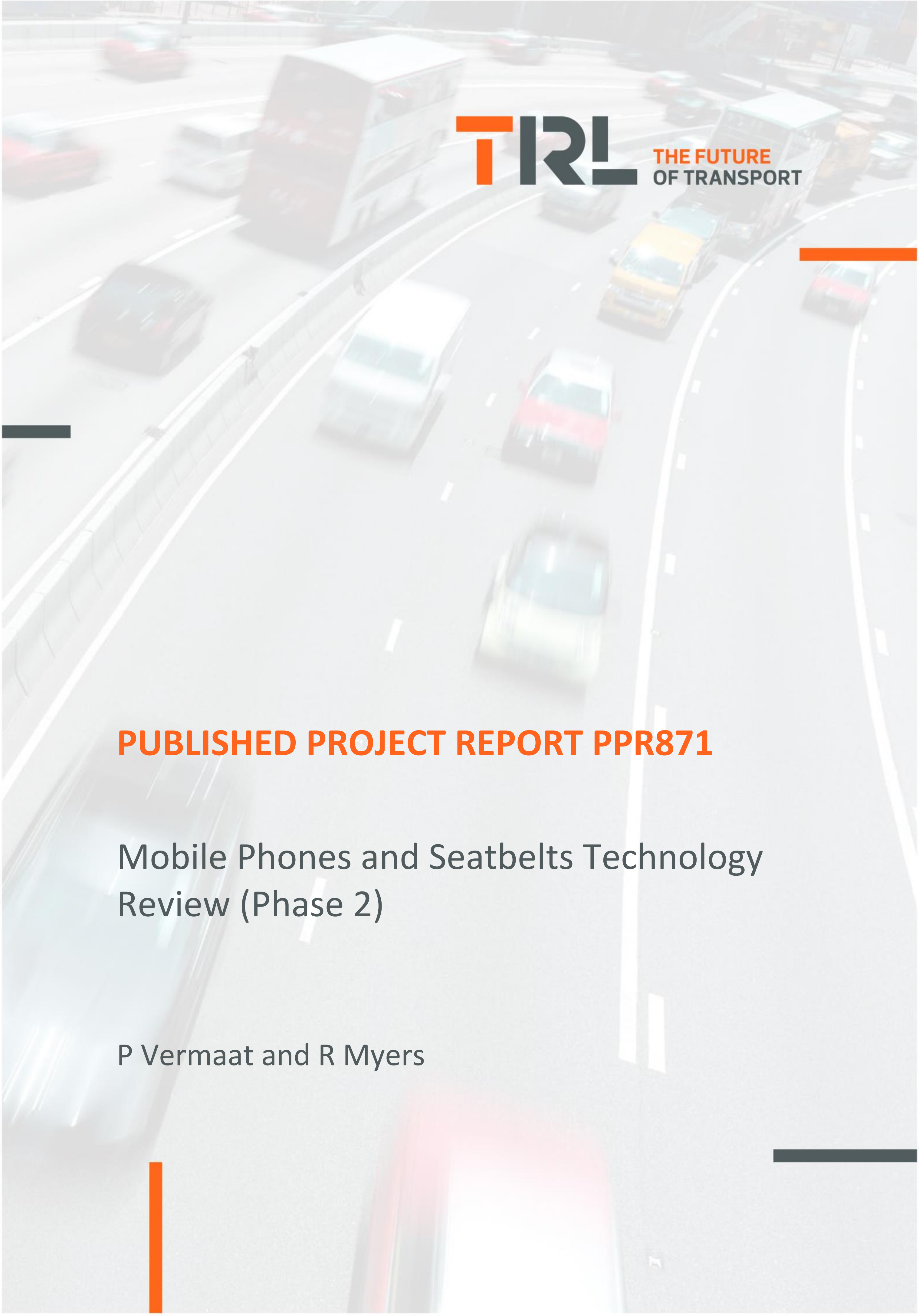

# PUBLISHED PROJECT REPORT PPR871

## Mobile Phones and Seatbelts Technology Review (Phase 2)

P Vermaat and R Myers

## Report details

| | |
|---|---|
| **Report prepared for:** | DfT and Highways England |
| **Project/customer reference:** | 1-321 Lot 4 |
| **Copyright:** | © TRL Limited |
| **Report date:** | |
| **Report status/version:** | |
| **Quality approval:** | |

| | | | |
|---|---|---|---|
| Su Buttress<br>(Project Manager) | | Brian Lawton<br>(Technical Reviewer) | |

## Disclaimer



## Contents amendment record

This report has been amended and issued as follows:

| Version | Date | Description | Editor | Technical Reviewer |
|---|---|---|---|---|
| 0.1 | 5/7/2018 | First draft for client comment | PV | BL |
| 0.2 | 10/8/2018 | Updated draft following client feedback | PV | BL |
| 1.0 | 08/11/18 | Final version for publication | PV, RM | |

# Table of Contents

## Executive summary

In Phase 1 of the Mobile Phones and Seatbelts Technology Review project, technologies which had the potential to measure the level of non-compliance by drivers and passengers in vehicles were identified. The non-compliances to be recognised were primarily, but not limited to, the use of mobile phones by drivers and the non-wearing of seatbelts by all vehicle occupants. This report presents the findings of Phase 2 of the project. The scope of this work was to:

1. Revisit deep learning to gain a deeper understanding of the realistic capabilities of image processing systems based on deep learning.
2. Further investigate the concept of assisted tagging.
3. Investigate camera requirements for creating the quality of images required for either of the above two scenarios.
4. Follow up on defence industry contacts to understand any defence technology which may, in the longer term, provide promising avenues.
5. Investigate the possibility of fusing data from multiple sources and of using 'big data' analysis techniques, engaging with the DfT's Data Science team where appropriate.
6. Further investigate the implications of gathering data for analysis, particularly the barriers to sharing.
7. Using the learning from the above points, investigate what any new approaches based on new technologies can deliver compared to current approaches.
8. Given the likely performance of new algorithms (point 1) and camera technologies (point 3), consider the metric requirements for future surveying using these enhanced technologies.

This project has found that the automatic detection of non-compliant behaviour using deep learning based algorithms has a reasonable likelihood of being successful, assuming high quality input images, though the development costs are likely to be substantial, driven principally by the human resource required over the 6-12 months of development time required, at least a quarter of which time will require the input of an expert in deep learning. It is not clear if the limited market for this type of solution will be sufficient to cover the development costs for commercial projects.

Assisted tagging whereby technology is used to significantly reduce the time taken by human operators to record non-compliant behaviours is very promising, though again the image quality needs to be sufficient. This is likely to incur a substantially lower cost, than a fully automated system; the actual costs will depend on the level of assistance envisaged, but are likely to be in the range of 25% - 50% of the cost of a fully automated system.

As both deep learning and assisted tagging rely on high-quality images, the use of cameras based on near infra-red pulsed illuminators should be considered prior to any further work on deep learning or assisted tagging. The requirements for fixed and mobile cameras are equivalent in terms of performance, though the practical deployment requirements of mobile cameras will make it difficult to achieve the same level of performance as fixed cameras.

It is possible that existing image database or live image streams could be used as a source for assisted tagging or deep learning based classification, though they would need to include a high level of image redundancy as many, possible the vast majority, of images are likely to be unsuitable due to image quality considerations.

Consultations with defence industry experts identified no technologies (which they are able to discuss) which could be implemented.

Making use of in-vehicle data is technically feasible for detecting seatbelt non-compliance, but faces probably insurmountable legal and commercial issues. The key legal issues involve the use of private data, made somewhat more difficult with the recent introduction of GDPR. It is likely that vehicles owners/driver would need to give consent to having their data used. The key commercial issue involves key players in the automotive industry being of the opinion that any data gathered from in-vehicle systems has value and will not be shared. The DfT would need to consult other arms of government to determine the viability of changing regulations to allow this.

The fusion of data from multiple sources is also technically feasible but, as with the previous point, the legal and privacy related challenges mean that this solution would need to be approached on a case-by-case basis as each combination of data sources will present a unique set of challenges.

Given the above findings, it would be premature to finalise metric requirements for future surveying in detail at this stage, though the potential to increase the reliability of the results is clear. The use of technology is expected to have a beneficial effect on survey accuracy, and certainly has the potential to broaden the scope of surveys to more areas, road types and populations. As the effect of new technology on measurement accuracy and reliability cannot be determined at this stage, a pilot study should be considered to evaluate the effect that new methods have on accuracy.

It is recommended that a research programme is developed in partnership with one or more suppliers with suitable expertise. As the only technologies which offer a realistic possibility of success are based on images capture, the programme should start with proving that images of sufficient quality can be reliably recorded. Analysis of recorded images will enable surveying of a wider range of sites than is possible at present. Following this, the programme should move to the development of an assisted tagging solution based on captured video images. As the assisted tagging solution should include an element of automatic detection of violations, it has the ability to be developed into a fully automated solution. Taking this stepwise approach maximises the probability of success, and at the end of each step an improved solution will be available, even if it proves impossible to ultimately achieve a fully automated solution.

# 1 Introduction

In Phase 1 of the Mobile Phones and Seatbelts Technology Review project, technologies which had the potential to measure the level of non-compliance by drivers and passengers in vehicles were identified. Using technology to measure non-compliance would make more regular surveys possible, making it easier to identify trends over time. The non-compliances to be recognised were primarily, but not limited to, the use of mobile phones by drivers and the non-wearing of seatbelts by all vehicle occupants. The findings of Phase 1 were presented in (Vermaat, et al., 2018); this report presents the findings of Phase 2 of the project. The scope of this work was to:

1. Revisit several deep learning based suppliers and developers to gain a deeper understanding of the realistic capabilities of image processing systems based on deep learning. This will include development timescales, costs, likely performance, risks and practical deployment issues. At least three UK suppliers, and a further two outside the UK (one in Spain, and in Australia), have been identified and will be approached.
2. Further investigate the concept of assisted tagging. If an automatic system (based, for example, on ANPR cameras) can derive clear still images of vehicle occupants, this would make human analysis significantly more efficient, and hence lower costs. We propose in Phase 2 to further develop this concept, investigating costs for developing a system, likely performance and timescales.
3. Investigate camera requirements for creating the quality of images required for either of the above two scenarios. Normal colour CCTV cameras are not suitable as the level of illumination required would dazzle drivers and create a safety hazard. Near infra-red cameras, similar to those used in ANPR systems, are likely to be suitable, but will require some development as the illumination levels used by ANPR cameras are not sufficient to allow capture of clear images of front seat passengers. The requirements for the two focus scenarios are similar, but there are differences, and these differences will be highlighted where appropriate. The differing requirements of fixed and mobile cameras will be included.
4. Follow up on defence industry contacts (for example, Qinetiq) to understand (non-classified details of) any defence technology which may, in the longer term, provide promising avenues.
5. Investigate the possibility of fusing data from multiple sources (for example, mobile phone meta-data, the DfT's journey time system, roadside technology) and of using 'big data' analysis techniques, engaging with the DfT's Data Science team where appropriate.
6. Further investigate the implications of gathering data for analysis, particularly the barriers to sharing (both legal and commercial), and what further opportunities might present themselves.
7. Using the learning from the above points, investigate what any new approaches based on new technologies can deliver compared to current approaches.

8. Given the likely performance of new algorithms (point 1) and camera technologies (point 3), consider the metric requirements for future surveying using these enhanced technologies. Here, we will consider the likely design requirements for future surveys (number of surveys on different road types at different times of day) which would facilitate regular national surveys that would enable non-compliance rates to be measured accurately and trends in them to be tracked reliably. TRL's transport statistics team will assist in this task, including identifying an acceptable level of precision given the low non-compliance rates.

All eight of the above points have been addressed and are described in this report.

The work mainly involved consulting the experts identified in the earlier part of the project, as well as additional subject matter experts subsequently identified.

Researchers at the University of Genoa were consulted to confirm that information gathered from commercial sources on deep learning was feasible.

Qinetiq provided input on the potential application of defence technology.

TRL's data processing and statistics experts were used to provide input on data fusion, performance measurements, metrics and precision.

# 2 Results

This section gives the results of the research on camera requirements, assisted tagging, deep learning, defence industry input and data fusion.

## 2.1 Camera technical requirements

**Key findings**

- As discussed in the Phase 1 report, standard CCTV cameras are unlikely to produce consistent images of adequate quality for analysis.
- Polarising filters can improve image quality in certain lighting conditions, but maintaining consistent image quality is challenging.
- Use of near-visible infra-red cameras and illumination based on ANPR camera technology could produce high quality images in all lighting conditions, including at night and at all traffic speeds.

In this section we assume that analysis will be of stored video images, either as continuous video or a sequence of still images. The analysis could be by human observer simply replicating what a roadside observer would do, by an automated image processing system (see the section on deep learning below), or some combination of the two (as described in the section on assisted tagging below). Storing video as opposed to still images provides a richer data set, potentially making a wider range of occupant behaviour detectable.

In any image processing system, the quality of source images is crucial to the eventual performance of the solution, and this imposes constraints on the camera requirements. The key requirements for an image/video capture camera are:

- The camera must produce a clear image of the occupants of the vehicle in all lighting conditions. This includes bright sunlight, night and from a variety of angles.
- The image from the camera sensor should have sufficient resolution and dynamic range. HD-video resolution should be sufficient.
- The image should be clear at all vehicle speeds.
- Any illumination used must not dazzle or otherwise distract the driver.
- The camera system should be as unobtrusive as possible to minimise behaviour change in the presence of cameras which could be mistaken for enforcement equipment.

As discussed in the Phase 1 report of this project, standard CCTV cameras are not suitable for all day operation of image processing systems which need a clear view of the vehicle occupants. There are a host of reasons for this, but the two most important are:

1. To achieve 24 hour operation requires a high level of artificial lighting which needs to fall on the faces of the vehicle occupants. This is not feasible for safety reasons.
2. Taking video images through the windscreen of a vehicle can lead to specular reflection of ambient light sources (light cloud cover on bright days is particularly

bad) which can be bad enough to completely dominate the image captured – see Figure 1.

**Figure 1: Windscreen specular reflection[1]**

It is possible to reduce the effect of specular reflection by using polarising filters as shown in Figure 2, but these require adjustment during installation to ensure maximum effectiveness.


**Figure 2: Polariser reducing specular reflections[2]**

The ANPR industry has addressed these limitations by developing specialised ANPR cameras which use Infra-Red (IR) artificial lighting which is invisible to the human eye though it is close enough to the visible spectrum that normal CCTV-type camera modules can be used, thus benefitting from the economies of scale these modules provide. However, the light levels utilised in ANPR cameras are too low for capturing images of vehicle occupants as they use the retro-reflective properties of number plates, reducing the light levels required by over an order of magnitude.

A number of approaches could be used to overcome the issues described above.

The ANPR camera approach could be extended by using much more powerful IR illuminators. This is both technically feasible and relatively low cost, and solves both issues. This solution

[1] From https://commons.wikimedia.org/wiki/File:Ponderosa_in_the_Windshield_(4249542107).jpg

[2] From https://commons.wikimedia.org/wiki/File:Polarising_Filter_Examples_(2252215670).jpg

would use existing ANPR camera technology fitted with more powerful IR illumination and a wider angle lens. As the camera would be the same physical unit as an existing ANPR camera, costs would be limited to providing additional IR illumination. Note that the resulting camera would no longer be suitable for use as an ANPR camera though, with additional development, a dual use camera would be feasible. The technical requirements for a suitable IR illumination system would be:

- Illuminators must be pulsed and synchronised with the camera shutter. To capture clear images of a moving vehicle typically requires a shutter speed of below 4ms (1/250 second) to eradicate motion blur. The illuminator should only be on while the shutter is open to keep average light levels to a minimum and reduce the thermal load on the LEDs used. Like a standard ANPR and CCTV camera, the output from the camera would be a continuous video feed i.e. the illuminator would be synchronised with the shutter.
- Illumination should be in a very constrained bandwidth, normally <25nm wide, and all LEDs used must operate in the same band. This requirement is easily met with low-cost LEDs.
- The camera must be fitted with a matched IR pass filter of the same centre frequency and matched bandwidth. This reduces the effect of natural light from the image, and hence significantly reduces specular reflections. Suitable filters are readily available and of low cost.
- The lighting used should operate outside the range of human vision (the typical spectral response of the human eye is shown in Figure 3), but not so far outside that normal CCTV camera modules and lenses no longer function properly. The limit of human vision is typically at about 700nm – 800nm wavelength (this varies somewhat from person to person), so illuminators in the range of 900nm – 1000nm are typically used in ANPR systems.

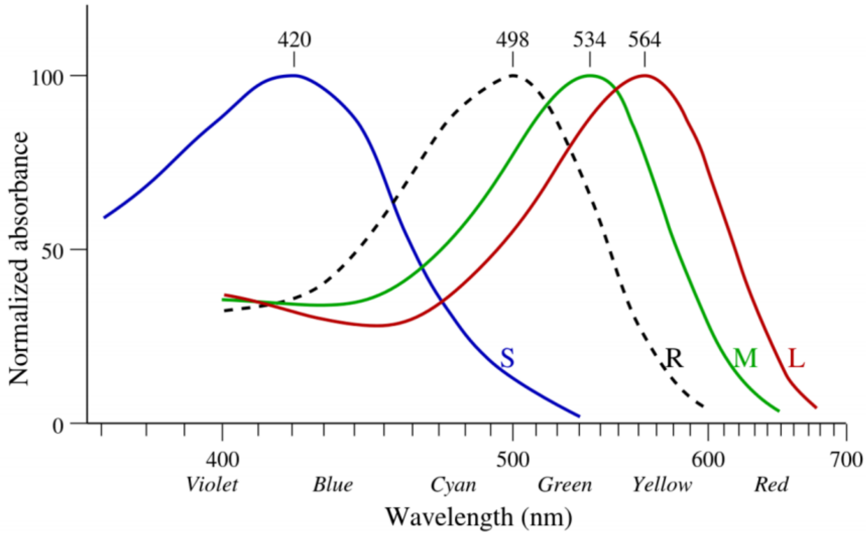


**Figure 3: Spectral response of the human eye[3], showing colour response and low-light monochrome response (dashed line)**

The response curve for typical LED illuminators is shown in Figure 4. Closely matched filters are available, which allows the image sensor to only receive light from a very narrow spectrum, eliminating nearly all natural light which ensures that only the light from the illuminator is used, which in turn can be closely controlled.

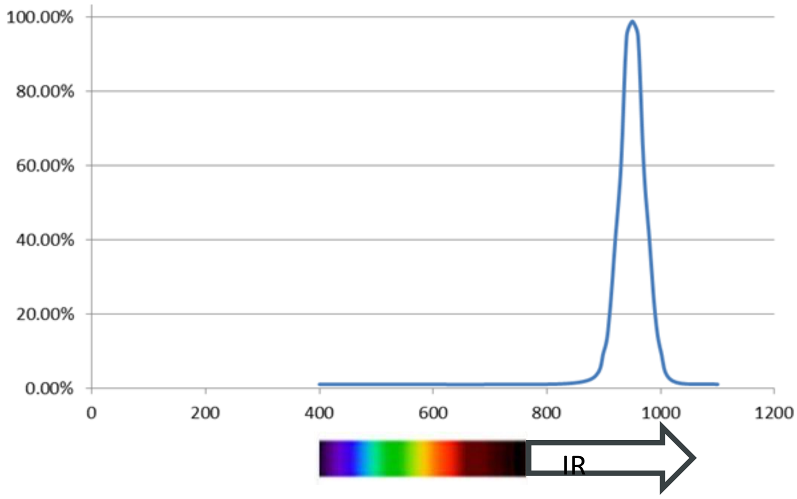


**Figure 4: Typical spectra of 950nm LEDs**

At least one UK-based ANPR provider has the capability of controlling multiple LED illuminators from its camera modules, making the solution feasible and at a cost which

[3] https://upload.wikimedia.org/wikipedia/commons/6/64/Sensory_Systems.pdf

would be very similar to that of an ANPR camera. A typical complete system would consist of:

- One ANPR-derived IR camera
- Additional IR illumination
- A digital video recorder (DVR)
- A power supply
- Deployment hardware (camera brackets and secure housing for the DVR).

Exact costs are not possible to evaluate at this stage as this would be very dependent on the core components used.

One camera sensor manufacturer has suggested that sensors could incorporate filters into their manufacture, including polarising filters. However, given the low cost of external bandpass and polarising filters, this approach is unlikely to be cost-effective unless many thousands of cameras are to be purchased.

Other approaches have been investigated, including tailoring camera response curves to attempt to mitigate the effect of specular reflections, but this has not revealed any feasible solutions.

The use of standard CCTV images gathered from ANPR camera installations is discussed in Section 2.2 below.

The above observations apply to both fixed and mobile cameras. Mobile cameras are more limited by size and weight, which may affect the level of IR illumination which can be achieved. Also the practical use of mobile cameras means that mitigation of reflections by means of polarising filters will be less effective as they need to be adjusted to the geometry of the camera.

## 2.2 Assisted tagging

**Key findings**

- An assisted tagging system could help to reduce the amount of human input required for seatbelt and mobile phone surveys, and has the potential to reduce costs.
- It is unlikely to be possible to capture data on rear seat passengers using any camera based approach.
- Up-front investment would be required in a coding tool and cameras. Software development is perceived to present little technical risk.

At present, observations are made at the side of the road using an audio recording then transcribed into a spreadsheet. This is time-consuming, costly, and limits the number of locations where surveys can be safely undertaken.

An assisted tagging system aims to limit human involvement in the data gathering process to the minimum possible. This is done by recording onto a digital video recorder (DVR) the video stream from a suitable camera (see the previous section for camera requirements). The recorded video is then pre-processed by an automated system before being analysed off-line by a human operator. The pre-processing could be as simple as using a movement detector to present only short portions of video which contain a vehicle to the operator, thereby making the operator more productive, or a more complex system which detects obvious compliances or non-compliances (see Section 2.3 for a method of doing this). It would only present those images which are not classified with a high degree of confidence to the human operator or coder for manual analysis. The images presented to the operator could be a short video stream, or still images.

A potential limitation of automation and assisted tagging is that it will only highlight pre-trained non-compliant behaviours, whereas the current observer-based system allows the survey team to add unusual findings in free text. While this is not possible for a fully automated system (see Section 2.3 below), an element of this could be included in an assisted tagging system where operators could be asked to enter unusual observations, probably at regular intervals such as at the end of each shift, rather than at any time.

### *2.2.1 Developing an Assisted Tagging system*

There are three major components which would form assisted tagging system: a camera, a human image coder and software to implement the process. This section covers the software and human components of the system (camera technical requirements are covered in Section 2.1).

The advantages of this approach over the current system are:

- It could practically form the basis of an automated system in future.
- Reduction in the level of human input required: coders stood at the side of the road is time and cost inefficient. Improving the efficiency could allow a larger sample size.
- Better understanding of the accuracy of the process: quality checks can be carried out and unusual results reviewed using original images.
- Ability to operate in areas where it would not be safe for a human observer to work, such as on high speed roads.
- Could be designed to work at night.
- Ability to review or cross-check measurements, hence improving reliability including inter-coder reliability.

The limitations are:

- Investment is required in camera systems, DVRs and a coding tool.
- Running costs (camera installations, software licences etc.).
- The need to install cameras may reduce the number of sites where it is feasible to collect data.
- It is unlikely to be possible to capture images of rear seat occupants in the sample due to the camera angle required.
- Data protection and privacy issues will need to be addressed and may incur some additional costs in data management.

• Substantial human input is still required (the system is not automated), though much less than at present.

### *2.2.2 Making use of existing camera systems in assisted tagging*

In Section 2.1, it was noted that existing CCTV cameras are unsuitable for image capture for analysis as, in most lighting and operational circumstances, the majority of captured images will be unsuitable for analysis, either by human operators or an automated system due to lighting, exposure and mounting geometry limitations. If, however, a sufficiently large database of images could be identified, it is possible that enough of the images would be usable, particularly if the unsuitable images could be automatically identified and eliminated. In this section we consider how image databases could be used.

A specific example of an assisted tagging solution would be to make use of images stored in the National ANPR Data Centre (NADC). The NADC takes ANPR results from all ANPR systems used by the police in England and Wales and stores them in a centralised database.

Most ANPR systems also capture an 'overview image' – this is a separate wider field of view image captured from a colour CCTV-style camera which is collocated with the ANPR camera. The overview image is intended to allow police to identify the vehicle and (possibly) the vehicle driver. Typical plate patch and overview images are shown in Figure 5.


**Figure 5: Plate patch (left) and overview (right) images from an ANPR camera**

The overview image shown also demonstrates the specular reflection problem that can occur with ambient-light cameras.

The current version of the NADC database stores only the ANPR result (plate characters, location, date and time of result) and an image of the number-plate, normally referred to as the 'plate patch', and as such is not suitable for use as a source of images for assisted tagging. However, the NADC is being upgraded and the new version will store overview images as well as the plate patch.

As discussed in the section on cameras, CCTV cameras have considerable limitations if they are to be used for the detection of seatbelt and mobile phone usage, and the NADC overview images are likely to suffer from the same quality issues. However, under favourable lighting conditions, it may be possible to collect usable images of acceptable quality from a small percentage of those captured and, given the number of ANPR results collected by UK police forces, enough of these images may be available to use in assisted tagging for baseline survey purposes.

It is therefore possible to envisage an assisted tagging scheme where overview images are extracted from the NADC for a particular location and time, and these are then either pre-processed by a detection algorithm before being presented to a manual tagging operative for verification, or presented to the operative without pre-processing for violation detection.

This solution would require cooperation from the NADC. To date it has not been possible to establish how difficult this would be, and there are significant security questions which would need to be overcome. However, this could be worth pursuing as the potential cost savings are considerable.

Another approach which could be adopted is to use existing CCTV camera streams to gather the video data, thereby removing the cost of camera purchase and deployment. There are a wide range of CCTV installations in city centres and along road networks. Image quality from these is somewhat variable with many different camera types used. Further research will be needed to quantify the potential use of these camera networks. Access to the video streams may be difficult due to the data privacy policies of the owners of the CCTV systems.

It is also understood that the DfT's journey time cameras are being replaced with new dual-purpose journey time and enforcement cameras. These may be usable in this context, but are likely to be subject to the same limitations as CCTV systems. In addition, it is likely that only the enforcement element would contain overview images which would be subject to the same security considerations as those on the NADC.

### *2.2.3 Exclusion of rear seat occupants from the sample*

The 2017 seatbelt survey showed that 6.9%[4] of rear seat occupants were unrestrained, compared to 3.8%[4] of front seat occupants (statistically significant, $p < 0.01$). This is not unique to Great Britain: a European Commission study (Janitzek & Achterberg, 2006) found the same result in all countries where data were available.

Excluding rear seat passengers from the sample would change the overall rate of non-compliance from 4.0%[4] to 3.8%[4]. This change is relatively small, since only 8%[4] of vehicles have any rear seat passengers. The extent to which this is a drawback depends on how important it is to monitor trends in rear seatbelt use specifically.

[4] Note: this compliance rate is based on raw, unweighted data, so may differ slightly from those published elsewhere.

### 2.2.4 Human inputs

Maximising the efficiency and quality of the image coding process requires a carefully designed, user friendly system. The main human factors considerations for identifying and developing an assisted tagging tool are outlined below:

- The image which coders view should be as large and good quality as possible: smaller or poor quality images will affect accuracy and speed.
- Repetitive strain injury is always a risk where physical actions are repeated many times. To reduce the risk, the software should allow a choice of inputting data by mouse or keyboard (or by other means), and allow the coder to vary key assignments.
- Any buttons on the interface should be a reasonable size and well separated, to minimise the risk of accidentally selecting the wrong one.
- The more data a coder is required to collect, the longer the task will take. This particularly applies to any free text data, such as that on general distraction.
- In general, familiarity with a task tends to improve the speed and accuracy of data collection, so it is desirable for the same coders to be used as far as possible.
- Regular breaks are essential: the task requires a high degree of concentration and contains little variation. A human factors researcher suggested that a break of 15 minutes after 45 minutes of coding would be ideal.
  - For comparison, UK Air Traffic Controllers carrying out high intensity tasks must take a minimum break of 30 minutes after 90 minutes of work. While the consequences of an error are obviously much less severe for image coding, it still provides an indication of the optimum time where people can be expected to maintain high levels of concentration.
- After coding a few images, the coder should be shown a summary of these images and given a chance to change their coding if there are obvious errors.
- A percentage of images should be reviewed by an independent coder to ensure consistency and inter-coder reliability.

It is estimated that it will take a coder an average of 16.2s per image to make both seatbelt and mobile phone observations of the driver and front seat passenger. Traffic counts (as required in existing surveys) would be automatically produced as part of the process to select single images. Full details of this calculation and the underlying assumptions are given in Appendix A.

In the surveys carried out as part of Lot 1 of this project, on average approximately 275[5] vehicles were recorded per site per 5 hour session. Three observers were required, one observing seatbelts, one observing mobile phones and a third carrying out a traffic count. Once the survey had been completed additional time was needed to transcribe the results.

[5] Including ‘unknown’ records where some variables could not be determined

If a survey of the same size was to be carried out using the assisted manual methodology, coding would take around two person-hours, including an independent check of 10% of observations, plus time to set up and take down the equipment used to gather the images for tagging.

### *2.2.5 Software requirements*

Potential suppliers have indicated that they see little technical risk in developing algorithms to select images from a video.

The software must provide two main functions:

- Conversion of video footage to static images focused on the driver (or possibly short video segments though, for automated analysis, still images are likely to be more amenable to processing)
- Provide a coding interface for these images, storing the outputs in a database

Assuming that the video is stored in digital form on hard-disk recorders, both functions are relatively easy to implement. The first part would require a vehicle detector algorithm, a simple task in image processing terms: off-the-shelf solutions are available. When a vehicle is detected, a short sequence of images would then be saved to a database, suitably tagged for further processing.

The second part requires the development of a parser to iterate through the database of image sequences, presenting the image to an operator for manual tagging for compliance.

Assuming high enough quality images – see Section 2.1 – a simple software tool like this could be developed in a few weeks. The actual time taken would depend to some extent on how the video is stored – hard-disk based video recorders would be the most suitable, and are widely used.

## 2.3 Deep learning

**Key findings**

- Experts consulted have a high level of confidence that the required variables could be detected by a deep learning system, though it is uncertain whether the very high level of accuracy required could be achieved.
- Development to production standards would take approximately 12 months.
- Specialised hardware would be required for a training system.

The earlier Phase 1 report identified image processing of video camera images using advanced deep learning techniques as potentially the most promising technology for violation detection in the short to medium term.

Object recognition using deep learning is an extension of earlier object classification techniques using Artificial Neural Networks (ANNs). These systems use a system of interconnected nodes which are ‘trained’ to recognise objects in images. The design of ANNs is inspired by biological brain structures, though it is simplified substantially. While

ANNs are successfully used in a wide range of applications, they have limitations in their performance when attempting to recognise a wide range of objects in complex images.

The invention of Deep Learning (DL) systems less than a decade ago has revolutionised the field of object recognition. DL extends the concept of the ANN by increasing the complexity of the network, usually using a Convolutional Neural Network (CNN). A CNN effectively cascades multiple ANNs and is closer in concept to the way biological brains are thought to work. DL requires prodigious computing power, at a level which has only become feasible in the last decade or so. Most DL training uses hardware which was originally developed for the rapid processing of computer graphics for gaming purposes.

Experts consulted for this report have confirmed that they have a high level of confidence that the wearing of seatbelts and use of handheld mobile phones could feasibly be detected using DL. The expected level of performance would initially be equivalent to that of a human operator and, given sufficient training, could even exceed the performance of a human.

The process of generating a CNN follows the same path as ANNs:

1. Gather images showing the offending behaviour, as well as the non-offending behaviour. Several thousand images of each type should be gathered.
2. Manually ‘tag’ the images, identifying which have examples of the behaviour to be detected and which do not. It is vital that this stage is done to a high level of precision, which may require secondary checking of the tagged images. This part is normally the most labour intensive and hence expensive part of the process. The biggest risk in this part of the process is finding enough examples of non-compliant behaviour, particularly lack of seatbelt wearing which is quite rare. The resulting tagged images are divided into a training set (typically 75-90% of the images) and a verification set which is used to measure the performance of the system.
3. Design the structure of the CNN. This requires considerable technical expertise as there is a multitude of ways in which a CNN can be structured in a case where the number of ‘objects’ to be classified is small. (In this case, there are only four types of object to classify: compliant, not wearing a seatbelt, using a handheld mobile phone[6], and both non-compliant behaviours.)
4. Using the training set of images, run the CNN training software. This is highly computationally intensive and even using specialised hardware may take many hours or even days.
5. Verify the performance of the CNN using the verification images.
6. If the performance does not meet a required level, repeat from step 3, revising the structure of the CNN as appropriate.

The training of deep learning networks requires many thousands of example images of each ‘object’ to be recognised, which is time consuming to gather and tag. This is exacerbated by the fact that in this case the non-compliant behaviour is quite rare, requiring the gathering

[6] We assume that the use of the mobile phone is visible in the captured image. If it is not, it cannot be detected.

of potentially tens or even hundreds of thousands of images to gather enough examples of the non-compliant behaviours. There are techniques which could be used to mitigate this, such as making use of artificially manipulated or 'synthetic' images.

Non-compliant behaviours are likely to change over time, in response to perceived enforcement activity, fashion, and new technologies, for example. This means that the deep learning networks will require regular updating to detect changing patterns of non-compliance. There is also a danger that perceived enforcement activity could lead to the emergence of more dangerous behaviours in an attempt to evade detection. This needs to be borne in mind when deploying systems such as these.

Experts consulted for this project have estimated that typical development timescales for a system like this would be some six months to produce a demonstration system, and a year to produce a system to production standards.

With the current state of hardware, specialised hardware would be required for the training system (companies already using DL will have this type of hardware available). The computational requirements for using the CNN once the training is complete are somewhat less, but for a real-time system would probably require more computational ability than a typical integrated ANPR camera. A standalone PC would probably be required.

In the current surveys, additional information the sex and age (in broad bands) of those identified as being non-compliant are also recorded. There is no reason why this could not be included in a deep learning system, though the relevant information would need to be gathered for training purposes. Training a system to categorise age would be particularly at risk of poor training as human observers are quite poor at this task, and this would be reflected in the likely performance of the DL output.

A further point to be noted is that, in current legislation, there are exemptions from seatbelt wearing laws for taxi and mini-cab drivers while on duty, as well as some medical exemptions, though these surveys are not used for enforcement activities.

Finally, it should be pointed out that deep learning is a complex technology and there is a risk that performance will not meet a sufficient level; this will be discovered only after the investment has been made in its development. Given that experts have expressed confidence in their technologies, however, an element of risk-sharing could be envisaged where a potential supplier invests some of their own time and effort in the technology in the hope of a significant return through future sales.

## 2.4 Defence industry input

| Key findings |
|---|
| • Defence industry sources did not identify any technologies in addition to those considered already. |

The defence industry has produced devices for detecting mobile phones based on detecting the radio waves transmitted by the devices. This is essentially the same technology which has been discussed in this report, and suffers from the same limitations, namely that it is not possible to detect who is using the phone (e.g. driver, passenger, passer-by, or local resident), or whether it was handheld or not. Discussion with researchers in the defence research field confirmed that no other detection technology which they were able to discuss was in use.

The military has been some of the earliest users of advance image processing techniques, typically used to detect enemy vehicles and people of interest. Deep learning algorithms are used in this field, but not to a significantly greater depth than is commercially available.

The military also makes extensive use of thermal imaging infra-red cameras. The ability to detect people in complete darkness makes these extremely valuable. However, as previously discussed, these cameras are expensive and, unless covert surveillance is envisaged, offer little advantage over normal visible light or near infra-red cameras.

A general discussion with defence industry sources has not uncovered any other existing or upcoming technologies which could be of use in detecting violating behaviour by vehicle drivers and occupants.

## 2.5 Data fusion

**Key findings**

- All options identified relate to mobile phone use only.
- Combining a mobile phone detector with a camera is technically feasible and could offer improved accuracy compared to a camera-only system.
- There are practical, technical and legal barriers to using systems based on tracking mobile phones through network data.

This section discusses the possibility for generating insights into seatbelt and mobile phone use from existing and/or new data sources. Inherently, mobile phones generate more data than seatbelts, as their purpose is to interact with other systems: as a result, it is unsurprising that the options identified relate to mobile phones only.

### *2.5.1 Potential data sources*

Five potential sources of data were identified which could potentially contribute to an understanding of seatbelt or mobile phone use on the network. These are outlined below.

No applications in addition to those discussed in the Phase 1 report (Vermaat, et al., 2018) were identified for the use of in-vehicle data in combination with any other data.

*Mobile network data*

As mobile phones move around the country, they change which cell they are connected to depending on the signal strength. Data from the network operator can be used to work out the approximate location, direction and speed of travel of a mobile phone, and specialist providers have developed tools to identify phones likely to be in a moving vehicle. Services which analyse data from this source are commercially available, though the datasets themselves may be tightly controlled.

It is not possible to use these datasets to identify whether a phone user is a driver, a passenger or a passer-by without further information (e.g. a unique identifier for phones present in each vehicle), as the locations it contains are very approximate and many phones on a particular road are likely to be using the same cell at any given time.

*Roadside camera*

As discussed in Section 2.3, it may be possible to automatically process images from a high quality roadside camera to detect mobile phone and seatbelt use: if this can be done reliably for all behaviours of interest, no data linking is required. However, if this cannot be achieved, cameras may still have an application in combination with other data sources:

- Image processing could provide data on the number of occupants in a vehicle or whether the driver is holding their hand to their head, which is likely to be less technically demanding.
- A small number of images could be manually reviewed, for example only images where a mobile phone voice call was detected coming from the vehicle (see Section 2.5.2).

*Roadside phone detector*

When mobile phones are actively transmitting (e.g. making a call, sending a text) they could be detected by a roadside unit. Investigating the signature of the transmission could tell whether it is a voice call, data transmission or text message, without collecting any personally identifiable data. It should be noted that some services (e.g. Skype, WhatsApp) transmit voice calls using the phone's data connection: these will not be detected as voice calls.

Using a directional antenna, it might be possible to determine which vehicle a transmission originated from, though not whether a user was a driver or a passenger, for example. We have not found evidence of this having been trialled in practice.

*Roadside IMSI harvester*

Each mobile phone has an International Mobile Subscriber Identity (IMSI), which is linked to the SIM (Subscriber Identity Module) card. When a mobile phone moves into range of a new cell tower, it broadcasts this number openly.

It may be technically feasible to create a device which emits a signal mimicking that emitted by a cell tower, causing all mobile phones in range to reveal their identity. Use of this in combination with a directional antenna may make it possible to identify which vehicle contains which phone including whether it is the driver's or a passenger's phone (though would not identify the contents of the phone call). This is conceptually similar to the StingRay device believed to be used by some police forces (Infosec Institute, 2014).

There are a number of potential political, financial and legal barriers to using this technology, as outlined below:

- Use of this type of device is highly controversial and seen by some as an invasion of privacy. Extension to a road safety survey may be politically unacceptable.
- The cost of the system is likely to be excessive, in the region of $400,000 per detector (Infosec Institute, 2014).
- It is not immediately clear whether use in this way would be legal. Relevant legislation includes the Wireless Telegraphy Act 2006 and the Regulation of Investigatory Powers Act 2000. We would advise anyone proposing to use a system like this to take specialist advice on its legality.

*GPS data (mobile phones)*

The majority of modern smartphones contain a GPS chip, which identifies the location of the phone relatively precisely. GPS processing has a high power requirement, meaning that the chips are only activated when required for a particular application (for example a maps application). This application may also detect when a call is being made.

### 2.5.2 Detector augmented camera

The data sources for this system would be a directional mobile phone detector and a camera. By recording precisely time-stamped observations of mobile phone use (from the detector) and passing vehicles (from the camera), images could be selected for review where at least one occupant of a nearby vehicle is making a call. A human coder would then review these images, to determine whether the call was made using a handheld device by the driver. As the coder would be presented with images taken only when a phone was in use, the false positive rate is likely to be lower than it would be using a fully manual system. On congested roads, it is more likely that an occupant of a nearby vehicle would be using a phone, hence more images would be tagged for analysis and the analysis could be more complex.

A system of this type is believed to be under commercial development for enforcement purposes.

### 2.5.3 Single occupant vehicle tracking using GPS (mobile phones)

This system would combine data from roadside cameras (with basic image processing technology) with GPS data from mobile phones. Unlike the system discussed above (Section 2.5.2), it would monitor vehicles along their journey, not just at the point where a camera is installed.

An outline of the process that this system could follow is outlined below:

- Firstly, a mobile phone with active GPS is identified moving along a road where a camera has been installed.
- The time that this vehicle passed a camera is matched to the image of the vehicle. The image is processed to identify whether there are any passengers; if there are, then calls from the vehicle are not analysed any further.

- If a call is made from the identified phone during the journey, it can therefore be assumed to have come from the driver.

There are several significant limitations of this approach:

- Mobile phone GPS has relatively high power consumption, so is only activated when absolutely required. This means it is likely that only a small minority of vehicles will contain a mobile phone with active GPS at any given time. These users are unlikely to be representative of the driving population as a whole. For example, a major use of GPS is in mapping applications: people using these might be more likely to be making longer or non-routine journeys.
- GPS data are not accurate enough to identify the location of a single vehicle on a busy road: this approach would work on only relatively quiet roads.
- Only single occupant vehicles would be monitored. These may not be representative of the driving population as a whole.
- Back seat passengers may not be captured by the cameras, meaning that calls by them may be incorrectly assumed to be the driver.
- It would not be possible to distinguish between handheld and hands free calls.
- We cannot be certain that the required dataset is held by the app developers. Even if the dataset is held, they make not be willing or legally permitted to share it.

Given these limitations, we do not see this system as a credible option for monitoring mobile phone use.

#### *2.5.4 Single occupant vehicle tracking using IMSI*

Some of the limitations above could be overcome by capturing the identity of the mobile phone directly, rather than relying on matching the time and location of a GPS record. This would work by combining data from three sources: roadside cameras with image processing technology, mobile network data and IMSI harvesters. The possible process is outlined below:

- A single occupant vehicle is identified by the camera system.
- The IMSI of this vehicle logged by the IMSI harvester is matched to the image.
- This IMSI is linked to the mobile network data. Any calls made from this phone during the same journey can be assumed to have come from the driver.

This approach overcomes the limitations caused by using GPS data. However, the other limitations still remain. In addition, new difficulties are introduced through the use of IMSI numbers: in particular, mobile networks may be reluctant to allow data containing these to be shared. This is in addition to the barriers to using an IMSI harvester identified in Section 2.5.1.

# 3 Implications

In this section we discuss the implications and potential of adopting the more promising approaches identified in this report.

## 3.1 Overall potential of new approaches

**Key findings**

- Deep learning systems have the potential to be as good as humans at coding images. However, this could only be achieved at a significant cost.
- Assisted tagging systems are lower cost and could form the first step towards a more automated system.
- Deep learning could be applied to an assisted tagging system to reduce operator workload further.
- Good quality images are essential for either approach.

The further analysis carried out in Phase 2 reinforces the view that technology has the potential to significantly reduce the number of person-hours needed to undertake surveys for non-compliant behaviours.

### *3.1.1 Application of Deep Learning*

Several experts consulted were confident that, given sufficiently good image quality, the detection of both lack of seatbelt wearing and use of handheld mobile phones would be possible to a performance level (accuracy and reliability) equivalent, or possibly even better than, human observers[7]. However, this performance would come at a significant development cost, with some six months of development required to develop a demonstration system.

As the market for this type of system is expected to be quite small, the development costs would need to be recouped through sales, resulting in high resale costs, or would need to be sponsored by a potential client.

There are risks associated with this solution. These include:

- Performance claims are necessarily speculative. These solutions would be attempting to detect very low incidence events, particularly in the case of seatbelts where non-compliance rates are very low (typically 1% - 3%), which means that the false positive rate would need to be very low, probably less than 0.1%.
- Collecting the training data required could be difficult and time consuming due to the low incidence of non-compliance. If a deep learning solution needs, say, 1,000 examples of non-compliant behaviour (this is likely to be a minimum), then assuming

[7] This observation is based on classification systems based on deep learning which have been able to classify objects based on single images as well as humans, and over time could exceed human operators who are affected by tiredness, boredom etc.

a population where 2% are non-compliant, this would mean that some 50,000 images would need to be collected and tagged. There are methods for gathering a smaller amount of data initially (say 10% - 20%) and synthesising the remaining required data, though this increases the risk of sub-optimal performance. However this may be a reasonable way forward for a 'proof of concept' demonstrator.

- Performance may be inconsistent from one site to another, depending on the variability in the training data used.
- The solution would depend on good quality images from the video cameras used, which cannot be guaranteed at this stage.

### 3.1.2 Assisted tagging

An assisted tagging scheme represents the lowest level of technology application considered, and as such also presents the lowest risk. There are a number of different ways this could be done:

- Use existing image databases and analyse the stored images. In the UK, the most relevant database is the NADC (National ANPR Data Centre) which will in future store overview images from ANPR cameras, and these could be used for analysis. It has, however, proved very difficult establishing contact with the relevant person in the Home Office to discuss the practicality of this proposal, as getting access to NADC data has traditionally been limited to law enforcement agencies. It is also unclear what fraction of images would be suitable as the camera sites are optimised for ANPR use.
- 'Tap into' existing CCTV camera streams and use simple vehicle movement detection software to store images of vehicles in a database which can then be tagged manually. This will significantly reduce the time for analysis. As above, it is unclear what fraction of images would be usable.
- Use the same approach as above, but use a dedicated camera installation. This allows the selection of optimal sites, camera types and geometries leading to higher performance.
- Using any of the approaches above, but pre-process the images using a deep learning based approach. However, in this case, the deep learning algorithm would be tuned to detect all those images which show compliant behaviour to a high degree of confidence, and present the rest to an operator for manual tagging. This will significantly reduce the workload on the operators, saving time and cost.

This solution could also be introduced in a step-by-step process, where initially one of the first three options is implemented and the automatic detection of compliant behaviour added later.

Assisted tagging represents a much lower risk approach than a deep learning based solution. The main risk is in the quality of video images collected, particularly for solutions based on existing CCTV or ANPR cameras.

## 3.2 Metric design and statistics

**Key findings**

- A pilot study could be carried out to manage the technical risk of introducing a new process, and to understand whether the results are comparable to the existing approach.
- New technologies offer the chance to improve representativeness by including, for example, night time and high speed observations of seatbelt and mobile phone use.
- Sites should be selected to represent the road user population as well as possible. It is always possible to improve the accuracy of a survey by including more sites: ultimately, the number of sites at which a survey should take place is a balance between affordability and representativeness. The existing DfT survey has a sample size large enough to detect large changes in compliance. However, increasing the sample size would make the analysis able to detect smaller changes.

This section considers how a survey of seatbelt and mobile phone use could be designed using new technology.

Three important considerations when planning a seatbelt and mobile phone survey using a new approach are:

- The technical ability of the new system to make accurate observations, and whether these can be compared to the existing methodology.
- How representative the sample is likely to be in terms of non-compliant behaviours of the road user population as a whole. Considerations for achieving a representative sample are discussed in Section 3.2.2.
- How large the sample size is. The larger the sample size, the smaller the change in compliance required to achieve statistical significance. If the sample size is too small, small changes in compliance may be concealed by random variation. Initial sample size calculations are included in Section 3.2.3.

### *3.2.1 Pilot study*

A pilot study could be carried out to establish the feasibility and technical benefits of any new technique which is being considered for adoption. This should aim to identify:

- Any practical issues with using the technique
- The technical ability of the technique in identifying the required variables in real-world conditions
- How comparable the results are likely to be with those resulting from the existing methodology

Carrying out a feasibility study would help to manage the technical risk of introducing a new approach: if the performance is inadequate changes can be made before any larger scale survey is commissioned. However, since the systems proposed in this report have not yet been developed, a level of investment would still be required before this point is reached.

As a first step, assisted manual tagging is the most promising of the data collection techniques discussed in this report. The remainder of this section therefore discusses how a pilot study of this technique could be carried out.

To provide confidence in the technical ability of the technique, it is essential that its use is trialled under a range of operational conditions. These should include:

- A range of mounting positions which may be used in practice
- A variety of road types and speeds
- Lighting conditions including in darkness with and without street lighting, and in bright sunlight

There are three numerical measures which can be considered when evaluating the technical success of the approach:

- The proportion of 'unknown' observations. This can be compared directly to the existing survey and provides a measure of whether the people carrying out the coding feel confident in their observations. A large increase in this proportion could suggest that the new technique is less reliable.
- Inter-rater reliability. In contrast to the other measures, there is no output from the existing surveys which is equivalent to inter-rater reliability. This means that it cannot be used to judge whether the approach represents an improvement on the current methodology. Its calculation still has value, as it can be used to identify differences in accuracy between road types or individual coders.
- Finally, it may be possible to compare overall non-compliance rates at the same site between both methodologies. Sites would be identified which are suitable for both human observers and camera installations (ideally existing survey sites). Images from the camera could then be analysed for a period when human observers were also present. Realistically, this approach could only provide an indication of the new system's performance on local roads during daylight, and the number of suitable sites is likely to be limited. However, it could also provide an indication of whether the results from the new technique are comparable to those of the existing survey.

The proportion of unknowns and detected non-compliance rates from a pilot study can be fed into the sample size calculations for a wider survey (see Section 3.2.3).

### *3.2.2 Representative site selection*

Once the performance of the technology is understood, a site selection process can be defined. The study design will depend on the objective of the survey and should be formed in consultation with Highways England and the DfT. Possible research objectives include:

- Gathering a representative indication of seatbelt and mobile phone use across the whole country, to enable trends over time to be monitored. This would require widely geographically distributed sites across a variety of road types and is probably the simplest option.
- Comparing the seatbelt and mobile phone use rates on the SRN to those of local authority roads. This in effect requires two surveys: one for local authority roads and one for the SRN. Each of these will require an adequate distribution of sites and an adequate sample size, meaning it is likely to be more costly than the option above.

- Comparing seatbelt and mobile phone use between areas. This effectively requires one survey per area to be compared, making it the most complex option.

Compared to the existing methodology, several features of an assisted manual tagging process mean that the results may be more representative of road user behaviour as a whole. These include:

- The ability to make observations at night, when journeys may be made for different purposes and by different demographics. This would also make it possible to run surveys at any time of year, which is not possible with the existing approach.
- The ability to make observations of passenger behaviour in moving vehicles
- Coverage of high speed roads including motorways

However, the requirement to install cameras at sites may lead to financial pressure to minimise the number of locations where the survey is taken. Reducing this number too far creates a risk that local conditions could skew the results. This is unlikely to be mitigated through increasing the sample size at each site.

Determining whether the sites selected for a survey are sufficiently representative is ultimately a judgement call. However, best practice includes considering the following factors:

- The proportion of traffic flow on the road network which each site type represents. For example, motorways account for approximately 21% of vehicle miles travelled in the UK (Department for Transport, 2017): for a national survey, approximately this proportion of observations should come from motorway sites.
- The time of year when results are collected. This may be managed either through consistency (always carrying out surveys at the same time of year) or distributed sampling (taking observations on dates distributed throughout the year).
- The day of the week when results are collected: road use patterns may vary considerably between days.

#### *3.2.3 Adequate sample size*

This section discusses the sample size requirements for a survey using the technologies discussed in this report. It is an extension the analysis in the Phase 2 report (Vermaat, et al., 2018) and evaluates the benefits of an increase in sample size, which may be made viable by a new methodology. The calculations presented here are based on testing the change in total compliance between surveys: this is equivalent to the first of the options outlined in Section 3.2.2.

Figure 6 shows the relationship between the number of observations made in a mobile phone survey and how sensitive it is (i.e. how large a change in the proportion of non-compliance needs to be for this change to be statistically significant).

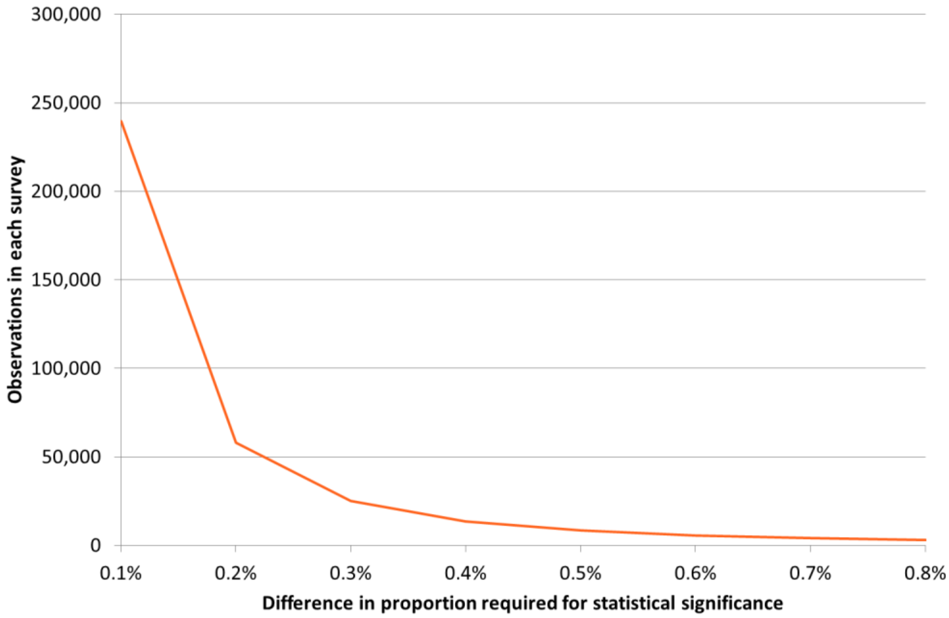


**Figure 6: Relationship between sample size and the change in compliance required to be statistically significant when comparing two mobile phone surveys**

The current DfT survey collects approximately 26,000 mobile phone observations[8], which means that, for a statistically significant difference to occur between two surveys, the rate of use must change by approximately 0.3% (e.g. from 1.6% to 1.9%). Increasing the number of observations improves the sensitivity; for example, if 58,000 observations were made then it would be possible to detect a change of 0.2%.

Figure 7 shows the relationship between the sample size and ability of the analysis to detect a small change for the seatbelt survey (both drivers and passengers).

[8] 2017 observations at static sites, rounded to the nearest thousand and excluding observations where mobile phone use could not be determined

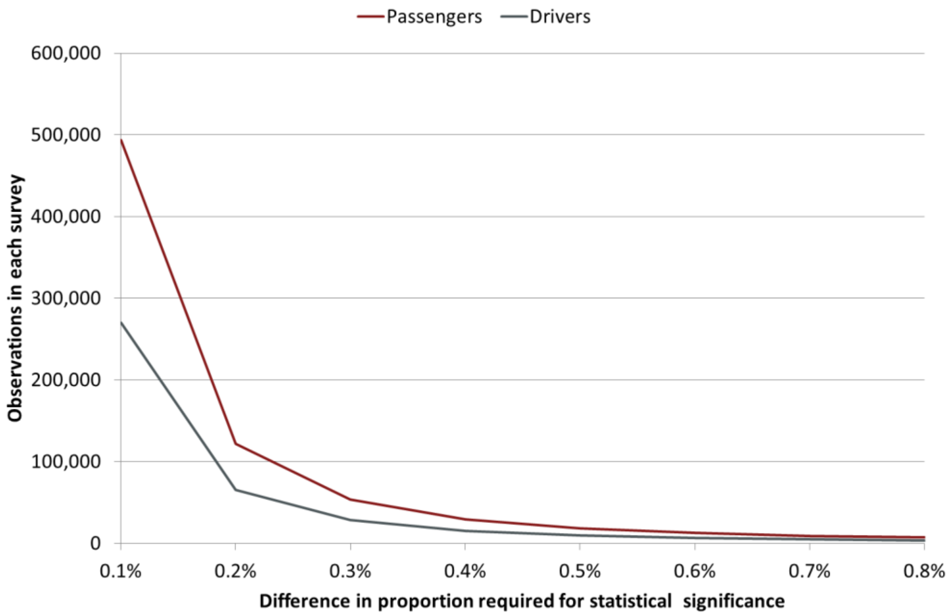


**Figure 7: Relationship between sample size and the change in compliance required to be statistically significant when comparing two seatbelt surveys**

The number of observations required to detect a change in passenger use is much higher than for drivers: this is because not all cars have passengers.

The existing survey achieves 23,000[9] vehicle observations, which is sufficient to detect a change of slightly over 0.3% for drivers and between 0.4% and 0.5% for passengers. There are fewer observations made in the existing seatbelt survey than the mobile phone survey as observers are required to collect more data from each vehicle.

The existing sample size is sufficient to give an indicative rate of non-compliance and its trend over time. However, the sample size may not be large enough for some potential uses, such as to evaluate the effect of initiatives on driver behaviour. For example, if an enforcement campaign caused a 10% reduction in driver phone use, this would result in the non-compliance rate changing by only 0.16%, which would not be a statistically significant change using the current methodology. It might also be desirable to look at regional or local trends: this would again require a larger sample size to detect the same change. Increasingly large sample sizes can produce statistically significant results from smaller and smaller changes; however, eventually the changes become too small to be of practical importance.

*Assumptions*

A number of assumptions[10] have to be made in the calculation above. Firstly, the estimates presented here are based on the assumption that non-compliance at the sites observed in

[9] 2017 observations at static sites, rounded to the nearest thousand and excluding observations where variables could not be determined

[10] Some additional statistical assumptions are made including a Type I error of 5% (and a 2-sided test), and a Type II error of 20% - these are common standards.

2014 is similar to non-compliance at any future sites. The 2014 mobile phone and seatbelt compliance surveys estimated that:

- 1.6% of drivers in moving traffic were using a handheld mobile phone
- 1.8% of car drivers were not wearing a seatbelt
- 3.3% of car front seat passengers were not wearing a seatbelt

Small changes in these rates between surveys will not make a large difference to the sample size required. It is possible that a larger change in the measured rate of non-compliance could result from introducing a new methodology: this would alter the sample size required. It is not possible to perform this calculation until the first results using the new methodology have been collected. It would be advisable to undertake a joint survey using both traditional and technologically assisted methods to take account of this effect.

The numbers of observations presented in this analysis relate to the number of valid observations (i.e. those where the required variables can be determined). They will need to be increased to allow for a proportion of 'unknown' observations following a pilot study.

## 3.3 Collection and processing of in-vehicle data for analysis

**Key findings**

- The majority of vehicles currently in operation are not capable of sharing any detailed data.
- Where data are collected, there is no standardisation of what is collected or how it is stored. There are proposals to enable a standard minimum dataset to be shared, but this is unlikely to include seatbelt data.
- It does not appear to be in the commercial interests of manufacturers to share data with authorities unless they are compelled to by law.
- The data collected may not accurately reflect those studied.
- The makeup of the sample may change over time.

### *3.3.1 Technical considerations*

At present, data (if any) gathered from vehicles vary substantially between manufacturers. The majority of vehicles transmit no data at all; however, new models are increasingly being offered with connected equipment, especially in premium vehicles. These can be used to provide a wide range of services, from control of the heating to live information on parking locations (BMW, n.d.). Manufacturers have reported considerable interest in these data from third parties, such as advertising agencies, but the industry is believed to see a reputational risk in releasing this information (Sharman, 2015).

Some manufacturers may make the data that they collect available to third party services. The head of connected car services at BMW has stated in an interview that he can see the need for this data sharing, but that it must be based on customer consent (European Automobile Manufacturers Association, 2018).

An ISO standard (ISO 20078-1) has been written which defines a process for sharing data from connected vehicles. The standard does not address the content of data to be collected and shared, leaving this to be decided for each application. It acknowledges that there may be a need for authorities to have access to some data. McCarthy, et al. (2017) notes that ownership of the data is controversial.

A TRL report for the European Commission (McCarthy, et al., 2017) explored in detail the possibilities for making data available to appropriate third parties. It identified a risk, in the absence of any governmental intervention, that car manufacturers could create a dominant market position by controlling the interface with vehicles.

The authors discussed several different technical architectures for data sharing and how they would impact on the provision of services. Of particular relevance to the subject of surveys is the need they identified for a standardised minimum dataset to be available from all connected vehicles. This is required to enable third party applications to work across the fleet, rather than for only a particular manufacturer, but could also enable standardised survey data capture. A similar approach has been adopted for event data recorders in the USA (Code of Federal Regulations, 2006), which collect information for use in collision investigation. The report also considers applications and use cases, but there is no suggestion that any collection of seatbelt data is currently planned.

### *3.3.2 Legal considerations*

The General Data Protection Regulations (GDPR) have tightened and standardised the regulations for storing and processing personal data in Europe. Importantly, they have also introduced substantially greater penalties for non-compliance than previous legislation: this may have caused organisations to adopt a more cautious attitude towards data processing. Anecdotally, one vehicle data services provider has indicated that the level of detail they receive from manufacturers has reduced as a result of these regulations.

The information in this section is based on the guidance published by the information commissioner (Information Commissioner's Office, 2018).

There are two broad approaches which could be taken to gathering data from vehicles. In both cases, data are first collected by the manufacturer. The two possibilities are outlined below:

- Personal data could be processed by the manufacturer to produce statistics. These can then be shared freely, since they cannot be used to identify an individual.
- Personal data could be supplied to a processing organisation, such as TRL or the DfT.
- Data could be collected by the manufacturer, but anonymised and processed by a trusted, independent third party, such as TRL, for the DfT.

If the manufacturer collects personal data based on consent, it is not possible for the data to be used for anything except that explicitly agreed by the individual. This would prevent either of the above suggestions from being implemented, unless the data subject actively consents to the processing. If the manufacturer collects data under a different basis, the GDPR does allow processing for ‘statistical purposes’, though individuals involved must still be provided with information on the new use of their data.

If the DfT or Highways England wishes to carry out data analysis themselves, there are additional considerations. Firstly, data controllers are required to identify a lawful basis for collecting and processing data. For a public authority, they must consider the 'public task' basis first for most of their processing, and are less able to rely on the 'legitimate interests' basis. 'Public tasks' are considered to be tasks performed in the public interest or as part of an official function, and must have a clear basis in law. Legitimate Interests is still available as a basis for processing for other legitimate processing outside of those tasks. The conditions for potentially relying on either basis would need to be fully reviewed. Secondly, arrangements for the sharing of data should be considered: these must balance the need to share the data with the interests of the data subject.

#### *3.3.3 Commercial considerations*

We are not aware of any legal obligation for manufacturers to share data to be used to gather government statistics. This could be a major barrier to deployment: if manufacturers are not willing to share data, nothing will be collected. In making decisions to share data, maintaining the trust of customers is likely to be a major consideration and will be weighed against any commercial benefit to sharing. Anecdotal evidence suggests that manufacturers are relatively unwilling to share data at present.

#### *3.3.4 Practical considerations*

Two main practical considerations will define whether the results from an exercise like this would be useful. These are:

- Whether the data collected provide an accurate reflection of the behaviour of those studied.
- Whether the individuals studied are representative of the whole driving population.

Drivers who do not wear their seatbelts have an incentive to try to fool in-vehicle sensors. These sensors are primarily installed as an input for seatbelt reminders: cars meeting the EuroNCAP recommendations will cause an audio-visual warning to sound when driven without the detection of seatbelt use (EuroNCAP, 2017). Drivers may choose to prevent the sensors from working effectively, for example by purchasing an additional seatbelt buckle to fool the system into thinking a seatbelt is worn.

In the medium term, seatbelt sensor data are likely to become available only from newer and/or premium vehicles. The users of these vehicles will almost certainly not be typical of the driving population as a whole. In addition, it appears likely that any data collection will require the agreement of users, vehicle manufacturers or both. If for some reason the make-up of the sample changes over time, for example due to a manufacturer increasing their market share, data may not be comparable between surveys. Taken together, these factors mean that it would be difficult to have confidence in the representativeness of any data collected using this method.

# 4 Conclusions and recommendations for next steps

From the discussions above, we draw the following conclusions:

1. The quality of images captured from standard CCTV or ANPR overview cameras has challenges which make them less than ideal for use in an analysis system. However, they should not be entirely discounted as they could provide a source of very low cost images for use in an assisted tagging system. Development of such a system would require 6-12 months of development time. Costs would depend on the level of risk the developers are prepared to accept.
2. Technologies exist which can provide images of somewhat higher quality, specifically using pulsed IR cameras like those used in ANPR systems, but with substantially higher power illumination. The requirements for fixed and mobile cameras are similar in terms of performance, though the practical deployment requirements of mobile cameras will make it difficult to achieve the same level of performance as fixed cameras.
3. The development of an assisted tagging approach is recommended as it offers the most potential to provide a low cost and low risk way of reducing the effort required in manual tagging. Costs would be driven by the level of assistance envisaged.
4. The automatic detection of non-compliant behaviour using specialised cameras (see previous point) and deep learning based classification has the potential to provide a fully automatic measurement solution. However there is a level of risk, so a step-by-step approach would be beneficial (see recommendations below).
5. Using automatic detection in an assisted tagging solution provides a low risk, stepwise way forwards.
6. The gathering of relevant in-vehicle data (mainly for detection of seatbelt non-compliance) has numerous significant challenges relating to data privacy, lack of clarity about data ownership, lack of standards in gathering the relevant data, and likely objections from vehicle manufacturers.
7. The fusion of data from multiple sources is technically feasible but, as with the previous point, the legal and privacy related challenges mean that this solution would need to be approached on a case-by-case basis as each combination of data sources will present a unique set of challenges.
8. The use of external sensors to detect mobile phone usage to the level of accuracy required does not seem feasible with current technology.
9. Current surveys use sample sizes (ca. 27,000) which can detect a change of 0.3% in non-compliant behaviour. The metrics required to reliably detect a small change in driver/vehicle occupant behaviour mean that large sample sizes are required. If technology is used to reduce the cost per measurement, this will make it possible to measure larger sample sizes, or to sample more frequently, or both. There is a level of uncertainty regarding the effect of new technologies on the metrics used, so a pilot study should be considered to identify suitable sample sizes and how

comparable the results from technology-driven surveys will be to existing manual survey methods.

10. The use of technology will allow a wider range of sites to be surveyed, allowing differences between geographical areas, road types and demographics to be determined.

We therefore recommend that:

1. A research and development programme is initiated to develop a roadmap to future automated data collection. This programme should start with the demonstration of a suitable camera solution. Any system based on image processing, either by human operators or by automatic means, will require clear, high-quality images.
2. An assisted tagging solution is developed, which includes intermediate solutions based on low levels of automation, moving on to automatic detection based on deep learning.
3. A deep learning based demonstration system is developed with a suitable industry partner, initially as part of an assisted tagging solution, with the aim to eventually progress to a fully automated system.
4. The feasibility of using CCTV images as a data source is investigated.
5. The DfT approaches the Home Office about the potential use of NADC images in an assisted tagging solution when these become available.
6. If suitable technology is developed, consider running a pilot study to develop and compare metrics between existing and new survey methods.

## References


BMW, n.d. *BMW ConnectedDrive Customer Portal.* [Online]
Available at: https://www.bmw-connecteddrive.co.uk/app/index.html#/portal/store
[Accessed 21 06 2018].

Code of Federal Regulations, 2006. *49 CFR 563.7 - Data elements.* [Online]
Available at: https://www.law.cornell.edu/cfr/text/49/563.7
[Accessed 21 06 2018].

Department for Transport, 2017. *Road traffic estimates in Great Britain: 2016.* [Online]
Available at: https://www.gov.uk/government/statistics/road-traffic-estimates-in-great-britain-2016
[Accessed 01 08 2018].

EuroNCAP, 2017. *Assessment protocol - safety assist.* [Online]
Available at: https://cdn.euroncap.com/media/32283/euro-ncap-assessment-protocol-sa-v802.pdf
[Accessed 28 06 2018].

European Automobile Manufacturers Association, 2018. *Video - "Access to vehicle data must be safe, secure and protect privacy".* [Online]
Available at: http://www.acea.be/news/article/video-access-to-vehicle-data-must-be-safe-secure-and-protect-privacy
[Accessed 21 06 2018].

Information Commissioner's Office, 2018. *Guide to the General Data Protection Regulation (GDPR).* [Online]
Available at: https://ico.org.uk/for-organisations/guide-to-the-general-data-protection-regulation-gdpr/
[Accessed 21 06 2018].

Infosec Institute, 2014. *StingRay Technology: How Government Tracks Cellular Devices.* [Online]
Available at: http://resources.infosecinstitute.com/stingray-technology-government-tracks-cellular-devices/
[Accessed 20 06 2018].

Janitzek, T. & Achterberg, F., 2006. *Seat Belt Reminders: Implementing advanced safety technology in Europe's cars.* [Online]
Available at: https://etsc.eu/wp-content/uploads/ETSC_Seat_belt_reminder_oct_06.pdf
[Accessed 20 06 2018].

McCarthy, M. et al., 2017. *Assess to In-vehicle Data and Resources.* [Online]
Available at: https://ec.europa.eu/transport/sites/transport/files/2017-05-access-to-in-vehicle-data-and-resources.pdf
[Accessed 21 06 2018].

Sharman, A., 2015. *Financial Times: BMW sounds alarm over tech companies seeking connected car data.* [Online]
Available at: https://www.ft.com/content/685fe610-9ba6-11e4-950f-00144feabdc0
[Accessed 21 06 2018].

Vermaat, P., Mohan, S. & Myers, R., 2018. *Review of Technology for Undertaking Mobile Phone and Seatbelt Surveys on the SRN ,* Wokingham, Berkshire: Transport Reseach Laboratory (TRL).

# Appendix A Estimated time required for video coding

## A.1 Assumptions

- The time taken to observe an image and record its contents in a coding interface is similar to the time taken to record the same observations by verbal means in the existing survey.
    - It may be slightly faster to select observations on screen than to verbalise them: in this case, the figures below will be an overestimate.
    - A poor quality image may significantly increase the amount of time required.
- Each variable (one unit of information, e.g. vehicle type, vehicle colour) takes an equal length of time to record. This is consistent with the audio recording from existing surveys, though the time taken may be longer when child seats observations are made. (The latter account for a relatively small proportion of observations, especially if rear seat passengers are excluded.)
- ‘Unknown’ observations are included in the count: it is assumed that the proportion of these will remain similar in any new methodology.

## A.2 Limitations

- It has not been possible to validate these results using a real coding interface
- Images of rear seat occupants have been excluded.

## A.3 Variables to be recorded

**Table 1: Variables to be recorded for each vehicle**

| | Vehicle without passenger | Vehicle with passenger |
|---|---|---|
| Number of variables | 6 | 10 |
| | Vehicle type | Vehicle type |
| | Colour | Colour |
| | Driver gender | Driver gender |
| | Drive age | Drive age |
| | Driver seatbelt | Driver seatbelt |
| | Drive phone | Drive phone |
| | | Passenger seatbelt |
| | | Passenger presence |
| | | Passenger age |
| | | Passenger gender |

Time has been allowed for an additional three units of information per vehicle, to give coders time to look at the image, check their work, and select the next vehicle.

## A.4 Inputs

**Table 2: Other model inputs**

| | | |
|---|---|---|
| Proportion of vehicles with passengers | 35% | |
| Time per unit of information | 1.32 | s |
| Time between breaks | 45 | minutes |
| Length of break | 15 | minutes |
| Contingency | 20% | |
| Quality check % | 10% | |

- 35% of vehicles observed in the 2017 static seatbelt surveys contained at least one passenger.
- Estimates of time per unit of information were made based on the average of 15 audio recordings for the existing surveys. Timings were measured from when the coder started to record a new vehicle to the description of the front seat passengers being complete. ‘New vehicle’ was included as a unit of information.
    - Time per observation was in the range 1.0s to 1.6s, with a mean of 1.32s.

## A.5 Outputs

**Table 3: Model outputs**

| | | |
|---|---|---|
| Information units per vehicle | **10.4** | units |
| Average observation time per vehicle | **13.7** | s |
| Time including contingency | **16.5** | s |
| Number per average working hour | **164** | observations |
| Number excluding quality checks | **149** | observations |

- The number of information units per vehicle is a weighted average of vehicles with and without front seat passengers.
- In total, the model predicts that an observer will record 149 observations of their own and independently quality check 15 made by other people per hour.

# Mobile Phones and Seatbelts Technology Review (Phase 2)

Following their identification in Phase 1 of this project (Vermaat et al, 2018), several technologies for monitoring mobile phone and seatbelt use by vehicle occupants were selected for further investigation. These included an 'assisted tagging' concept, where human coders review partially processed images captured from cameras, data fusion concepts and deep learning technology. Research into these areas was carried out through engagement with subject matter experts. Recommendations include a research and development programme to work towards future automated data collection, starting with the demonstration of a suitable camera solution. Initially, this could be deployed as part of an 'assisted tagging' process which has the potential to deliver more useful survey data at lower costs than existing methods.

**Other titles from this subject area**

| | |
|---|---|
| **In press** | Feasibility study: Seatbelt and mobile phone use surveys on the SRN. R Myers, L Croft, L Durrell, S Chowdhury. 2017 |
| **PPR870** | Review of Technology for Undertaking Mobile Phone and Seatbelt Surveys on the SRN (Phase 1). P Vermaat, S Mohan and R Myers. 2018 |

**TRL**
Crowthorne House, Nine Mile Ride,
Wokingham, Berkshire, RG40 3GA,
United Kingdom
T: +44 (0) 1344 773131
F: +44 (0) 1344 770356
E: enquiries@trl.co.uk
W: www.trl.co.uk

ISSN

ISBN

**PPR871**