UDRIVE

European Naturalistic Driving Study

The potential of naturalistic driving studies with simple Data Acquisition Systems (DAS) for monitoring driver behaviour

Deliverable 52.1

DOI: 10.26323/UDRIVE_D52.1

UDRIVE

European Naturalistic Driving Study

EUROPEAN COMMISSION

SEVENTH FRAMEWORK PROGRAMME

FP7-SST-2012.4.1-3

GA No. 314050

eUropean naturalistic Driving and Riding for Infrastructure and Vehicle safety and Environment

Deliverable No.	UDRIVE D52.1	
Deliverable Title	The potential of Naturalistic Driving studies with simple Data Acquisition Systemss (DAS) for monitoring driver behaviour	
Dissemination level	Public	
Written By	Tsippy Lotan (Or Yarok) Nicole van-Nes (SWOV) Gila Albert (Or Yarok) Oren Musicant (Or Yarok) Steve Reeds (Loughborough University) Ruth Welsh (Loughborough University)	30-04-2017
Checked by	Marta Peireira Cocron (TUC)	15-05-207
Approved by	Marika Hoedemaeker (TNO) *QA manager* Nicole van Nes (SWOV) *Project Coördinator*	29-05-2017 06-06-2017
Status	Final	06-06-2017

Please refer to this document as:

Lotan, T., Van Nes, N., Albert, G., Musicant, O., Reeds, S., & Welsh, R. (2017). *The potential of Naturalistic Driving studies with simple Data Acquisition Systems (DAS) for monitoring driver behaviour.* UDRIVE Deliverable 52.1. EU FP7 Project UDRIVE Consortium. https://doi.org/10.26323/UDRIVE_D52.1

Acknowledgement:

The authors would like to thank Einat Grimberg for helping in finalizing the deliverable and Marta Sofia Pereira Cocron (TUC) for performing the quality assurance on earlier and the final draft

Disclaimer:

This project has received funding from the European Union's Seventh Framework Programme for research, technological development and demonstration under grant agreement no 314050.

Executive Summary

This report addresses the important question regarding the potential of simple and low-cost technologies to address research questions such as the ones dealt with in UDrive.

The resources and efforts associated with big naturalistic studies, such as the American SHRP II and the European UDrive, are tremendous and can not be repeated and supported frequently, or even more than once in a decade (or a life time..). Naturally, the wealth and richness of the integrated data, gathered by such substantial studies and elaborated DAS, cannot be compared to data collected via simpler, nomadic data collection technologies. The question that needs to be asked is how many Research Questions (RQs) can be addressed, at least to some extent, by other low-cost and simple technologies? This discussion is important, not only in order to replace the honourable place (and cost!) of naturalistic studies, but also to complement and enable their continuity after their completion.

Technology is rapidly evolving and almost any attempt to provide a comprehensive and complete state of the art of existing technologies (as well as their features and cost) is doomed to fail. Hence, in chapter 1 of this report, we have created a framework for presentation, on which the various important parameters associated with the question at hand, are illustrated, positioned and discussed. This framework is denoted by “Framework for Naturalistic Studies” (FNS) and serves as the back bone of this report. The framework is a conceptual framework and hence, is flexible in the sense that its dimensions, categories and presentation mode are not rigid and can be adjusted to new features and new technologies as they become available. The framework is gradually built using two main dimensions: data collection technology type and sample size. The categories and features of the main dimensions are not rigidly fixed, and their values can be ordinal, quantitative or qualitative. When referring to parameters that are not numerical – even the order relation among categories is not always clear. In this way – the FNS can be, at times, viewed as a matrix rather than a figure with order relation among categories presented along its axes.

On the two main dimensions of the FNS – data collection technology type and sample size – other dimensions are incorporated. These dimensions include: cost, data access, specific technologies and research questions that can be addressed by the various technologies. These other dimensions are mapped and positioned in the plot area of the FNS. Other presentations, in which the axes and the plot area are interchanged, or 3-dimensional presentations are performed, can be incorporated to highlight specific angles of the involved dimensions.

The various technologies for data collection were mapped on the FNS. The technology groups include: mobile phone location services, mobile phone applications, telematics devices, built-in data loggers, dash cameras and enhanced dash cameras, wearable technologies, compound systems, eye trackers and Mobileyetype technologies.

After this detailed illustrations of analyses that can be conducted using simple low-cost technologies are described. It is demonstrated how temporal and spatial analysis can reveal important aspects on the behavioural patterns of risky drivers. Also one stand alone smartphone app can be used to monitor and evaluate smartphone usage while driving.

Most of the simple systems relate to specific behaviour that is monitored (i.e. speeding, lane keeping etc.). Additionally, certain thresholds or triggers are used to single out risky situations, which are related to that behaviour. However, once those instances are detected, no information on the circumstances leading or accompanying this behaviour are available. Typically, visual information (discrete or preferably continuous) is needed in order to fully understand the circumstances. Hence, upgrading simple (single-task oriented) technologies by other technologies (most typically by cameras), can significantly improve researchers' ability to obtain information on the circumstances, which accompany the detected risky behaviour.

One of the most conceptually straightforward integrated systems is a system, for which the basic technology detects the desired behaviour (e.g. harsh braking) and triggers a simple continuous dashboard camera to

save the relevant information, which occurs together with that behaviour. Many RQs can be addressed using this type of combined systems.

THE POTENTIAL OF SIMPLE DATA ACQUISITION SYSTEMS TO MONITOR DRIVING BEHAVIOUR

Table of Content

UDRIVE
European Naturalistic
Driving Study

Introduction

This report identifies and illustrates a range of in-vehicle data collection devices that could be used for future naturalistic studies. Recent large scale Naturalistic Driving studies such as UDRIVE and SHRP2 have used complex, tailor made and expensive data acquisition devices that require a large amount of technical input and data storage capability. In the meantime, technological developments have continued. This report explores data collection technologies currently available on the market and examines their potential future use in forthcomingNaturalistic Driving studies.

As technological breakthroughs are happening constantly, the capability to increase the extent and quality of data (collection) becomes more evident. This is best illustrated with mobile phones. These devices used to be relatively crude and although cameras were integrated into the phones, the resolution, quality and quantity of recordable datawas significantly limited and of low quality. With the advancement of smart phones, the ability for these devices to record large quantities of high definition video (from multiple cameras) has drastically increased.

The goal of the present study was to provide insight into the possibilities of simple and low-cost technologies and discuss their ability to provide high quality data, which was previously the preserve of expensive, bespoke devices. The study identifies new technologies, whichcould be implemented in future data collection techniques, in order to gain further insight into certain driver behaviours. This is important in order to improve the safety and sustainability of our traffic system.

Initially , naturalistic driving studies have been conducted with 4 wheeled road vehicles such as cars and trucks. As research priorities changed, the need for data from other modes of transportation such as cycling, motorcycling and even walking became more salient. In addition, other participants related metrics are increasingly more available, such as the recording of sleep patterns or monitoring heart rates.

The current report includes six chapters. The first chapter introduces the conceptual Framework for Naturalistic Studies (FNS) for the presentation of the various parameters associated with the technologies, which are extensively used throughout this report. The framework is used to illustrate the various parameters and their features, which aree associated with naturalistic studies.

The second chapter outlines the variety of technologies, which are currently available. Each ‘group’ of data acquisition device is discussed individually and related to one another based on the Framework for Naturalistic Studies. The various technologies are graphically presented on the Framework, where each data acquisition group is positioned in relation to data complexity, cost of the technology and the potential sample size - in case it was used in trials. The technologies included in this chapter are not complete, due to the range of available options and rapid pace of the technological changes. However, the included information could make part of the decision making process for researchers and practitioners. A summary of the available data acquisition technologies is presented in tabular form in Appendix A.

Chapters 3 and 4 depict two detailed examples of how simple and low-cost technologies can be used to collect meaningful data and perform analysis of specific research questions. Chapter 3 refers to data collected via two technologies: an in-vehicle-data-recorder (IVDR) and the Mobileye system. Additionally, chapter 3 presents a detailed description of potential analyses (both temporal and spatial), which were performed using this data. In chapter 4, a demonstration of the potential of a smartphone application to collect meaningful data for secondary task measurement and evaluation is presented.

Chapter 5 tries to answer the following valuable question: what research questions (RQs) can be addressed by simple and low-cost technologies, such as the ones that are surveyed in chapter 2 The RQs in UDrive are grouped into: Risky behaviour (as portrayed through speeding, distance and lane keeping and aggressive driving), engagement in secondary tasks, interactions between drivers and VRUs, risky behaviour related to Power Two Wheelers (PTW) and ECO driving. The various RQs along with their variations are mapped on the suggested framework according to technology group type, sample size, cost and data access. Cases, in which a combination of technologies is needed to address specific RQs are also explored and discussed.

Finally, chapter 6 discusses the value of simple low-cost technologies for future Naturalistic Driving studies in relation to the complex, tailor made and expensive data acquisition devices that have been used in recent large scale studies.

1 Framework for Naturalistic Studies

Data collection technologies can be presented according to various parameters in many ways. In this chapter, we adopt a unique graphical representation that is used throughout the report. The framework is built on two major dimensions: data collection technology and sample size.

1.1 Data collection technology – the Y axis

The Y axis illustrates broad groupings of data collection technologies; the scale does not have units but could be considered to represent the expected data level or the complexity of the device being used. Figure 1-1 presents this concept. The technologies are not explicitly listed in the figure, instead a short text description is used to outline some of the basic features of this type of data collection method. As an example, the scale can cover low data levels near the origin such as simple GPS based telematics boxes right up to bespoke, advanced systems designed and built for a particular study.

The scale used throughout this document is split into three broadly sections which are bounded by two 'break points', these are further described in the break point section.

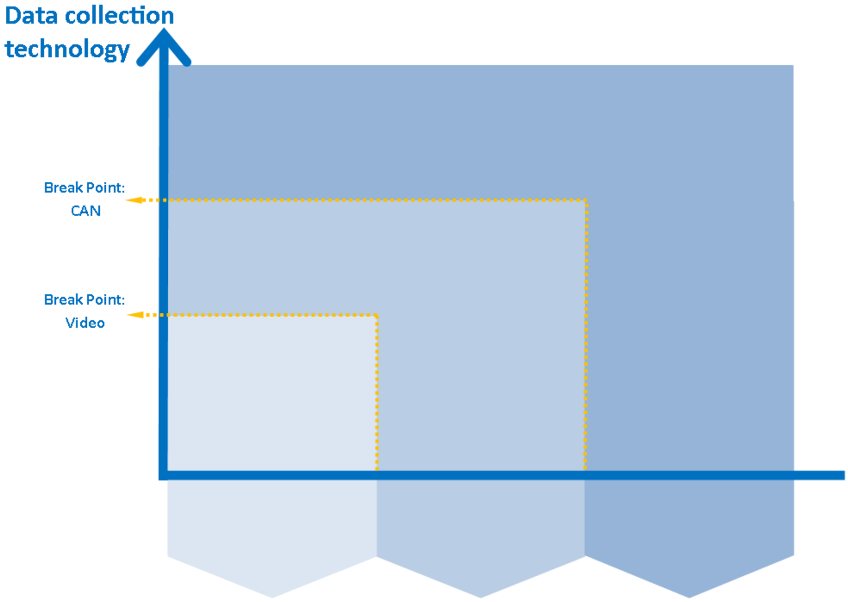


Figure 1-1: Framework for Naturalistic Studies (FNS)– *the Y axis*

1.2 Sample size – the X axis

The scale of the X axis, as presented in figure 2, relates to sample size. This scale does not have units and is based on the expected size of the study for which the technology is currently best suited or where it is typically used. Because of the lack of scale this axis can be assumed to mean a number of different things; for example, sample size in terms of the number of data acquisitions units needed or the total number of participants involved or even the expected vehicle mileage or vehicle years to be recorded.

The scale is roughly split into four groups consisting of small, medium, large and extra-large categories. Under these headings is an example of the sample size in terms of the expected mileage and/or vehicle/participant fleet. These numbers should not be taken as a definite, it would be possible, for example, to run a very small scale study using technology that might be best suited for large sample sizes – this decision should remain the prerogative of the researcher and the needs defined in the study deign.

Note that the Extra-Large sample size is defined to include more than 10,000 participants corresponding to possibly large portions of the population. Note also that the scale goes from large to small, unlike usual ordinal scales. This is used to represent a conceptual framework in which quantity and complexity play interchangeable roles, and there is a tradeoff between quantity and complexity (typically very large quantities correspond to simple technologies, while small quantities can be associated with very complex systems).

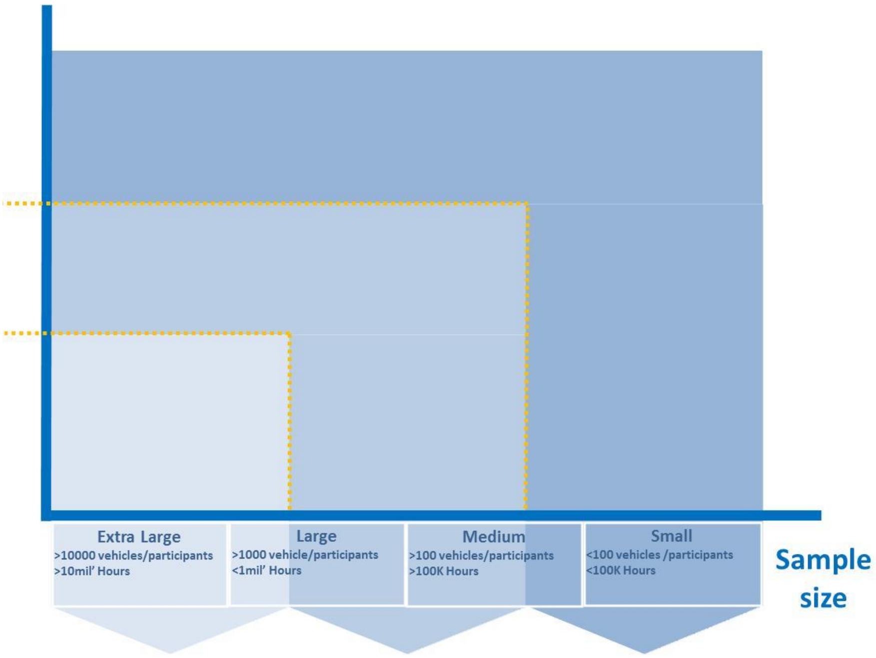


Figure 1-2: Framework for Naturalistic Studies (FNS)– the X axis

1.3 The plot area – costs and technologies

The plot area contains two main forms of information as illustrated in figure 3. The first and simplest is the trend line broadly titled 'costs'. This line is designed to illustrate that there is a difference between a small scale study using low-cost equipment and a large scale trial using more complex equipment – it is not designed to provide explicit costs on a particular device; this information is available in the table of technologies in Appendix A.

Within the plot area the technology groups will be plotted. These plots will take the form of a circular or ellipsoid area and are again designed to broadly illustrate where a particular group lies in terms of suitability to technological requirement, sample size and cost. It is likely that some technology groups will be quite large and will cover a number of different categories; this will be shown where the plotted area covers one or more regions in the plot area. It is also likely that areas of different technology groups will overlap.

As an example, illustrated in Figure 1-3 are telematics devices which are simple, low-cost and can be present at large quantities of participants. On the other end of the scale, dedicated data acquisition systems (such as the DAS used in UDrive) are expensive, sophisticated and can be implemented on a relatively small number of participants. Recall that sample size on the X-axis goes from large to small.

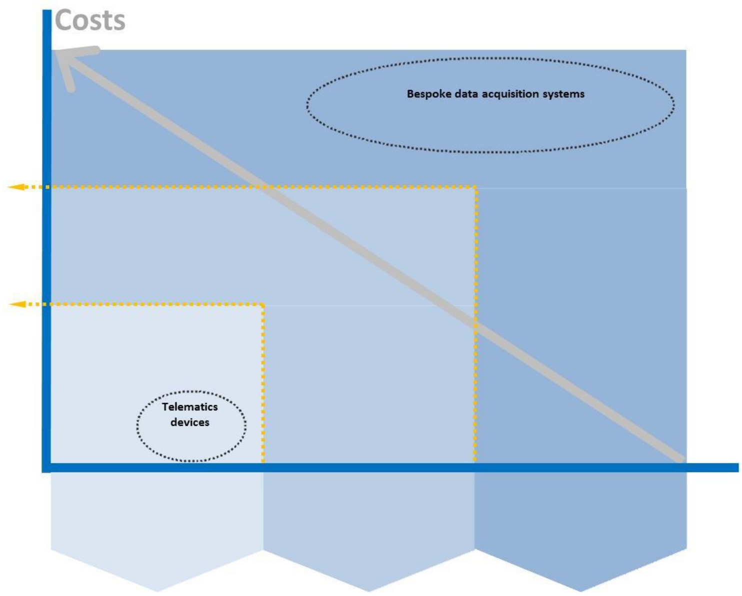


Figure 1-3: ***Plot area***

1.4 Break points

Break points are provided in the plot area to broadly differentiate between the major groups of data collection technologies. These break points are not based on exact figures of cost or sample size but they do represent the major technological advancements which separate different groups of equipment. For example, very simple telematics data loggers occupying the very cheapest and simplest technologies are separated from the next major group by their lack of video capture. Similarly devices that record video and basic driving data are normally separated from the next major group by their lack of detailed Controller Area Network (CAN) access.

Some devices do transcend these break points and will be illustrated on the plot area as such; however in general these break points describe data acquisition technologies as they exist currently. Although not represented in the figures, it is worth noting that the break points only concern the collection of data. Of course, collection is only half of the story in terms of a driving trial as it does not consider any storage or analysis phase, the figure would appear differently if the analysis of the data was to be considered. In this case the break points would most likely be reversed i.e. for analysis, video is a much more demanding medium when compared to CAN data.

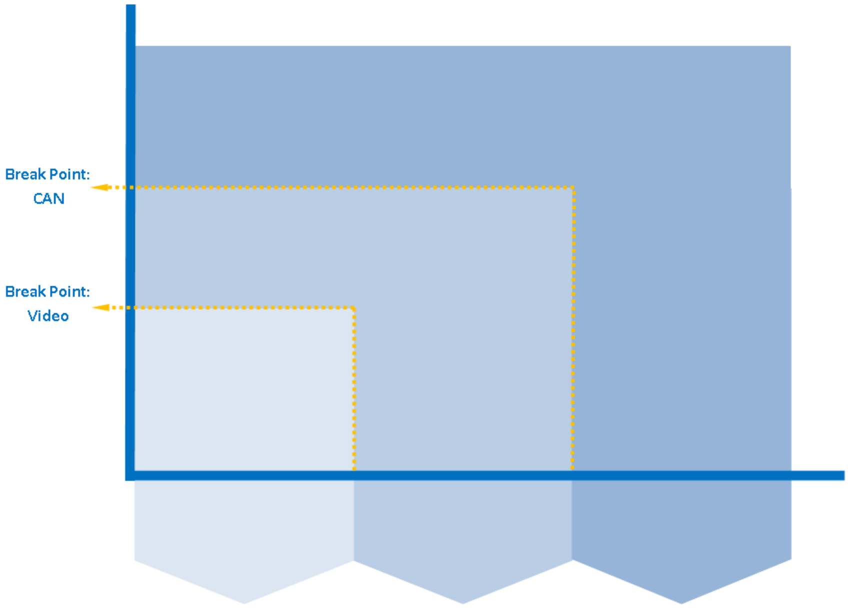


Figure 1-4: ***break points on the plot area***

Above these two general break points lies a wider range of different categories of data collection level, for example the introduction of 'event based' recording and the inclusion of Radar or distance measurement. For simplicity these have been removed from most of the figures in this report but are included in Figure 1-5 below and in the final figure in chapter 2 (Figure 2-23). As can be seen from Figure 1-5 , the break points discussed earlier do not map exactly over the 4 categories along the Y axis; It is not possible to be specific on the data capture available for each technology so these two categories should be used as guidance as to what *might* be expected but not on what *will* be expected.

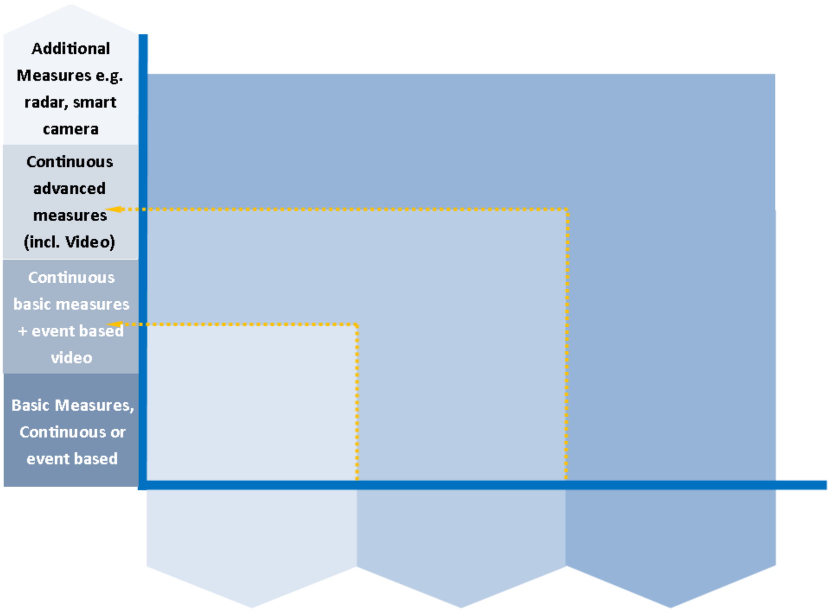


Figure 1-5*: Enhanced technology level and break points on the Y axis*

1.5 Data Access

The final piece of information that can be included in the figure is data access. Data access can be viewed on several dimensions and several difficulties related to data access could arise. Here we refer to two dimensions of data access: permission from the user to access the data and complexity of accessing the data. The first dimension refers to the setting in which data is collected. Looking at the X axis, it can be seen that for technology residing near the origin of the axis, access to the data is not always assured. This lack of access could be for a number of reasons (using third party devices or web-based servers) which will be outlined where necessary in chapter 2. As a general rule of thumb it can be said that the smaller the sample size – the more likely it is that the scheme is of a dedicated data collection type and hence access to the data is guaranteed. One exception to this is in access to CAN data or OEM supplied crash recorders/black boxes. In these cases it is likely that the data will be commercially sensitive to the manufacturer and access will likely be restricted, particularly in its raw form, to analysts. This presentation also highlights the fact that nowadays a lot of data is being collected, however access to large quantities of it is not possible due to privacy protection and lack of integrated efforts to access it.

The second dimension of data access refers to the effort needed to actually access the data and make it available for analysis. It can be seen that as technology level increases in its complexity (the Y axis) - it is more likely to be difficult to access it and make it workable for analysis. Consequently, accessing video data requires access to secured storage, and dedicated algorithms for image and pattern recognition. Similarly, access to CAN data requires agreement with manufacturer's restrictions and protocols.

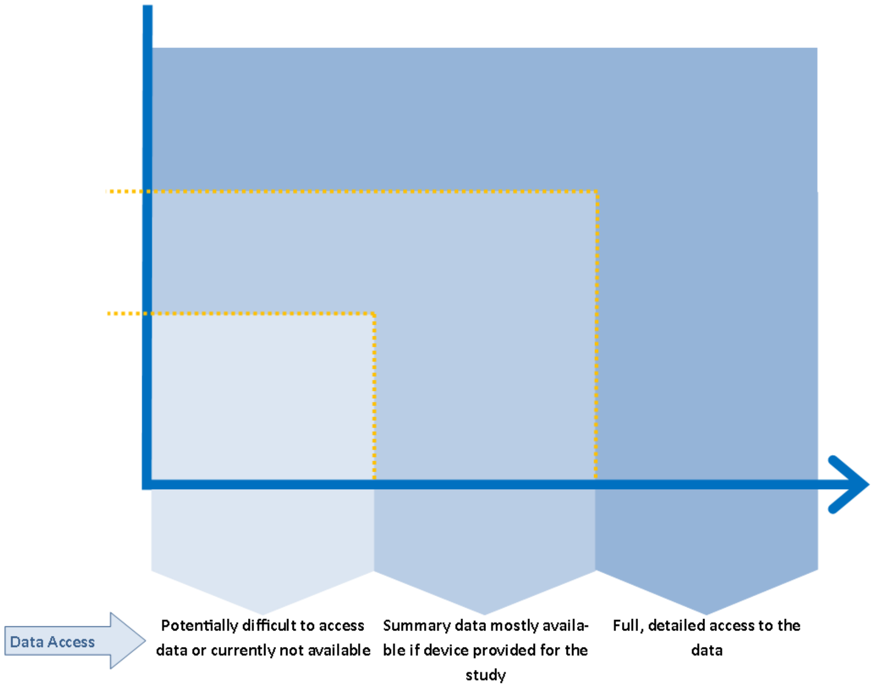
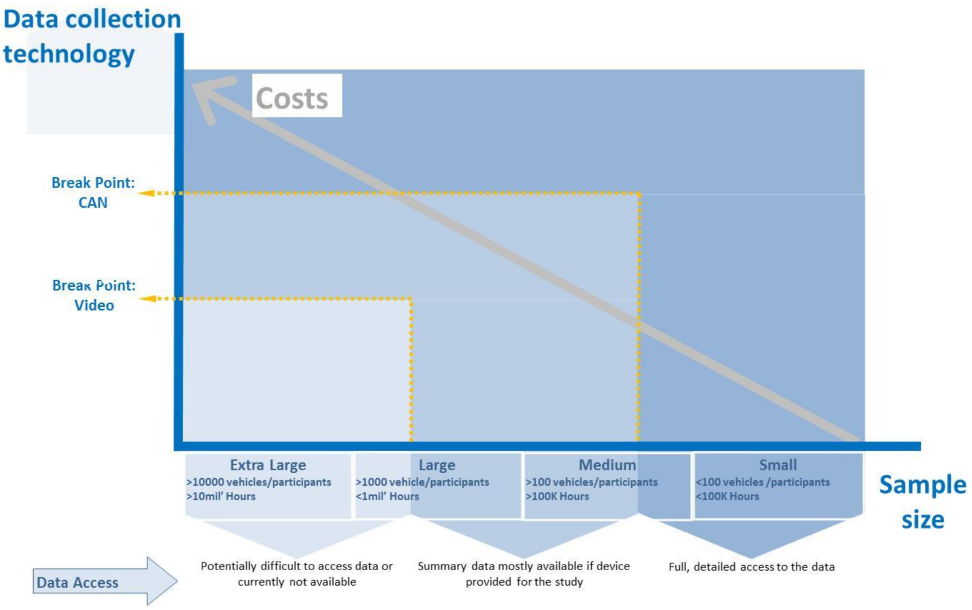
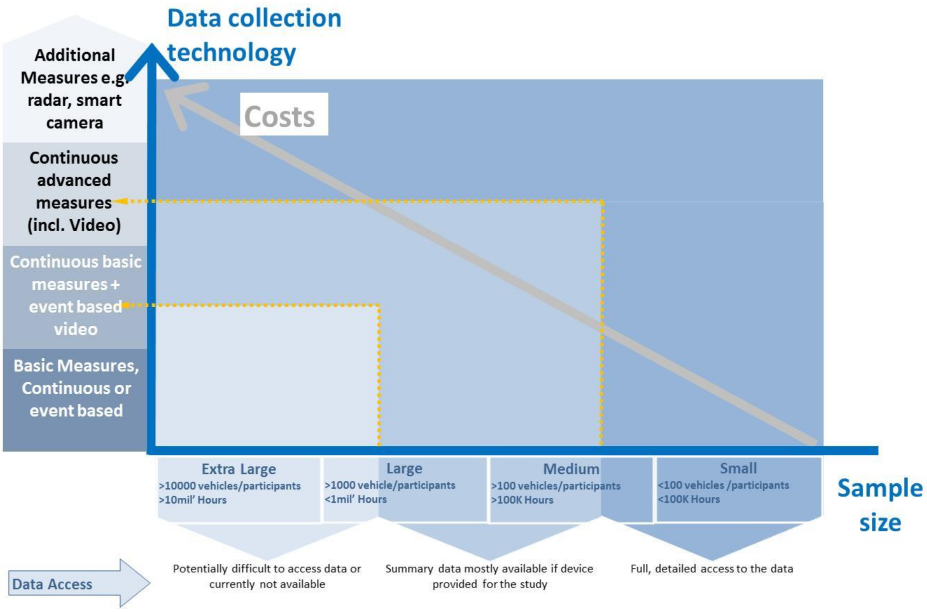


Figure 1-6: ***Data access level on X axis***

The overall framework is presented in Figure 1-7 in two forms: (a) with the Y-axis corresponding to the CAN data and video data categories, and (b) with the Y-axis corresponding to levels of enhanced technologies. We denote these Figures by: "Framework for Naturalistic Studies" (FNS) and will be using them throughout this report. Note that the arrow sign does not appear in Figure 1-7 (a) where there is no clear definite order between CAN and video data.



(a) The FNS with the Y-axis corresponds to the CAN data and video data categories



(b) The FNS with the Y-axis corresponding to levels of enhanced technologies

Figure 1-7: The Framework for Naturalistic Studies (FNS)

2 Types of technologies for data collection

2.1 Mobile phone location services

Most smart mobile phones (along with many tablet devices) have the ability to find and record location. This is normally done in one of two ways but both can provide a diary of movement for that particular device. The benefit of using a personal communications device for the monitoring of movements is that it is now normal for participants to carry these devices (particularly smart phones) and as such the quality and volume of data will increase.

The resolution of data recorded can be variable depending on the location finding service employed (whether through WiFi or GPS) however simple data points such as timings, trip durations, distance of trips and crude speed measures are recorded relatively reliably, certainly within an acceptable error limit. The system is commonly provided and supported through Google maps and as such some level of map matching is provided. Figure 2-1 shows how travel data can be represented for a series of journeys.

Figure 2-1 shows that routes on the map can be colour coded to show the different travel modes used and locations visited, while timings, distances and durations are illustrated in the side bar. Figure 2-2 also shows just one element from a larger range of journey where extra detail can be seen. Additionally a bar graph shows the full extent of journeys undertaken.

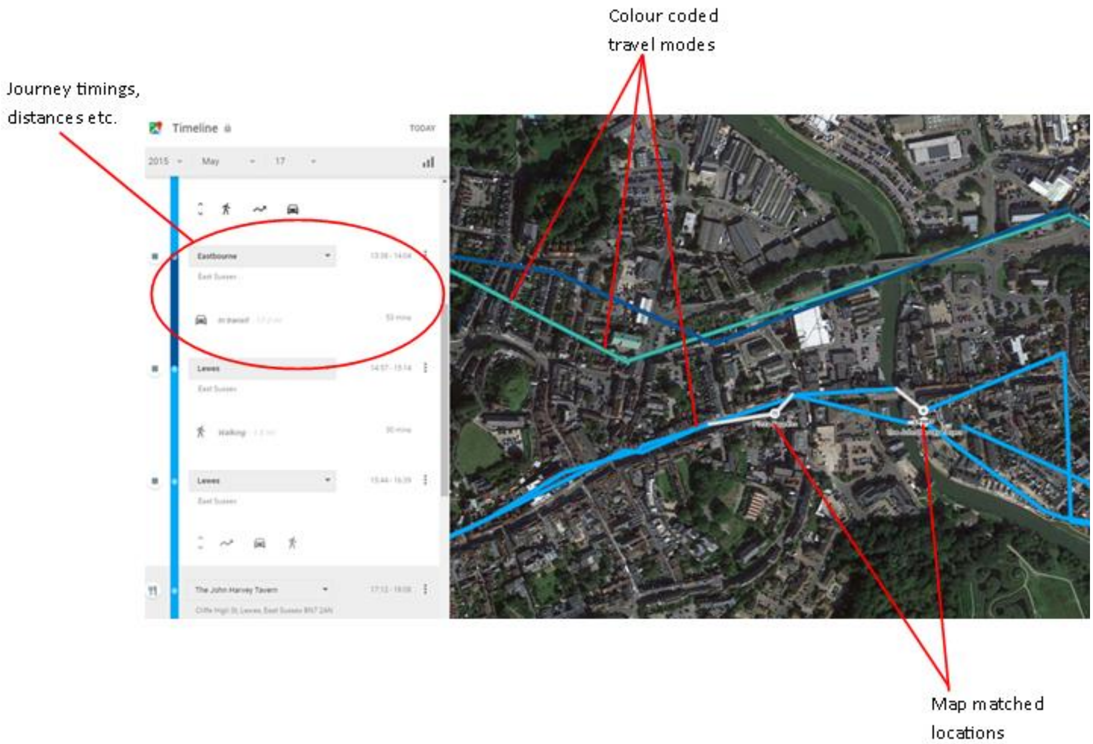


Figure 2-1: ***Example of Google location timeline***

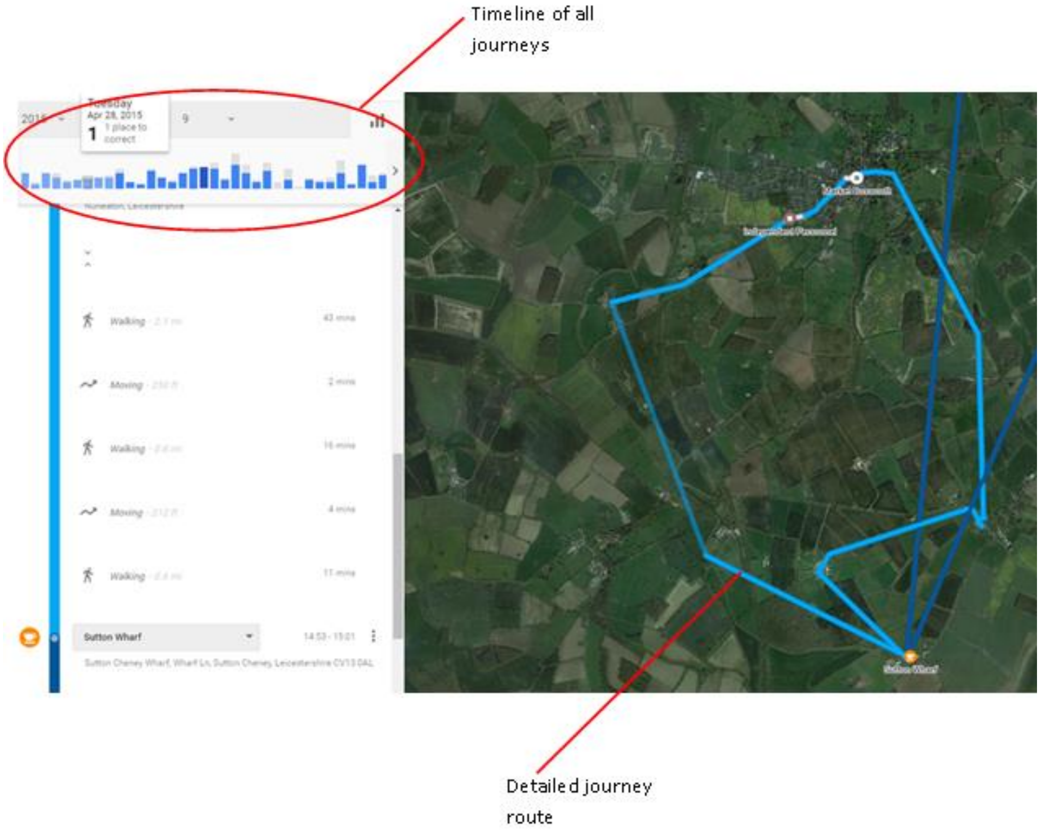


Figure 2-2: ***Example of Google location timeline on a detailed route***

There can be some issues with the quality of the data and this can be dependent on the type of journeys or user being recorded. For example, journeys away from WiFi (remote rural journeys for example) can miss detailed destination points but provided there is a good GPS fix the actual journey profile should be relatively reliable. Conversely in heavy populated areas, particularly those featuring urban canyons or Metro/underground transport the WiFi waymarks (i.e. the locations where the device identifies an available WiFi connection) may be more distinct that the GPS trace which could become lost.

The location of these types of devices on the main figure is illustrated in Figure 2-3

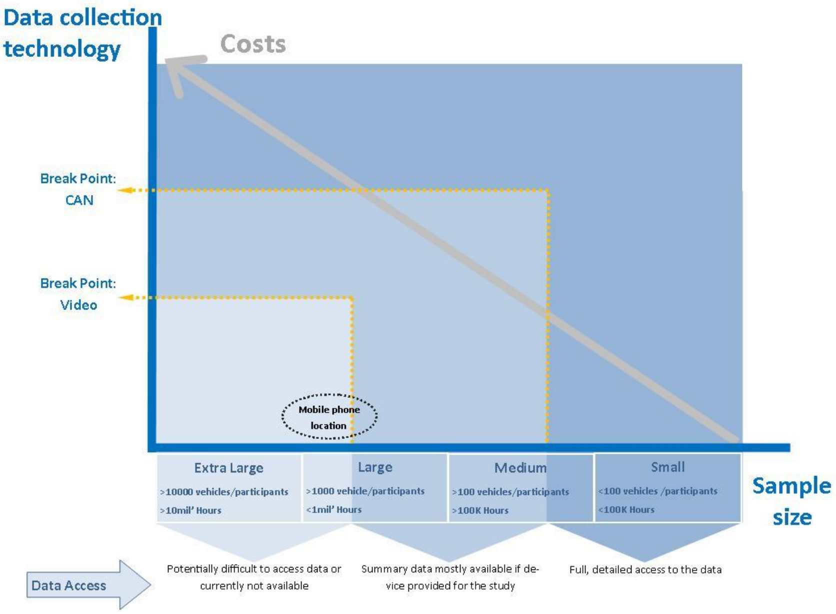


Figure 2-3*: Location of Mobile phone location services on the FNS*

2.2 Mobile phone applications

Probably the widest category of data acquisition tools; those based on a smart phone (android, windows or apple) platform. There is no simple way to group this category as the range of capabilities is only limited by the application developers imagination; as such these tend to include different driver aids.

A lot of these applications are based around the video capabilities inherent with modern smart phones and therefore provide another level of data over and above the simple GPS and location based data loggers.

As an example a few application developers provide a simple forward collision warning system or headway monitoring based on the interrogation of the video stream from a smart phone mounted in a vehicle (dashboard or windscreen) cradle. The computing demand of these systems normally means that there is little if any video data recorded. Figure 2-4 show a range of screen captures from these applications that show the augmented video channel and the type of data recorded.

Figure 2-4: ***Examples of augmented video stream from smart phone applications***

In addition to the augmented video based applications that offer driver support there is crossover between traditional dash cameras (see below) and the use of smart phones. Many developers provide these systems to passively capture the video of the road ahead, this data is often enriched with the inclusion of speed, time, driving duration, simple maps and identifications of 'events' however the video stream is not interrogated leading to a simpler design and the capability to record more video data. Figure 2-5 shows some examples of simple smart phone based dash cameras.

Figure 2-5: ***Examples of smart phone based dash cameras.***

Increasingly there are a range of mobile phone applications that can be used to monitor or even block mobile phone use while driving. Potentially these applications could provide information to a test side on mobile phone use while driving, including access to social media sites providing information on distraction behaviours. Currently these applications are designed as a preventative measure, for example they prevent a company driver from using a phone while driving, but furthermore, it indicates that it is possible to record this information through this method.

The position of mobile phone applications on figure FNS is shown in the following figure:

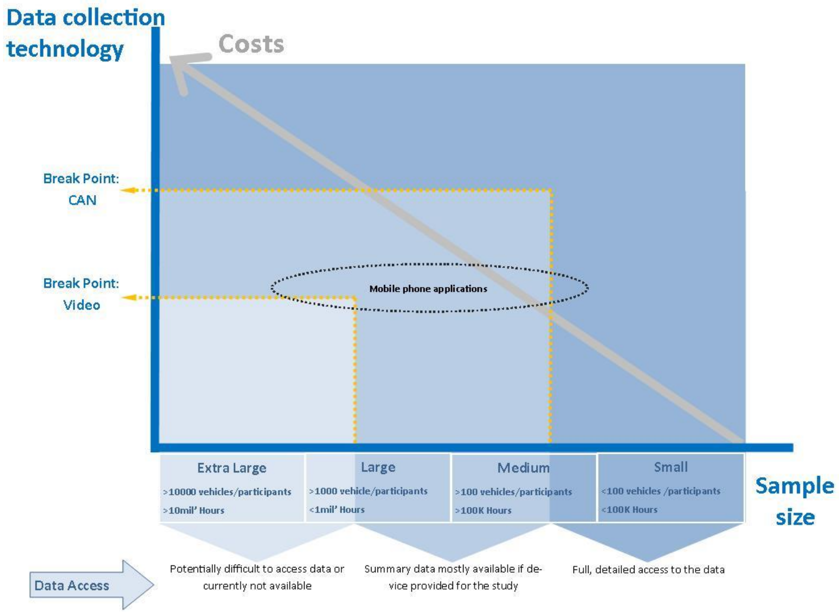


Figure 2-6*: Location of mobile phone applications on the FNS*

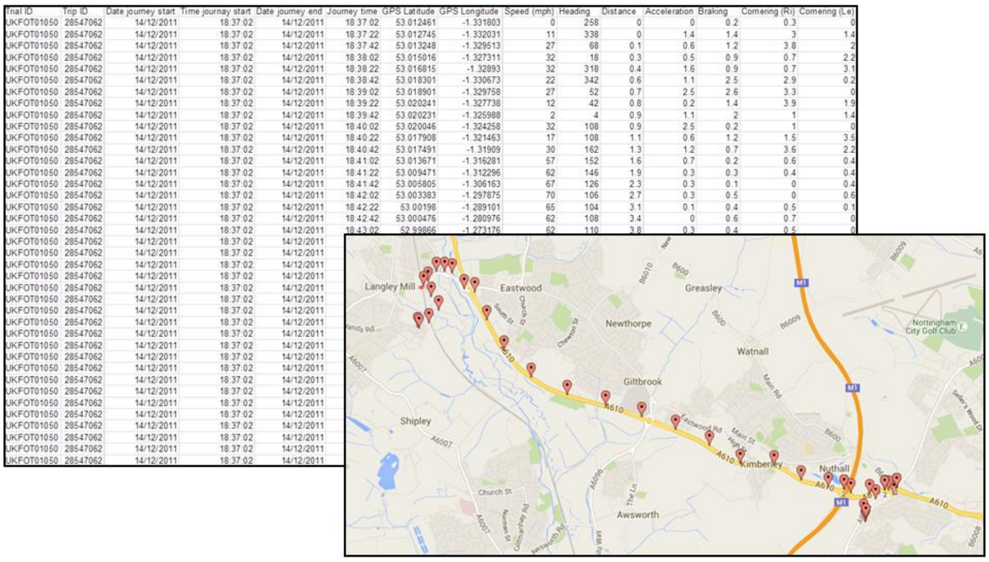
2.3 Telematics devices (including Green box, IVDR)

The use of telematics devices for vehicle insurance is relatively new use for a mature technology. These devices are normally a simple GPS data logger that can be installed onto the vehicles switched live power circuit (in order for them to turn on/off when the vehicle does) recording a relatively simple range of variables. The power of these systems is that they are low-cost, reliable and normally unobtrusive allowing for continuous uninterrupted data collection from vehicles.

Although the technology differs between devices there is some commonality between the data recorded. As GPS location is sampled (normally in the range of once, twice or three times a minute) so a vehicle speed, heading, journey time, journey distance and journey duration can be recorded. This simple data is normally augmented by some form of accelerometer that identifies lateral and longitudinal accelerations; typically only if a threshold is breached such as emergency braking or harsh cornering.

Data can normally be transferred by GPRS over a mobile communications system so no interference with the device is necessary. Limitations with these types of devices occur when multiple drivers use the same vehicle (there is normally no ID system) and with the sample rate of the data in that it is difficult to assess 'normal' driving unless an acceleration threshold is breached. Figure 2-7 shows a sample of the raw data extracted from a device used in a UK field trial and a visualisation of the data using mapping software.

The marked points (the google 'pins') on the visualisation are GPS 'fixes' which correspond to each line in the raw data file, the frequency of the GPS points is determined by sampling rate (typically one data record per second or fewer) and contain the latitude and longitude readings, the instantaneous speed at that point along with the instantaneous lateral and longitudinal accelerations and other calculated metrics such as cumulative distance and journey time. Different systems provide data in different ways and Figure 2-7 shows only an example of what is possible.



Trial ID	Trip ID	Date journey start	Time journay start	Date journey end	Journey time	GPS Latitude	GPS Longitude	Speed (mph)	Heading	Distance	Acceleration	Braking	Cornering (Ri)	Cornering (Le)
UKFOT01050	28547062	14/12/2011	18:37:02	14/12/2011	18:37:02	53.012461	-1.331803	0	258	0	0	0.2	0.3	0
UKFOT01050	28547062	14/12/2011	18:37:02	14/12/2011	18:37:22	53.012745	-1.332031	11	338	0	1.4	1.4	3	1.4
UKFOT01050	28547062	14/12/2011	18:37:02	14/12/2011	18:37:42	53.013248	-1.329513	27	68	0.1	0.6	1.2	3.8	2
UKFOT01050	28547062	14/12/2011	18:37:02	14/12/2011	18:38:02	53.015016	-1.327311	32	18	0.3	0.5	0.9	0.7	2.2
UKFOT01050	28547062	14/12/2011	18:37:02	14/12/2011	18:38:22	53.016815	-1.32893	32	318	0.4	1.6	0.9	0.7	3.1
UKFOT01050	28547062	14/12/2011	18:37:02	14/12/2011	18:38:42	53.018301	-1.330673	22	342	0.6	1.1	2.5	2.9	0.2
UKFOT01050	28547062	14/12/2011	18:37:02	14/12/2011	18:39:02	53.018901	-1.329758	27	52	0.7	2.5	2.6	3.3	0
UKFOT01050	28547062	14/12/2011	18:37:02	14/12/2011	18:39:22	53.020241	-1.327738	12	42	0.8	0.2	1.4	3.9	1.9
UKFOT01050	28547062	14/12/2011	18:37:02	14/12/2011	18:39:42	53.020231	-1.325988	2	4	0.9	1.1	2	1	1.4
UKFOT01050	28547062	14/12/2011	18:37:02	14/12/2011	18:40:02	53.020046	-1.324258	32	108	0.9	2.5	0.2	1	0
UKFOT01050	28547062	14/12/2011	18:37:02	14/12/2011	18:40:22	53.017908	-1.321463	17	108	1.1	0.6	1.2	1.5	3.5
UKFOT01050	28547062	14/12/2011	18:37:02	14/12/2011	18:40:42	53.017491	-1.31909	30	162	1.3	1.2	0.7	3.6	2.2
UKFOT01050	28547062	14/12/2011	18:37:02	14/12/2011	18:41:02	53.013671	-1.316281	57	152	1.6	0.7	0.2	0.6	0.4
UKFOT01050	28547062	14/12/2011	18:37:02	14/12/2011	18:41:22	53.009471	-1.312296	62	146	1.9	0.3	0.3	0.4	0.4
UKFOT01050	28547062	14/12/2011	18:37:02	14/12/2011	18:41:42	53.005805	-1.306163	67	126	2.3	0.3	0.1	0	0.4
UKFOT01050	28547062	14/12/2011	18:37:02	14/12/2011	18:42:02	53.003383	-1.297875	70	106	2.7	0.3	0.5	0	0.6
UKFOT01050	28547062	14/12/2011	18:37:02	14/12/2011	18:42:22	53.00198	-1.289101	65	104	3.1	0.1	0.4	0.5	0.1
UKFOT01050	28547062	14/12/2011	18:37:02	14/12/2011	18:42:42	53.000476	-1.280976	62	108	3.4	0	0.6	0.7	0
UKFOT01050	28547062	14/12/2011	18:37:02	14/12/2011	18:43:02	52.99866	-1.273176	62	110	3.8	0.3	0.4	0.5	0
UKFOT01050	28547062	14/12/2011	18:37:02	14/12/2011										
UKFOT01050	28547062	14/12/2011	18:37:02	14/12/2011										
UKFOT01050	28547062	14/12/2011	18:37:02	14/12/2011										
UKFOT01050	28547062	14/12/2011	18:37:02	14/12/2011										
UKFOT01050	28547062	14/12/2011	18:37:02	14/12/2011										
UKFOT01050	28547062	14/12/2011	18:37:02	14/12/2011										
UKFOT01050	28547062	14/12/2011	18:37:02	14/12/2011										
UKFOT01050	28547062	14/12/2011	18:37:02	14/12/2011										
UKFOT01050	28547062	14/12/2011	18:37:02	14/12/2011										
UKFOT01050	28547062	14/12/2011	18:37:02	14/12/2011										
UKFOT01050	28547062	14/12/2011	18:37:02	14/12/2011										
UKFOT01050	28547062	14/12/2011	18:37:02	14/12/2011										
UKFOT01050	28547062	14/12/2011	18:37:02	14/12/2011										
UKFOT01050	28547062	14/12/2011	18:37:02	14/12/2011										
UKFOT01050	28547062	14/12/2011	18:37:02	14/12/2011										
UKFOT01050	28547062	14/12/2011	18:37:02	14/12/2011										
UKFOT01050	28547062	14/12/2011	18:37:02	14/12/2011										
UKFOT01050	28547062	14/12/2011	18:37:02	14/12/2011										
UKFOT01050	28547062	14/12/2011	18:37:02	14/12/2011										
UKFOT01050	28547062	14/12/2011	18:37:02	14/12/2011										

Figure 2-7: ***Example of raw data stream and visualised location data from simple telematics device***

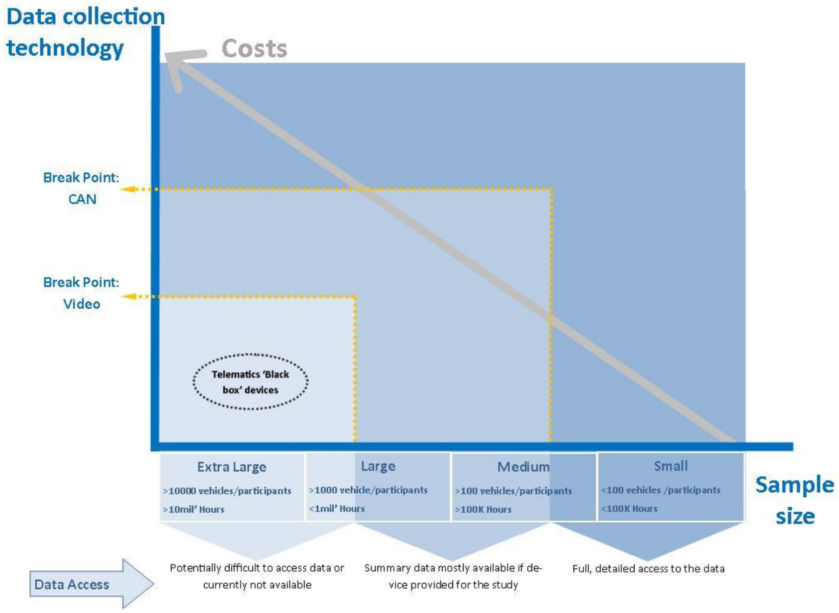


Figure 2-8***: Location of Telematics on the FNS***

2.4 Built-In data loggers

As with most technology innovation, particularly one that relates to a rapidly moving and potentially transformative shift such as autonomous vehicles, the data recorded in the built-in data collection systems are both incredible valuable and commercially sensitive. These devices tend to be Tier one supplied (from the major suppliers of parts to OEMs) and fitted as standard by the Original Equipment Manufacturer (OEM).

Data recorded by these devices will be incredibly extensive but will probably be limited to the lead up to a collision; variable groups that are likely to be recorded include: Vehicle dynamics (road speed data, wheel speed data, suspension data, active control system data), Body control (yaw rates, roll rates, active body control measures), Drivetrain control (engine metrics, gearbox data, power and torque information), Driver data (steering input, pedal operation, vehicle settings, protective system data such as occupancy for airbag control) and other data such as measures the vehicle has taken to mitigate or avoid a collisions such as radar data, city safe systems, ACC information, pre-safe technologies and stability control operation.

Historically, for the purposes of vehicle based data collection, built-in systems have been seen as a solution to many road safety questions as it normally represents the vanguard of a technological advancement. Indeed there are examples of built-in data collection since the inception of the airbag where data on the severity of the impact, provided in the form of a crash pulse was recorded in the tier one supplied airbag control module. Some development in the extraction of data was made during this period however due to a number of unresolvable issues no full scale, widespread use of the data was used.

The same story was repeated again approximately 10 years after airbags with the launch of other secondary safety technologies such as stability control. Again, issues such as physical access to the recording device (i.e. can an analyst actually get to the 'black box' without taking the car apart?), electronic access to the data (can an analyst plug in a device, meet security requirements and download the data?) and interpretation of the data (can an analyst decrypt the data or do manufacturers allow decryption?) bugged the process and made this approach very difficult to achieve.

There have been some successes in using 'black box' data from built-in devices, however these normally have one a major strength; an OEM or tier one supplier as part of a project consortium. The most recent was the UDriver project which used Renault manufacturer support to aid the collection and recording of sensitive CAN data; this cooperation did not extend to the full access to the raw data files as this disclosure is still well within the remit of commercial sensitivity. Without manufacturer support it is extremely unlikely that any access, even filtered or sanitised, will be granted to the content of the data collection systems as this information will invariably relate to the underlying algorithms that allow the vehicle to perform autonomously.

As autonomous vehicles progress and automation of other processes filter down to most everyday vehicles the need to access the data may become more pertinent. Currently there is still viability for collision investigations as autonomous collision avoidance technology is neither mature enough nor at high enough fleet penetrations to entirely remove the collisions, however as these becomes less common and the physical data correspondingly scarce the untapped source of vehicle data may be required in order to monitor driver or vehicle behaviour in near-miss events.

The position of built-in data loggers on the FNS is shown in the following Figure 2-9.

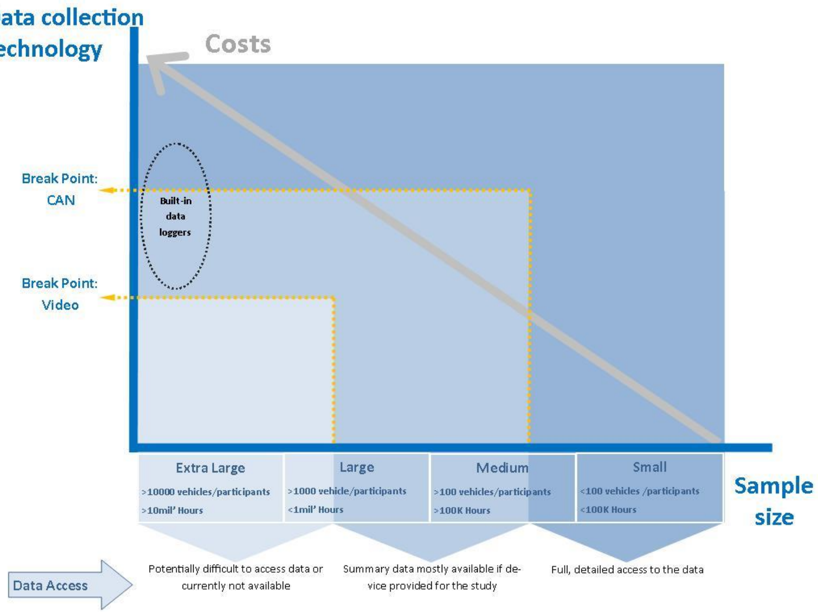


Figure 2-9*: Location of built-in data loggers on the FNS*

2.5 Dash cameras and enhanced dash cameras

The first major advanced step in data collection is the introduction of video, this has already been addressed in the above section however it is worth discussing specific dash cameras as they provide some significant advantages over devices for which the camera is not its sole design consideration.

Dash cameras have become increasingly popular in recent years as the technology has improved and the costs have reduced. Simple dash cameras can be purchased from as little as €10 with the cost increasing in relation to picture quality and additional enhancements. Data recorded is normally stored on the device (or on removable media) due to size restraints so ownership of the data is not ensured unless devices are owned or operated by the study.

Enhancements to the video stream normally include GPS locations, speed, time, data, journey time, duration and accelerations. These enhancements results in a very rich data output as can be seen in Figure 2-10

Figure 2-10***: Still images from enhanced Dash Camera data streams.***

Figure 2-11***: Location of NextBase Dash camera behind rear view mirror***

The position of dash cameras and enhanced dash cameras on the FNS is shown in the following figure:

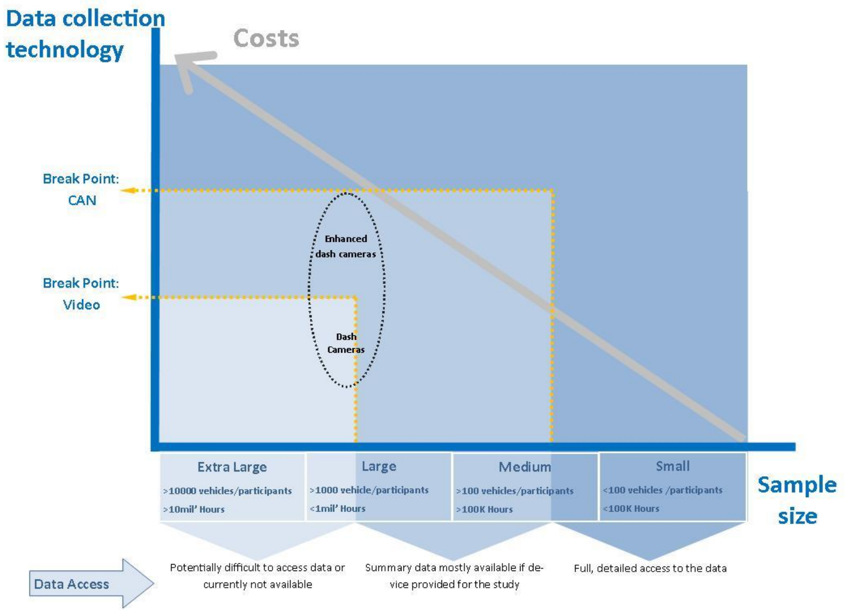


Figure 2-12: Location of dash cameras and enhanced dash cameras on the FNS

2.6 Wearable technology

Not all recording devices need be physically fitted to a vehicle, nor indeed does it necessarily need to record the movement and behaviour of a road vehicle. The most recent category of data collection fulfils this brief and can be described as 'wearable technology' (sometimes called activity monitors or activity trackers). This technology is predominantly tasked with monitoring the activity of individuals and can cover all modes of transport including when people are sedentary or even asleep. This technology has been around for a number of years (The VTTI driver fatigue study of the late 1990s used this technology) although it is only recently that its use has become widespread.

The first commercially successful activity monitor was made available in 2014 but since 2015 there has been a proliferation of different technologies covering a range of different metrics. The most commonly seen devices currently are the FitBit, Garmin and Jawbone rivals which have become increasingly popular as wearable technology has become more mainstream over the previous couple of years.

These devices generally record the general activity of a wearer/user by using an accelerometer to detect movement. This acceleration is run through an algorithm to detect the mode of travel (walking, running, cycling) and combined with heart rate measurement. Most available devices use a GPS receiver to locate this activity and there is a range of proprietary software to access and analyse the data.

Activity monitors are particularly useful if studying a participants travel across all modes or looking into the how sleep, rest and activity patterns affect other elements of a participants behaviour. Heart rate can also be used to determine conditions of driving or other stressful activities outside of physical activity, where you would normally expect an elevated heart rate. There should be relatively good acceptance of these devices for use in trials as they are relatively mature and popular, additionally usage/consumer surveys indicate high uptake with people actively wanting to use this technology[1]. These two points also make the cost of these devices (particularly the simpler versions) relatively inexpensive.

As the technology advances it is probable that other metrics will be added to the sensing repertoire. There are already items of wearable technology that can collect biometric data such brainwave (EEG) and muscle bio-signals (EMG) data however, these are currently some way off large scale rollout in trials, they could however provide invaluable information in the field of health care and wellness if applicable.

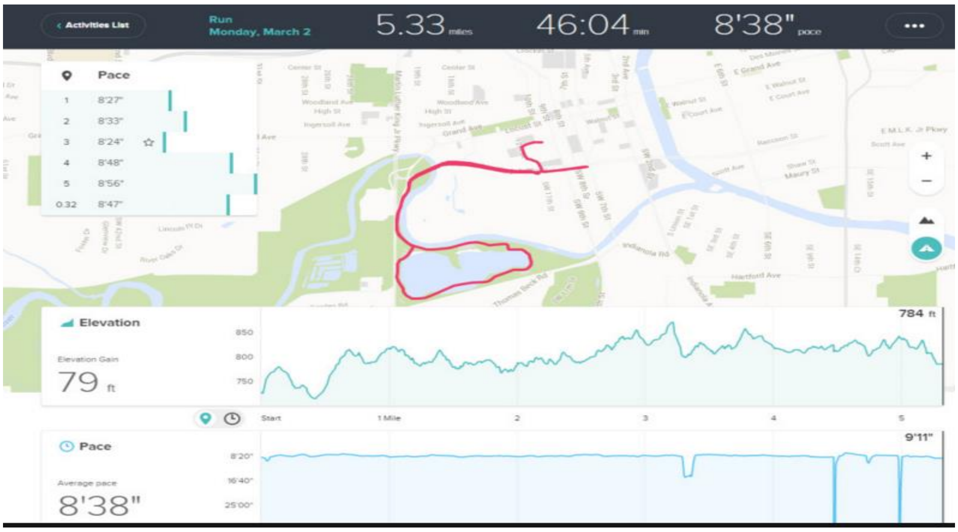


Figure 2-13: Example of activity tracking data

There are other variations of activity monitor that either use a smart phone in isolation or use a manufacturer wrist, chest or head band to sync to a smart phone application. This method can enhance the data by providing more data and variables.

Another example of an activity tracker is those that use existing GPS devices (smart phones, GPS watches, GPS computers) to record activity for set periods which can then be viewed at another time. Two examples of these are included in the table in chapter 3 (Strava and Endomondo) and they record similar data to the wrist band based devices outlined above. Again, having a data logger present in a device that is routinely carried such as a smart phone is advantageous as it reduces the chances of missing data, it also reduces equipment costs as participants may already have the required computing power in their possession.

There are a number of ways in which the data can be enhanced, for example in the world of cycling heart rate monitoring, pedal cadence and power output can also be added to the data trace, with more development and innovation appearing all the time.

[1] Source: GlobalWebIndex - digital consumer insight.

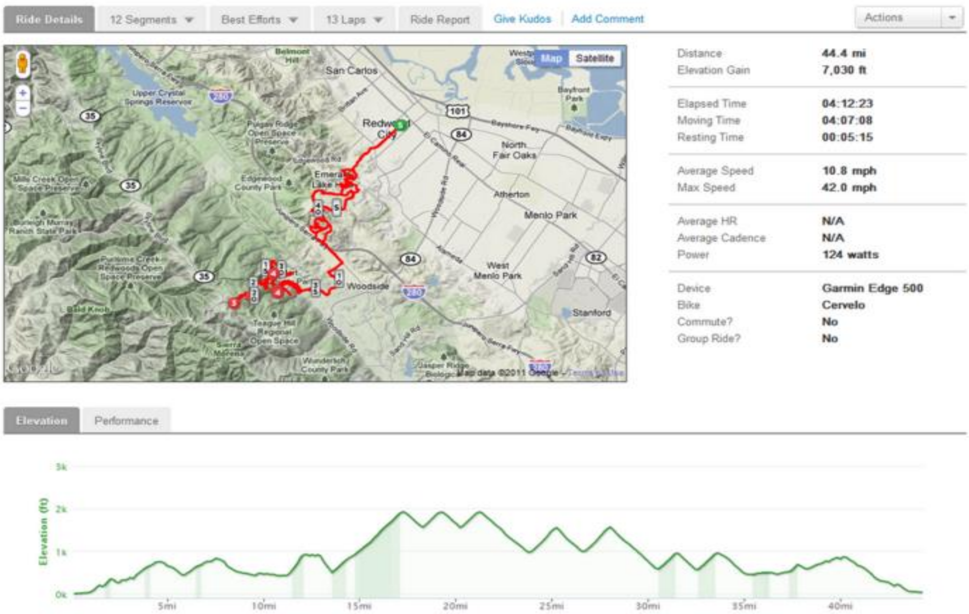


Figure 2-14: Example of cycling based tracking data

The position of wearbles technology on figure A is shown in the following figure:

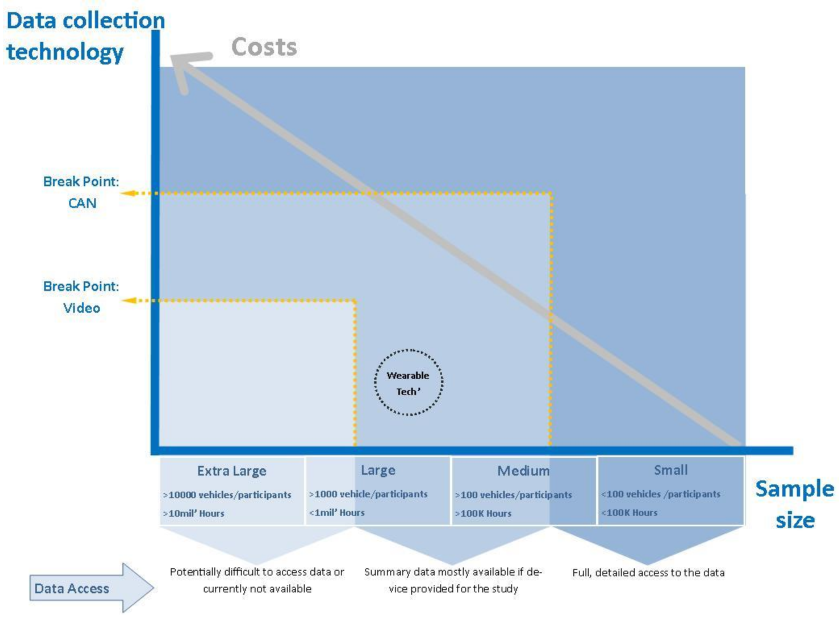


Figure 2-15: Location of wearable technology on Figure A

2.7 Compound systems

Relatively new to the market are 'compound' systems. These are broadly a collection of different data collection devices or data decoders that can be fitted to a vehicle, connected by Bluetooth and run using a powerful smart phone.

These devices are still in their infancy but a great deal of development over recent years and continuing in the future could bring huge gains to the recording or complex but disparate data sources such as video and CAN. Previously recording both of these data sources together necessitated a complex industrial specification data logger and even the recording of one source such as CAN was notoriously difficult and unreliable.

Compound systems rely on a decoder box which can be fitted to the CAN port of most passenger vehicles built since 1996. This box can be configured to read the CAN data and can then send it via Bluetooth to a smart phone within the vehicle. The application that reads and stores this CAN data can be run in the background while other apps, such as a location services or a dash cam apps can be run simultaneously, the data is then amalgamated into one source providing a better overview of the whole vehicle behaviour.

Figure 2-16 shows an example of a Bluetooth CAN adaptor while Figure 2-17 illustrates how vehicle data from the CAN adaptor can be combined with a dash camera based system.

Figure 2-16: Bluetooth CAN adaptor

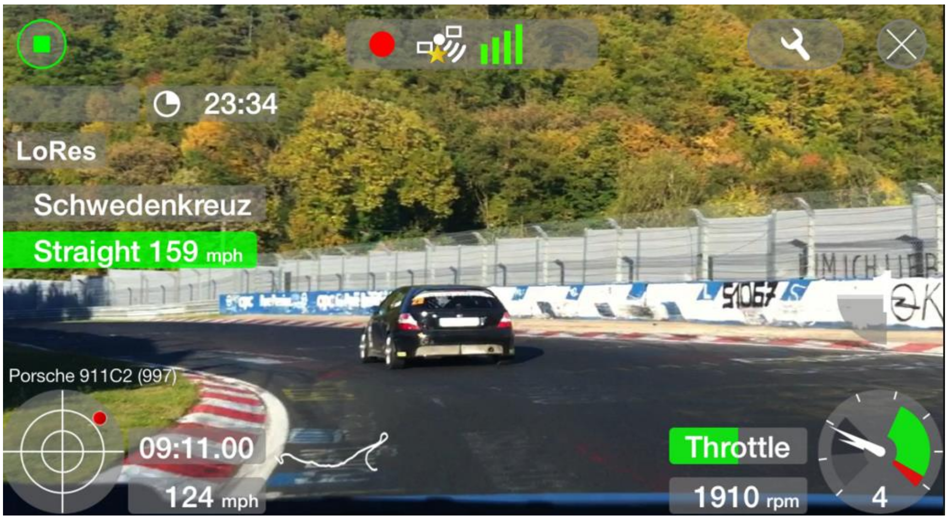


Figure 2-17: example of dash camera application running with CAN adaptor data

The benefit of these systems is that they could provide almost endlessly configurable data, with each innovation more and more connectivity is developed allowing a range of other sensors and data sources to be integrated. The system outlined in the initial table also allows other third party cameras (GoPro etc.) to be synced with the smartphone app allowing much greater video coverage.

Costs vary based on system complexity but systems begin at around €100 per unit based on a smartphone application, inbuilt smartphone video camera and the CAN adaptor. There could be unforeseen issues with the system in terms of Bluetooth connectivity and compatibility with certain vehicles/sensors types which could make its suitability to long term trials, where a vehicle may be running independently for extended periods, problematic.

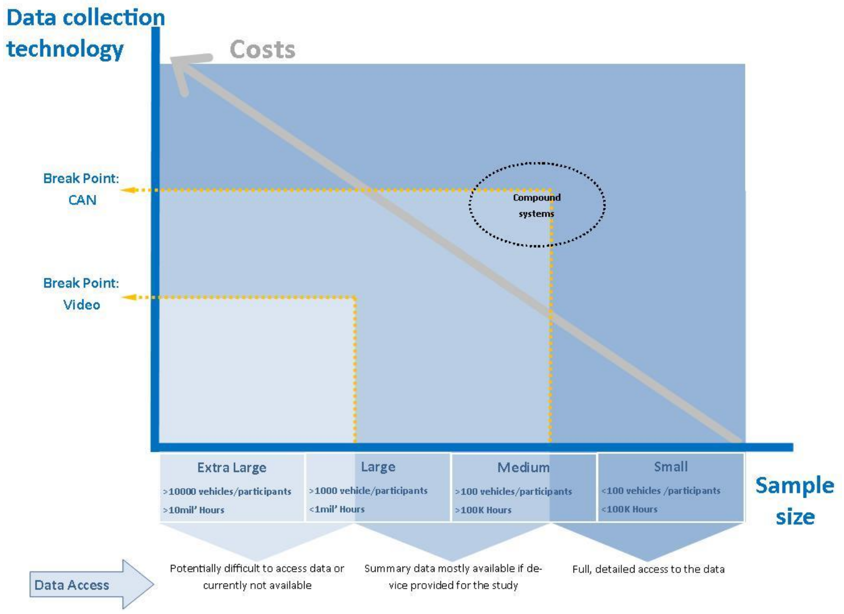


Figure 2-18*: Location of compound systems on the FNS*

2.8 Eye trackers

Eye trackers form a section of technology that have been slowly developing over a couple of decades, as the technology has steadily improved so has the availability of this technology in more challenging environments; whereas before these systems were the preserve of laboratory or simulator testing it is now more practical to use them within vehicles on real roads. Most current technologies use optical systems to detect and track a participant's eye behaviour and may include metrics such as blink rate, saccades and eye closure, all of which can be used to determine driver alertness/sleepiness.

There are a number of different systems available but most can be split into two main groups; those that have to be worn (normally in the form of a pair of glasses) and those that are 'non-intrusive' for example, having their cameras mounted on the dashboard or within the instrument cluster of a vehicle.

These technologies are still currently quite expensive and can suffer integration issues with other forms of data collection technologies, as such the location drawn on the FNS is designed to represent a 'best fit' between small sample sizes and the top 'specialist' level of data, although this puts it above the CAN and Video break points it is unlikely that these data sources will be included in the final eye tracking data stream.

Typically, for non-intrusive eye tracking systems, a period of calibration is required before use. Other system types require some form of interaction with the device i.e. putting on a pair of glasses fitted with the technology. The benefit of these issues is that they provide a wide range of detail eye behaviour metrics which is not currently available through any other form of technology.

Figure 2-19: wearable eye tracking glasses

Figure 2-20: non-intrusive eye tracking equipment (desk mounted)

Placing eye trackers on the FNS poses a few issues as it does not fit naturally within the defined structure, this is likely because it is a specialist piece of technology which is normally used alongside other data recording devices and not normally in isolation. The placement shown in Figure 2-21 is not truly representative of the break points as eye trackers do not typically use CAN or video in the traditional sense, as such the break points have been removed from this figure. What the placement does attempt to show is that these devices occupy the upper limits of the technology and cost scale; due to this they are also best used for studies with small or even very small samples as they can be intensive both in terms of time, labour and processing.

The case of head mounted eye trackers could be considered as a special case of wearable technology while keeping in mind that usually an eye-tracker is an add-on to other systems.

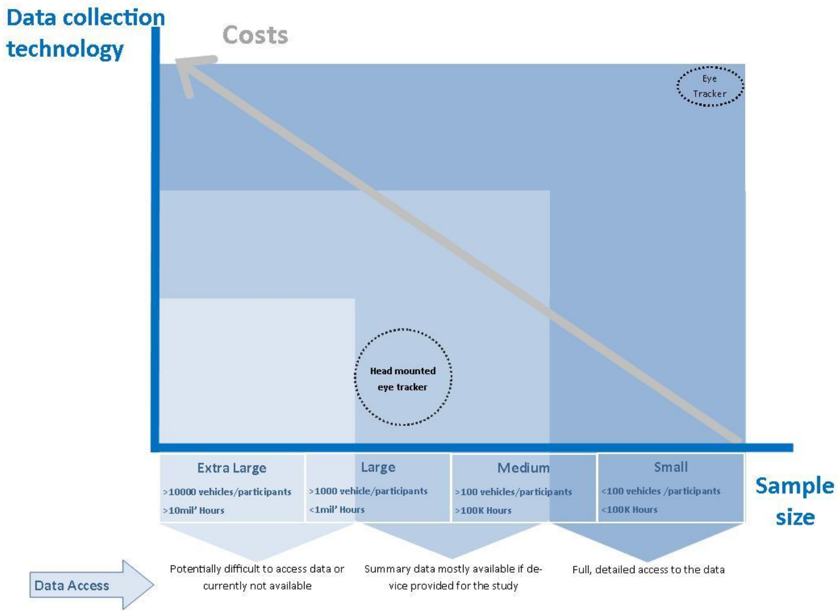


Figure 2-21: Location of eye tracking equipment on the FNS

2.9 Mobileye

Mobileye is driver support system provided both as a third party system (i.e. non-OEM fitment) and as an integral part of modern vehicles crash warning or vehicle automation system.

The device acts as a driver support system in that it provides warnings and real time information on a variety of driver safety and performance metrics. A number of different systems are available which all broadly cover warnings for pedestrian/cyclist collisions, lane departure, forward collision, headway monitoring and speed limits. The system operates using a windscreen mounted video camera, the output of which is interrogated to identify the features listed above such as the rear of a vehicle ahead or a pedestrian crossing the road.

The technology has previously been made available in large scale Field Operational Trials and as such has a track record of integration into data acquisition systems. As a standalone device, not built in to a larger system the output data may be provided in the form of trigger points (i.e. incidences where the system activated a warning and what that warning consists of). The Mobileye technology is integrated into the UDrive DAS and hence is specifically referenced here.

Special features of the Mobileye system include speed monitoring and even speed violation monitoring (based on visual identification of posted speed signs).

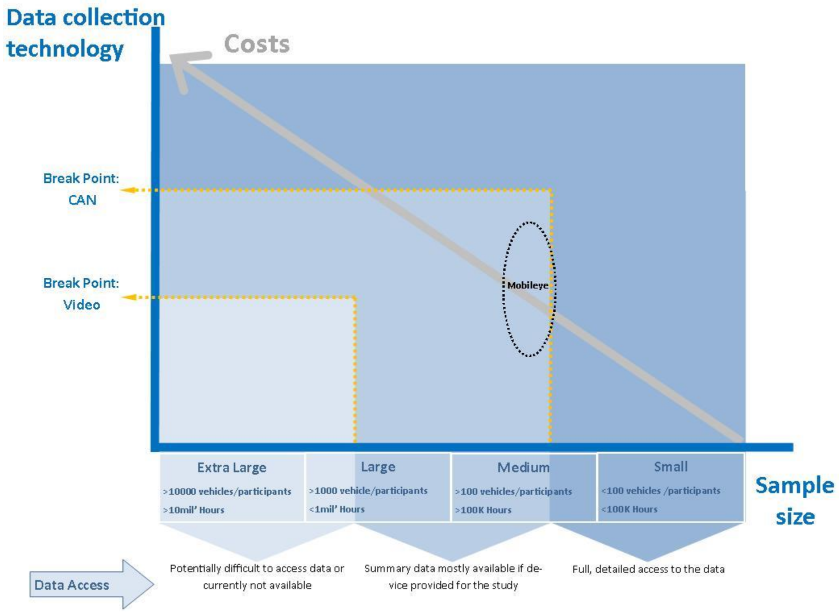


Figure 2-22: Location of Mobileye on the FNS

2.10 An overview of technologies mapped on the FNS

Figure 2-23 illustrated the mapping of the various technologies that were surveyed in this section on the the FNS.

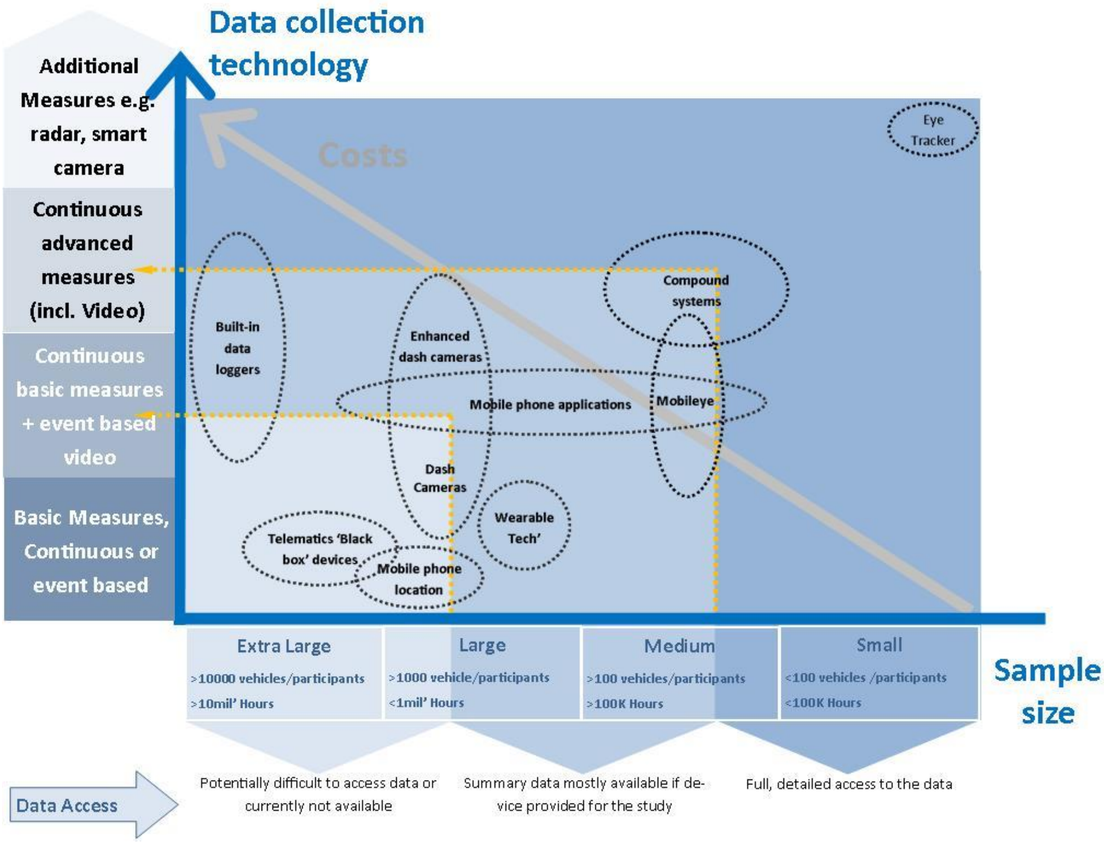


Figure 2-23: Location of the various technologies on the FNS

A tabular form of the various technologies surveyed in this chapter appears in Appendix A. Given the vast variety of technologies and their rapid development, it is impossible to keep track and provide a comprehensive list. In some areas, such as mobile phone applications, there are many hundreds of similar products. In this case the most popular applications were surveyed with only the most relevant for a naturalistic study included in the table in Appendix A. Additionally technology moves on at a rapid pace; it is therefore very likely that a table seeking to include every available device would be out of data almost immediately after production.

An important consideration with all of the technologies outlined in this report is reliability and accuracy. As with most pieces of scientific equipment, increasing the research and development can lead to greater reliability and accuracy, however this normally impacts the user in terms of increased purchase costs.

Specifically for the application in large scale real world trials many of the technologies outlined in this report occupy the cheaper end of the instrumentation spectrum. These devices may rely on other devices to operate, a mobile phone for example, or be built on or developed from an existing platform; as such not all technologies are created equal. For example, accelerometers which form a cornerstone of driving data vary significantly both in terms of reliability and more importantly accuracy.

Pilot trials for UK FOTs have shown that identical devices placed next to each other in a vehicle may record noticeably different data. This may not be a problem so long as drivers are not compared within the cohort but could prove problematic if ranking performance or driving styles. The reason for these differences becomes clear when comparing cost – the devices used for this pilot were around €30 per unit making the accelerometer only a small proportion of this value. Compare this to industry standard accelerometers which may cost upwards of €600 and the difference in expected outputs becomes clear. Reliability is also a factor with cheaper devices; an attractive, low purchase price may allow a much larger trial to be conducted however the benefits of this can be reduced significantly if reliability of the devices leads to high drop-out rates. As an example of this was a satellite navigation system used in another UK trial which proved to be unreliable over an 18 month trial period. This factor lead to varying levels of data completeness and reduced power for detailed analysis – a 'cost' to the project significantly greater than the small saving made on unit price at the beginning of the trial.

Unfortunately it is not possible to provide specific or detailed information on reliability and accuracy for each of the technologies included in the report. Some technologies will follow the general rule of "you get what you pay for" but others may have specific data accuracy considerations associated with them. More detailed information on accuracy, reliability, sensitivity and data quality assurance is available in UDRIVE deliverable D22.1 [Welsh, R., Reed, S., Lenard, J., Kotiranta, R. (2017) UDRIVE deliverable D22.1 Guidelines for data quality assurance of the EU FP7 Project UDRIVE].

In general reliability and accuracy cannot be assured; even market leading, industry quality equipment can provide poor quality results or prove unreliable if used incorrectly, the converse is equally true of cheaper equipment if used intelligently and carefully applied to a trial. Very often it is not possible to determine how a piece of equipment will perform before a trial has begun and it is therefore extremely important to build in pilot testing to real world studies for any technology type.

3 The potential of IVDR G-Based technology integrated with vision based technology to identify and document risky driving behaviour.

3.1 Background

The purpose of the analysis presented in this section is to illustrate the potential of two relatively low-cost technologies to identify, record and analyse risky events related to unsafe driving behaviour. The two technologies are the Mobileye system and an IVDR (also known as Green Box) system manufactured by Ituran.

The data used for the demonstration presented here was recorded by 13 vehicles of the Or Yarok fleet over a period of 15 months.

The data sources used for the demonstration include: events recorded by the two technologies, trip data (e.g. start time and end time of the trip) and the GIS layer of roads in Israel.

The events generated by the two systems are temporally synchronized. This is achieved by letting the Ituran system read the display protocol of the Mobileye system, and generate a time-wise synchronized flow of events.

It is important to mention that, unlike in traditional naturalistic studies, the feedback from the two systems was available to drivers, either through a warning alert, a signal on event occurrence, and on a web platform. Hence the analysis could be influenced by drivers changing their behaviour as a response to that feedback. However, for future studies – the option of not providing real-time feedback can be easily implemented.

Descriptive statistics

The Mobileye and the Green Box technologies identified and recorded a total of 130,010 driving events, generated by 13 vehicles, over the time period lasting from August 2014 till October 2015. Not all vehicles were equipped throughout this whole time period. Additionally, not all vehicles had the same version of the Mobileye system. Furthermore, the same vehicle could be used by different drivers.

The events recorded by the Green Box system correspond to G-force type of events and include: braking, acceleration, lane change, speed bump, bypass, turns, turn while braking, turn while speeding and speeding.

The events recorded by the Mobileye system correspond to vision based type of events and include: headway keeping, forward collision warning (FCW), lane departure warning (LDW), pedestrians in danger zone (DZ) and pedestrians' collision warning (PCW).

Out of the 130,010 events recorded, 129,173 had also a known location (GPS coordinates). The analysis presented here excludes events under the "other" category (such as: seat belt not buckled).

For each event we also have information on speed gathered through GPS and classification into urban & inter-urban road calculated using a simple map-matching technique that associates the event (according to its GPS coordinates) to the nearest road.

Figure 3-1 presents the events count by event type. The events are sorted according to their frequency of occurrence. In this data set - headway, braking and speed violation are the most common driving events. The analysis presented in this chapter focuses on the "vision based pedestrian" events (marked by solid circles in the Figure) which include two types: 270 pedestrian collision warnings (PCW) and 3,829 Pedestrian Danger Zone (Pedz DZ).

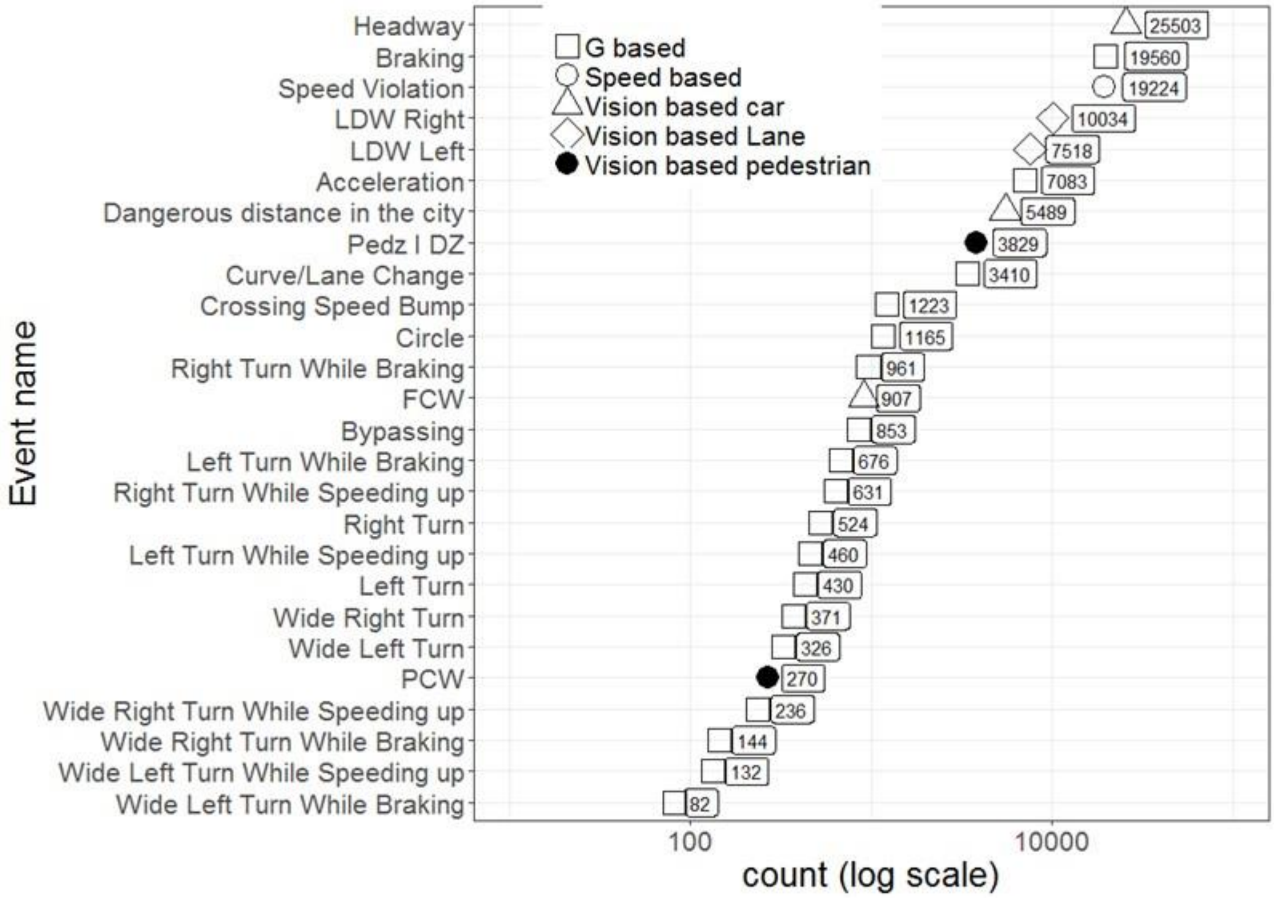


Figure 3-1: Events count by events type

The average speed per event type and its confidence interval (CI) are presented in the Figure 3-2. The events are ordered along the y-axis as in the previous graph, according to their frequency. Figure 3-2 also separates between events performed in urban and in inter-urban roads. Clearly, for all event types – the speed is higher in inter-urban roads.

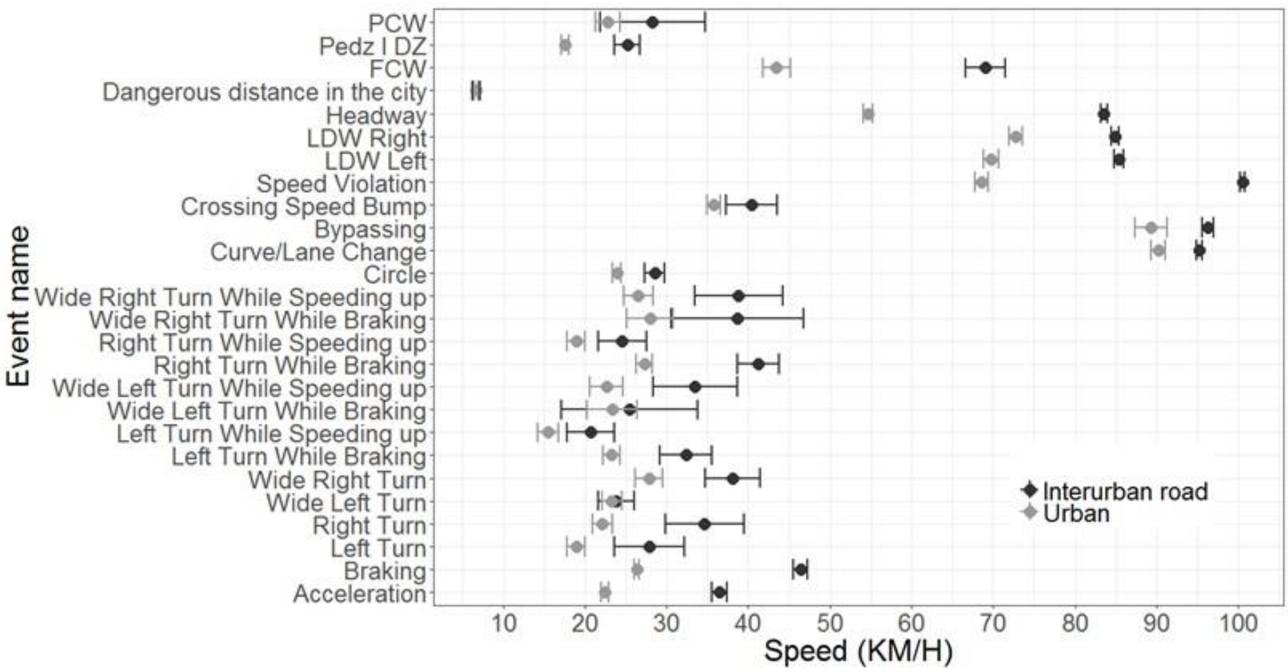


Figure 3-2: Mean and CI of speed by event type and road (urban / inter-urban)

The Mobileye version which identifies pedestrians is relatively new, hence not all vehicles had this feature from the start of the data collection. More specifically: two vehicles had this feature from the start of the data collection phase, and 11 vehicles has this feature only after July 2015 (that is, at most for 4 out of the 15 mounths of data collection). This explains the low occurrence of pedestrians' related events. It is also important to note that the sample and data base presented in this section is by no means representative of any population, not even the safe drivers population of Or Yarok. The analysis presented in the next sections is aimed to show the potential of the database to monitor and record risky behaviours and not representative results for specific populations.

3.2 Temporal analysis

Time between events: The time stamp of each event is available with 1sec resolution. Furthermore, the events generated by the two systems are synchronized. This is achieved by letting the Ituran system read the display protocol of the Mobileye system, and generate a time-synchronized flow of events.

Figure 3-3 describes the observed proportion of time between all types of events. If events were independent, then the time between them should follow an exponential distribution. The observed distribution (presented in Figure 3-3 between 0 and 60 sec) is different. Clearly, there is over representation of the 1sec 'time between events'. This analysis is currently deficient (limited) as it includes events from all trips of the same vehicle – this means that our calculation might include events in sequential trips, which normally should be ignored. Interestingly, despite this limitation, events tend to occur in close time proximity, for example, 15.16% of the events are 10 (or less) seconds apart. This phenomenon can be related to dependence among events and is next further explored.

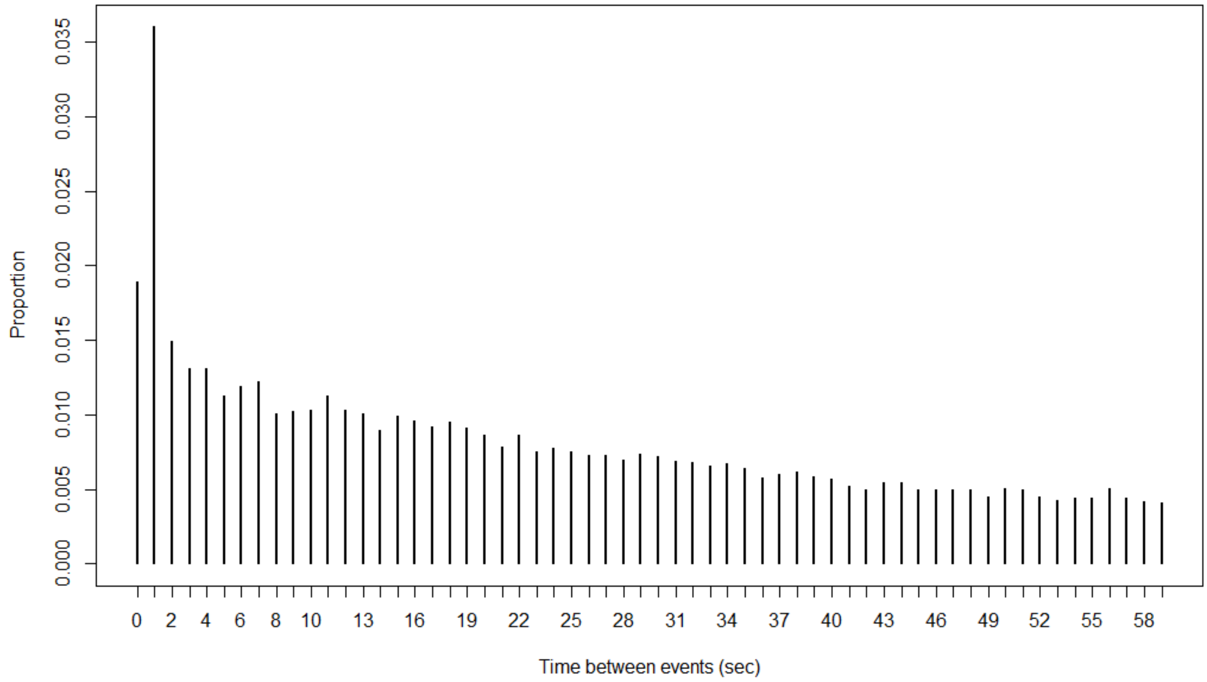


Figure 3-3: Distribution of time between events

To better understand inter-dependency between events we look at events occurring at the same time (the same second). The most common combination is two instances of speed violation, followed by headway and speed violation combination.

Next we repeat the same analysis when the time between events is one second. There are 6,513 events occurring one second apart from each other in our database. In this analysis, the pair of two speed violation combination is much less common. The most common combination is two instances headway with more than 1,000 occurrences. A possible explanation is that the driver is in a situation of close following to the lead vehicle and continues to keep short distance from it.

To understand events related to pedestrians' warnings, we next look into combination of events including one or two of the vision based pedestrian warnings. Such combinations are for example "PCW and braking" or "acceleration and DZ". A combination of events suggests a short time and potential dependence between them. The determination of what is a "short time" that implies correlation between events depends of prior knowledge of how events develop in real driving situations. For example, how much time passes between receiving a pedestrian collision warning and the resulting braking event? In this analysis, we considered an exploratory approach. If events tend to occur together (e.g. PCW and braking) the count of their combination within a short time-frame (e.g. 0-4sec) should be larger than the count of their combination in a wider time-frame (e.g. 5-9sec). Figure 3-4 presents all combinations of pedestrian events (PCW and DZ) with any of the other events. The count of combinations of these events in the short time frame (0-4sec) is plotted compared to the count in the longer time-frame (5-9sec). In order to detect higher counts - the plot also presents the 45deg line. The pairs of combinations appearing above the line correspond to higher occurrence close to each other. The figure suggests that FCW and Ped DZ occur more frequently in short time frames (12 occurrences vs. 1). Occurrences of two PCW are also more common in the short time-frame of 0-4sec.

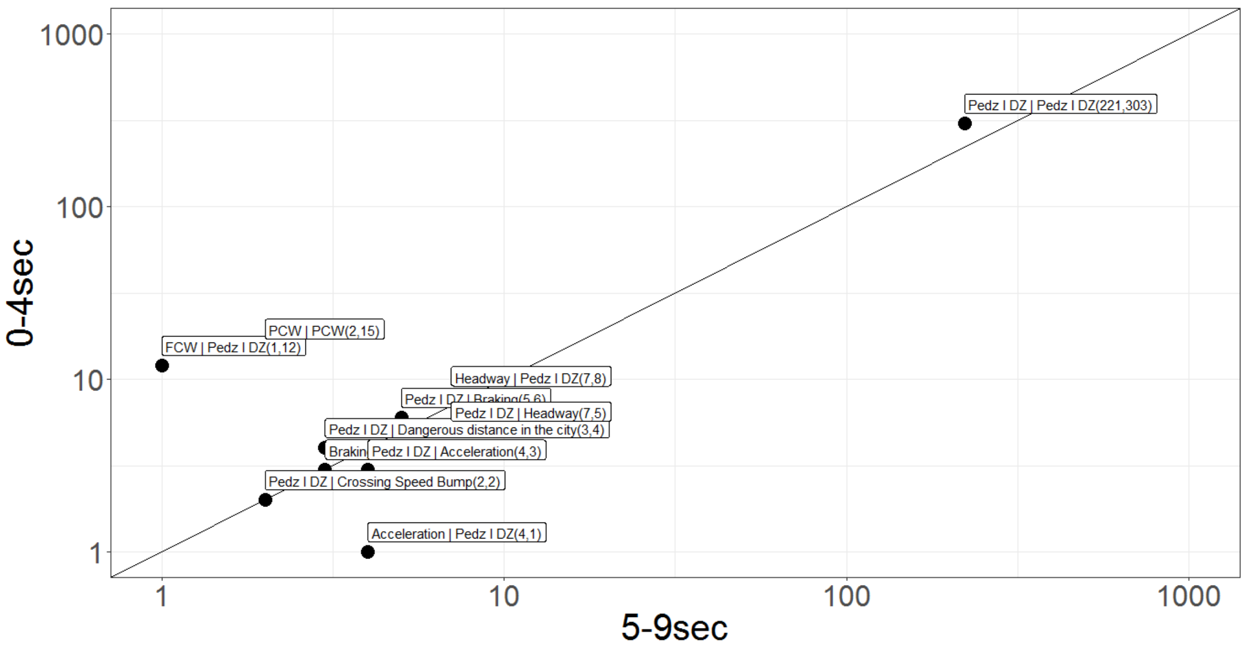


Figure 3-4: Count by events combinations occurring within 0-4 seconds and 5-9 seconds

Time-distance analysis

Another approach to analyze events according to their time distance from each other is to use a cluster analysis. For the current analysis, we chose a hierarchical clustering procedure. Naturally, as the researcher allows larger time distance between cluster members, the clustering procedure suggests fewer clusters. Figure 3-5 presents the number of clusters produced per time distance (x-axis). The hierarchical procedure suggested 109,425 clusters for time distance equal to zero. Most clusters include a single event, 1,616 clusters included two events that were reported at the same time. Plots as presented in Figure 3-5 are useful to detect the recommended number of clusters - when the graph levels, additional time distance unit does not change the number of clusters significantly. The levelling of the graph indicates on the number of clusters recommended for analysis. However, the graph in Figure 3-5 continues to drop even after 10 and 15 seconds. Thus, we look at the change in the number of clusters per time distance (the derivative of the graph in Figure 3-5). This data is provided in Figure 3-6. According to the data, allowing one second time distance among cluster members reduces the number of clusters by 3229 (from 109,425 clusters). Allowing for two seconds' time distance reduces the number of clusters even more by 1637. This trend continues until time distance of eight seconds when the graph starts to level. There is another change in trend at 13 seconds. According to this analysis, the 8 seconds' time distance might be a good choice for differentiating between clusters. We therefor chose to continue with this selection of 8 seconds.

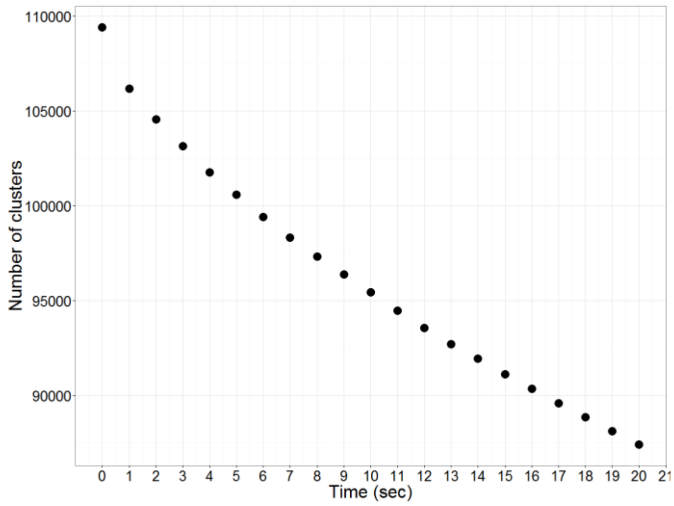


Figure 3-5: Number of clusters by time distance within cluster members

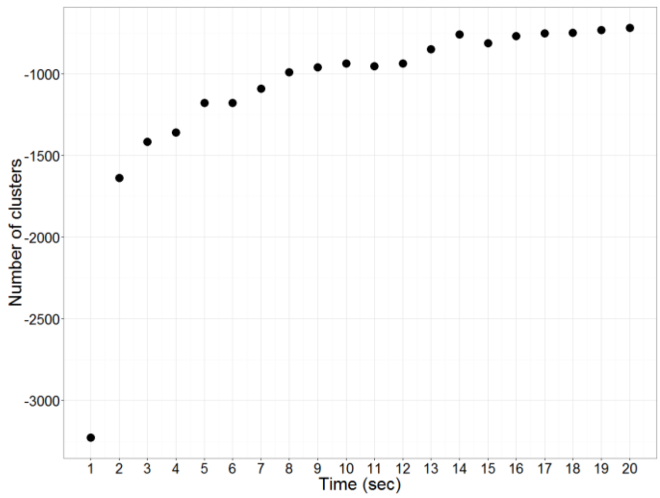


Figure 3-6: Change in number of clusters by time-distance within cluster members

When analysing the information in the “time distance clusters”, most clusters (88%) include a single event. 11% of the clusters include two events and the remaining 1% includes between three to six events (where three is the most likely). Figure 3-7 presents the probability of each of the event types as a cluster member under three conditions: (1) without any condition (black colour) (2) under the condition that a PCW event occurred within this time-frame (dark grey) and (3) under the condition that Pedz DZ event occurred within this time-frame (light grey). According to the results presented here, the probability of **other** events occurring in general (no condition) is higher than the probability given pedestrian related warnings. This is an indication that at least in the near period (7 sec) surrounding the pedestrian warnings, drivers are relatively carful and avoid receiving warnings from any other kind.

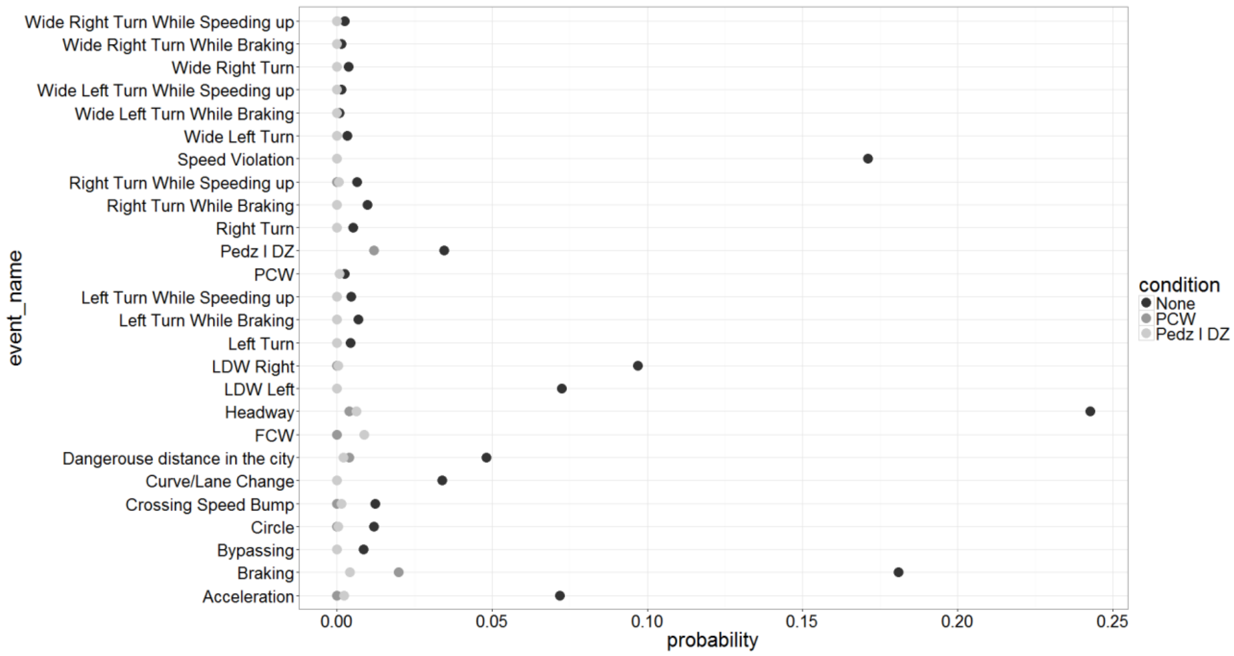


Figure 3-7: The probability of various events in general and in the time proximity of pedestrians' related events

3.3 Spatial analysis

The pedestrians' events were clustered according to their GPS coordinates by applying the mean-shift clustering procedure. Figure 3-8 presents some of these clusters on a map. The green points represent clusters with only Pedz DZ zone events and the blue points represents clusters that include at least one PCW event. As can be seen, the same road section can include several pedestrian related clusters. The radii of the points in the figure are proportional to the number of events (PCW+ Pedz DZ zone). The number of events within each cluster varies between 1 and 79 with a mean of 2.78.

Figure 3-8: Clusters presented on a map where cluster radii are proportional to events count

UDRIVE European Naturalistic Driving Study

The size of the clusters in terms of the distance between cluster members is determine by a predefined threshold set by the researcher. On one hand, too low threshold may unnecessarily differentiate among similar points, on the other hand, allowing large clusters may cloud our ability to analyse the infrastructure characteristics of the clusters. Figure 3-9 presents maximal distance between points in a cluster on the Y-axis and the number of events on the X-axis. In the extreme, the distance between cluster members is 93 meters and in the minimum zero (corresponding to a cluster with one event).

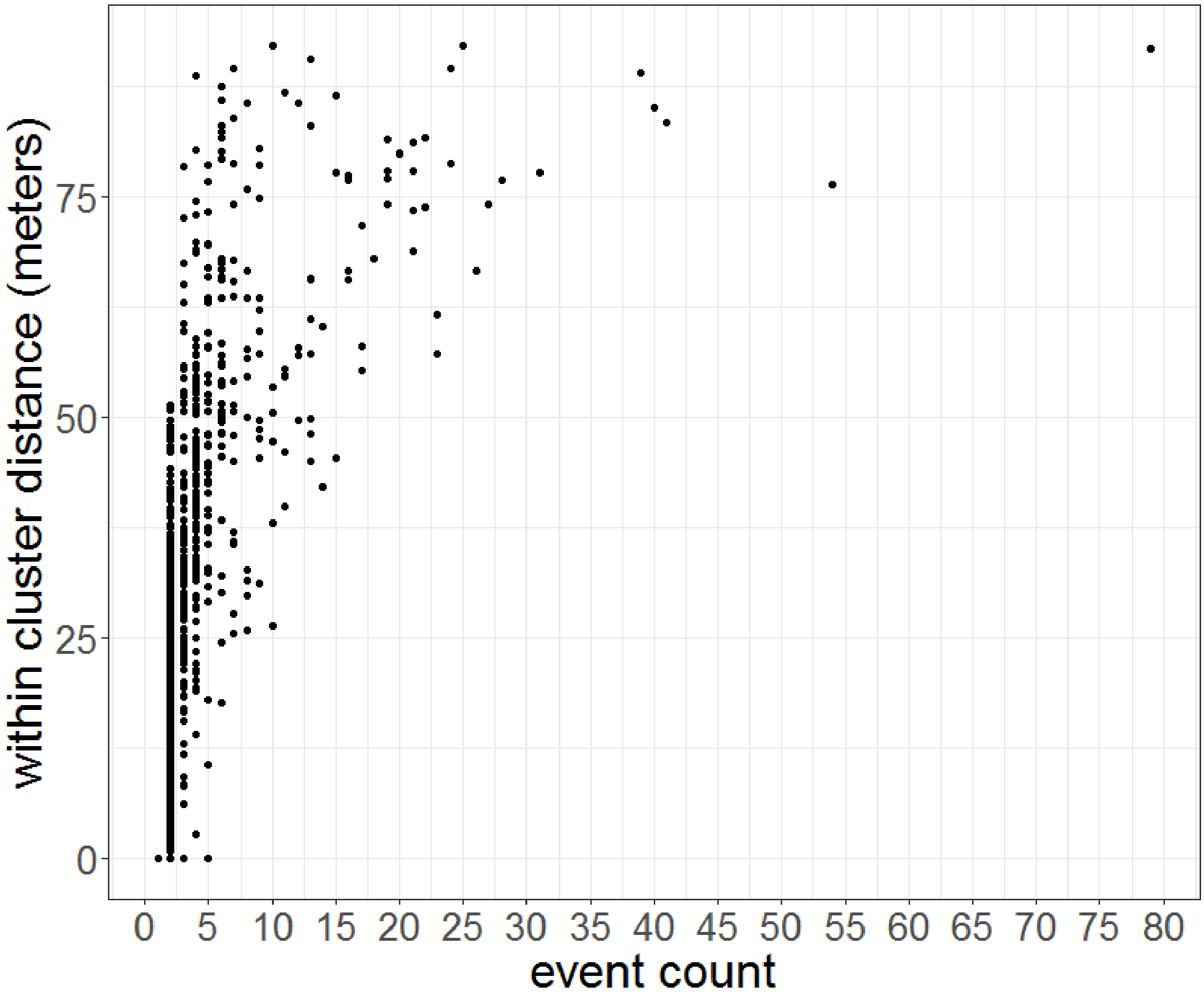


Figure 3-9: Cluster size vs. event count

3.3.1 Accounting for exposure

The map presented in Figure 3-8 presents the clusters with different sizes (radii) per the number of events in the cluster. Surly, as drivers traverse a specific location more times, the probability of events occurrence (at that location) increases. Thus, a consideration of exposure in the specific clusters is needed in order to provide a more realistic picture of risk. Figure 3-10 presents the clusters where their radii are proportional to the total count of events divided by an exposure index. The exposure index is an estimate of the numbers of times that vehicles traversed that specific location. As can be seen in the figure, the various clusters have similar radii and so none of them "stand out" as a location that is more prone to pedestrians' event. This map therefore conveys a different message than the map in Figure 3-8. The statistical correlation between the events count and events rate is 0.02. This low correlation score strengthens the importance of wisely selecting the index to be used for decision making on risky locations. For example, the largest cluster, appearing at the bottom left in Figure 3-8, corresponds to the parking facility of the Or Yarok offices, which clearly corresponds to a lot of exposure with pedestrians. When accounting for exposure, as can be seen in Figure 3-10, the size of this cluster reduced significantly, reducing the potential risk associated with this location.

Figure 3-10: Clusters presented on a map where cluster radii are proportional to events rate and exposure is accounted for

As events count and events ratio are non-correlated indices, they can jointly characterize clusters. Figure 3-11 presents the frequency of clusters per each combination of events ratio and events count. These indices have a skewed distribution, and their scales are different. Thus, both indices (counts and ratios) were log transformed and normalized (Z scores). Clusters in the first quarter (up-right) are clusters in which both indices values are above the average. In extreme cases clusters are located 2 or more standard deviations away from both indices' average value. These clusters may be good candidates for further examination.

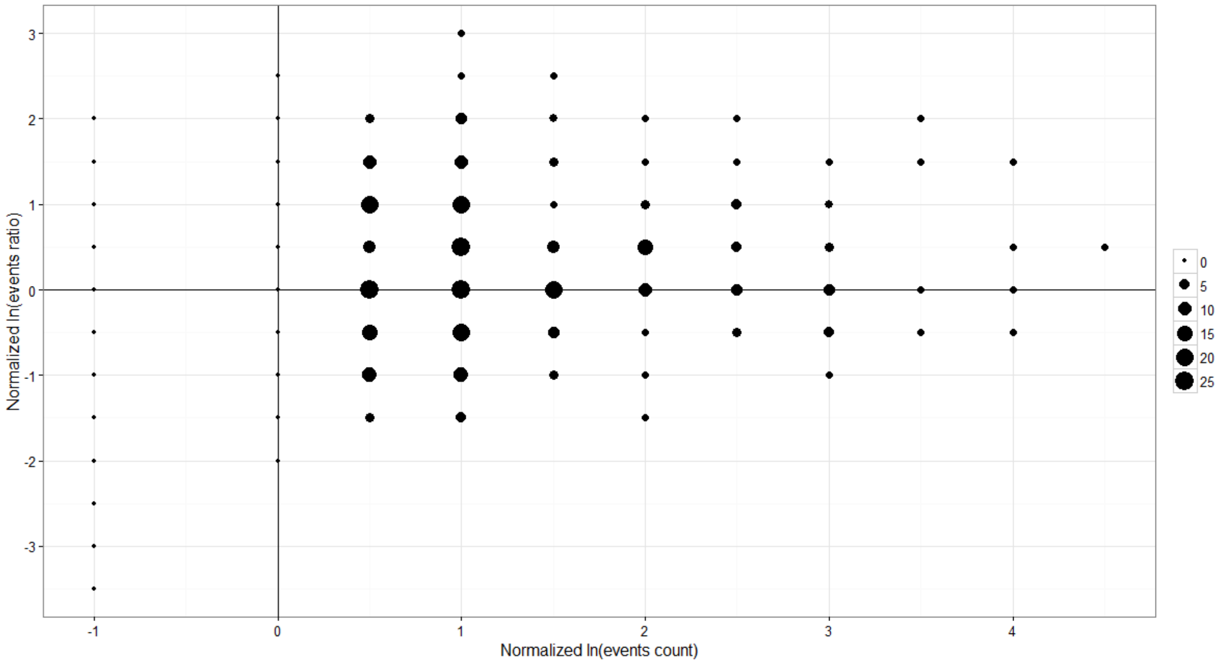


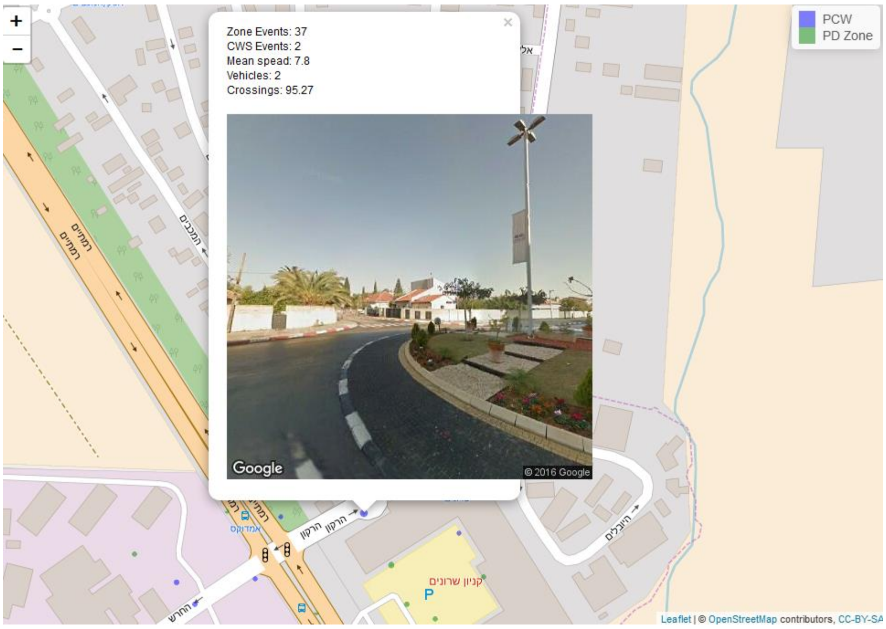
Figure 3-11: Frequency of clusters by normalized ln (events count) and ln (events ratio)

Next we present some examples for clusters in which both events count and events ratio were relatively high:

Example #1: The center of the cluster is a roundabout in an industrial area near a large mall. In this cluster the exposure variable indicates ~37 crossings and 27 DZ events.

Example #2: The center of the cluster is a roundabout near a large mall located in an urban neighborhood. In this cluster the exposure variable indicates ~95 crossings and 39 pedestrian events (2 of them correspond to PCW).

UDRIVE
European Naturalistic Driving Study



Example #3: The center of the cluster is on an urban road in Tel-Aviv. In this cluster, the exposure variable in dicates ~15 crossings and 12 pedestrians' presence events.

3.3.2 Estimation of the exposure index

In order to account for exposure, the main question that needs answering is how many times drivers cross specific clusters' location. To answer this question, we analysed drivers' trajectories (N=897,268). Trajectories include a set of several points (latitude and longitude) along the trip trajectory. For each given trajectory (set of coordinates) and for each cluster, the question is what is the probability that this cluster is included in the trip trajectory, i.e. the distance between the cluster coordinates and the trajectory is close enough. Using GIS software this distance matrix can be calculated. This analysis needs to determine whether a value in the distance matrix is small enough to be classified as belonging to the cluster. Figure 3-12 details the empirical cumulative distribution between a point and a trajectory. To estimate this distribution the following procedure was applied:

1. Per each trajectory constructed from more than 2 points (coordinates)
    a. Take one random point off the trajectory not including the origin and destination points.
    b. Calculate the distance between that point and the trajectory.
    c. Add the distance to the pull of distance values
2. Use the pull of distance values to estimate the empirical cumulative distribution

In circa 90% of the cases the distance is less than 200 meters, but in some (rare) cases the distance from one point to its own trajectory is more than 1000 meters (the graph was trimmed for clarity).

To estimate the number of times drivers crossed a certain cluster location, we used the empirical distribution (denoted by ***F***) in Figure 3-12 as follows:

1. Cluster ***c*** given with its coordinates
2. For each trajectory ***t***
    a. ***d*** =distance between cluster ***c*** and trajectory ***t***
    b. Exposure = Exposure+ 1- ***F(d)***
3. Log cluster exposure

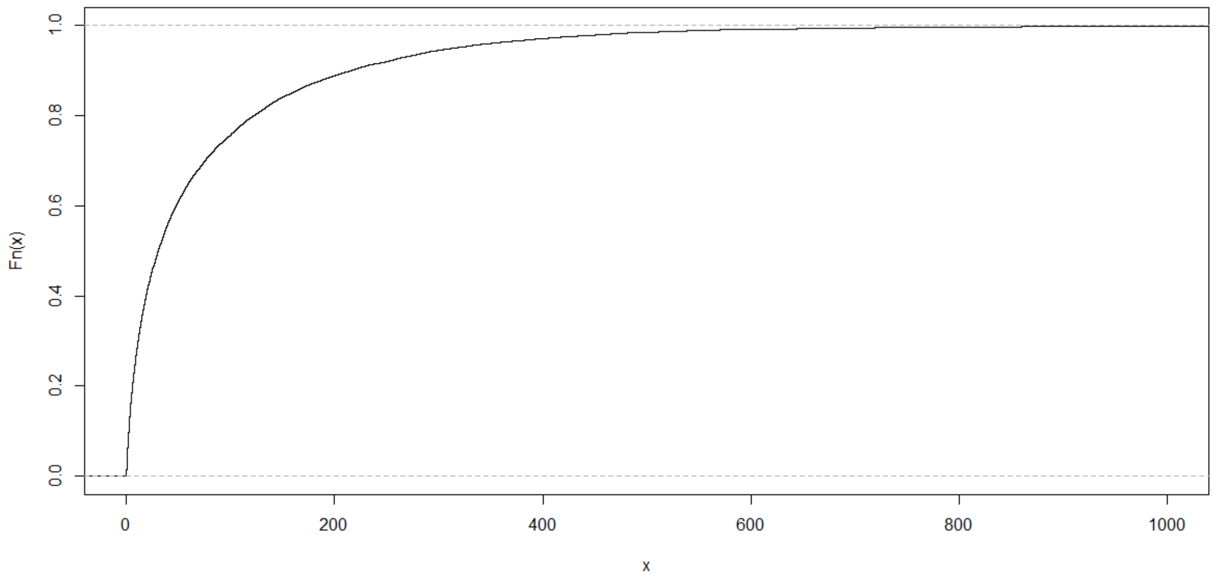


Figure 3-12: Empirical cumulative distribution of point-trajectory distance.

3.3.3 Differentiating between PCW and DZ events

The analysis above considers the PCW events the pedestrian danger zone (DZ) events as equal indicators of risk. While a PCW event indicates a potential real conflict with pedestrian to which the driver must actively act to prevent a crash, a DZ event indicates that pedestrians are identified in the close environment and hence the driver should be aware of their presence. That is, while the PCW events correspond to actual risk, the DZ events indicate on increased risk potential. The count of the DZ events is therefore more suitable as an explanatory variable in a statistical model for PCW events. Figure 3-13 arranges clusters according to the count of DZ events on the X-axis. The Y-axis presents the mean (and confidence intervals) count of PCW events for clusters with similar count of DZ events.

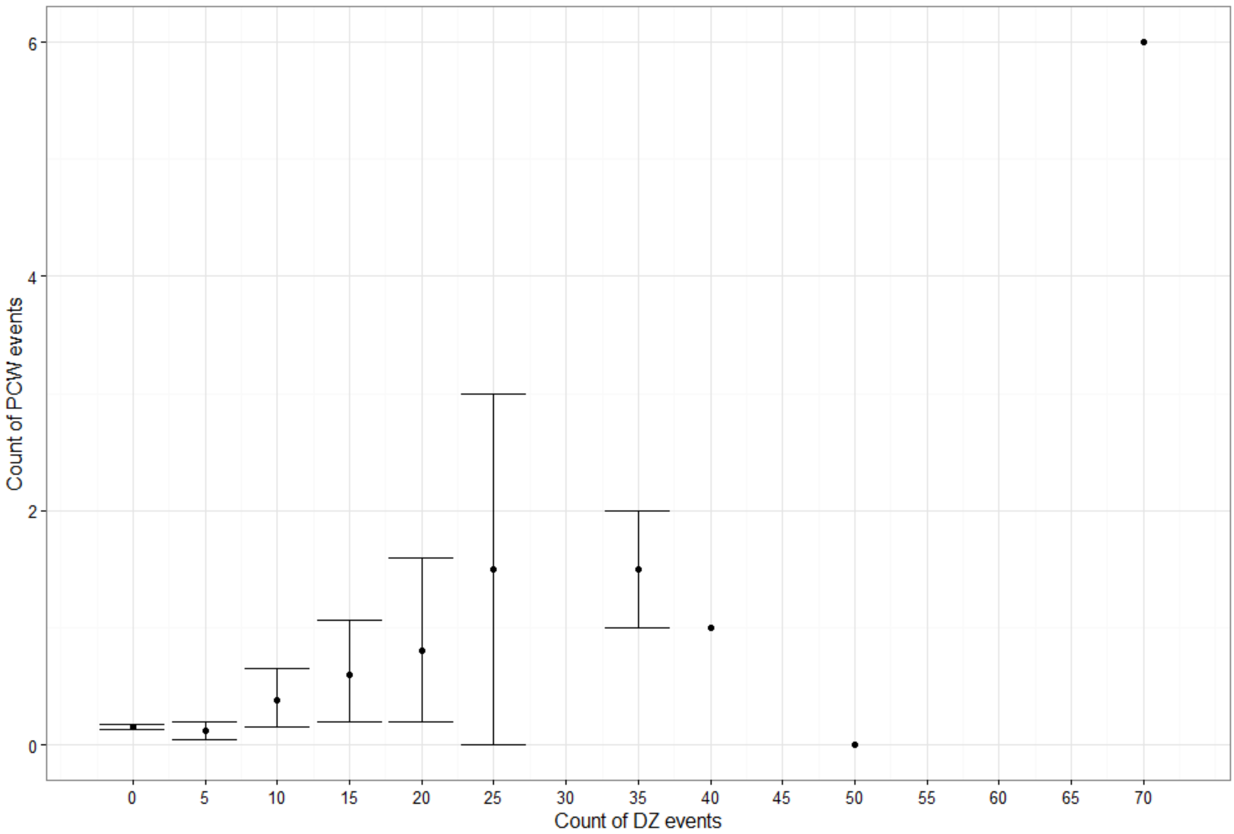


Figure 3-13: Count (mean and CI) of PCW events against count of DZ events

Note: the X-axis is arranged in groups of 5 (0 corresponds to [0-4] , 5 to [5-9] etc.)

Further, a negative binomial model for estimation of the count of PCW events in our spatial clusters was developed;

It is assumed that:

$$\mathrm{PCW} \sim \mathrm{NB}(\mathrm{SPF}, \mathrm{a})$$

$$\mathrm{SPF} = \beta_0 \mathrm{C}^{\beta_1} \mathrm{e}^{\beta_2 * \mathrm{Dz}}$$

Where PCW follows a negative binomial distribution with expectation that is determined by a Safety Performance Function (SPF) and the dispersion parameter a. **C** indicates the expected number of crossings ("exposure") and Dz the number of danger zone events. β_0, β_1 and β_2 are coefficients. The fitted parameters by the maximum likelihood estimation procedure are described below:

	Estimate	Std.Error	Pr(>|z|)
$\mathrm{Ln}(\beta_0)$	-3.03	0.23	< 2e-16 ***
β_1	0.30	0.06	7.29e-08 ***
β_2	0.03	0.01	0.00186 **
a (dispersion parameter)	1.299	0.528	

Signif. codes: '***' 0.001 '**' 0.01 '*' 0.05

The results indicate that as the number of crossings (the coefficient is β_1) is larger and as the number of DZ events is larger - more PCW events are expected.

Surely each cluster has different characteristics that are not know to us. The actual risk of a specific cluster may deviate from what can be predicted by this model. Thus, using this model to estimate the risk of a specific cluster is perhaps a naïve approach. On the other hand, the variability in the observed PCW counts may be large (especially for low counts) and thus estimating the cluster risk based only on the observed data may be too rough. The empirical Bayes method proposes a way to use both observed data and model predictions in the estimation of risk. The expectation for the PCW in cluster i per the empirical Bayes method is:

$$\mathrm{E}(\mathrm{PCW}_i) = \mathrm{SPF}_i * \mathrm{w}_i + \mathrm{PCW}_i * (1 - \mathrm{w}_i)$$

$$\mathrm{w}_i = \frac{1}{1+\mathrm{SPF}_i/\mathrm{a}}$$

The map is again presented (Figure 14) yet this time the cluster radii are proportional to the expected count of PCW events as determined by the empirical Bayes model. As this method incorporates exposure both in terms of “crossings” and in terms of exposure to pedestrians’ presence - it may be used as a robust method to determine the radii of the clusters and point out the clusters that are safety wise “more important”. Note that in this presentation the cluster in the bottom left part of Figure 3-14 is again large (similar to Figure 3-8 and unlike Figure 3-10) as pedestrians’ presence plays an important role.

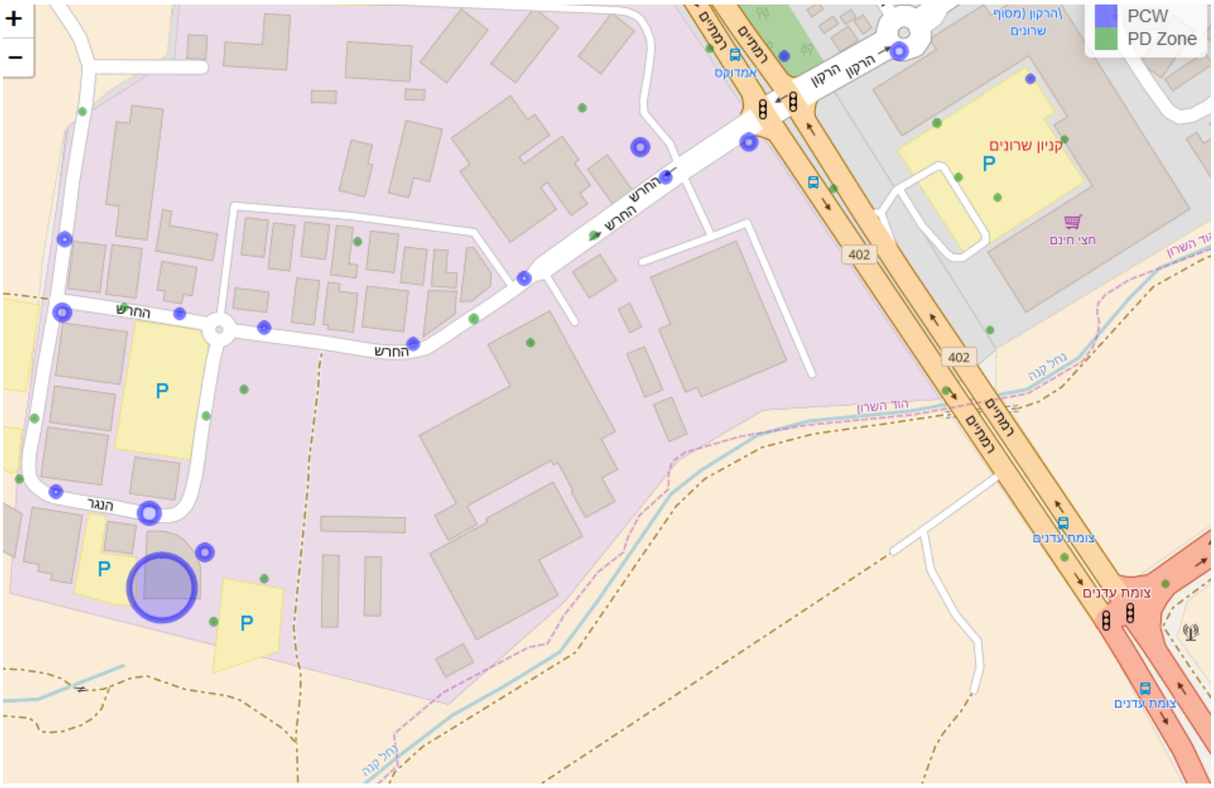


Figure 3-14: Clusters presented on a map where cluster radii are proportional to expected count of PCW events

3.4 Conclusion

The study presented in this chapter is by no means complete, however, it provides a good demonstration of the potential of data collected by simple and low-cost technologies and provides ample data for important analysis. It demonstrates how two relatively simple low-cost technologies, an IVDR and the Mobileye system, can be used to monitor and analyse driving behaviour, and to map noteworthy spots .

The analysis is based on data collected from 13 vehicles of the Or Yarok fleet over a period of 15 months, and includes events recorded by the two technologies, trip data (e.g. start time and end time of the trip) and the GIS layer of roads in Israel.

The events generated by the two systems are temporally synchronized. This is achieved by letting the Ituran system read the display protocol of the Mobileye system, and generate a time-wise synchronised flow of events.

The Mobileye and the IVDR technologies identified and recorded a total of 130,010 driving events. The events recorded by the IVDR system correspond to G-force type of events and include: braking, acceleration, lane change, speed bump, bypass, turns, turn while braking, turn while speeding and speeding. The events recorded by the Mobileye system correspond to vision based type of events and include: headway keeping, forward collision warning (FCW), lane departure warning (LDW), pedestrians in danger zone (DZ) and pedestrians' collision warning (PCW). It is important to note that although the Mobileye system is vision-based and includes a camera, no visual documentation of the driver's view is provided by it.

In the analysis presented in this chapter, unlike conventional naturalistic studies, the feedback from the two systems was available to drivers, either through a warning alert, or a signal on event occurrence. Hence the analysis could have been influenced by drivers altering their behaviour as a response to that feedback. However, for future studies – the option of not providing real-time feedback can be easily implemented.

This study shows that IVDR and the Mobileye system can be used to get a descriptive view of safety critical events, G-force based and vision based, their frequency of occurrence and the speed at which they occurred.

Next, the analysis of the temporal characteristics shows insights in sequences of events. Sequences of events can reveal an important picture of drivers' behaviour. For example, a high frequency of consecutive headway events corresponds to consistent behaviour of close following.

The Mobileye system is also useful to study events related to interactions between drivers and pedestrians, as conveyed through the Mobileye visual based events PCW and DZ. It was found, for example, that at the close period (7 seconds) surrounding the pedestrian potential conflict indication (PCW), drivers were relatively careful and avoided performing events from any other kind.

Another very interesting analyses possibility is the spatial analysis. Using location information (GPS coordinates), it is possible to geographically map the events and investigate infrastructure characteristics related to the occurrence of these events. For demonstration, we focused on events related to pedestrians (PCW and DZ). First, we located the events on the map and clustered them according to the distance from one another. This simple analysis already helped reveal locations that are candidates for potential conflicts. Next, we performed the same analysis, but accounted also for exposure, namely for how many times drivers actually traversed those clusters. Such analysis provides a different and valuable view of potentially risky locations. Finally, the study illustrates that it is possible to develop a model predicting occurrence of PCWs to the analysis, which give further insights in risky locations.

4 The feasibility and potential of smartphone apps to monitor secondary tasks and distracted driving

Motivation

Smartphone usage while driving, a prominent type now-days of secondary tasks, has become a major concern in the area of road safety and is likely to remain a key issue as smartphone usage continues its global rise. Recent studies clearly indicate that activities related to mobile phone, such as texting, dialling, conversation, involve all types of distracted driving: cognitive, manual, visual, and audible. Of the many types of smartphone usages, texting was found to be the most risky behaviour (Dingus, 2014, Klauer et al., 2014, Victor et al., 2014; Delgado et al., 2016; Kinnear and Stevens, 2016). Clearly, texting while driving negatively affects lane position control, reaction time, speed and headway deviation (Fitch et al., 2013; Owens et al., 2011; Yager, 2013; He et al., 2014). Recent studies show that texting results in drivers taking their eyes off the road for an average of 23 seconds in total and increase the risk of a crash or near-crash by two times (Hedlund, 2011; NHTSA, 2015). Kinnear and Stevens (2016) also point out that the impact of distraction on safety depends on the exposure time; that is the duration of distraction and how frequently it occurs. No doubt, in an era in which the extent of smartphone addiction increases to an alarming extent, its implications on road safety are terrifying.

While it is widely agreed that distracted driving poses a serious risk, one of the main challenges is how to evaluate and objectively measure it. Traditional studies are based on self-reports (see for example NHTSA, 2015 and the references there). However, while having well-recognized advantages, such as its ease of use and the ability to collect large data sets relatively easily it suffers from limitations regarding its validity as an indicator of actual behaviour, and it does not use technology. More advanced methods are based on controlled simulator studies (Yannis et al., 2013; Yager; 2013; He et al., 2015) and on driving in an instrumented car equipped with in-vehicle systems (Owens et. al. 2011; Reimer et al., 2016) but, their capability to collect large data sets is very limited. There is no doubt that naturalistic studies have the greatest potential to evaluate and objectively measure distracted driving, but they are extremely complex to conduct and require extensive resource allocation.

While being a major cause of risk and a key source of distracted driving, smartphones apps may also serve as a means to monitor, control and reduce risky driving behaviour (Albert et al., 2016). The greatest advantage of smartphone apps as countermeasures is their low cost and wide availability. It is expected that this trend will become even more widely spread and used. For the specific detailed driving behaviour analyses in SP4 smartphone apps may not be suitable. However, they could be very valuable in collecting selected performance measures for future monitoring and can be easily collected on a large scale.

Objective

The objective of this section is to demonstrate how data collected from an existing app can be used to measure and analyse distracted driving as portrayed through smartphone use while driving. The smartphone app is downloaded to smartphones and is automatically activated while driving. It continuously monitors smartphones usage while driving and measures important indicators of distraction such as texting while driving. The uniqueness of this approach is that the measuring of the phenomena of distracted driving is evaluated using the same device which causes the distraction – the smartphone.

4.1 Method

In order to demonstrate the capabilities and the potential of collecting large-scale ND throughout a smartphone app we refer to a recent field study titled "Drive Mode" which has been conducted in Israel in January – October, 2016, and its comprehensive results are still being probed and analysed. Two hundred and sixty Israeli young drivers aged 17-24 (average age = 19.4 years old, SD=1.7) who hold a valid driver's license for at least three months and drive, on average, at least two hours a week, participated in this study. Among them, 147 participants (64%) were male and 113 (36%) were female. As can be seen in Figure 4-1 which presents participants residence locations on the map of Israel, young drivers from all areas of Israel participated in the study. This indicates the potential of conducting such study with no geographical constraints as long as participants own a smartphone.

Each participant was driving for a period of four months in his or her family owned car. However, not all participants started at the same point of calendar time. Participants installed a smartphone app called ProtextMe®, which was adjusted and configured especially for this study, on their own smartphones. For each participant, in the first month of the "Drive Mode" study, only monitoring of actual smartphone usage while driving has been performed by the app. In the following two months participants were introduced to various interventions aimed to mitigate smartphone distraction; that is, for each participant actual smartphone usage has been monitored under a specific intervention. During the last month of the study no intervention has been presented and only monitoring has been performed.

Figure 4-1: Participants' residence locations

4.2 Monitoring technology

ProtextMe® is a smartphone app available for free download in Google play for Android mobile operating system[2]. The app has been adjusted and configured especially for this study. Each participant has a unique user name and a password to log in to the application.

In its research oriented version the app is continuously monitoring smartphones usage while driving. The monitoring is automatically activated once the app is activated, through a smart detection of driving via several indicators (e.g. Bluetooth connection, GPS and NFC identification) thus indicating both the start and the end of the journey. The monitoring includes (depending upon smartphones' type and operating system): foreground apps, time stamps and speed. Most importantly: the monitoring captures the number of times the young driver is actually touching the smartphone screen. If for some reason, the app starts monitoring when the young driver is not driving, then he/she are asked to turn off the app by marking the "I am not driving" button. In order to respect privacy and increase acceptance among users, all content was absolutely not monitored. All the collected information is on-line transferred to a central server, where all the data is stored. The multiple collected parameters includes: identification of apps used by drivers while driving, which apps are appearing in the foreground while driving, incoming/outgoing calls durations, speed and more.

4.3 Collected data

During the first phase of the "Drive Mode" study (matching the first 4 weeks of monitoring) almost 10,000 hours of driving have been collected corresponding to a total of 34,450 trips (with a minimum duration of 3 minutes each) which have been undertaken by the 260 participants. Approximately two millions (!) of notifications were captured and a total of 450,028 screen touches have been performed. As mentioned previously, this is an on-going research. However, in order to demonstrate the type of results which can be gathered by the app and used for evaluating smartphone use while driving we present here part of the analysis related to the monitoring phase (i.e., the first month for each participant in "Drive Mode" study).

4.3.1 Trip characteristics

During the first monitoring phase of the study, i.e., the first month for each participant, a total of 11,528 trips (with a minimum duration of 3 minutes each) were undertaken by 254 participants (6 participants out of the 260 had no monitored trips in this phase) and 3,304 hours of driving have been collected. In 9,274 trips (80%), the GPS feature on the smartphone was switched on, and therefore also speed data have been collected. Table 4-1 presents summary statistics of the trips performed by these 254 participants.

[2] https://play.google.com/store/apps/details?id=ok.driver_care_auto

Table 4-1 Characteristics of the trips undertaken by the participants

	Average	SD
No. of trips per participant	45.4	47.3
Trip duration (minutes)	18.2	9.0

These statistics regarding driving patterns and trip characteristics are in line with figures reported from other naturalistic studies which dealt with young drivers in Israel (Toledo et. al. 2014, Albert et al., 2014), indicating that a simple and low-cost tool such as a smartphone app can provide valid results regarding driving exposure.

4.4 Distraction evaluation based on screen-touches

In this section distracted driving is evaluated based on the characteristics of actual smartphone screen-touches while driving. Moreover, we refer, when applicable, to the speed in which the screen-touches occurred which may serve as an indication to extreme risky driving behaviour, to foreground apps and to time stamps. Clearly, while the driver touches the smartphone screen, his/her eyes are off the road.

4.4.1 How many times do young drivers touch their smartphone screens?

According to the study's results, young drivers touch their smartphone's screens on average 1.6 times per minute of driving (SD=1.4, min=0.0, max=6.7). The distribution of the number of screen-touches per minute of driving among all participants is presented in Figure 4-2.

As can be seen 109 participants (that is, 43%) touched the smartphone screen, on average, not more than once per a minute. It should be noted that many types of smartphone usages involve touching the screen (even performing phone calls which is legal), and therefore "zero screen-touches while driving" may be unrealistic. However, higher averages, as revealed by most of the participants may lead to worrisome consequences from a safety point of view.

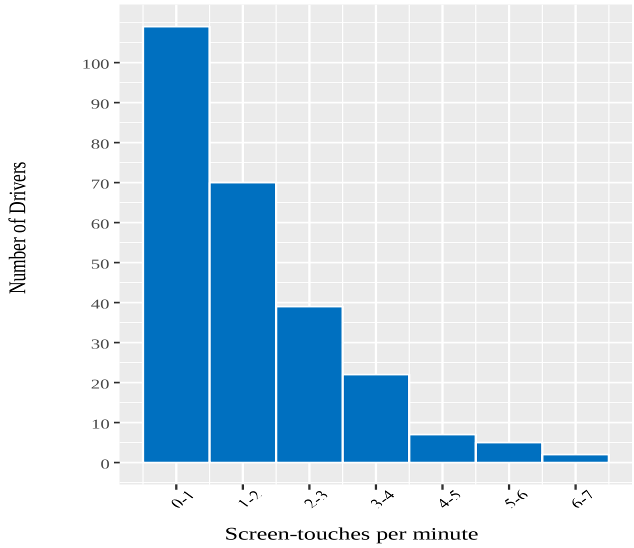


Figure 4-2: Distribution of average screen-touches per minute of driving (N=254)

4.4.2 What are the prevalent usages of smartphones while driving?

According to our findings, a total of 324,726 screen touches have been performed during the study period. The monitoring enables to capture also the type of smartphone usage in which the screen-touch occurs. As expected, the most common usage relates to apps. However, due to technological issues such as android operation system and smartphone types, in 53% of the screen touches, the monitoring could not specify the type of app. In the rest of the 47% of the screen touches, that is 150,868 screen-touches, the specific use has been monitored, and the distribution among the various smartphone usages is shown in Figure 4-3.

As is notably seen from the figure, the majority of the screen-touches (45%) were performed while using WhatsApp, a popular free messaging app. In Israel, the usage of WhatsApp is widely spread; a survey from 2013 showed that WhatsApp is installed on approximately 92% of all smartphones and about 86% of users reported using it daily or almost daily (Globes, 2013). These figures may be even higher for young adults. Waze, the most common navigation app in Israel, accounts for 9% of the screen-touches and Facebook, the well-known social network accounts for 8% of the screen-touches. Alarmingly, the YouTube app, which enables to watch videos, accounts for 7% of the screen-touches.

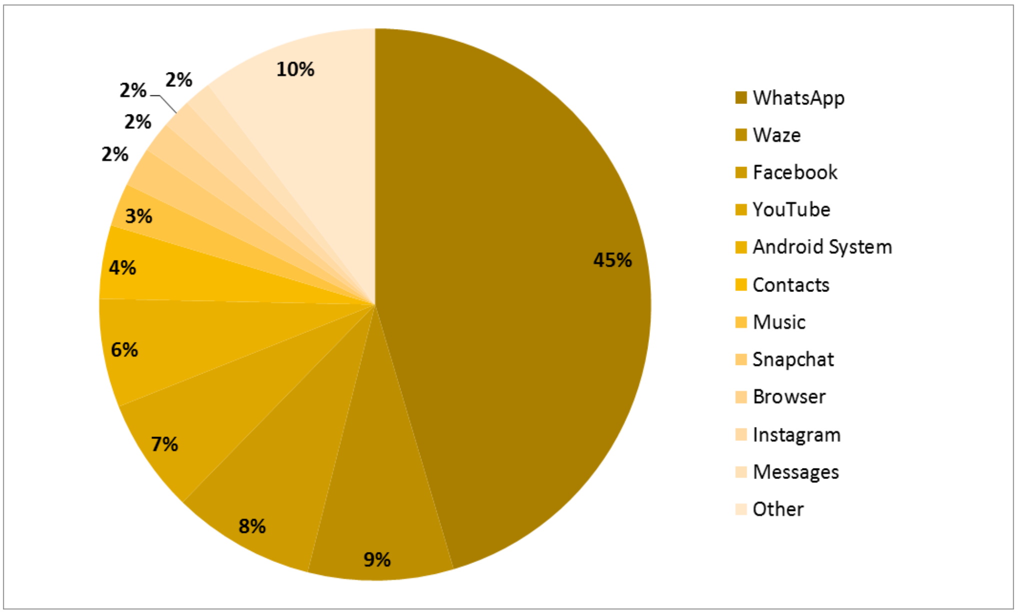


Figure 4-3 Distribution of screen-touches according to the various smartphone usages

It is also interesting to note that the "traditional" usage of phone for making and receiving phone calls is not that common among young drivers. According to our findings, during the study period, a total of 4,232 outgoing calls were made and a total of 1,961 calls were received. The duration of these calls is short: the average is 84 seconds (both for an outgoing call for an incoming call).

4.4.3 Does trip duration affect smartphone screen touches?

As reported earlier, a total of 11,528 trips were undertaken by the participants in the first monitoring phase of the study, with an average duration of 18.2 minutes for a trip. Only 1,710 trips (15%) were longer than 30 minutes, 4,937 trips (43%) took 10-30 minutes, and 4,881 (42%) were shorter than 10 minutes. Figure 4-4 presents the distribution of screen-touches by fraction of trip duration for long trips, medium, and short trips. As expected, longer trip durations account for higher number of screen touches.

Some interesting insights are revealed from Figure 4-4; As can be seen, regardless the trip duration, screen-touches occurred throughout the trips. However, it seems that at the beginning of a trip (first tenth of its duration) more screen-touches are performed. This pattern can be explained by typical behaviour associated with trip start such as typing an address in the navigation app. A more careful analysis supports this: during the first tenth of trip duration 22% of the screen touches were performed in in Waze compared to 9% in total.

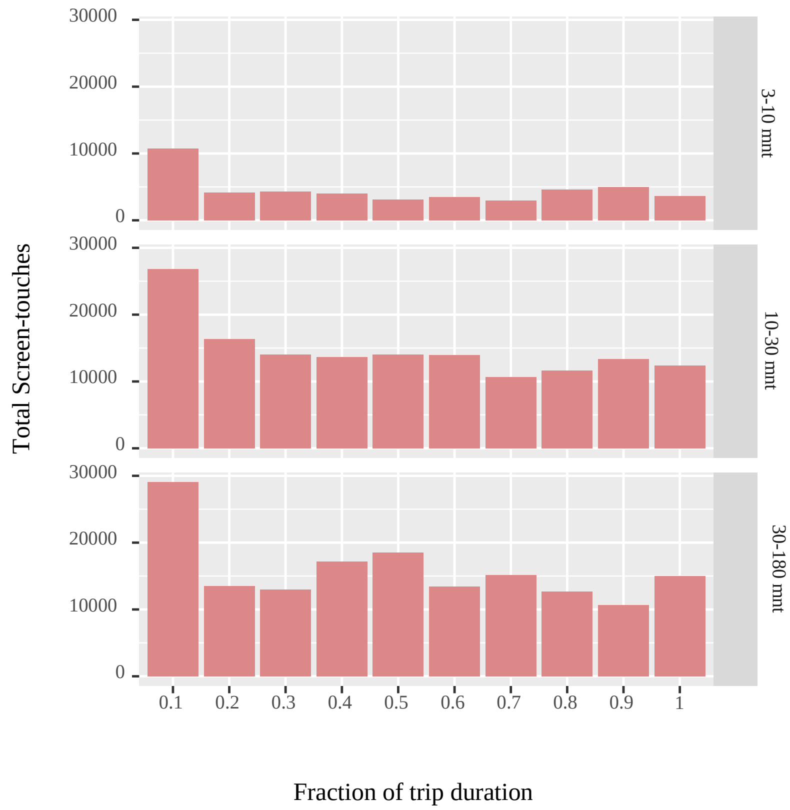


Figure 4-4: Distribution of screen-touches by fraction of trip duration for short, medium and long trips

4.4.4 Do young drivers touch their smartphone screen while the vehicle is in motion?

As mentioned, in 9,274 trips (80% of the total number of collected trips), the GPS feature on the smartphone was switched on, and therefore also speed data was monitored and documented. Those trips reflect 2,555 hours of driving (77% of the total duration driving time) and include 249,403 screen-touches (77% of the total number of screen-touches).

Those 9,274 trips were performed by 241 participants (that is, 13 participants drove with GPS switched off in all their trips). Figure 4-5 shows the distribution of the number of screen-touches by speed in those trips. As can be seen, about half (51%) of screen-touches are performed when the vehicle speed is less than 5 km/h, that is, probably in traffic light, traffic jam, etc. However, about half of the screen-touches occurred while the vehicle is in motion, a behaviour which is extremely dangerous from a safety point of view. As is alarmingly revealed from the figure, screen touches were also performed at high speeds and even at speeds higher than 100 km/h.

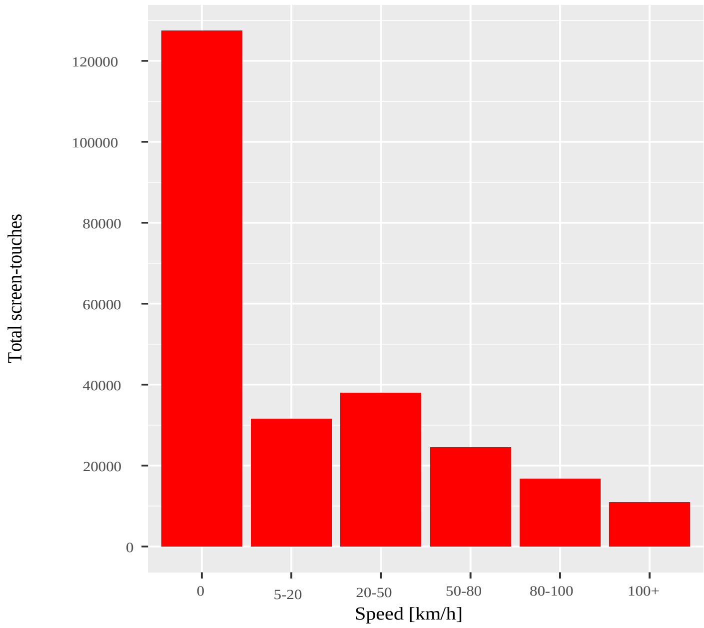


Figure 4-5 Distribution of screen-touches by speed

4.5 Conclusion

This study demonstrates how a smartphone app can provide valuable information and data regarding the occurrence of various smartphone uses while driving. The uniqueness is that the measuring and evaluation of the phenomena of distracted driving is performed using the same device which causes the distraction – the smartphone. This is done by installing an app which captures secondary tasks by continuously monitoring smartphones usage while driving. The easy implementation of such platform, which requires only access to smartphone indicates its potential to be widely implemented without geographical constraints.

As shown by the information gathered relating to trips characteristics, this simple and low-cost tool can provide valid results regarding distracted driving, as manifested by number of screen-touches performed by drivers while driving and their characteristics (e.g. frequency, speed, relevant apps and duration). As with any new approach its verification and validation with actual smartphone usage while driving should be further probed. Clearly, technologies are evolving rapidly, and so are their capabilities to address features such as driver's identification and detailed monitoring of smartphone usage. Hence it is expected that in the near future even more relevant information could be collected including additional information from other sensors embedded in smartphones (such as: GPS location indicators, accelerometers and camera).

4.6 References

Albert, G., Lotan, T., Toledo, T., Grimberg, E., Lasebnik, M. (2014) Are young drivers as careful as they deem? In Vehicle Data Recorders and Self Reports Evaluations. *European Transport Research Review*, 6(4), 469-476.

Albert, G., Musicant, O., Oppenheim, I., Lotan, T. (2016) Which smartphone's apps may contribute to road safety? An AHP model to evaluate experts' opinions. *Transport Policy*, 50, 54-62.

Delgado, M. K., Wanner, K. J., McDonald, C. (2016). Adolescent cellphone use while driving: an overview of the literature and promising future directions for prevention. *Media and communication*, 4(3), 79-89.

Dingus, T. A. (2014). Estimates of prevalence and risk associated with inattention and distraction based upon in situ naturalistic data. *Annals of advances in automotive medicine*, 58, 60.

Fitch, G. A., Soccolich, S. A., Guo, F., McClafferty, J., Fang, Y., Olson, R. L., Perez, M. A., Hanowski, R. J., Hankey, J. M., Dingus, T. A. (2013). *The impact of hand-held and hands-free cell phone use on driving performance and safety-critical event risk*. Report No. DOT HS 811 757.Washington, DC: National Highway Traffic Safety Administration.

Globes (2013) (http://www.globes.co.il/news/article.aspx?did=1000893814 (In Hebrow) Accessed Sep 29th, 2016.

He, J., Choi, W., McCarley, J. S., Chaparro, B. S., Wang, C. (2015). Texting while driving using Google Glass™: promising but not distraction-free. *Accident Analysis & Prevention*, 81, 218-229.

Hedlund, J. (2011). *Distracted driving: What research shows and what states can do.*

Kinnear, N. Stevens, A. (2016). *The battle for attention – driver distraction, a review of recent research and knowledge*, TRL.

Klauer, S. G., Guo, F., Simons-Morton, B. G., Ouimet, M. C., Lee, S. E., Dingus, T. A. (2014). Distracted driving and risk of road crashes among novice and experienced drivers. *New England journal of medicine,* 370(1), 54-59.

NHTSA (2015). *Driver electronic device use in 2014. Traffic safety facts research note*. National Highway Traffic Safety Administration.

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

Reimer, B., Mehler, B., Reagan, I., Kidd, D., Dobres, J. (2016). Multi-modal demands of a smartphone used to place calls and enter addresses during highway driving relative to two embedded systems. *Ergonomics,* 59(12), 1565-1585.

Toledo T., Farah H,. Morik S. Lotan T. (2014). Driving exposure of Israeli young male drivers within a graduated driver licensing system, *Transportation Research* F, vol 26, part A, 180-189.

Victor, T., Dozza, M., Bärgman, J., Boda, C. N., Engström, J., Markkula, G. (2014). Analysis of naturalistic driving study data: safer glances, driver inattention, and crash risk.

Yager, C. (2013). An evaluation of the effectiveness of voice to-text programs at reducing incidences of distracted driving. College Station, Texas: Texas A &M Transportation Institute, the Texas A&M University System.

5 Which research questions can be addressed by what technologies?

The resources and efforts associated with major naturalistic studies, such as the American SHRP II and the European UDrive, are substantial and can not be repeated and supported frequently, maybe noteven more than once in a decade (or a life time..).

Naturally, the wealth and richness of the integrated data, gathered by such extensive studies, can not be compared to data collected via simpler, sporadic data collection technologies, such as the ones described in this report. However, many Research Questions (RQs) can be addressed, at least to some extent, by other low-cost and simpler technologies. This discussion is important, not only in order to replace the honourable place (and cost!) of naturalistic studies, but also to complement and enable their continuity after they are completed.

The underlying understanding is that there is no real substitute to a car equipped with 7 cameras and ample data loggers and sensors. Still, given the rapid advancement of technology – many interesting and relevant RQs can be addressed with much less sophisticated technologies. It is the aim of this current chapter to explore this issue. It is important to keep in mind that part of the strength of large naturalistic studies (such as UDrive) is not merely the actual technologies, which are used, but their integration, and hence, the ability to (almost) fully explore the circumstances of the driving instances. Consequently, when addressing the potential of simple (isolated) technologies, as those described in chapter 2, the option of combined technologies will be explored as well. Most of the systems, which are described in chapter 2 refer to a specific driving behaviour that is being monitored (i.e. speeding, lane keeping etc.) and the use of certain thresholds or triggers to single out risky situations related to that behaviour. However, once those instances are detected – no information on the circumstances leading or accompanying this behaviour are available. Typically, visual information (discrete or preferably continuous) is needed to fully understand the circumstances of that behaviour. Hence, when considering the potential of simple low-cost technologies to address RQs – the possibility to combine and integrate such systems will also be discussed, even if currently such an option is not available. One of the most conceptually straightforward integrated systems that comes to mind is a system, in which the basic technology detects the desired behaviour (e.g. harsh braking) and triggers a simple continuous dash-board camera to save the relevant circumstances, which accompany that behaviour.

It is also important to keep in mind that many of the systems described in the sections of this chapter are designed to serve as alerting systems or specific-goal oriented systems. For the sake of data storage, typically, a dedicated algorithm and routines need to be developed and added on to the already existing system.

5.1 The UDrive Research Questions

UDrive addresses many RQs. Due to the richness of the collected data, it would have been possible to address even more questions.

The major UDrive RQs can be roughly categorised into the following groups:

1. Risky behaviour
2. Engagement in secondary tasks
3. Interactions of drivers with cyclists and pedestrians
4. Risky behaviour related to Power Two Wheelers (PTW)
5. Eco-driving

Next, each of these groups will be discussed.

5.2 Risky behaviour

Risky behaviour has many manifestations. In the following section, we will refer to driving behaviour as depicted through speed choice, close following, lane keeping and abrupt or aggressive driving.

Speed

Among all forms of risky driving, speed choice plays an important role and is addressed in many of the UDrive RQs either as a main or as a contributing factor. Speed choice and speed distribution can be easily measured and recorded by most technologies surveyed in chapter 2, such as: mobile phone applications, telematics devices and built-in data loggers. In order to address RQs related to excessive speed or speed that exceeds the speed limit, information regarding the actual speed limit is required. This can be achieved through dedicated applications (e.g., updated map information or navigation systems) or more advanced features (e.g., the ability to read posted speed signs, as performed by the MobilEye system). In order to understand the circumstances of speeding, such as cases in which speeding takes place to avoid risky situations – a visual (and preferably continuously visual) video data is required. In most simple nomadic systems, this information is not available, but could be retrieved by integrating a simple dashboard camera that continuously records and is automatically triggered to save the time segment around the speeding incidence. The location of the various speed choice RQs on the FNS are illustrated in Figure 5-1.

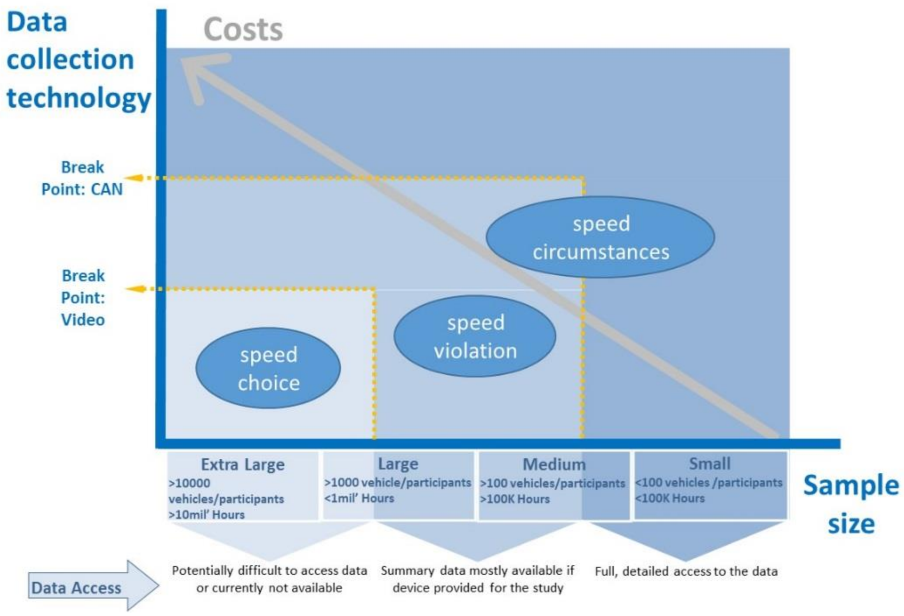


Figure 5-1: Location of speed choice RQs on the FNS

Headway and lane keeping

Headway and lane keeping behaviour are indicators of risky behaviour. As described in chapter 2 and tabulated in

Table 6-1, these behaviours can be detected and recorded using several technologies, ranging from simple mobile phone applications, through enhanced dash-cameras, to the more sophisticated compound systems. The specific features in all of the above mentioned types of technologies need to include visual capacities combined with some form of pattern recognition. There is a clear trade-off between the simplicity and the reliability of the technology. A well-calibrated Mobileye type system is more robust and reliable than a mobile phone application, which requires the phone to be attached to the front windshield. Distance keeping and forward collision instances depend on the distance from the leading vehicle. Not all systems that are built to identify these instances can handle each vehicle type (such as PTWs and bicycles) and operate in all lighting conditions (such as darkness, rain and fog).

For example: *Do car drivers keep PTW riders at a different distance (i.e., time headway) than other motorised traffic on straight road sections?* This question can be addressed by a nomadic version of Mobileye that measures relative distance and can identify PTW and cars in all lighting conditions.

Regarding lane keeping behaviour, systems that identify lane departure are typically based on visual recognition of lane markings and hence, are sensitive to the quality of the markings and lighting conditions. Moreover, naturally, there is a need to differentiate between intentional and unintentional lane departures. In order to identify unintentional lane departures – the easiest way is for the system to connect to the vehicle's signal indicator, provided that drivers indeed signal their intentional lane departures. If drivers do not use their signals to mark intentional lane departures, then the detection of unintentional lane departures becomes much more complex. For that reason, a detailed view and analysis of the driver's behaviour needs to be monitored and interpreted. For example, if the driver is falling asleep or is engaged in a secondary task – then lane departure can be defined as unintentional based on the verification of that specific condition. This requires the installation of cameras, which face the driver as well as a rigorous annotation procedure.

Figure 5-2 illustrates the location of the various distance and lane keeping RQs on the FNS.

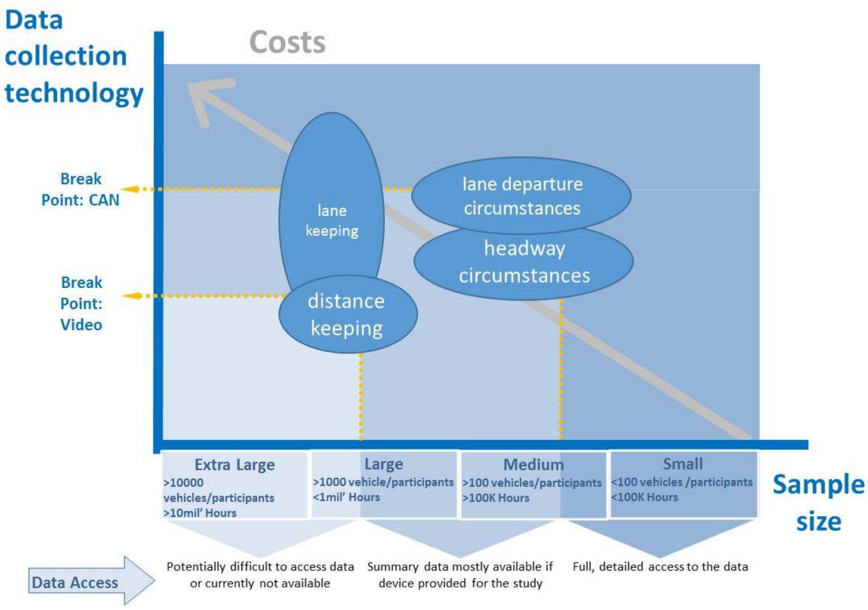


Figure 5-2: Location of distance and lane keeping RQs on the FNS

Abrupt and aggressive behaviour

Abrupt and aggressive behaviours are typically associated with G-force based events corresponding to actions such as: harsh braking, excessive accelerations, abrupt lane changes and sharp cornering. These behaviours can be easily detected by simple technologies such as: telematics devices, simple mobile-phone apps and data loggers. Once these types of behaviour are detected, a more elaborated technology involving visual documentation of the driving environment is required to understand the circumstances of the detected aggressive behaviour, as well as the driver's driving manoeuvres, which typically follow the aggressive driving behaviour.

Figure 5-3 illustrates the location of the different kinds of aggressive behaviour related RQs on the FNS

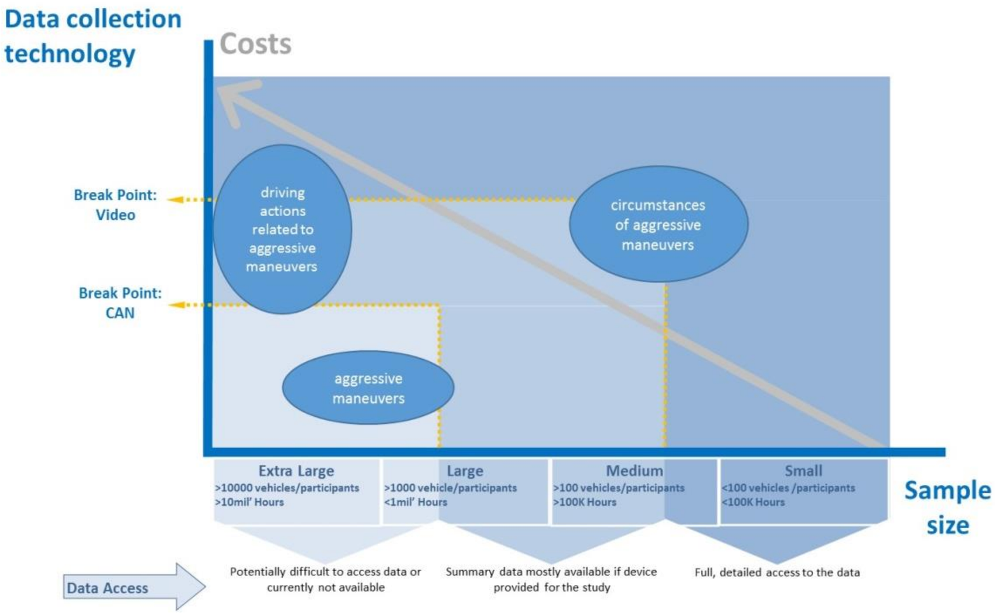


Figure 5-3: Location of aggressive behaviour on the FNS

5.3 Secondary tasks

Distraction, inattention, mind wandering and engagement in secondary tasks are all related to one of the most prominent and disturbing issues in road safety today – the fact that drivers do not devote themselves fully to the driving task. This phenomenon is nurtured by the existence of advanced communication platforms, such as mobile phones and society's expectation to continuously be available. Curiosity, inability to resist glancing towards incoming notifications and the need to constantly check the favourite communication networks (e.g. email and social networks) for updates – have created behavioural patterns that are on the verge of an addiction.

Unlike other risky behaviours such as speeding and driving under the influence of alcohol, successful counter-measures, best practices and effective policies are still far from being established and widely implemented. Consequently, the safety community worldwide struggles to evaluate the extent of the phenomenon, measure it objectively and derive effective counter-measures and policies to mitigate it.

One of the major obstacles to evaluate the extent of the problem, is the lack of validated evidence to actual engagement in secondary tasks. In that context, it is quite evident that naturalistic studies can help establish such evidence.

UDrive, as a large naturalistic study, deals with secondary tasks extensively. Many RQs related to distraction and secondary tasks can be addressed via the UDrive database and its platform of analysis. The UDrive RQs, which address secondary tasks vary. They include simple inventory type of questions such as: *What is the prevalence of secondary task activity in normal driving? or What specific types of secondary tasks do drivers typically engage in*. Further, UDrive addresses more complex RQs regarding the relationship between driving task complexity and secondary task engagement, such as: *To what extent do driving task complexity and secondary task complexity influence the decision to engage in secondary tasks while driving?* . Furthermore, there are questions that address the adaptation of driving behaviour while engaging in secondary tasks, i.e.: *To what extent do drivers adapt their safety margins while performing secondary tasks?*.

Clearly, the most valid way to determine that a driver engaged in a secondary task – is to actually watch him/her doing it. This requires some form of visual presentation and interpretation of the driver’s actions and gaze behaviour. A dashboard camera facing the driver can provide such information. Naturally, the quality of the camera and the camera's ability to monitor and document actual glance behaviour under various conditions (e.g. darkness, a driver wearing sun-glasses), plays a major role in the ability to detect involvement in secondary tasks. Once such a camera is available and operational – the next challenge is to screen the excessive data it scans into meaningful information regarding the engagement in secondary tasks. Technically, it is possible to continuously collect and store the continuous visual data (even simple dashboard cameras can easily store up to 10 hours of recorded data) and then manually scan it. This is clearly time and storage consuming and not very efficient. A more efficient configuration is to integrate the camera with some trigger mechanism, so that the visual data is stored only for instances that are candidates for risky behaviour. For those cases only, time segments around the risky behaviour (typically approximately 15 seconds before and 10 seconds afterwards) are marked and stored. Possible triggers for such screening can be G-force-based (e.g. excessive braking), visually related (e.g.: lane departure) or based on other risky behaviour detection (e.g. proximity to pedestrians).

An eye-tracking system is an even more sophisticated system, as it can verify that the driver was actually looking at the object that was engaged with the secondary task. These systems, in their more sophisticated versions, can be programmed to identify specific glance fixations and hence, can effectively scan and store relevant data for further analysis.

Other approaches to detect involvement in secondary tasks can be geared towards the specific device, which is related to the secondary tasks. Such an approach was demonstrated in chapter 4, where a simple smart-phone app was installed and adjusted to collect data on actual smart-phone usage while driving. When visual data on the driver is not available, a special effort to identify the driver needs to take place. As with most smart-phone applications, the driver is typically in charge of the phone and can decide to stop the data-collection mode at all times.

Location of the various RQs related to secondary tasks on the FNS is illustrated in Figure 5-4.

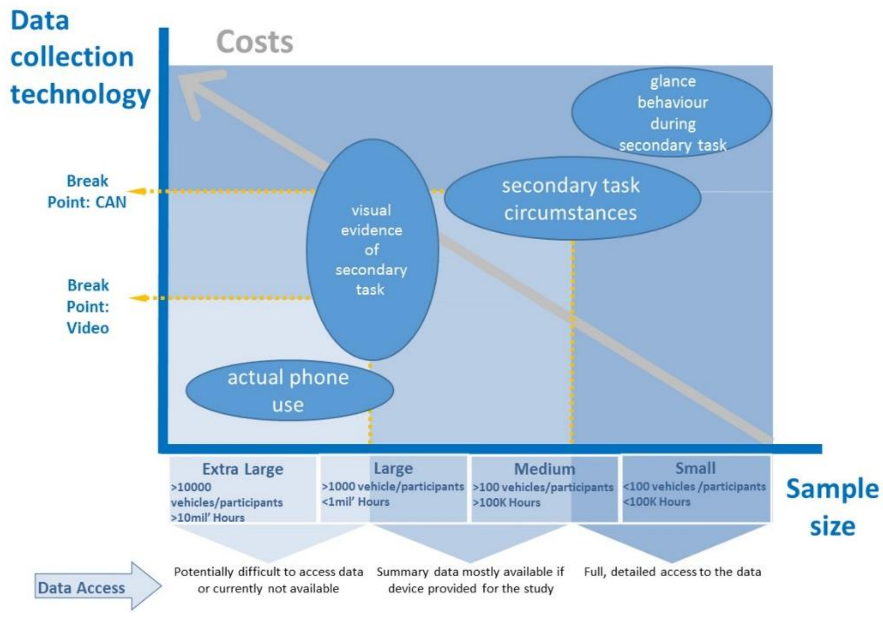


Figure 5-4: Location of RQs related to secondary tasks on the FNS

5.4 Interactions of drivers with cyclists and pedestrians

The detection, monitoring and recording of drivers' interactions with VRUs, and specifically pedestrians and cyclists, require dedicated effort and integration. At its simplistic form, this effort can be achieved through simple and enhanced dash cameras. However, without triggers for VRUs presence, scanning of the raw data generated by the cameras requires significant efforts and resources related to both storage and manual scanning of the data.

In chapter 3, an analysis of VRU presence and conflicts, which were detected through two technologies was presented. The technologies were: a G-force based IVDR and a Mobileye system. The core of the analysis relied on the ability of the Mobileye system to detect VRUs on two major levels: presence of VRUs and a potential conflict with VRUs. The analysis included temporal and spatial analysis of VRU-related events with focus on sequences of events leading to conflicts. Association of the location of the VRU-related events to infrastructure characteristics was also demonstrated. However, in the analysis presented in chapter 3, no information on the detailed circumstances of the conflicts was available. For example, when a major conflict with a pedestrian was identified, it was possible to view the speed of the vehicle during that specific time segment, the location of the conflict, as well as the braking and acceleration behaviour of the driver before and after the exact time of the conflict. However, it was not possible to identify, which circumstances led to the conflict: i.e., was the conflict due to a pedestrian jumping unexpectedly into the road or due to the distractiveness of the driver who didn't notice the pedestrian on time?. Therefore, relying on data generated by the types of technologies used in chapter 3, it can provide important aggregated data, but cannot account for detailed disaggregate understanding and analysis of specific cases.

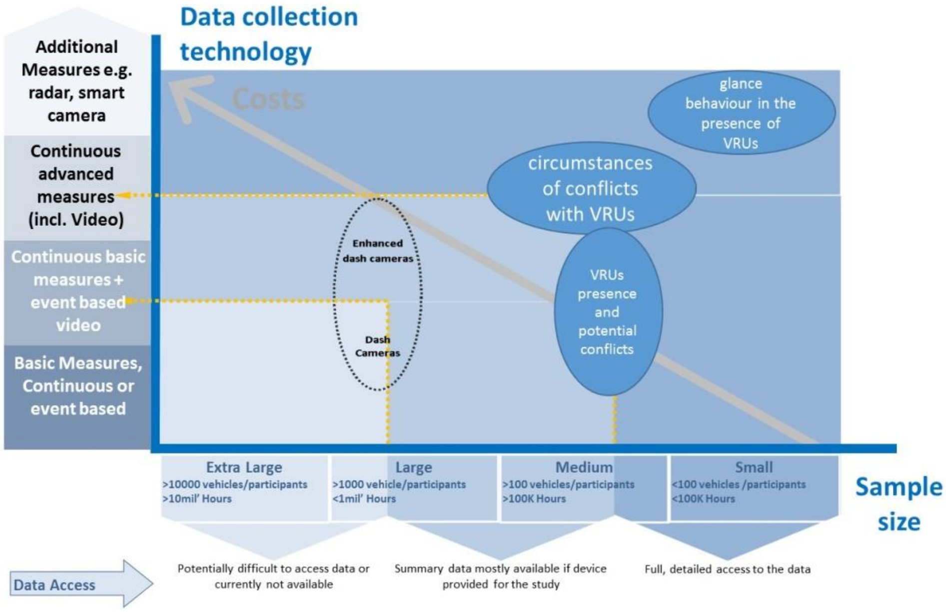
If the data collection configuration presented in chapter 3 had been augmented by a dash camera, enhanced cameras and integrated into a compound system, then a more detailed, informative and disaggregate analysis could have been performed.

It is important to mention that in order to monitor and analyse VRU related data – a technology that can detect VRUs presence and (relative) distance from the driver is highly recommended. Note, that in order to address RQs on normal driving behaviour, it is enough to have the ability to detect VRUs presence (not necessarily potential conflicts).

In order to address specific detailed RQs, such as: *Where do car and truck drivers look when making a right turn manoeuvre in urban areas, and which factors (i.e., infrastructure, traffic situation, distraction) influence this behaviour?*, clearly, a dedicated integrated compound system is required. This system needs to include, in addition to the basic data collection features such as speed and location, also data related to the identification of right turns, but most importantly: a visual view of the drivers' head gaze and glances (either through a camera facing the driver and/or via an eye tracker). Even more complex DAS is needed if there is a need to monitor drivers' behaviour regarding detection of VRUs in their blind spot. This would require a dedicated camera and/or eye tracker, which will be integrated and synchronised with the other sensors of the DAS.

Of special interest is the following RQ dealing with pedestrians' conflicts: *Would warnings generated by pedestrian detection systems could have had the potential to reduce conflicts?* This question is of high importance to the evaluation of the effectiveness of advanced technologies and their potential to improve pedestrians' safety. Since warnings were not given to UDrive participants, this is an hypothetical question and hence, needs combined data sources and a decision making procedure. In UDrive, this question was addressed by performing detailed analysis of potential conflicts with pedestrians and analysing drivers' speed, acceleration and gaze behaviour prior to the conflict. If, forinstance, a driver was engaged in a secondary task and hence, overlooked a pedestrian jumping into the road, then that is a clear case of potential effectiveness. However, many cases were more ambiguous than the latter and could not be easily determined. A smart compound system including a pedestrian detection mechanism integrated with a view of the driver and the road ahead can be used to address such complex RQs.

The location of the various RQs related to VRUs are illustrated in Figure 5-5. Note, that the figure presents three types of RQs (marked by the blue-background ovals), which focus on technologies that can identify VRUs at least to some extent. The oval corresponding to dash cameras and enhanced dash cameras appears in doted lines and corresponds to the cameras' ability to monitor the drivers' environment and encounters withVRUs, but in most cases, not very efficiently.



8.

Figure 5-5: Location of VRU related RQs on the FNS

5.5 Risky behaviour related to Power Two-Wheelers (PTW)

Safety and risky behaviour of and towards PTWs is a major issue in road safety. UDrive addresses PTWs both from the cars' and trucks' perspective but also from the perspective of riders of the PTWs. This was enabled through an extensive and dedicated DAS, designed especially for PTWs.

Many of the technologies described in chapter 2 and tabulated in Appendix A are not (yet) suitable for PTWs. Even the simple telematics boxes can not be mounted and collect data on PTWs, as lateral and longitudinal accelerations are strongly influenced by PTWs' manoeuvres. Similarly, smart-phone apps that need to be fixed to the car, or mounted onto the car's window – cannot be used for PTWs. On the other hand, wearable technologies are very suitable for riders of PTWs and some are already used extensively. For example, portable cameras are typically used by riders as a means to document their rides for the possibility that the evidence will be needed in case of a crash.

Two major RQs relating to PTWs are addressed in UDrive:

RQ_PTW_N1: *Which circumstances related to rider, infrastructure and trip have an impact on SCE occurrence?*

RQ_PTW_N2: *Do car drivers keep PTW riders at a different distance (i.e., time headway) than other motorised traffic on straight road sections?*

RQ_PTW_N1 is a very general research question and can be addressed on many levels. Hence, depending on the depth and complexity of the data needed for the analysis – a wide variety of technologies can be used to address this question. In order to determine SCEs, as was extensively explored in UDrive and other naturalistic studies, several triggers can be used. Moreover, many of the triggers can be available through simple technologies (e.g. telematics devices, mobile-phone applications, VRU detection, etc.). If dash cameras or enhanced dash cameras are used – then it is also possible to visually verify the occurrence of

SCEs and their circumstances. Clearly, dash cameras and enhanced dash cameras mostly provide views of the front view (and in some cases of the driver). More dedicated efforts can include use of portable cameras or cameras positioned to collect a wider angle of data (such as sideway views and blind spots).

Regarding data collected from the PTWs themselves – other than portable cameras, which riders typically wear on their helmets – more dedicated effort and adjustment of technology to fit PTW is needed in order to address RQs such as distance keeping of PTWs from the lead vehicle, detection of VRUs and more.

RQ_PTW_N2 is clearly asked from the point of view of the car drivers and hence, the technologies surveyed in chapter 2 are relevant. In order to address this specific question – a technology that monitors headway and can differentiate between cars and PTWs is needed. In UDrive, this is performed by using the protocols of the Mobileye system, which is part of the UDrive DAS. Hence, this RQ could be addressed by a nomadic Mobileye (or similar) system. The identification of straight road sections can be conducted by using speed as an indicator or a more detailed map indication.

The location of the various RQs related to PTWs are illustrated in Figure 5-6. Note that the figure presents three types of RQs (marked by the blue-background ovals) that could be addressed by car and truck based technologies. The light-blue oval corresponds to portable cameras worn by PTW's riders that can provide visual view of their rides.

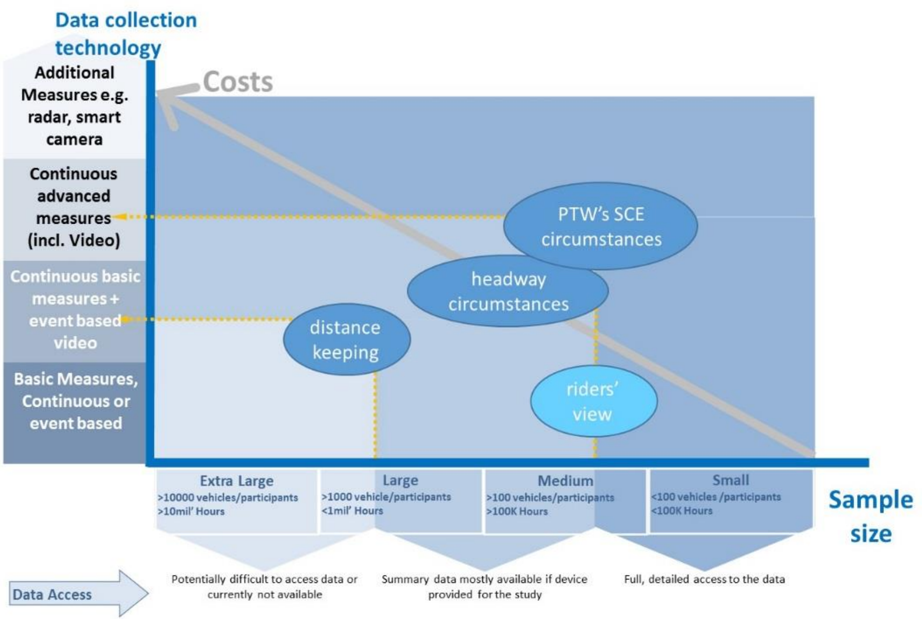


Figure 5-6: Location of PTW RQs on the FNS

5.6 ECO Driving

Most RQs related to ECO driving deal with the efficiency of fuel consumption under various scenarios of driving conditions and driving behaviour. Recently, a lot of scientific evidence has shown positive and high correlations between safe driving and fuel-efficient driving. The main parameters for both safe and efficient driving are related to speed, accelerations and smooth driving.

The UDrive RQs, which address ECO driving deal mostly with speed distribution and speeding, braking and acceleration behaviour, driving style as well as the relations between ECO driving and the behaviour of specific drivers. In order to answer these RQs, typically only G-force technology is required. Since no video data is required in order to answer the RQ's, it makes the analysis much more straightforward and simple. However, in order to validate the results, typically large quantities of data are required.

Hence, answering most RQs related to ECO driving by simple technologies, can be achieved through built-in data loggers and telematics devices. The location of these RQs on the FNS is presented in Figure 5-7**Error! Reference source not found.**. Note that the order of the categories on the Y-axis has changed to represent the relevance of CAN data in very large samples at a relatively low-cost.

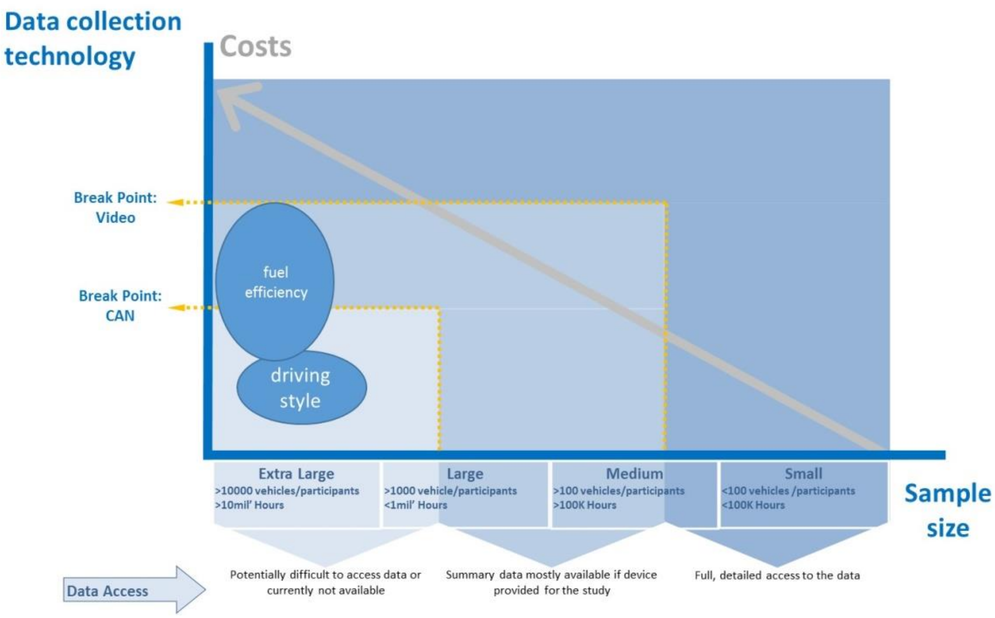


Figure 5-7: Location of ECO driving RQs on the FNS

6 Conclusion and Discussion

6.1 Conclusion

The underlying notion is that there is no real substitute to a car equipped with 7 cameras and ample data loggers and sensors, all integrated into a workable platform. Still, given the rapid advancement of technology,a vast amount of interesting and relevant RQs can be addressed with much less sophisticated and costly systems. It is important to keep in mind that part of the strength of big naturalistic studies (such as UDrive) is not solemnly the actual technologies, which are used, but their integration and hence, the ability to (almost) fully explore the circumstances of the driving instances. Consequently, very often, the potential of simple (isolated) technologies, as described in chapter 2, to address complex RQs, is limited.

Most of the simple systems, which are described in chapter 2, relate to specific behaviour that is monitored (i.e. speeding, lane keeping etc.). Additionally, certain thresholds or triggers are used to single out risky situations, which are related to that behaviour. However, once those instances are detected, no information on the circumstances leading or accompanying this behaviour are available. Typically, visual information (discrete or preferably continuous) is needed in order to fully understand the circumstances. Hence, upgrading simple (single-task oriented) technologies by other technologies (most typically by cameras), can significantly improve researchers' ability to obtain information on the circumstances, which accompany the detected risky behaviour. This option was considered and demonstrated for several RQs in chapter 5.

However, it is important to keep in mind that an upgrade of the technological features requires integration and the establishment of a unified platform for researchers. The compound systems and configurable automotive data loggers, mentioned in chapter 2 and Appendix A, provide examples for such systems.

One of the most conceptually straightforward integrated systems is a system, for which the basic technology detects the desired behaviour (e.g. harsh braking) and triggers a simple continuous dashboard camera to save the relevant information, which occurs together with that behaviour. As demonstrated in chapter 5, many RQs can be addressed using this type of combined systems. In that context, it is worth mentioning that while in previous times, setting up cameras in cars was considered an intrusion to privacy and strongly objected to, recently, many drivers willingly choose to add cameras to their cars (and PTWs to their helmets). In that way, drivers are able to generate evidence that could be used to protect them in case of a crash.

It is also important to keep in mind that many of the systems, which are described in chapter 2 are designed as alerting systems or specific goal oriented systems and hence, are not configured to collect and store data. In case that data needs to be stored, typically a prerequisite for a valid monitoring, a dedicated algorithm and routines need to be developed and implemented (such as the adaptation of the Ituran system to read, synchronize and store the Mobileye data reported in chapter 3, as well as the adaptation of the ProtextMe app to monitor and store smartphone usages while driving, as reported in chapter 4).

A major strength of the elaborated DAS used in UDrive is its extensiveness and hence, its ability to address many RQs and provide a comprehensive view of the driving environment and circumstances. When looking at smaller systems, both in terms of size and cost, clearly a more limited view can be provided. However, this can still be very productive and relevant to many RQs, but not to all. Consequently, a well-defined understanding of the research (or customer's) needs can help make the decision, which DAS is more suitable for which project.

When going one-by-one over the RQs of UDrive, almost all of them could be addressed, at least to some extent, by simpler technologies (or by a combination of few), such as the ones described in chapter 2. In addition, many of the RQs could be addressed to a greater extent if upgraded by cameras and/or CAN access. This is a step forward, which paves the way to continue the analyses performed in UDrive in several directions, using the UDrive data as baseline, or guideline, and to complement the available information with new data.

6.2 Discussion

The aim of the present report was to outline the value of simple low-cost technologies for future Naturalistic Driving data collection in relation to the complex, tailor-made and expensive data acquisition devices that have been used in recent large scale naturalistic studies.

The conceptual Framework for Naturalistic Driving Studies (FNS) presented in chapter 1 outlines the key variables for a naturalistic study design: the 'complexity level' of the data collection technology and the sample size. Figure 6-1 (taken from chapter 2) presents anoverview of the various simple and low-cost technologies according to the dimensions of the FNS.

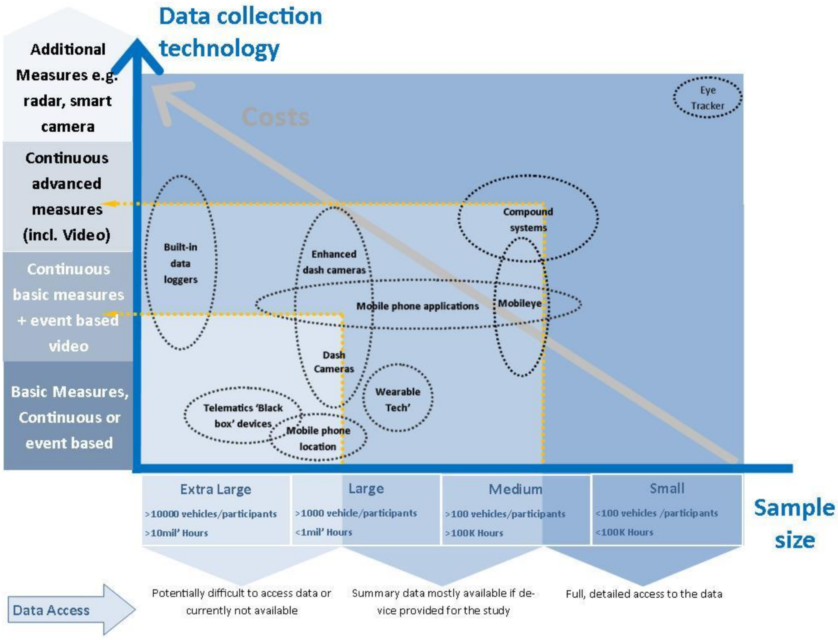


Figure 6-1: Summary of simple low-cost DAS

Figure 6.2 (adapted from the Figures in chapter 5) depicts, which equipment and sample size are required to answer certain RQs, such as the RQs addressed in UDrive.

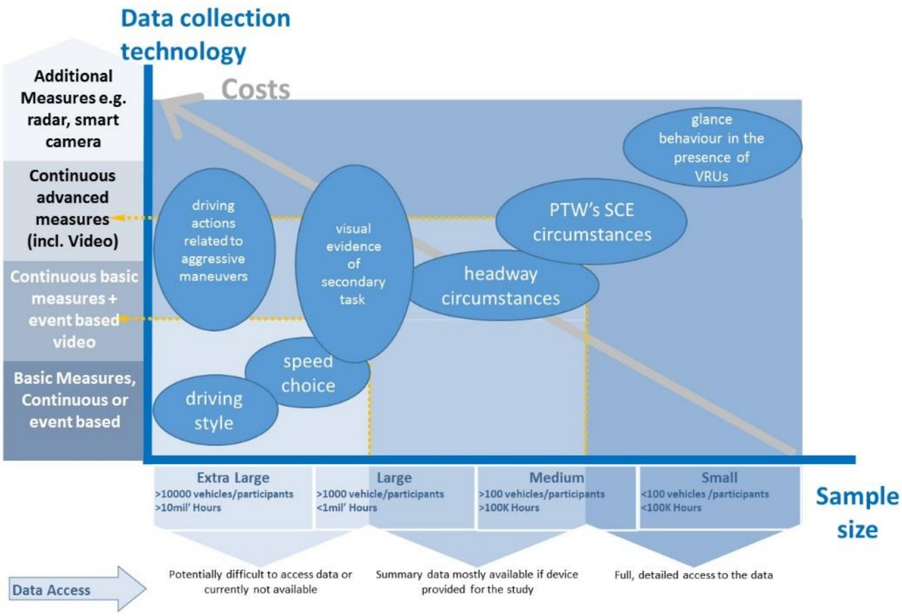


Figure 6-2: Various research questions mapped on the FNS

Trying to merge, super-impose and correlate the 2 Figures, we can roughly suggest the following insights;

Simple DAS, such as the ones appearing in the lighter blue areas of the 2 Figures, correspond to low-cost simple technologies that can be easily implemented on large samples. These DASs can address RQs that appear approximately on the same light-blue areas of the FNS and correspond mostly to ECO driving, aggressive driving behaviour and speed choice.

The darker-blue areas of the 2 Figures correspond to more complex DASs that can address complex RQs, such as the circumstances of the driving behaviour. These RQs typically require visual continuous information, as well as integration with the CAN data of the vehicle. Consequently, they can be implemented on relatively small sample sizes.

There is a clear trade-off between 'complexity level' of the technology and sample size. Consequently, the strength of large scale ND studies like SHPR2 and UDRIVE is the in-depth data collection. However, their limitation is the relatively small sample size and the extensive resources needed for analysis.

Looking into the future, it is reasonable to expect that as technology will rapidly progress, it will be possible to collect large samples of continuous detailed observations. With the advancement of car technology and the shift towards partial and full automation, it is very likely that built-in data loggers will continuously collect detailed high-level information.

Already in the near future, vehicles will include smart cameras, which will monitor behaviour both inside and outside the vehicle. It is also reasonable to expect that cameras will become more advanced in screening the visual data and hence, a meaningful output will enable drivers, researchers and planners to focus on the desired behaviour (which is of interest to them). For instance, a smart internal camera could register head pose, involvement in a secondary task and fatigue. It may well be the case that in the near future, this data will be collected by OEM.

This somewhat futuristic, but realistic view is presented in Figure 6-3. With the advancement of technology, it is expected that advanced measures will be cheaper and more accessible, hence the measures on the Y-axis of Figure 6-3 are shifting down and pave the way for more advanced measures suitable for autonomous control too (such as: V2V, V2I and A.I. based measures). Naturally, the cost will also go down, as marked by the yellow arrow, and most importantly: the top part of the Figure will be left to handle new and more advanced RQs relating, most probably, to high level system approach type of control.

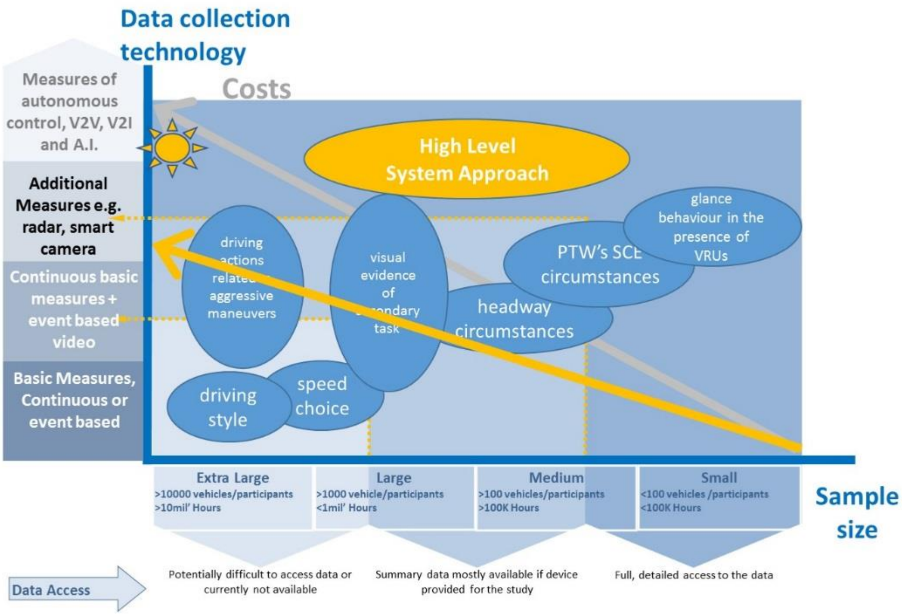


Figure 6-3: The FNS adjusted to technological advancements

Appendix A A Tabular summary of data collection technologies

Table 6-1 summarizes the data collection technologies according to the following parameters:

Technology group

The technology group column corresponds generally to the families of technologies surveyed in Chapter 2, according to the concept of the Y-axis of the FNS as defined in figures 4 and 5. In general terms, the associated of each technology group can be seen to step at two points; these points are broadly defined by access to video data and access to CAN or vehicle data and are shown with the 'breakpoint' lines.

Video data: Access to this level of data can be considered as medium. Video brings with it a data richness that simple GPS loggers cannot compete with. The constraints of video, such as high data usage, high storage demand, location of devices and processing power means that not all simple GPS based devices can achieve this level.

CAN or vehicle data: The next data level can be considered to have a break point of CAN access or detailed vehicle data as a minimum. Devices below this level can broadly be considered 'nomadic' in that they can be moved from vehicle to vehicle with minimal effort or minimal set up time. Access CAN or vehicle data means that the data collection devices need an extra level of complexity in order to record this detailed data; they can be considered the next level up from video based devices.

Description

The second column contains a brief description of the type of device.

Link

This column contains a web link (where possible) to demonstrate the device in question. The link is not always to the manufacturer's website as other sites may contain more pertinent information. Bespoke devices, such as those designed and built for a specific project, have a reference to a project report where possible.

Data recorded

The 'data recorded' column endeavours to cover all the data variables recorded by each device. Clearly as the data level of the devices increase (and particularly where video or CAN is included) then it is not always possible to be specific with all the data variables recorded. For example video data is an incredibly rich resource which can be analysed and reanalysed in a number of different ways for different research questions, as such the column is filled in simply with 'video'. This same approach goes for CAN data where it is not possible to define each particular CAN variable recorded.

Price

The cost of the data collection equipment, service or data access is shown in this field. It is not always possible to provide completely accurate costings for each data source as some are specific to vehicles or will need to be tailored to each specific user thereby incurring additional costs. In each case a best estimate is provided or a subscription rate is provided.

Care has to be applied where a costing has been provided as 'free'. There are often restrictions applied to this and where relevant these have been identified by a short section of text in the relevant price box.

The costs are included in their native currency where available. As projects can be run over a number of countries with different currencies (even within the European 'Euro' zone) no effort has been made to convert to one standard. Additionally some devices may incur different pricing strategies across different countries which are not covered here. All cost should therefore be considered as a guideline.

Notes

Any additional notes for each device, data, cost or other factor can be found in this column.

Data access

The data access column attempts to define in broad terms what the expected access to the collected data would be in full scale operation. Where data access is known or guaranteed, for example in a study where the data logging equipment is bought and installed by the test site then a then a "✓" sign is shown in the data access box. There is no further attempt to define how the data access is guaranteed however it is likely that if full ownership of the equipment is assured then full access to the data would follow. Where questions exist as to the access of the data then a series of "?" signs appear in the data access box.

An entry of "???" indicates that there are currently a lot of uncertainties as to the data access. This is particularly evident if the data collection source is associated with a third party. For example Google and car insurance companies both collect data on peoples (or drivers) movements for their own or individual user purposes. It is however unlikely that without very stringent data protection agreements that researchers will be able to access detailed information on these people unless the project is specifically designed with the stakeholder involvement. In this case a clause is written into the data access box to illustrate this case.

Where "??" is entered it normally means that there is some publically available data but that detailed information is reserved for the data collector or specific individual using the service. For example with mobile phone applications such as Strava there will be general data fields and aggregated data for individuals available for any other Strava subscriber, however detailed and possibly sensitive information is restricted to that particular user. It is again possible for this level of data access to be reduced to '?' if the software was provided through a study for participants use or if the application developer was involved in the study.

Table 6-1: Summary of data collection technologies

Technology Group	Description	Link	Data recorded	Price	Notes	Data access
Mobile phone location services	Mobile phone location services	http://trendblog.net/cant-remember-last-night-google-location-history-can-help-you/	Time, Date, GPS location, Duration, Mileage, Mode type, User demographics	Free	Enabled devices record GPS and WiFi location and log over time. Some mpa matching and mode algorithms	??? (no visible data for any other user)
Mobile phone applications	Mobile phone application for monitoring driving style; linked to insurance policies.	http://www.aviva.co.uk/drive/	Time, Date, GPS location, Speed, Acceleration, Driving duration, Mileage, User demographics, Vehicle data	Free	Drivers can earn discounts from their results for the following years insurance.	??? (no visible data for any other user)
		http://www.drivology.co.uk/howitworks/	Time, Date, GPS location, Speed, Acceleration, Driving duration, Mileage, User demographics, Vehicle data	Free	Drivers can earn discounts from their results for the following years insurance.	??? (no visible data for any other user)
		http://www.ionroad.com/	Time, Date, GPS location, Speed, Distance, Forward collision warning, Headway, Lane departure, Video, User demographics, Audio	1 €	Pro version and Lite version (free)	? (if provided for use in study) ??? (if requesting data).

Telematics devices (including IVDR, Green box and black box)	Telematics (Blackbox) device for insurance.	http://www.theaa.com/insurance/car-insurance.jsp	Time, Date, GPS location, Speed, Acceleration, Driving duration, Mileage, User demographics, Vehicle data	Free (with insurance policy)	As with phone devices but hardwired into vehicle.	??? (no visible data for any other user)
		https://www.directline.com/car-insurance/telematics	Time, Date, GPS location, Speed, Acceleration, Driving duration, Mileage, User demographics, Vehicle data	Free (with insurance policy)	As with phone devices but hardwired into vehicle.	??? (no visible data for any other user)
Mobile phone applications with data logger	Mobile phone application – black box	http://roadrecorder.eu/	Time, Date, GPS location, Speed, Video, User demographics, Audio	Free	Works with other apps working in the background	? (if provided for use in study) ??? (if requesting data).
	Mobile phone application – for teen drivers	http://apps4driving.com/	Time, Date, GPS location, Speed, Acceleration thresholds, mobile phone use, User demographics, vehicle data	Free		? (if provided for use in study) ??? (if requesting data).
	Mobile phone application for recording routes but with an additional crash sensor.	http://www.realrider.com/	Time, Date, GPS location, Speed, Acceleration, Driving duration, Mileage, User demographics, Crash detection algorithms, Service and maintenance logs	3.50€ monthly/ 36€ yearly.		?? (some data are visible if logged on)

Wearable Technology	Mobile phone app for recording and measuring cyclist performance	https://www.strava.com/	Time, Date, GPS location, Speed, Riding duration, Mileage, User demographics (+ performance metrics and external sensors, Heart rate, power etc. for 'premium' version)	Free (4.80€ per month/ 48€ yearly for premium)		?? (some data are visible if logged on)
	Mobile phone app for recording and measuring activity (40+ activities)	https://www.endomondo.com/	Time, Date, GPS location, Speed, Duration, Mileage, User demographics, Mode type, (+ external sensors, Heart rate etc.)	Free (5.60€ per month/ 28€ yearly for premium)	Can be connected to a range of other activity monitors	?? (some data are visible if logged on)
	Wearable GPS connected heartrate and activity monitor	https://www.fitbit.com/uk/charge2	(with software) Time, Date, GPS location, Speed, Duration, Mileage, User demographics, Mode type, Heart rate	155€		? (if provided for use in study)
		https://buy.garmin.com/en-GB/GB/c10002-p1.html	(with software) Time, Date, GPS location, Speed, Duration, Mileage, User demographics, Mode type, Heart rate	168€		? (if provided for use in study)
Dash Cameras and Enhanced Dash Cameras	Mini and High resolution cameras	http://www.amazon.co.uk/Mini-DV-Smallest-Resolution-Camcorder/dp/B003URUXVQ/ref=pd_sim_sbs_23_2/277-9993987-	Video, Time, Date	6€	Limited recording time – no external sensors	✓

		5492606?ie=UTF8&dpID=31DMoEdqF5L&dpSrc=sims&preST=_AC_UL160_SR160%2C160_&refRID=10HR4QBQ0E0KR99WDF55				
		http://store.driftinnovation.com/cameras/drift-stealth-2/	Video, Time, Date	186€	WiFi enabled	✓
	High definition dash camera with enhancements	http://shop.roadhawk.co.uk/roadhawk-hd-2-dash-cam.html	Video, Time, date, GPS location, Speed, threshold acceleration, event button, Audio	300€	Multi camera options (at additional cost). Configurable.	✓
		http://lukashd.com/eng/product_info/blackbox_14_1.html	Video, Time, date, GPS location, Speed, Event button, Event detection, Motion detection, Audio	140€		✓
		http://www.nextbase.co.uk/product/in-car-cameras-series/in-car/in-car-cam-402-g-professional/	Video, Time, date, GPS location, Speed, Event button, Event detection	144€	Provides 'event' data capture.	✓

Compound systems	Bluetooth based OBD decoder that can be linked to phone applications	http://www.plxdevices.com/Kiwi-3-OBD-Car-to-Smartphone-Connection-p/897346002832.htm	Nothing itself but in conjunction with below:	92€		
	Mobile phone app – in conjunction with above	http://www.palmerperformance.com/store/index.php?cPath=21_29	Date, Time, Distance, Accelerations, GPS location, Inclinometer, MPG, Power, Speed, Fuel level/range, Temperatures, Vehicle data	Free	Some info may not be available depending on OBD access – however it may be expandable	✓
		http://www.gps-laptimer.de/	Date, Time, Distance, Accelerations, GPS location, Inclinometer, MPG, Power, Speed, Fuel level/range, Temperatures, Video, Audio, Throttle, Brake, Vehicle data	8.4€ –28€ depending on level	Some info may not be available depending on OBD access – however it may be expandable	✓
configurable automotive data logger	Stand alone, configurable GPS based data logger	https://www.race-technology.com/wiki/index.php/DL1PRO-WP/Introduction	Date Time, GPS Location, Acceleration(s), Yaw, Pitch Roll, CAN access, up to 12 analogue inputs, additional inputs (RS232) for other data sources	360€ - 720€	Highly configurable logger that can be expanded	✓
	Video expansion box for above system	https://www.race-technology.com/wiki/index.php/Hardware/VIDEO4	4 x video channels. Overlay of data recorded above	720€	Bespoke video overlay system designed to work with above logger	✓
bespoke industry level	Stand alone CAN/LIN logger	http://vector.com/vi_canlog_en.html	CAN data, LIN data.	€???	Can be linked with other sensors (GPS etc	✓

data logger					to enhance data)	
	High definition GPS and accelerometer package with 2 channel video overlay	http://store-uk.vboxmotorsport.co.uk/index.php?route=product/product&product_id=42	Time, Date, GPS location, Speed, Acceleration(s), Video, Audio, CAN, configurable variables	1800€	Lots of additional sensors and upgrades.	✓

List of abbreviations

Abbreviation	Meaning	Comments
A.I.	Artificial Intelligence	
CAN	Controller Area Network	
CI	Confidence Interval	
DAS	Data Acquisition System	
DZ	Danger Zone	Relates to mobileye warning
FCW	Forward Collision Warning	
FNS	Framework for Naturalistic Studies	
FOV	Field Of View	
GIS	Geographical Information System	
IVDR	In Vehicle Data Recorder	
LB	Location Based	Relates to control
LDW	Lane Departure Warning	
ME	Mobileye	
NFC	Near Field Communication	Relates to the ProtextMe application
OEM	Original Equipment Manufacturer	
PCW	Pedestrian Collision Warning	Relates to Mobileye warning
PSM	Propensity Score Matching	
PTW	Power Two Wheeler	
RQ	Research Question	
SCE	Safety Critical Event	Concerns crashes and near-crashes
TTC	Time To Collision	
VRU	Vulnerable Road Users	
V2I	Vehicle to Infrastructure	
V2V	Vehicle to Vehicle	

List of Figures

List of Tables