital and regular post yielded many elderly respondents, especially elderly men. Thus, the age composition of respondents derived through mass information varies significantly from both the sample invited through Digital Post (DK, χ^2=302.3, df=2, p<0.001) and regular post (S, χ^2=246.8; df=2; p<0.001). Even though the samples from Digital and regular post both show the same trends of having fewer young people and more elderly men among the enrolled participants than in their respective country's population, the age distributions of Danish and Swedish respondents are significantly different (χ^2=8.2; df=2; p=0.02).

The enrolled participants were sent an AQ four times. The total nonresponse (not opening the questionnaire) can be seen in Table 3.

Nonresponse seems lower in the first distribution than in the following but somewhat stabilizes around 22% in the last parts of the longitudinal study. Yet, there remain statistically significant differences between the three sampling methods, as the nonresponse rates are all significantly different when comparing them within each round (Round 1: χ^2= 22.76, df= 2, p<0,001. Round 2: χ^2= 449.44, df=2, p<0,001, Round 3: χ^2= 6.47, df=2, p= 0.04. Round 4: χ^2= 7.74, df=2, p=0.02).

		Round 1	Round 2	Round 3	Round 4	Total
Digital Post [DK] N= 22,128	No.	151	2,434*	1,211	1,266	5,062
	%	2.73	44*	21.89	22.89	22.88
	95% CI	2.30-3.16	42.69-45.31*	20.80-22.98	21.78-24.00	22.32-23.43
Regular post [S] N=3,880	No.	24	194	185	191	594
	%	2.47	20	19.07	19.69	15.31
	95% CI	1.49-3.45	17.48-22.52	16.60-21.54	17.19-22.19	14.18-16.44
Mass information [B] N=4,732	No.	5	196	279	292	772
	%	0.42	16.57	23.58	24.68	16.31
	95% CI	0.05-0.79	14.45-18.69	21.16-26.00	22.22-27.14	15.26-17.37

**) Round 2 and the reminder overlapped with the summer vacation period in DK.*

Table 3: Nonresponse to AQ. Number of respondents not opening the AQ in each round of distribution (every three months)

Dropout, i.e., item nonresponse, in the AQ can only be examined in the answers of respondents who have begun by answering affirmatively that they had experienced an accident in the previous three months— respondents without any accidents were only provided with one question, and item nonresponse is thus nugatory. The concurrent dropout rates can be seen in Table 4. The dropout rate when inviting through Digital Post is significantly higher than the dropout rate with the other invitational modes (S, z=-3.97, p<0.001 and B, z=4.40, p<0.001). Regular post and mass information show no significant differences on the level of dropout.

		Affirms accident in AQ	**Completed accident questionnaires**	**Dropout**
Digital Post [DK] N= 5,532	No.	699	619	80
	%	12.64%	80.56%	11.45%
	95% CI	11.76-13.51	86.20-90.92	9.09–13.81
Regular post [S] N=970	No.	400	380	20
	%	41.24%	95%	5%
	95% CI	38.14-44.34	92.86-97.14	2.86–7.14
Mass distribution [B] N=1,183	No.	274	263	11
	%	23.26%	95,99%	4.01%
	95% CI	20.76-25.57	93.66-98.31	1.69–6.33

Table 4: Dropout rates on the accident questionnaire

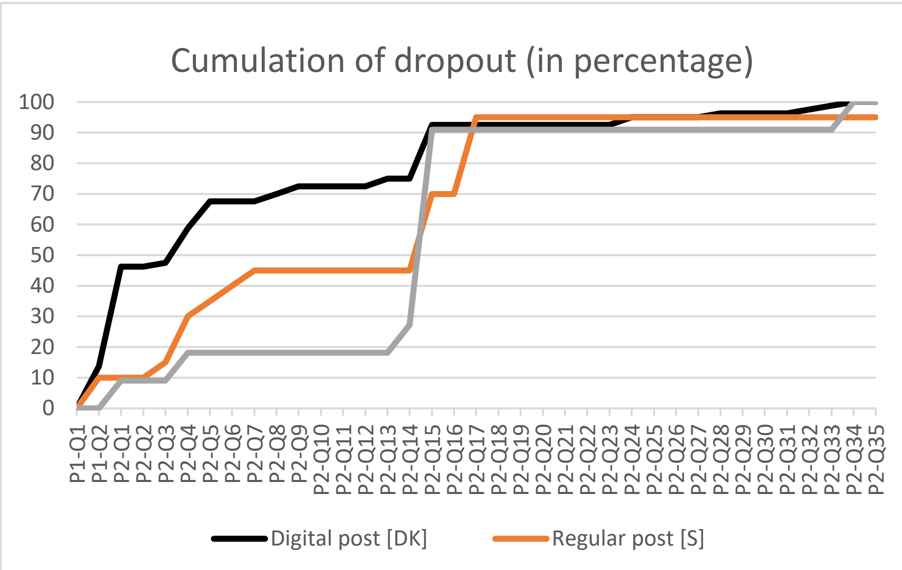


Figure 13: The distribution of dropout in the accident questionnaires. *The curves show the accumulation of dropout throughout the accident questionnaires, in percentage of the total dropout. The Y-axis shows the percentage, and the X-axis shows which question elicited the dropout, e.g., P2-Q15 is Part 2 Question 15 (P= Part, Q=Question). For the complete phrasings of all questions, the reader is referred to Møller et al. (2017). Some respondents (3 Danish and 1 Swedish) are missing from the figure because they had multiple accidents and thus dropped out on their second or third repetition of the questionnaire.*

P2-Q15: **Please state if the following conditions apply to your accident.** You can choose all the conditions you find applicable. [Multiple answers can be chosen].

1) I was listening to music
2) I was talking to someone (this includes use of cell phones)
3) I was immersed in thoughts
4) I was looking at my cell phone
5) I was in a hurry
6) I was tired
7) I was affected by alcohol or other intoxicating substances (e.g., medication or drugs)
8) I was distracted by something (e.g., another road user's behavior, a sign, or a shop)
9) I was ill/unwell
10) I was looking at my radio or navigation system *
11) My speed exceeded the speed limit at the location of the accident *
12) There was a mechanical error/technical fault on my means of transport *
13) I thought the other road user was aware of my presence*
14) I was not aware of the other road user*
15) I disregarded another road user's priority*
16) The other road user disregarded my priority*

P2-Q17: **Have you been in contact with one of the following due to your accident?** [Multiple answers can be chosen].

1) The police
2) Your own general practitioner
3) The emergency room or the hospital
4) Your insurance company
5) I have not been in contact with the abovementioned
6) I have not been in contact with the abovementioned yet, but I plan on contacting the police

Figure 3: The phrasing of P2-Q15 and P2-Q17 and all possible answers provided.

**) Notice that the answers given to the respondent in question P2-Q15 vary depending on his/her previous answer, e.g., respondents who had an accident while biking are not presented with the option: 11) "My speed exceeded the speed limit at the location of the accident."*

The item nonresponse in the AQ is distributed across the questionnaire as shown in Figure 2. From Figure 2, it is clear that item nonresponse is most common in the first half of the questionnaire. For the respondents invited through Digital Post, almost half of the item nonresponses occur within the first three questions. These are straightforward questions about the specifics of the accident such as P1-Q2: *"How many accidents did you experience?"* and P2-Q1: "*What means of transport were you using?"* After the initial item nonresponse in the accident-specific questions, two questions in particular seem to prompt dropout: P2-Q15 and P2-Q17. These questions are related to the conditions present at the time of accident and the contact to authorities or administrative bodies; the phrasings can be seen in Figure 3.

As the study was longitudinal, the respondents were presented with an AQ four times in total. Completion of the previous AQ was not a prerequisite for filling out the next AQ, and thus only a portion of the respondents answered all four AQs. The number of full-compliance respondents can be seen in Table 5 together with responses to the AQs.

	Digital Post [DK]	**Regular post** [S]	**Mass information** [B]
Sent AQ	22.128	3.880	4.732
Not opened AQ (nonresponse)	5.062	594	772
Item nonresponse (dropout)	80	20	11
Completed AQ	16.986	3.266	3.949
Response rate	76.76%	84.18%	83.45%
Full-compliance respondents	2.321	147	587
Percentage full compliance of enrolled	41.96%	15.15%	49.62%

Table 5: Overview of response and completion of accident questionnaires

The proportion of full-compliance respondents invited through regular post is very small compared to the other invitational modes.

Table A2 in Appendix A shows the demography of the full-compliance respondents. The total gender distribution in the full-compliance respondents invited through Digital Post is significantly different from that of the population; there are fewer men and more women in the final sample of full-compliance participants than women (46% men compared to 49%, z-test; -2.987, p=0.0028). The other invitational modes do not show any significant differences when comparing the gender distribution with the populations.

The age distribution is significantly different (see test statistics in Appendix A) from that of the general population both for men and women in the sample invited through Digital Post; for instance, young men comprised 7% of the sample compared with

12% in the population, and elderly women comprised 15% compared with 24% in the population. The age distribution within each gender category is also significantly different (see test statistics in Appendix A) from that of the overall population in the sample that was mass invited; for instance, elderly men and women comprised 1% and 0% of the sample, respectively, compared with 20% and 24% of the population. No significant difference is found on the age distribution of the full-compliance respondents invited through regular post compared to the population.

A comparison of the demographic information of respondents with unfinished AQs (dropout) with our full-compliance respondents is shown in Table A3 in Appendix A. None of the overall gender distributions of the dropout respondents appear to be significantly different from those of the full-compliant respondents when conducting z-tests. Due to the small sample size, we cannot calculate whether the age distribution of the dropout respondents with regular post and mass information as the invitational modes are different from that of the full-compliant respondents. A χ^2-test on the respondents invited via Digital Post does not show any significant difference in the age distribution of the dropout and full-compliance respondents.

4. DISCUSSION AND CONCLUSION

First, we will exemplify how the invitational mode and its implications on the demography of respondents could affect the results of traffic accident surveys. Second, we will discuss the methodological considerations of our study and compare our results with of those of others. Finally, we will provide the reader with recommendations drawn from this project regarding survey invitations and give an overview of the three invitational modes' shortcomings and advantages.

4.1 WHAT DOES THIS ENTAIL FOR SELF-REPORT STUDIES OF TRAFFIC ACCIDENTS?

Let us start by addressing what the findings entail for the knowledge gained of traffic accidents. We can exemplify this with a closer investigation of the composition of the full-compliance respondents invited via Digital Post.

We find that, among the full-compliance respondents, there are fewer men (46%) than women (54%), and the proportion is significantly different from that of the population. Men are more likely to be injured in a traffic accident than women; in fact, 62% of the injured in traffic accidents are men (Vejdirektoratet 2016). More than three times as many men than women are involved in single-vehicle accidents (Vejdirektoratet 2016). It is therefore obvious that an underrepresentation of men in a sample may induce biased results if nothing is done to compensate for the skewedness in gender distribution. For instance, nine out of ten drunk drivers are men (Rådet for Sikker Trafik 2019); thus, a composition of respondents with an overweight of women would require a very large sample if a reasonable number of accidents with drunk drivers should be reported. Not only the overall skewedness of gender distribution is problematic; the respondent's age must also be considered. The composition of full-compliance respondents lacks younger males and females, which proves troublesome, as young people are especially overrepresented in accident statistics; 20% of the injured are 18– 24 years of age, but only 8% of the population is 18–24 years old (Vejdirektoratet 2016). As the younger males are lacking full-compliance more than women, the results would be further biased, as twice as many young men than young women are injured in traffic accidents (Vejdirektoratet 2016). Single-vehicle accidents are thus very hard to gain knowledge of in the sample of full-compliance respondents, as 25% of all injuries in single-vehicle accidents are men aged 18–24 (Vejdirektoratet 2016).

When looking at the elderly among the full-compliance respondents, we find there are fewer women but more men than in the general population. We previously stated that there are more men than women injured in traffic accidents—this is true for all age groups with the exception of the elderly. At ages 75–80, there are more women injured in traffic accidents than men (Vejdirektoratet 2016); accidents with the elderly is thus underrepresented in the sample invited through Digital Post.

The number of traffic accidents recorded from this sample would thus, due to the scarcity of young people and elderly women, lead to an underestimation of accident frequency if extrapolated to count for the total population, if nothing were done to correct the skewedness. Further, if the invitational mode were to be used in a study of, for instance, single accidents or drunk driving, the number of people contacted for enrollment would need to be relatively high if occurrence of these themes should be fairly frequently reported by the final composition of full-compliance participants.

4.2 NATIONAL DIFFERENCES?

The different invitational modes were carried out in three different European countries: Denmark, Sweden, and Belgium. It could be argued that differences in the response propensities across the countries might influence the results. Yet, if we take a look into the commonly known response propensities such as income and educational level, the three countries are not that different. They are all in the top 5 of the median gross hourly earnings (EuroStat 2014). When measuring the BNP per inhabitant in Purchasing Power Standards, they are all within the seven countries in the EU with the highest purchasing power (EU Information Centre 2016). Educational level is quite similar: 45.9% of the population aged 30–34 have a tertiary education in Belgium; in Denmark, it is 48.8% and in Sweden 51.3% (EuroStat 2017). The three countries' unemployment levels in 2016 were 7.8% in Belgium, 6.2% in Denmark, and 6.9% in Sweden (EU Information Centre 2017). Of course, differences exist, but it is our assumption that the differences in the overall national response propensities are so small they will not change the overall study conclusions. For many reasons it was not possible to carry out the study in only one of the counties; for instance, Digital Post is only available in Denmark. Besides, contamination of randomly sampled respondents with mass information could also become a problem with concurrent sampling within only one country.

4.3 INTERNET AND EMAIL

Sixteen percent of the respondents invited via Digital Post open the EQ but do not fill in any information. This is higher than when invited through regular post (9%) or mass information (4%). Digital Post entails that respondents were using an electronic device connected to the internet while reading the information letter, which makes it faster and less troublesome to click on the link that opened the EQ. This is not the same for the regular post; here, the respondent had to find an electronic device with an internet connection and then manually type in the link provided in the letter before they could access the EQ. This could be considered more troublesome and time-consuming; thus, it is less likely that respondents would do this just to satisfy their curiosity if they have no intention of signing up for the survey. That such a mode shift results in lower response rates is congruent with the findings of Bech & Kristensen (2009). When examining the item nonresponse among the respondents invited through mass information, it is clear that unwillingness to share their email address is the most common reason for the 272 participants with item nonresponse. This is in stark contrast to the other invitational modes, where only 5 respondents among those invited through Digital Post and 0 respondents invited through regular post had item nonresponse. Still, the completion rate of the EQs opened with the mass information invitational mode was almost 80%, which is high compared with, e.g., Bandilla et al. (2012) who find that only 45% are willing to provide their email. To enhance completion rates, it would be advised to stress anonymity and the legitimacy of the survey in close proximity to where the respondent is asked for their email when sampling through mass information. That the respondent had to provide their email to enroll in the study was expected to cause selection bias; Bandilla et al. (2012) found that men are significantly more willing to provide their email address than women. Tivesten et al. (2012) also provide an expectation of gender differences; they have significantly more females (38%) responding to their questionnaire on traffic accidents than men (34%). Yet when examining the gender composition of our enrolled respondents, we find no significant difference from that of the general populations. The only invitational mode that leads to an overall gender composition among the full-compliance respondents significantly different from that of the population is Digital Post, with 54% women.

4.4 METHODOLOGICAL CONSIDERATIONS OF THE ACCIDENT QUESTIONNAIRE

The dropout rate in the AQ is significantly higher when inviting through Digital Post (11%) than the two other sampling modes (4-5%), and this sampling mode is much more susceptible to item nonresponse scattered throughout the first ten items. A study of the answers that prompt respondents to drop out is of interest, as it can shed light on the methodological aspects of survey. Less than 10% of item nonresponse occurs in the second half of the AQ, regardless of invitational mode. This indicates that respondents that have begun answering do not drop out due to the length of the questionnaire, even though the 35 questions have an estimated completion time of 20 minutes. This is in line with previous findings regarding self-reports of traffic accidents that finds respondents to be very dedicated to describing their accident (Lahrmann et al. 2014). The current study adds to the findings with the evidence that this dedication is not exclusive to bicycle enthusiasts but also appears in random sampling.

Two specific questions seem to influence dropout due to their content. The first question is related to the circumstances during the accident. There are two obvious reasons why this question could prompt dropout: structure and content. The number of possible answers is quite long (the exact length depending on the specific accident details), and thus the question can seem overwhelming for some respondents. Some of the content of the answer options can be regarded as private, e.g., the use of alcohol/narcotics and speeding. This could also trigger some respondents to drop out. The question regarding contact to authorities and administrative bodies also seems to specifically prompt dropout. This item does not have any obvious structural flaws—the reasons for dropout here are expected to be purely related to the content, which could also, to some extent, be considered private, e.g., contacting the police or insurance.

The nonresponse in the AQs (respondents not opening the questionnaire) is affected by the longitudinal design. The nonresponse is substantially smaller in the first distribution of AQs (maximum 3%) and more or less stable in the following three distributions. The exception to this is the second distribution to respondents invited through Digital Post, where the nonresponse is twice as big in the second round (44%) as in the third and fourth distribution. This is most likely due to the collision of both the questionnaire and reminder distribution with the summer holidays. Distribution dates were not uniform across sampling modes, so this was not the case for the regular post sample or the mass information sample. In future studies, it would be recommended to keep all distribution dates out of the holiday season when planning a longitudinal study.

Mass information yields a sample with worse representativeness than random sampling with regard to age; the elderly are underrepresented. This is in contradiction to the findings of Tivesten (2012), who found response rates on AQs for the elderly to be higher than for the young or middle-aged. The full-compliance respondents yielded through mass information show a higher proportion of both younger males and females than in the general population, but with the elderly of both genders clearly lacking. This shows the need for tailoring the distribution of invitations specifically to the elderly when mass information is applied. Besides interest organizations for pedestrians and bicyclists, organizations and associations with many elderly members should be contacted when applying a mass invitation design. It might also be advised to use a paper mode for the questionnaire (and not just the invitation) to improve elderly representativeness, as Bech & Kristensen (2009) find that response rates to postal questionnaires in a sample of older respondents are significantly higher than in web-based questionnaires (42% versus 17%). As work by Denscombe (2006) indicates, there seems to be no significant difference in the information respondents provide whether or not the questionnaire they answer is provided digitally or on paper, and there should be no reason not to provide elderly respondents with a paper version of the questionnaire and other respondents an online version.

4.5 THE PREFERRED INVITATIONAL MODE DEPENDS ON RESEARCH GOALS

As stated by Andersen et. al (2018), it is important to keep in mind what the goal of one's study is when deciding on a survey design including invitational mode. The overall advantages and disadvantages of the different invitational modes can be seen in Figure 3.

	Advantages	Disadvantages
Random sample with Digital Post	+ Highest response rate in enrollment + High percentage of full-compliance respondents + Highest number of respondents and completed questionnaires	÷ Not free ÷ High dropout in enrollment ÷ Highest dropout in accident questionnaires ÷ Elderly females underrepresented
Random sample with regular post	+ Low dropout in enrollment + Lowest dropout in accident questionnaires	÷ Not free ÷ Low number of enrolled participants ÷ Poor at gaining knowledge of young people
Volunteers with mass information	+ Can be carried out free of costs + High percentage of full-compliance respondents + Very good at gathering data from young men and women	÷ High dropout rate in enrollment due to the question about providing email address ÷ Very poor at gaining knowledge of elderly people

Figure 14: Summarization of the advantages and disadvantages of the three invitational modes.

It is quite clear that, if one needs to conduct a study on a random sample, the Digital Post invitational mode shows better results in terms of response rates, full compliance, and representativeness than regular post. If one’s research goals can be reached by using volunteers and not necessarily a random sample, it is also evident that mass information yields results fairly equal to regular post regarding dropout and response rates to AQs. Yet, there are substantially more full-compliance respondents in the sample invited via mass information, and in that respect, it surpasses regular post as an invitational method. Representativeness regarding age is low in the sample; there are very few elderly, but the invitational mode proved impressive at gaining young respondents.

The abovementioned conclusions and recommendations are, of course, just some of the many considerations related to survey design. Much is still unknown, both of invitational aspects (e.g., content of the information letter, greeting, and status of the signatory) as well as response propensities related to respondents and accident-specific variables. A heightened awareness of the need to report response rates in papers concerning self-reports of traffic accidents would be a first step toward gaining more information to improve the methodology of self-report regimes.

Acknowledgements

This work is a part of the EU Horizon 2020 project InDeV (In-Depth understanding of accident causation for Vulnerable road users, grant agreement No. 635895).

REFERENCES

Andersen, C. S., Madsen, T. K. O., Agerholm, N., & Møller, K. M. (2018). *Chapter 3: Self-reporting of accidents and near-accidents.* I E. Polders, & T. Brijs (red.), *How to analyse accident causation?: A handbook with focus on vulnerable road users* (p. 71–92). Hasselt University.
Available: https://www.indev-project.eu/InDeV/EN/Handbook/Handbook-InDeV.pdf?__blob=publicationFile&v=1

Bandilla, W., Couper, M. P., Kaczmirek, L. (2012). The Mode of Invitation for Web Surveys. *Survey Practice* Vol. 5, Issue 3, 2012.
Available: https://www.surveypractice.org/article/3096-the-mode-of-invitation-for-web-surveys

Bech, M., & Kristensen, M. (2009). Differential response rates in postal and Web-based surveys in older respondents. *Survey Research Methods*, 3(1), 1–6. http://dx.doi.org/10.18148/srm/2009.v3i1.592

Danmarks Statistik. (2017). *Folketal*.
Retrieved from: http://www.dst.dk/da/Statistik/emner/befolkning-og-valg/befolkning-og-befolkningsfremskrivning/folketal

Denscombe, M. Web-Based Questionnaires and the Mode Effect - An Evaluation Based on Completion Rates and Data Contents of Near-Identical Questionnaires Delivered in Different Modes. *Social Science Computer Review*. Volume 24 Number 2 Summer 2006 246–254.

Dillman, D. A., Christian, L. M., & Smyth, J. D. (2014). *Internet, phone, mail, and mixed-mode surveys - the tailored design method* (4th ed.) John Wiley & Sons, Incorporated. ISBN: 9781118921302 (e-book).

ERSO, European Road Safety Observatory 2017. *Annual Accident Report 2017.*
Available:https://ec.europa.eu/transport/road_safety/sites/roadsafety/files/pdf/statistics/dacota/asr2017.pdf

EuroStat (2017). Educational attainment. Homepage:
https://ec.europa.eu/eurostat/statistics-explained/index.php?title=Educational_attainment_statistics
Last visit 01-10-2019

EuroStat (2014). Median gross hourly earnings. Homepage: https://ec.europa.eu/eurostat/statistics-explained/index.php?title=File:Median_gross_hourly_earnings,_all_employees_(excluding_apprentices),_2014_YB17.png
Last visit 10-01-2019.

EU Information Centre (2017). The Danish Parliament, EU Information Centre. *Unemployment*. Homepage: https://www.eu.dk/da/fakta-om-eu/statistik/arbejdsloeshed
Last visit 01-10-2019.

EU Information Centre (2016). The Danish Parliament, EU Information Centre. *BNP/PPS*. Homepage: https://www.eu.dk/da/fakta-om-eu/statistik/bnp
Last visit 01-10-2019.

Kamaluddin, N. A, Andersen, C. S., Larsen, M. K., Meltofte, K. R. , & Várhelyi, A. (2018). Self-reporting traffic crashes – a systematic literature review. *European Transport Research Review*. 10:26. https://doi.org/10.1186/s12544-018-0301-0

Kaplowitz, M.D, Lupi, F., Couper, M. P., Thorp, L. (2012). The Effect of Invitation Design on Web Survey Response Rates. *Social Science Computer Review*. 30(3) 339-349. https://doi.org/10.1177/0894439311419084

Kirkwood, B. & Sterne, J. (2003). *Essential Medical Statistics. 2nd edition.* Blackwell Science Ltd, Oxford UK.

Lajunen T, Özkan T (2011) Chapter 4 - Self-Report Instruments and Methods A2 - Porter, Bryan E. In: *Handbook of traffic psychology*. Academic Press, San Diego, pp 43–59.

Lahrmann, H., Madsen, J. C. O., Madsen, T. K. O., Olesen, A. V., Thedchannamoorthy, S., & Bloch, A-G. (2014). *Projekt Cykeljakke: Den sikkerhedsmæssige effekt af en gul cykeljakke.* Aalborg: Aalborg Universitet. Institut for Byggeri og Anlæg. (DCE Technical Reports; Nr. 167). Available:http://vbn.aau.dk/files/197570851/Projekt_cykeljakken_den_sikkerhedsm_ssige_effekt_af_en_gul_cykeljakke.pdf

Møller, K. M., Andersen, C. S., Várhelyi, A., Schönebeck, S., Reumers, S., Hosta, P., & Szagała, P. (2017). *Accident Information from six European Countries Based on Self-reports*. Bruxelles: Publications Office of the European Union. Available:http://vbn.aau.dk/files/262153504/Accident_Information_from_six_European_Countries_Based_on_Self_reports.pdf

Møller, K. M., Madsen, T.K.O., Olesen, A.V., & Lahrmann, H. (2015) *A case study on agreement between self-reported bicycle accidents and hospital and police records.* Selected Proceedings from the Annual Transport Conference at Aalborg University 2015. Available:http://www.trafikdage.dk/td/papers/papers15/324_KatrineMeltofte_fagfaellebedoemt.pdf

Rådet for Sikker Trafik (2019). *Spirituskørsel.* Homepage: https://www.sikkertrafik.dk/raad-og-viden/i-bil/spirituskoersel Last visit 01-10-2019.

Sax, L.J., Gilmartin, S.K., & Bryant, A.N. (2003). Assessing Response Rates and Nonresponse Bias in Web and Paper Surveys. *Research in Higher Education* 44: 409. https://doi-org.zorac.aub.aau.dk/10.1023/A:1024232915870

StatBel (2017). *Population figures*. Retrieved from: https://statbel.fgov.be/en/themes/population/structure-population#figures

Statistiska Centralbyrån. (2017). *Population statistics*. Retrieved from: http://www.scb.se/en/finding-statistics/statistics-by-subject-area/population/population-composition/population-statistics/

Porter, S.R. & Whitcomb, M.E. (2003). The Impact of Contact Type on Web Survey Response Rates. *The Public Opinion Quarterly*, Vol. 67, No. 4 (Winter, 2003), pp. 579–588 Published by: Oxford University Press on behalf of the American Association for Public Opinion Research Stable. Available: https://www.jstor.org/stable/3521694

Tai, X., Smith, A.M., McGeer, A.J., et al. (2018). Comparison of response rates on invitation mode of a web-based survey on influenza vaccine adverse events among healthcare workers: a pilot study. *BMC Medical Research Methodology* ,18: 59. https://doi-org.zorac.aub.aau.dk/10.1186/s12874-018-0524-8

Tivesten, E., Jonsson, S., Jakobsson, L., & Norin, H. (2012). Nonresponse analysis and adjustment in a mail survey on car accidents. *Accident Analysis and Prevention,* 48, 401-145. https://doi.org/10.1016/j.aap.2012.02.017

Vejdirektoratet (2016). *Kønsforskelle i trafikulykker.* Published by Vejdirektoratet, København Denmark. Available:http://www.vejdirektoratet.dk/SiteCollectionDocuments/kønsforskelle_i_trafikulykker_web.pdf

APPENDIX

			Below 25	25-65	Above 65	Total
Digital Post	men	n	240	1,764	727	2,731
		%	8.79	64.59	**26.62**	49.37
		% of male population	12.13	67.41	**20.46**	49.33
	women	women	283	2,022	496	2,801
		%	10.10	**72.19**	**17.71**	50.63
		% of female population	11.26	**64.84**	**23.90**	50.67
regular post	men	men	28	339	109	476
		%	**5.88**	71.22	22.90	49.07
		% in male population	**11.24**	66.98	21.78	49.80
	women	women	38	359	97	494
		%	7.69	**72.67**	**19.64**	50.93
		% in female population	10.34	**64.27**	**25.39**	50.20
mass information	men	men	84	494	21	599
		%	14.02	**82.47**	**3.51**	50.63
		% in male population	10.63	**69.69**	**19.69**	48.71
	women	women	132	451	1	584
		%	**22.60**	**77.23**	**0.17**	49.37
		% in female population	**9.77**	**65.92**	**24.31**	51.29

Table A1: Demographic information of enrolled participants. Numerical differences larger than 5 percentage points are marked in bold.

Test statistics:
Gender
Digital Post percentage men: z-test=0.05, p=0.96
Regular post percentage men: z-test=-0.45 p=0.65
Mass information percentage men: z-test= 1.32, p=0.186

Age
Digital Post women: χ^2= 71.55 df=2 p<0.000
Digital Post men: χ^2=78.89, df=2, p<0.000
Regular post men: χ^2=13.72 df=2, p=0.001
Regular post women: χ^2= 15.23 df=2 p<0.000
Mass information men: χ^2=100.20 df=2 p<0.000
Mass information women: χ^2= 249.63 df=2 p<0.000

			Below 25	25-65	Above 65	Total
Digital Post	men	n	72	742	259	1,073
		%	**6.71**	69.15	24.14	46.23
		% in male population	**12.13**	67.41	20.46	49.33
	women	n	105	958	185	1,248
		%	8.41	**76.76**	**14.82**	53.77
		% in female population	11.26	**64.84**	**23.90**	50.67
regular post	men	men	6	51	20	77
		%	7.79	66.23	25.97	52.38
		% in male population	11.24	66.98	21.78	49.80
	women	women	6	46	18	70
		%	8.57	65.71	25.71	47.62
		% in female population	10.34	64.27	25.39	50.20
mass information	men	men	42	234	4	280
		%	15.00	**83.57**	**1.43**	47.70
		% in male population	10.63	**69.69**	**19.69**	48.71
	women	women	76	230	1	307
		%	**24.76**	**74.92**	**0.33**	52.30
		% in female population	**9.77**	**65.92**	**24.31**	51.29

Table A2: Demographic information of full-compliance participants who have finished all accident questionnaires compared with values for the population in general.
Numerical differences larger than 5 percentage points are marked in bold.

Test statistics

Age

Digital Post men:	χ^2 =33,51, df=2,	p<0,000
Digital Post women:	χ^2=79,33, df=2,	p<0,000
Mass information men:	χ^2=60,19, df=2,	p<0,000
Mass information women:	χ^2=146,91, df=2,	p<0,000

			Below 25	25-65	Above 65	Total
Digital Post	men	n	4	17	5	26
		%	**15**	65.38	19.23	32.50
		% in full compliance	**6.71**	69.15	24.14	46.23
	women	n	9	35	10	54
		%	**16.67**	**64.81**	**18.52**	67.50
		% in full compliance	**8.41**	**76.76**	**14.82**	53.77
regular post	men	men	2	7	1	10
		%	**20**	**70**	**10**	50.00
		% in full compliance	**7.79**	**66.23**	**25.97**	52.38
	women	women	1	8	1	10
		%	10	**80**	**10**	50.00
		% in full compliance	8.57	**65.71**	**25.71**	47.62
mass information	men	men	1	4	0	5
		%	**20**	80	0	45.45
		% in full compliance	**15**	83.57	1.43	47.70
	women	women	0	6	0	6
		%	**0**	**100**	0	54.55
		% in full compliance	**24.76**	**74.92**	0.33	52.30

Table A3: The distribution of age and gender on the participants who dropped out before finishing the accident questionnaires. Numerical differences larger than 5 percentage points are marked in bold.

APPENDIX D PAPER IV

Perceived Fault and Accident Causation Factors in Self-Reported Traffic Accidents

Katrine Meltofte Møller

Anne Vingaard Olesen

Laila Martinussen

Harry Lahrmann

Manuscript

PERCEIVED FAULT AND ACCIDENT CAUSATION FACTORS IN SELF-REPORTED TRAFFIC ACCIDENTS

Katrine Meltofte Møller — *PhD.-fellow, Civil Engineering Department, Aalborg University, Denmark*
Anne Vingaard Olesen — *Associate professor, Civil Engineering Department, Aalborg University, Aalborg, Denmark*
Laila M. Martinussen — *Senior researcher, DTU Management Engineering, Technical University of Denmark, Kgs. Lyngby, Denmark*
Harry Lahrmann — *Associate professor, Civil Engineering Department, Aalborg University, Aalborg, Denmark*

Abstract

For one year, 5,532 Danish citizens self-reported their traffic accidents, reporting 532 accidents in total. These self-reports contain, e.g., information on accident causation factors and who the respondent finds at fault for their accident. This research shows that respondents are more likely to blame their counterpart than themselves for their accident, either wholly or partly (33%). Mode of transport affects the perceived fault; bicyclists are less likely to find themselves at fault than car drivers. The admission of several accident causation factors are investigated, and their lacking congruency with other data sources is discussed. The paper proposes that self-reported causation factors should be used as an insight into respondents' beliefs toward traffic culture.

1. INTRODUCTION

When trying to improve traffic safety, a common method is investigating the cases where something went wrong, i.e., accidents. This investigation could focus on either an aggregated form or in an in-depth analysis of each accident (an overview presented by, e.g., InDeV [2016]). The former can be applied in safety campaigns, the groundwork for legal changes, and the development of assistive equipment in cars. For instance, if many accidents happen because of distraction, campaigns could be implemented to heighten attention, laws could change to ban the distracting features (as with handheld mobile phones), and technical equipment in car cabins could be developed to monitor the driver's focus and alert them if something is amiss. The foundation for all this is the knowledge of common accident causation factors.

In Denmark, traffic accident knowledge is based on police reports. Unfortunately, it is well known that police reporting is incomplete and biased - not just in Denmark (Janstrup, Kaplan et al. 2016) but in other countries as well (Elvik & Mysen 1999, Derriks & Mak 2007). To complement police records, self-reports are increasing in frequency (Kamaluddin et al. 2018, Madsen et al. 2018, Lajunen & Özkan 2011). Whereas police records are made by objective, trained police personnel, self-reports are - as the name suggests - made by the accident-involved persons themselves. One of the major strengths of self-reports compared to police records is that self-reporting schemes provide an opportunity to ask the accident-involved persons about the circumstances leading up to the accident, for instance, as carried out by Tivesten (2013), who compared self-reported accident descriptions with other data sources.

However, the subjective aspect of the self-report methodology can be also one of its weaknesses. Problems with the respondents' memory/recollection as well as social desirability issues (Wåhlberg, Dorn, & Kline, 2010) are often discussed as problems with regard to the validity and reliability of the self-reported information (af Wåhlberg 2002, Lajunen & Özkan 2011, Christianson & Lofthus 1987). Deterioration of memory is not relevant in the current study, which applies a recall period of a maximum of three months; a literature review by Kamaluddin et al. (2018) found that a recall period of a maximum of one year is often considered preferable to minimize erroneous recall. But social desirability is the key issue and a central focus of many of our further discussions.

Social desirability can be seen as composed of two factors: an intentional deception referred to as "other-deception" or "impression management" and an unintentional deception referred to as "self-deception" (Paulhus 1984). Other-deception is the deliberate intention to provide others with a better image of oneself than what the respondent believes to be true, meaning that the respondent lies and falsifies their self-report (Lajunen & Özkan 2011). Other-deception increases in public settings, which indicates that anonymity reduces the effect (Paulhus 1984). Self-deception is unintentional; the respondents give answers that they believe to be honest, yet the answer might be objectively untrue (Lajunen & Özkan 2011). Thus, self-deception cannot be reduced by anonymity or other methodological features. Self-deception is a way to cope with emotions like anxiety and the willingness to maintain high self-esteem; the respondents are more likely to claim positive characteristics than deny negative characteristics (Paulhus 1991).

That social desirability affects answers when applying a self-report methodology is obvious. However, there seems to be no consensus on what this entails for the use of self-reported accident information in traffic safety research, even though the implications on measures of driver behavior seem relatively well established (Lajunen & Özkan 2011, de Winter & Dodou 2010, af Wåhlberg 2010).

To address this gap, this paper will focus on self-reported accident data related to accident causation and fault, two subjects that are traditionally investigated by objective personnel (police and/or accident investigators) and seem highly susceptible to social desirable responding.

Two research questions (RQ) are stated:

> RQ1: Which factors (demographic, modal, and accident related) influence the perception of own fault in traffic accidents?
>
> RQ2: What can we learn from the respondents' choice of accident causation factors; can they be taken at face value, or should they be interpreted as something else?

This paper will show how accident causation factors recorded in self-reports diverge from factors found by other methods. We will focus on the discrepancies one would suggest could be related to social desirable responding. As an exploratory study, we will then move on to show how the self-reported data can be applied more appropriately in a study of beliefs and traffic culture - here exemplified by an investigation into how accident causation factors influences who the self-reporter holds at fault for their accident, something that has not previously been studied.

2. METHOD

2.1 SAMPLING OF RESPONDENTS

In Denmark, all inhabitants (national citizens, foreigners with a residence permit, or people otherwise required to pay taxes in Denmark) are issued a CPR number (Civil Registration Number, a unique personal identifier), (CPR-administration 2018). The CPR numbers of 40,000 random individuals above the age of 18 were obtained through Statistics Denmark. The sample was then contacted through Digital Post (an online, secure mailbox through which communication from public authorities is sent to Danish citizens). They were each sent an information letter about the study that contained a link to an online enrollment questionnaire. Since 10% of the Danish population is exempted from Digital Post (e.g., due to age or mental health), approximately 36,000 people out of the original sample of 40,000 received the information letter in their digital mailbox. Thereafter, 5,532 persons (15%) registered for participation in the study.

2.2 QUESTIONNAIRE DESIGN

Every third month between 2012–2013, the enrolled participants received an email with an online accident questionnaire that asked if they had experienced any traffic accidents in the previous three months. If answering in the positive, they were then given several questions about their accident. To reduce the influence of "other-deception," it was stressed that the survey was anonymous and the answers confidential.

The participants who had an accident with a counterpart were asked about their fault in the accident. The question was phrased: "*Please choose the option with which you agree the most.*" Possible answers were: "*I think the accident was mainly the other road user's fault,*" "*I think the accident was mainly my own fault,*" and "*We were equally to blame for the accident.*" The latter two will be combined (due to small sample sizes) as an indication that the respondent partly or wholly found themselves at fault. Respondents were also asked about the presence of several possible accident causation factors. The question was phrased: "*Please state if the following conditions apply to your accident. You can choose all the conditions you find applicable.*" The respondent was then provided with a list of multiple-choice options.

Only pedestrians and vehicle drivers were included in the study; passengers would not be able to respond meaningfully on fault and causation factors. Only multiparty accidents were included in the study.

For more information on the complete questionnaire, information letter, other self-reported accident information, and detailed demographic information about the participants, please see Møller et al. (2017).

2.3 STATISTICAL TESTING

Summary statistics were calculated as frequencies and percentages with 95% confidence intervals (CI). Tests of significance were conducted as χ^2 tests. We used a binary logistic regression model with respondent admission of being partly or wholly at fault for the accident as the outcome and a specific accident causation factor as the independent variable. The measure of association of the logistic regression model was the odds ratio (OR). In subsequent analyses, the OR was controlled for gender, respondents' own means of transport, and their age in 10-years age groups.

The OR was used to measure the strength of the association between admission of a given accident causation factor and admission of being at fault in the accident, which could be seen as an indirect measurement of how strongly respondents believed their behavior was connected to the causation of the accident. In other words, if a respondent admitted to having been speeding/hurrying/talking etc., what are the odds that he/she will find him/herself at fault in the accident? Thus, the OR can be seen as a measurement of how respondents perceived their behavior as contributory to the accident taking place.

The statistical analysis was conducted using STATA (version 15.1; STATA Corp., College Station, TX), and the significance level was α =0.05.

3. RESULTS

3.1 ADMISSION OF CAUSATION FACTORS

In total, the 5,532 respondents reported 532 accidents, of which 300 were with a counterpart; this is the dataset applied here. Due to sparse data on transport modes other than a car or bicycle (cf. Table 1), further analysis will only be conducted on car drivers and bicyclists.

Table A1 in appendix presents the causation factors that the respondents chose as applicable for their accident. Most frequently, the respondents believed that their counterpart was aware of them (41% of the accidents), or the respondents stated that they were not aware of their counterpart (31%). Accident causation factors such as difficulties with orientation and distraction (12%), auditory tasks (10%), and stress (13%) were also common.

3.2 ADMISSION OF FAULT

The respondents were asked how they perceived the fault of the accident. One-third admitted wholly or partly that the accident was their fault, c.f. Table 2, and two-thirds perceived the counterpart to be primarily at fault.

Table 2: Respondents' perceived fault for the accident, N=300.

Perceived fault	N	%	95% CI
Primarily my own fault	75	25%	20.10–29.90
We were equally to blame	23	7.67%	4.66–10.68
Primarily the counterpart's fault	202	67.33%	62.03–72.64

3.2.1 ADMISSION OF FAULT AND TRANSPORT MODE

From Table 3, it is clear that the admission of fault varies with the respondents' mode of transport; car drivers finds themselves primarily at fault in 29% of the accidents, whereas the same is true for only 14% of the accidents with responding bicyclists; thus, car drivers finds themselves primarily at fault twice as often as bicyclists do. The difference in car drivers' and bicyclists' perceived fault is significant when applying a χ^2 test (p=0.03314).

Bicyclists show a significantly different (p=0.033) perception of fault depending whether their counterpart's mode of transport was another bicyclists or a car (cf. table 4); they are more likely to blame the counterpart if the counterpart was a car.

Car drivers show a significantly different perception of fault depending on their counterpart (p=0.009); they are more likely to blame themselves if their counterpart was a bicyclist, cf. Table 5.

Table 3: Association between admission of fault and mode of transport.

Mode of transport (self)	My own fault		Equally to blame		The counterpart's fault		Total
	N	%	N	%	N	%	N
Car (drivers only)	61	28.64%	14	6.57%	138	64.79%	213
95% CI	*22.56– 34.70*		*3.24–9.82*		*58.37–71.20*		
Bicycle	10	14.08%	8	11.27%	53	74.65%	71
95% CI	*5.99–22.18*		*3.91–15.18*		*64.53–84.76*		
Other	4		1		11		12
Total	75	25%	23	7.67%	202	67.33%	300

Table 4: Bicyclists' perceived fault of the accident, dependent on their counterparts' mode of transport

Counterpart	Partly or wholly my own fault			The counterpart's fault			Total
	N	%	95% CI	N	%	95% CI	N
Car	4	13.79	1.24–26.34	25	86.21	73.66–98.77	29
Bicycle	12	37.5	20.73–54.27	20	62.5	45.73–79.27	32
Other	2	16.67	-4.42–37.75	8	66.67	39.99–93.34	12
Total	18	25.35	15.23–35.47	53	74.65	64.53–84.77	71

Table 5: Car drivers' perceived fault of the accident, dependent on their counterparts' mode of transport

Counterpart	Partly or wholly my own fault			The counterpart's fault			Total
	N	%	95% CI	N	%	95% CI	N
Car	60	33.52	26.60–40.44	119	66.48	59.56–73.40	179
Bicycle	9	69.23	44.14–94.32	4	30.77	5.68–55.86	13
Other	6	28.57	9.25–47.89	15	71.43	52.11–90.75	21
Total	75	35.21	28.80–41.63	138	64.79	58.37–71.20	213

3.2.2 ADMISSION OF FAULT AND DEMOGRAPHIC VARIABLES

We find no significant difference in the admission of fault depending on gender (p=0.883) or age (p=0.159) when conducting χ^2 tests, cf. Tables A2 and A3 in Appendix A.

3.3 ADMISSION OF FAULT AND CAUSATION FACTORS

A calculation of adjusted ORs was performed, c.f. Table 6. The factors related to attention are most strongly (OR=55.45) associated with the respondents' perception of being wholly or partly at fault for the accident. Second most are the factors related to orientation and distraction (OR=20.55). Two themes do not show significant values: vehicle errors and auditory tasks. The factors related to interpretation and assessment have an OR lower than one, meaning that respondents are more likely to blame the counterpart if they believe the counterpart was aware of them. It has not been possible to calculate the OR for Inappropriate driving behavior, as all 12 respondents who acknowledged one of the accident causation factors in this theme (cf. Table A2 in Appendix A) also believe themselves to be fully or partially responsible for the accident. This is also the case for the OR for illness.

Table 6: The OR for admission of fault (respondent wholly or partly at fault) and accident causation factors.

	Odds Ratio	P>|z|	95% CI
Fatigue	6.94	0.006	1.76–27.28
Attention	55.45	0.000	24.93–123.32
Interpretation/assessment	0.10	0.000	0.05–0.20
Inappropriate driving behavior			*
Orientation and distraction	20.55	0.000	7.38–57.26
Vehicle errors	7.72	0.122	0.58–102.83
Stress	4.56	0.000	2.13–9.78
Auditory tasks	2.18	0.059	0.97–4.92
Illness			*

**Due to zero cells, the OR was not estimable for Inappropiate driving behavior (12 observations) or Illness (4 observations).*

4. DISCUSSION

4.1 PERCEIVED FAULT

Being at fault is shown to have implications for the accident-involveds' mental health; persons who are not at fault in their accidents demonstrate more mental and emotional disturbance, though their physical health is the same than at-fault persons (Littleton et al. 2012). Given that perspective, it is concerning that only one-third of the participants take full or partial responsibility for their accident. This could perhaps be explained by the methodological drawbacks of a voluntary self-report regime, as persons who have experienced a traffic accident that makes them feel guilty and remorseful might evade recollection of the event by not reporting the accident, as argued by Meunier et al. (2018). The methodology applied in this study (asking a random sample instead of a sample consisting only of persons who have been known to previously have a traffic accident) seems to lessen this effect; in the study by Meunier et al. (2018), 20% perceived themselves as either fully or partly responsible for their accident; in our study, the percentage is 33%. Another reason for the relatively small amount of perceived at-fault respondents could be that respondents are not objective in their perception of fault, a finding congruent with those of Preston & Harris (1965), who found considerable discrepancies between culpability judged by the police and culpability judged by their study participants. The mechanisms behind this are not well established in the sparse literature on perceived fault but could perhaps be related to the mechanisms of self-serving bias or superior conformity of self, as seen by the fact that people tend to see themselves as better drivers than the average population (Félonneau 2013). The current study does not aim to assess the correctness of culpability, which could be necessary for different types of studies (as argued by Wåhlberg & Dorn 2007), but instead focuses on what can be learned from the accident-involved respondent's own view of culpability. The overall finding that more respondents than not find the counterpart at fault is congruent with other self-report studies (Meunier et al. 2018, Mayou & Bryant 2003).

We found no significant association between perceived fault and age or gender. Gender differences were expected, as Siren & Meng (2013) found that older men tend to rate their self-assessed driving skills higher than do older women, and Gheorghiu & Havâneanu (2011) found that women tend to feel more guilt than men. We expected to find a similar trend regarding age and gender, as it is known that these factors are associated with different types of faults while driving (Karacazu & Arzu 2011), but we could not find grounds to accept this hypothesis. Hence, our results are in contrast to a self-report study with Australian women, where 62% of young women found themselves to be partially or fully responsible for their latest accident, whereas 52% of middle-aged women stated the same (Dobson et al. 1999).

The findings on the effect of the respondent's own mode of transport on perceived fault are congruent with those of Mayou & Bryant (2003), who found that approximately 65% of car drivers and approximately 72% of cyclists do not blame themselves for their injuries.

When examining how fault is associated with the respondent's own mode of transport as well as their counterpart's, we find that both factors influence who the respondents holds at fault for the accident. Thus, both car drivers and bicyclists are significantly more likely to find the car driver at fault in accidents between bicyclists and cars.

We have found no other studies on the connection between self-reports of fault in real accidents and the counterpart's mode of transport as comparison, but a study on car drivers' attitudes toward bicycles show that car drivers tend to agree that a collision between a car and a bicycle, in general, is the car driver's fault (Goddard 2017).

4.2 CAUSATION FACTORS

Lajunen & Summala (2003) found that self-reported drinking and driving is a factor highly susceptible for social desirability bias when testing the Driving Behaviour Questionnaire. We find that being under the influence of alcohol and other substances is self-reported in only one of the accidents. When using official Danish accident statistics as a comparison, we see a very large difference: 13.2% of all accidents with injuries in 2015 had one or more parties who were under the influence of alcohol (blood alcohol level more than 0.5‰), (Vejdirektoratet, 2016). Even when the findings in this study are compared with other self-reported data, the level is low; Meunier et al. (2018) found drugs and medication as a causation factor in 7.3 % of the accidents.

Speeding is another causation factor linked to inappropriate driving. Speeding was self-reported in only one of the accidents, while in comparison, speeding was found in 42% of the accidents investigated by a Danish in-depth analysis team (AIB 2014). Even though differences are expected between the Danish Road Traffic Accident Investigation Board's (AIB's) in-depth studies (e.g., due to the severity of the accidents investigated by the AIB) and the self-reported level of admitted speeding, the self-reported percentage is still very low. A partial explanation, besides the obvious influence of other-deception or impression management, could be that some respondents are simply not aware they were speeding. This is congruent with findings in a study that compared self-reports of speeding in traffic accidents with an in-depth analysis of the accidents. Here, it was found that 73.6% of drivers underestimated their traveling speed, which is consistent with other studies of speed estimates (Versteegh, 2004).

While both speeding and intoxication could be considered socially undesirable behavior, the respondent could also choose an option that indicated that his/her vehicle had a mechanical fault or technical error. This option is not related to the respondent's behavior or state of mind at the time of the accident and could be considered the most desirable option, as it, to some extent, partly exonerates the respondent of the responsibility for the accident. There are only self-reported vehicle faults in three (0.01%) of the accidents. As a comparison, a Danish in-depth accident analysis study found errors and faults in vehicles in 19.93 % of the accidents (AIB 2014). It is possible that respondents are simply not aware of the faults in their vehicles unless a skilled vehicle expert examines the vehicle after the accident, which would be a methodological discrepancy between the obtainment of information in the AIB's in-depth analysis and a self-report regime. However, even when compared with other self-reported data, there seems to have been relatively few self-reports of mechanical failures; Meunier et al. (2018) found car failure as a causation factor in 5.6 % of the accidents, and a study of Danish bicycle accidents found that 3% of bicycle accidents had a vehicle factor present (Janstrup et al. 2017). That so few of the participants in this study acknowledge the social desirable factor of mechanical error is surprising and illustrates that the concept of social desirability does not necessarily present itself clearly in results.

The discrepancies between the frequency of the self-reported accident causation factors and those found in other studies (either self-reported or not) highlight that the methodological differences have a huge impact on the level of measured factors and that the frequency of self-reported factors should not be taken at face value.

Causation factors differ in nature and intricacy. Whereas the previously discussed factors related to speeding, drunk driving, and vehicle faults are relatively easy to understand, assess, and compare, some are more complicated. There are no commonly used definitions of, for instance, *distraction* and *attention* with a practical application as accident causation factors; what does it mean to be attentive, and which questions should a respondent provide answers to for a researcher to conclude that he/she was lacking attention? A review of distractions related to car driving found that a distraction from within the cabin was present in approximately 30% of all accidents (Møller et al. 2010), though the definition of *distraction* is not completely comparable with the one applied in the current study. The applied measures of accident causation factors such as attention, distraction, etc. were derived through inspiration from the AIB's publications (AIB 2014) and are, to some extent, applied elsewere in project InDeV (Madsen et. al 2018). We have found no studies on the relation between accident causation factors and self-reported fault of the accident with which to compare the findings.

5. CONCLUSION

When comparing the frequency of causation factors, self-reports seem to be unreliable and affected by many possible errors (e.g., social desirability, knowledge of the accident causation factors' presence, and estimation of real driving speed). A standardization of a method for examining accident causation factors would greatly enhance the possibility of comparing results across studies. Still, the results from this paper add to the existing literature that warns of validity issues regarding self-reports (e.g., af Wåhlberg [2002] and af Wåhlberg et al. [2010]) in the context of accident causation factors, as we have found a very low agreement of the frequency of self-reported causation factors compared with other studies. However, this does not mean that self-reported accident causation factors cannot be used in traffic safety research, but that one must show careful consideration of their application. This paper illustrates a new way of utilizing self-reported accident causation factors in the calculation of the odds ratio of perceived fault, and the results from this study provide new knowledge of how accident causation factors are associated with perceived fault. This could be valuable input in, e.g., the development of safety campaigns or other traffic safety work that aims at changing road users' beliefs and attitudes toward different causation factors, as the calculation of the odds ratio provides a measurement of how strongly the respondents believe their actions are contributory to the accident. That perceived fault is found to be associated with both the respondents' own means of transport as well as their counterparts' tells of a traffic culture where the expectations of road users' behavior and responsibility is not uniform across transport modes but to some degree exonerates bicyclists of fault in their accidents.

ACKNOWLEDGEMENTS

Data for this research was gathered as part of the EU Horizon 2020 project InDeV (In-Depth understanding of accident causation for Vulnerable road users, grant agreement No. 635895).

REFERENCES

af Wåhlberg, A.E. (2002). On the validity of self-reported traffic accident data. *Manama: E140 Proceedings of the 2nd Safety On Road International Conference*. Available: http://pubcare.uu.se/digitalAssets/19/a_19175-f_af_Wahlberg_The_validity_of_self_reported_accident_data.pdf

af Wåhlberg, A. E., Dorn, L., & Kline, T. (2010). The effect of social desirability on self reported and recorded road traffic accidents. *Transportation Research Part F: Traffic Psychology and Behaviour,* 13 (2), 106–114. https://doi.org/10.1016/j.trf.2009.11.004

af Wåhlberg, A. E., & Dorn, L. (2007). Culpable versus non-culpable traffic accidents; What is wrong with this picture? *Journal of Safety Research,* 38 453–459. https://doi.org/10.1016/j.jsr.2007.01.013

af Wåhlberg, A.E. (2010) Social desirability effects in driver behavior inventories. *Journal of Safety Research,* Vol 41 (2), 99–106. https://doi.org/10.1016/j.jsr.2010.02.005

AIB – The Danish Road Traffic Accident Investigation Board (2014). *Why do road traffic accidents happen?* København: Vejdirektoratet. Available:http://www.hvu.dk/SiteCollectionDocuments/HVUdec14_UK_HvorforSkerUlykkerne.pdf

CPR-administation (2018). *Can I get a civil registration number?* Available: https://www.cpr.dk/english/moving-to-denmark/ [last visited 10-16-2018].

Christianson, S-Å & Lofthus, E. (1987). Memory for traumatic events. *Applied Cognitive Psychology*, Vol 1, 225–239.

Derriks, H. M., & Mak, P. M. (2007). *IRTAD special report – Underreporting of road traffic casualties*. The Netherlands: Ministry of Transport, Public Works and Water management. Available:https://www.who.int/roadsafety/publications/irtad_underreporting.pdf

de Winter, J. C. F. & Dodou, D. (2010). The Driver Behaviour Questionnaire as a predictor of accidents: A meta-analysis. *Journal of Safety Research*. Volume 41 (6) 463–470.
https://doi.org/10.1016/j.jsr.2010.10.007

Dobson, A., Brown, W., Ball, J., Powers, J., & McFadden, M. (1999). Women drivers' behaviour, socio-demographic characteristics and accidents. *Accident analysis and prevention*, 31(5), 525–535.
https://doi.org/10.1016/S0001-4575(99)00009-3

Félonneau, M.-L., Causse, E., Constant, A., Contrand, B., Messiah, A. , Lagarde, E. (2013). Gender stereotypes and superior conformity of the self in a sample of cyclists. *Accident Analysis & Prevention*, Volume 50, 336–340.
https://doi.org/10.1016/j.aap.2012.05.006

Gheorghiu, A. & Havârneanu, C. (2011). Driving behavior of a sample of young Romanian drivers. *Procedia – Social and behavioural Sciences*. Volume 33 page 697–701.
https://doi.org/10.1016/j.sbspro.2012.01.211

Goddard, T. B., (2017). *Drivers' Attitudes and Behaviors Toward Bicyclists: Intermodal Interactions and Implications for Road Safety*. Doctoral thesis. Portland State University, ProQuest Dissertations Publishing, 2017. ProQuest no. 10286009.
Available: https://search-proquest-com.zorac.aub.aau.dk/docview/1937577402?pq-origsite=primo

Janstrup, K. H., Mette Møller, Ninette Pilegaard 2017. Cyklistuheld – hvilken betydning har vejen, køretøjet og trafikanten. *Proceedings from the Annual Transport Conference at Aalborg University 28.-29. august 2017*. ISSN 1603–969.

Janstrup, K. H., Kaplan, S., Hels, T., Lauritsen, J., & Prato, C. G. (2016). Understanding traffic crash under-reporting: Linking police and medical records to individual and crash characteristics. *Traffic Injury Prevention*, 17(6), 580–584.
https://doi.org/10.1080/15389588.2015.1128533

Kamaluddin, N. A, C. S. Andersen, M. K. Larsen, K. R. Meltofte, A. Várhelyi (2018). Self-reporting traffic crashes – a systematic literature review. *European Transport Research Review.* 10:26.
https://doi.org/10.1186/s12544-018-0301-0

Karacasu, M. & Arzu, E. (2011). An Analysis on Distribution of Traffic Faults in Accidents, Based on Driver's Age and Gender: Eskisehir Case. *Procedia - Social and Behavioral Sciences,* Volume 20, 776–785. https://doi.org/10.1016/j.sbspro.2011.08.086

Lahrmann, H., Madsen, J. C. O., Madsen, T. K. O., Olesen, A. V., Thedchannamoorthy, S., & Bloch, A-G. (2014). *Projekt Cykeljakke: Den sikkerhedsmæssige effekt af en gul cykeljakke.* Aalborg: Aalborg Universitet. Institut for Byggeri og Anlæg. (DCE Technical Reports; Nr. 167). Available: http://vbn.aau.dk/files/197570851/Projekt_cykeljakken_den_sikkerhedsm_ssige_effekt_af_en_gul_cykeljakke.pdf

Lajunen, T. & Summala, H. (2003). Can we trust self-reports of driving? Effects of impression management on driver behaviour questionnaire responses. *Transportation Research Part F: Traffic Psycology and Behaviour* , vol 6 (2) , 97–107. https://doi.org/10.1016/S1369-8478(03)00008-1

Lajunen, T. & Özkan, T (2011) *Chapter 4 - Self-Report Instruments and Methods*. In: Handbook of traffic psychology. Academic Press, San Diego, pp 43–59.

Littleton, S.M., et al. (2012). The influence of fault on health in the immediate post-crash period following road traffic crashes. *Injury,* Volume 43, Issue 9, 1586–1592. https://doi.org/10.1016/j.injury.2011.03.002

Madsen, T. K. O., Várhelyi, A., Polders, E., Reumers, S., Hosta, P., Bibiloni, D. J., Lahrmann, H. S. (2018). *Assessment of Safety of VRUs Based on Self-Reporting of Accidents and Near-Accidents.* Brussels: Publications Office of the European Union. Available: https://www.indev-project.eu/InDeV/EN/Documents/pdf/3-2.pdf?__blob=publicationFile&v=2

Mayou, R., & Bryant, B. (2003). Consequences of road traffic accidents for different types of road user. *Injury*, 34:3, pp.197–202. https://doi.org/10.1016/S0020-1383(02)00285-1

Meunier, J.-C., Dupont, E., Mersch, J. & Van den Berghe, W. (2018). *My Life After the Crash - MyLAC: An international study on medical, psycho-social and economic consequences of road injuries.* Brussels, Belgium: Vias institute. Available: https://fevr.org/wp-content/uploads/2018/05/DEF_MyLACReport.pdf

Møller, K. M., Andersen, C. S., Verhelyi, A., Schönebeck, S., Reumers, S., Hosta, P., & Szagala, P. (2017). *Accident information from six european countries based on self-reports.* Bruxelles: Publications Office of the European Union. Available: http://vbn.aau.dk/files/262153504/Accident_Information_from_six_European_Countries_Based_on_Self_reports.pdf

Møller, M., Troglauer, T., & Hels, T. (2010). *Distraktion i forbindelse med bilkørsel.* Kgs. Lyngby: Technical University of Denmark, Transport. DTU Transports rapportserie, Nr. 2010/. Available:http://orbit.dtu.dk/files/5286069/distraktorrapport_10-02-2010.pdf

Paulhus, D. & Reid D. B. (1991). Enhancement and denial in Socially Desirable Responding. *Journal of Personality and Social Psychology*. Vol 60, No 2, 307–317. doi:10.1037/0022-3514.60.2.307

Paulhus, D. (1984). Two-component models of socially desirable responding. *Journal of Personality and Social Psychology*, Vol 46, No 3, 598–609. doi: 10.1037/0022-3514.46.3.598

Preston, C. E. & Harris, S. (1965). Psychology of drivers in traffic accidents. *Journal of Applied Psychology,* 49(4), 284–288. doi:10.1037/h0022453

Siren, A. & Meng, A. (2013). Older drivers' self-assessed driving skills, driving-related stress and self-regulation on traffic. *Transportation Research Part F: Traffic Psycology and Behaviour* (2013). Vol 17, pages 88–97. https://doi.org/10.1016/j.trf.2012.10.004

Tivesten, E. & Wiberg, H. (2013). What can the drivers' own description from combined sources provide in an analysis of driver distraction and low vigilance in accident situations? *Accident Analysis and Prevention,* (52), 51–63. https://doi.org/10.1016/j.aap.2012.12.016

Vejdirektoratet. (2016). *Uheldsstatistik. Året 2015.* Copenhagen, Denmark: Vejdirektoratet. Available: www.vejdirektoratet.dk/DA/viden_og_data/publikationer/Lists/Publikationer/Attachments/909/Uheldsstatistik%20året%202015.pdf

Versteegh, S. (2004). The accuracy of driver accounts of vehicle accidents. *2004 Road Safety Research, Policing and Education Conference. Session: Crash and Injury Analysis.* Perth, Australia. Available: https://core.ac.uk/download/pdf/12769779.pdf

APPENDIX

Table A1: Key sample variables concerning respondents and accident data, N= 300.

Variable	Description or phrasing of question	N	%	95% CI	
Gender					
	If respondent is male = 1, female =0	156	0.52	0.4631	0.5769
Age					
18–29	If 18–29 years =1, otherwise = 0	41	0.14	0.0976	0.1758
30–39	If 30–39 years =1, otherwise = 0	61	0.20	0.1575	0.2491
40–49	If 40–49 years =1, otherwise = 0	66	0.22	0.1729	0.2671
50–59	If 50–59 years =1, otherwise = 0	75	0.25	0.2007	0.2993
60–69	If 60–69 years =1, otherwise = 0	42	0.14	0.1005	0.1795
Above 70	If >70 years =1, otherwise = 0	15	0.05	0.0252	0.0748
Means of transport					
Bicycle	If respondent is driving a bicycle = 1, otherwise = 0	71	0.24	0.1883	0.2850
Car	If respondent is driving a car = 1, otherwise = 0	213	0.71	0.6584	0.7616
Driver of truck, van, or bus	If respondent is driving a truck, van, or bus = 1, otherwise = 0	5	0.02	0.0021	0.0312
Pedestrian	If respondent is a pedestrian = 1, otherwise = 0	5	0.02	0.0021	0.0312
Motorcycle	If respondent is driving a motorcycle = 1, otherwise = 0	3	0.01	-0.0013	0.0213
Rollerblades/ Skateboard	If respondent is using rollerblades or a skateboard= 1, otherwise = 0	1	0.00	-0.0032	0.0099
Moped	If respondent is driving a moped = 1, otherwise = 0	2	0.01	-0.0026	0.0159
Accident causation factors					
Fatigue	If yes to fatigue var. = 1, otherwise =0	13	0.04	0.0202	0.0665
var. 1: Fatigue	*I was tired* =1, otherwise =0	13	0.04	0.0202	0.0665
Attention	If yes to one or more of the attention var. = 1, otherwise =0	110	0.37	0.3118	0.4215
var. 1: Attention	*I was immersed in thoughts* =1, otherwise =0	19	0.06	0.0356	0.0911
var. 2: Attention	*I was not aware of the other road user* =1, otherwise =0	94	0.31	0.2605	0.3661
Interpretation and assessment	If yes to the interpretation and assessment var. = 1, otherwise =0	123	0.41	0.3540	0.4660
var. 1: Interpretation and assessment	*I thought the other road user was aware of my presence*=1, otherwise =0	123	0.41	0.3540	0.4660

Inappropriate driving behavior	If yes to one or more of the inappropriate driving var. = 1, otherwise =0	120	0.04	0.0177	0.0623
var. 1: Inappropriate driving behavior	*My speed exceeded the speed limit at the location of the accident* =1, otherwise =0	1	0.00	-0.0032	0.0099
var. 2: Inappropriate driving behavior	*I disregarded another road user's priority* =1, otherwise =0	11	0.04	0.0153	0.0581
var. 3: Inappropriate driving behavior	*I was affected by alcohol or other intoxicating substances (e.g., medication or drugs)* =1, otherwise =0	1	0.00	-0.0032	0.0099
Orientation and distraction	If yes to one or more of the orientation and distraction var. = 1, otherwise =0	36	0.12	0.0830	0.1570
var.1: Orientation and distraction	*I was distracted by something (e.g., another road user's behavior, a sign, or a shop*) =1, otherwise =0	30	0.10	0.0659	0.1341
var. 2: Orientation and distraction	*I was looking at my radio or navigation system* =1, otherwise =0	1	0.00	-0.0032	0.0099
var. 3: Orientation and distraction	*I was looking at my cell phone* =1, otherwise =0	5	0.02	0.0021	0.0312
Vehicle errors	If yes to the vehicle error var. = 1, otherwise =0	3	0.01	-0.0013	0.0213
var. 1: Vehicle error	*There was a mechanical error/technical fault in my means of transport* =1, otherwise =0	3	0.01	-0.0013	0.0213
Stress	If yes to stress var. = 1, otherwise =0	39	0.13	0.0917	0.1683
var. 1: Stress	*I was in a hurry* =1, otherwise =0	39	0.13	0.0917	0.1683
Auditory tasks	If yes to one or more of the auditory var. = 1, otherwise =0	30	0.10	0.0659	0.1341
var. 1: Auditory task	*I was listening to* music =1, otherwise =0	14	0.05	0.0227	0.0707
var. 2: Auditory task	*I was talking to someone (this includes use of cell phones)* =1, otherwise =0	18	0.06	0.0330	0.0870
Illness	If yes to illness var. = 1, otherwise =0	4	0.01	0.0003	0.0264
var. 1: Illness	*I was ill/unwell* =1, otherwise =0	4	0.01	0.0003	0.0264
No causation factor	*None of the above* =1, otherwise =0	196	0.65	0.5992	0.7075

Table A2: Gender and perceived fault

	Partly or wholly my own fault			The counterpart's fault			Total
	No.	%	95% CI	No.	%	95% CI	
Male	51	32.69	25.33–40.05	105	67.31	59.95–74.67	156
Female	47	32.64	24.98–40.30	97	67.36	59.70–75.02	144

Table A3: Age and perceived fault

	Partly or wholly my own fault			The counterpart's fault			Total
Age	N	%	95% CI	N	%	95% CI	N
18–29	12	29.27	15.34–43.20	29	70.73	56.80–84.66	41
30–39	17	27.87	16.62–39.12	44	72.13	60.88–83.38	61
40–49	18	27.27	16.53–38.02	48	72.73	61.99–83.47	66
50–59	26	34.67	23.90–45.44	49	65.33	54.56–76.10	75
60–69	16	38.10	23.41–52.79	26	61.90	47.21–76.59	42
70–79	9	64.29	39.19–89.39	5	35.71	10.61–60.81	14
80–89	0	0		1	100		1

APPENDIX E INTERVIEW GUIDE

INTERVIEWS:
GUIDELINES AND EXECUTION

Thirty-five respondents were invited via regular post for participation in a follow-up interview after their completion of the questionnaire from the Emergency Room dataset. These respondents were selected among the 411 respondents as they:

A. Had stated that they had been in contact with the police due to the accident
B. Had given permission for their answers to the self-report questionnaire to be combined with other databases
C. Resided in the area of Aalborg (postal codes 9000-9270)

Interest in interview participation could be stated by either email or telephone. Ten respondents acknowledged wanted participation, but only eight interviews were conducted and recorded: It was not possible to find a suitable date for one potential interviewee and one interview was held but the voice recording was corrupted and therefore not applicable in this research.

Aim

There were several aims with the interviews:

- To investigate if questionnaire answers are congruent with interview answers. If discrepancy exist, then to gauge which is most coherent or otherwise true.
- To ascertain if the interviewee had an accident that should be reported by the police
- To investigate if interview answers are congruent with the police record of the same accident. If discrepancies exist, then to gauge if the interviewees' story of their accident seem coherent and otherwise true
- To investigate motivation for participation in the self-report regime
- To learn of any respondents' difficulties in using google maps to locate their accident or in describing the maneuvers right before the accident
- To investigate interviewees views of the questions asked in the questionnaire about fault and accident causation factors; are questions considered provocative and do interviewees believe to answer truthfully

The interviews are conducted as semi-structured. This is construed as an interview in which interviewees' own storytelling of their accident presents the main structure of conversation and questions are asked to expand on what is offered by the interviewee or, if necessary, to change the subject to something more related to the stated research goals. This entails that even though an interview guideline has been constructed, it is only regarded as possible sketch for conversation. Thus, the topics might be addressed in another order if the story from the interviewee evolves in another direction, and any other information the interviewee provides during the interview is explored further if it seems of interest and relevance. The questions in the interview guidelines are not meant to be verbatim reproduced, instead attempts at a natural conversation in everyday language, though not on everyday topics, is made.

One of the aims is to discover any discrepancies between the verbal description of the accident and the one provided via the questionnaire answers. As the element of natural conversation is destroyed if the interviewer consistently looks in notes to compare with questionnaire answers, it is needed to be well-versed in the interviewee's questionnaire answers beforehand. If a discrepancy is discovered during the interview, this is explored further: The interviewee is asked if he/she can expand, if he/she might have chosen something else in the questionnaire, if the interviewer have misunderstood etc. In this situation, it is considered important to keep interviewees from getting the impression that they are saying anything wrong, but instead simply that the interviewer needs to understand it thoroughly.

In some aspects, the interviews have begun before the voice recorder is turned on. As all interviewees are greeted at the main entrance and lead to the room reserved for the interviews, there is a short walk in which interaction between interviewer and interviewee begins. The semi-structured nature of the interview benefits from interviewees, who themselves offer stories from their accident and are willing to share their experience and views. Therefore, an informal setting is attempted in order to help interviewees relax and talk freely. The walk to the meeting room is thus utilized for small talk; i.e., thanking them for participation, asking about any trouble finding the building, offering own experience of wayfinding on campus etc. When entering the meeting room, interviewees are offered coffee/tea/water, which is already placed on the table. They are asked, if it is all right to turn on the voice recorder, and interviewer explains that everything is anonymous and what this entails. Interviewees are asked permission for use of their interview in writing (without their name, but direct quotes).

The recorded part of the interviews are begun with an informal introduction of the interviewer. A brush-up of the scope of the interview offered: To investigate what we can learn from their accident by interview compared to questionnaires and other data sources and to evaluate of some of the questionnaire items could improve further.

Execution

All interviews were held at varying meeting rooms at Thomas Manns Vej 23, Department for Civil Engineering, Aalborg University. Interviews lasted from ca. 30 minutes to 1hour.

Before the interview, the following were prepared:

- The interviewees answers to the self-report questionnaire was retrieved, printed and memorized
- The accident situation that matches the one chosen by the interviewee in the self-report questionnaire (if any) was printed along any other situations in the same main category
- A map (with aerial photo) of the location, that interviewee had pinpointed in the questionnaire, was printed

The following was brought to the interview:

- Voice recorder and extra batteries
- Printout of map, accident location and questionnaire answers
- Printout of corresponding accident situation and other situations within the same main category
- This sheet of guidelines
- Blank paper and two pens
- Coffee/tea/water for interviewee and interviewer

Interview guidelines

Introduction

- First, we will talk about what happened in your accident – I have brought some maps and papers that we might look at. We will also talk a bit of some of the questions and your answers in the questionnaire. And lastly we'll talk about how you understand some of the questions, we asked you, what motivated you to participate in the survey and if anything felt odd or difficult to answer.

The story of the accident *(as triggering of recall and prompting storytelling)*

- I will ask you to remember the day the accident happened. Can you give me a short description of your accident?

Location of accident

- Could you express with words, as precisely as possible, where the accident happened?
- I have brought a map of the location you pinpointed back when you answered the questionnaire – does it more or less consist with what you just described? (If no, I can just turn on my computer and we will find the correct location together).
- Could you try to *(with the pen and the paper map provided)* to put an X as closely possible to where you collided/hit/fell/…?

- Now let us try to make a small sketch of the accident *(blank paper and pen is provided)*. I have taken some illustrations with me where you can see an example. It could be a good idea to start by drawing how the road/intersection was where your accident happened. Then you can try to show with arrows where you were driving to and from – and likewise for your counterpart. Then mark with an X where you collided. If you know any of the street names, or can place north or south, it would be great.
- Could you try to describe the drawing you have just made?
- *This theme is only abandoned when interviewer have enough information to correctly establish an accident situation according to police recording practice.*

Accident details

- What happened right before the accident? Where were you going to and from? Why? What did you think of before the accident happened?
- How was the lighting?
- How was the weather?
- How was the surface?

Accident causation

- What do you think was the reason for the accident?
- I have noted that you in your questionnaire stated that causation factors xxx and xxx was present…do you think the accident could have been avoided if some of these circumstances had not been present?

Contact to authorities

- Whom of these have you had contact to *(own GP, Emergency Doctor Service, police, insurance, road authority)?* Can you describe what your contact consisted of? Why did you contact them?
- Do you know if there is filed a police report?

Questionnaire and participation

- We asked you in the questionnaire to state whom you thought was primarily at fault in the accident – the options were *you*, *your counterpart* or *equally at fault*. Do you think that this is a difficult question to answer? Is it a provocative or imprudent question? Do you think you would answer it correctly if you were at fault?
- In the questionnaire you were asked to locate your accident on google maps – do you remember the experience? If so, did you find it easy? How do you in general experience locating places on a map; what can cause troubles? Are you familiar with google maps? Do you normally find it easy to use?

- Why did you want to participate in the self-report regime? Did you consider saying no – if so, why?
- If we need anything further from you, would it be all right if we contact you again? If so, how do you prefer?

SUMMARY

The application of self-reports of traffic accidents is increasing within traffic safety work and research, yet many of the methodological aspects related to self-reporting are unknown. This thesis examines the correctness of self-reported information and its congruency with police reports of the same traffic accidents. The effect of various sampling methods and survey designs on e.g. response rates and the demographic composition of respondents is furthermore investigated. The thesis also presents ideas for the possible application of the self-report methodology within the area of traffic safety as well as highlighting fields of application where self-reports should not be deployed.

ISSN (online): 2446-1636
ISBN (online): 978-87-7210-497-3

AALBORG UNIVERSITETSFORLAG