Evaluation of rail safety interventions – Developing a framework for evaluation of technologies for GB rail

Dr Brendan Ryan, Arran Philippou

Human Factors Research Group, University of Nottingham, University Park, Nottingham,

NG7 2RD

17 November 2015

Summary

This project was motivated by the realisation in the industry that many suicide prevention initiatives are implemented without giving sufficient consideration to all of the necessary steps in planning for and, in particular, collecting sufficient data to evaluate the likely impacts of these interventions. The provision of blue lighting at two stations, on two separate routes, has been reviewed and reported in Ryan and Philippou (2015). Evidence drawn from this review has been described and interpreted. The findings have been used in producing a framework that can be used to highlight important steps that are needed in taking a more systematic approach to the planning, implementation and evaluation of the impact of the intervention. This report describes the development and content of this framework. To test the potential for wider use of this type of structured approach, the framework has been used to describe the types of data and research methods that are needed in two examples: (i) the study of a new blue light initiative and (ii) the study of the use of a new Smart CCTV camera system on the railway. Findings from use of the framework with these examples are discussed.

1. Introduction

1.1. Background

The University of Nottingham, who have experience of working on the evaluation of preventative measures for railway suicide on the RESTRAIL project (www.RESTRAIL.eu) were asked to conduct a short study of the processes involved in the implementation of blue lighting at stations, in order to develop a framework (a simple structured process and guidelines) that could be used to help industry staff. Practically, the following scenario is being considered:

- A station manager is fitting / considering fitting blue lights at a station.
- What do they need to do to know that the lights will be fit for purpose? (i.e. in relation to design and implementation)
- What do they need to do to know that they work, in their specific context or situation? (i.e. for evaluation)

This project also gives consideration to what is needed for the evaluation of other new technologies for the prevention of rail suicide. This has regard to examples of other technologies that are in development currently and clarifies the types of questions that need to be considered in the evaluation of these technologies. A core framework has been proposed, outlining the most important data requirements and data collection / analysis activities. The framework has been tested by applying the framework to examples of two emerging technologies (i.e. blue lighting at stations and a Smart CCTV camera). These case examples help to produce guidance for application of the framework. Recommendations are provided for the additional consultation and development that is needed for more testing and customising of the framework for future work in different contexts in the industry.

Global
Top 100
University

The University of
Nottingham
UNITED KINGDOM · CHINA · MALAYSIA

1.2 Some initial requirements for the framework

A set of requirements for the framework were proposed at the start of the project. These were that
the framework should be:

- Easy to use, understandable
- Capable of being used by people who are not experts in evaluation or research
- Relate to technologies (e.g. blue lighting, CCTV, Virtual assistant) as well as other safety
  interventions (physical barriers)
- Likely to incorporate a step-by-step guide, leading people through what needs to be done,
  before, during and after implementation of the intervention
- Based around things that can be done quite quickly (e.g. in response to recent incidents),
  easily deployed and without a lot of additional work for operational staff.

It should be acknowledged at the outset that this is intended to be a practical framework and guide
to help improve the ability of operational staff to plan for, introduce and start to evaluate the
effectiveness of preventative measures at locations under their control. It is likely that more
ambitious, in-depth programmes for implementation and evaluation could be possible by drawing
on expertise (externally or internally in the industry) on each occasion that a programme is initiated.
At present, there is a need within the industry to try to react to minimise the effects of incidents at
stations. This project has been developed to provide guidance to those involved in the design and
implementation of suicide prevention interventions, so that better information about the likely
effectiveness of these interventions can be collected.

1.3 Developing the framework

The review in Ryan and Philippou (2015) has been a useful exercise to understand the range of
factors that influence decisions in the design and development of a new technology and the
opportunities for collecting relevant data (e.g. data types, what, where, how) on the implementation
and outcomes of a safety intervention in this type of context.

These findings are now used to produce guidance for operational staff to collect relevant data in
similar situations. A simple framework is presented to provide guidance for evaluation, as part of
programme planning.

There are a number of published frameworks for programme evaluation (e.g. PRECEDE, PROCEED,
RE-AIM, Fink, 2015). These have the capacity to be used with complex, large scale programmes.
Even where the programme / intervention is of a narrower scope (e.g. an item of new technology)
there is a multitude of potential evaluation questions that can be considered and relevant data types
that can be collected to demonstrate the effectiveness, cost, quality and value of the programme. In
this piece of work a simple structure has been presented, based on important components of
evaluation programmes (Fink 2015). The six stages of the framework are outlined briefly below. In
Appendix 1, the framework has been used as a structure to summarise the progress that has been
made in two blue light studies that have been reviewed in Ryan and Philippou (2015). In this current
report, two examples are presented (in Section 3) to demonstrate how the framework can be used
for the purpose of evaluation of two technology based safety interventions.

2. An outline of the framework for collection of data for the evaluation of safety interventions

Important components of the framework are shown in Figure 1. More explanation of the things that should be considered in applying the framework is given in Sections 2.1 to 2.6. Later in Section 3, two examples are presented that give further guidance on the use of this framework in practical situations.

Figure 1 – Components of the framework for collection of data for the evaluation of safety interventions

1. Setting objectives
   - How does the preventative measure work?
   - What exactly are you trying to show in your study?

2. Identifying evidence that is needed to show that the preventative measure works as it should
   - What evidence is needed to fulfil your study objectives (including the effectiveness, quality and value of the intervention)?
   - When is the right time to collect the evidence that you need?

3. Selecting locations and people to study
   - Where should the study take place?
   - Who should be involved?

4. Collection of data
   - What are the best methods to collect the data that you need?
   - How do you collect these different types of data?

5. Analysis of data
   - What do you need to do with the data that you collect?
   - How do you deal with different types of data (e.g. quantitative data, descriptive / qualitative data)?

6. Concluding and reporting
   - How do you draw appropriate conclusions from the data that you have collected and analysed?

2.1. Setting Objectives

How does the preventative measure work? What exactly are you trying to show in your study?

The objectives of the study must be articulated precisely. These objectives will be essential in providing a clear focus for how and why the safety intervention will be applied, as well as helping to identify the measures and targets that can be used to demonstrate success of the project. This will require you to give some consideration to the design and specification of the proposed technology, the context into which the technology will be placed and what you can anticipate about how the technology will be used.

You may have some longer term objectives (e.g. reducing the numbers of suicides) and some intermediate objectives (e.g. to show that an intervention, such as a sound warning, is changing behaviours of passengers in the way that you thought it would).

2.2. Identifying evidence that is needed to show that the preventative measure works as it should

What evidence is needed to fulfil your study objectives (including the effectiveness, quality and value of the intervention)?

You need to be confident that the things that you are measuring (e.g. the numbers of incidents at a station, the behaviours of people at the station) are linked clearly to the objectives of the study. It is important that any changes in the things that you are measuring are indicative of real changes in the study environment.

You should identify the types of evidence that help to demonstrate that the new technology prevention measure has the following qualities:

Effectiveness
- that it has the impact that it was expected to have
- that it works in different settings and contexts
- that its effects can be sustained over time

Quality
- the technology solution is based on good theory or findings from earlier systematic analysis of the problem

Value
- that the benefits outweigh the costs
- the technology meets a need that is not satisfied by other means
- the technology provides a sufficient improvement within available resources.

When is the right time to collect the evidence that you need?

There are various types of evidence that can be considered, at different time periods (e.g. initial baseline – prior to an intervention; during the programme – collecting data related to the implementation and initial findings on changes in behaviours or perceptions of users; after the programme –to determine likely effects and sustained effects of the programme). Outcome data can include statistics such as the numbers of suicides, attempted suicides and other pre-suicidal

Global
Top 100
University

The University of
Nottingham

UNITED KINGDOM · CHINA · MALAYSIA

behaviour. Descriptive data about the potential impacts of the intervention could include the following: descriptions of behaviours / changes in behaviours at stations, such as how people move about the stations, where they stand, whether there are areas that are not used; also opinions of staff and other users at the stations about the likely effects of the technology, their acceptance or trust in the technology, their responses to the intervention.

2.3. Selecting locations and people to study

Where should the study take place?

This includes identification of the locations at which technology interventions will be implemented and trialled, but also other locations that should be included within the study for the purpose of data collection (i.e. for control data in studies to compare the study location with another similar location where the technology has not been implemented). Studies can range from separate, single station case studies (e.g. in the early stages of trialling a product) to multi-site studies (e.g. at later stages of product development). It is important to be able to explain the reasoning for inclusion of each station in the study. Stations might be chosen because of questions of need (e.g. there has been evidence of a problem at a station, other solutions are not practicable, there has been an opportunity to implement and test the preventative measure).

You may need to get expert support for selecting an appropriate sample of stations to include in a larger study, so that the types of stations that you choose to test the technology represent different characteristics of stations on the railway and can tell you how the technology can work in different contexts. It is important to be aware of statistical problems when selecting stations for inclusion in an intervention study (e.g. there can be a problem of regression to the mean, where implementing a preventative measure following a number of recent incidents may appear to be effective, but any change in numbers of incidents may have resulted from normal fluctuations in the data, rather than as a result of the intervention). This type of problem can be minimised by collecting data from a control station that is matched on relevant criteria to the station that is being studied.

Who should be involved?

It is also important to identify the individuals or groups who can provide relevant information on the implementation of the preventative measure and outcomes from the intervention (e.g. effects on incidents, opinions and perceptions of the intervention)

2.4. Collection of data

What are the best methods to collect the data that you need?

There are various methods that you can consider for the collection of this type of information, such as interviews, questionnaire surveys, rating scales, observations at sites and collection of statistics from industry databases. You are likely to choose one of more of these for your evaluation study.

How do you collect these different types of data?

Your plan for collecting data should be guided by your objectives for the evaluation (section 2.1) and your decision on the evidence that is needed (section 2.2). You will have identified relevant

Global
Top 100
University

The University of
Nottingham
UNITED KINGDOM · CHINA · MALAYSIA

locations for study and any relevant groups of people to consult in Section 2.3. You will need to consider the practicalities of collecting relevant data types, using the different types of methods (e.g. could you expect to observe the expected behaviours in a set of short observations from the platform, or is video recording over a longer time period and more detailed analysis required? How long would it be necessary to carry out observations at a location in order to be able to see sufficient examples of the types of behaviours that you are expecting to occur?).

2.5. Analysis of data

What do you need to do with the data that you collect? How do you deal with different types of data (e.g. quantitative data, descriptive / qualitative data)?

Qualitative and quantitative data are likely to be collected, though the precise nature of the data is dependent on the type of study design that is chosen and the methods that have been used to fulfil the study objectives. There are various stages involved in transcribing, inputting, collating and preparing the data set, prior to carrying out analysis of the descriptive content (e.g. the things that people have told you in interviews / questionnaires or the things that you have seen in observations at the stations) and quantitative data (e.g. counts of the numbers of acceptable / unacceptable responses of passengers to sound warnings, or counts of the numbers of movements into restricted areas at a station).

Your analysis of the quantitative data may help you understand the types of changes that have occurred as a result of the technology intervention, and whether these seem to be sustained over time. Your analysis of the descriptive data will provide you with a greater understanding of the types of impacts that the technology has had (e.g. the ways in which behaviours have changed and how staff and passengers have responded to the intervention). Your analysis of the descriptive data will also help you to determine whether the technology intervention has been implemented in the way in which it was planned (and conceived).

2.6. Concluding and reporting

How do you draw appropriate conclusions from the data that you have collected and analysed?

Conclusions need to be drawn on the basis of reliable data that have been collected in relation to the programme. All relevant information from the project should be considered, but it is necessary to avoid the temptation to base conclusions on data that are anecdotal in nature. It is likely that several different sources of data will be used during the study and it is necessary look for consistency and inconsistency in evidence from different sources.

The structure and content of any report will be influenced by the intended purpose and readership of the report. It is important to draw conclusions on important findings relating to implementation of the programme (in different types of settings) and any outcomes from the programme (bearing in mind that it might take some time for some of the longer term outcomes of an intervention to be determined).

3. Examples of use of the Framework

Two examples of use of this framework are now presented as a means of illustrating the things that should be considered by those who are planning evaluation activities. The first of these relates to a new blue lighting scenario. For the purpose of this scenario it is assumed that there is not yet an understanding of the likely effectiveness of blue lighting as a means of preventing rail suicide, or clarity on the mechanism of how blue lighting may work (see further discussion on this in Ryan and Philippou, 2015). However, there are likely to be circumstances where a manager may want to test a new or promising preventative measure. It is important that this kind of innovation is possible, though it is also important that this is approached in the right way and that effort is made to learn as much as possible about the effectiveness of the measure. The six basic steps are now considered relating to this scenario.

3.1. Example 1 – Evaluation of a blue light intervention

3.1.1. Objectives

How does the preventative measure work?

A number of potential mechanisms for how the lighting may work have been identified in the preliminary research for the blue lighting intervention. It would appear that the blue lighting intervention in Japan was based around use of overhead lighting to the whole platform area. The intention of this blue lighting safety intervention is to try to replicate the reported success of the work on Japan, in order to reduce suicides at the station. Therefore, the study would involve fitting similar lighting to the specification that was used in Japan. The longer term objective for the project would be to understand the likely effect of the lighting on rail suicide. However, it is likely that a considerable time of several years would be needed to collect data on the numbers of fatalities at stations, in order investigate this type of impact of the lighting. Furthermore, there are many social, environmental and contextual factors that could influence the number of fatalities at a station. Therefore, for the purpose of this study, an intermediate objective will be set.

What exactly are you trying to show in your study?

For this particular example, the study objective would be to investigate the effect of overhead blue lighting on behaviours of people at stations and responses of passengers at stations. It is known that the effects of some safety interventions can be temporary and indeed it would be reasonable to expect that the blue lighting effect might be temporary if it works by one of the suggested mechanisms (i.e. something unfamiliar). Therefore, a second objective of the study is to investigate whether there is a sustained effect of the blue light intervention on the behaviours and responses of people at stations over time.

Global
Top 100
University

The University of
Nottingham
UNITED KINGDOM · CHINA · MALAYSIA

Other examples of intermediate objectives could include goals of enhancing and monitoring the cooperation of stakeholders, and implementation of the lighting in the manner intended at the outset of the programme (short term). These types of intermediate objectives can be identified by considering the logic of the intervention (i.e. what will be the initial impacts of the intervention, and then subsequent impacts, leading to the overall objective over time). More information on logic mapping is given in Hills and Junge (2010). Some more specific examples of intermediate objectives for blue light interventions are listed in Appendix 2.

3.1.2. Identifying evidence that is needed to show that the preventative measure works as it should
What evidence is needed to fulfil your study objectives (including the effectiveness, quality and value of the intervention)?

There are limited examples of evidence so far that demonstrate the effectiveness of the blue light intervention. It has been recognised that it will be difficult to demonstrate the effect of the lighting on the number of incidents at the stations at a single station, especially over a small period of time. In this example, studies at the station can be done to *observe behaviours* of passengers and *ask questions* to collect accounts from staff at stations about *peoples' opinions and how passengers have responded* to the introduction of the lights. The behaviours of passengers could include observations of where people stand (in relation to the lighting), for how long, whether their attention appears to be drawn to the lighting, and whether their behaviour appears to be influenced by the lighting.

It is also possible to question people who work in the location to collect their perceptions of how peoples' behaviours may have changed and to collect any third party reports that they have received after the fitting of the lighting. It may be possible to ask similar questions to passengers, but the decision to question passengers should be considered carefully to avoid drawing attention to potential problems of suicides at the station. Further details of how to apply the methods (including observation checklists and questions for collecting the data) are explained in Section 4, below.

When is the right time to collect the evidence that you need?

In order to investigate changes in passengers' behaviours over time, observations will be carried out before and after fitting of the lighting. A series of observation periods (e.g. up to hour periods) will be planned before provision of lighting, targeted to enable observation of people at relevant times and lighting conditions. This could include peak periods with crowding at the station, off-peak times where there may be more opportunities for seclusion in parts of the station. Observation periods may also need to be planned to account for night and day time conditions, dependent on the proposed functioning of the lights (i.e. will the lights be operational in daytime and night time?). A similar set of observations will be conducted after the provision of the lighting.

Observations at a particular site may need to be restricted to periods of relatively short duration, for practical reasons. In order to investigate whether any effect of the lighting is sustained, it will be

necessary to repeat the observations at a suitable period after fitting of the lighting (e.g. six months after the first post- lighting observations).

It should be noted that the before / after study design that is suggested here is appropriate for the study of changes in behaviour at this location. It is important to recognise that changes in behaviour may have arisen from any other changes in circumstances at the station. Therefore, it will be necessary to determine whether any such changes have occurred at this location. This type of before / after study design may not provide reliable results where the focus is on the number of incidents at a location and problems of regression to the mean (see earlier in section 2.3) may be apparent. Under these circumstances reference measures would be needed from a similar control station that is matched to the station on characteristics such as the size, function, location (rural, urban, inter-city), manned / unmanned, footfall.

Evidence to determine the value of the intervention

The blue lighting is implemented with the intention of meeting a need that is not satisfied by other prevention measures at the moment. Consideration should also be given to the cost of blue lighting as a suicide prevention measure (including the costs of the design, fitting and maintenance of the lighting), in conjunction with the potential benefits that might be expected by this type of intervention. It can be difficult to assess benefits such as the reduction in fatalities at the early stage of development of a preventative measure, but it may be able to identify other benefits of safer behaviours as a result of the lighting (such as fewer people standing in places of risk).

Evidence to determine the effectiveness of the intervention

Whilst this is not a priority for this example (but may be relevant for longer term objectives), some consideration may need to be given to ensuring that outcome data (on the numbers of incidents) can be collected to evaluate the likely success of the lighting. It will therefore be important to ensure that systems are in place to record reliable information on all major and minor events at the station.

3.1.3. Selecting locations and people to study
Where should the study take place?

In this current example, the study will be carried out at a single station where blue lighting is going to be fitted.

Who should be involved?

It is proposed that station staff will be consulted about how passengers have reacted to the introduction of the lighting. In this case, it is possible to carry out short interviews / discussions after the introduction of the lighting programme.

It is important to ensure that there is dialogue between organisations (Network Rail and train operating companies), to ensure that any differences of opinion can be resolved without

Global
Top 100
University

The University of
Nottingham
UNITED KINGDOM · CHINA · MALAYSIA

compromising the objectives of the study. It may therefore be possible to monitor examples of inter-organisational communication and consult with relevant staff in all relevant organisations about their perceptions of the issues that have been discussed.

3.1.4. Collection of data
What are the best methods to collect the data that you need?

This can begin early in the process of design, including information on the proposed lighting from research publications, industry sources and specifications, site visits and inspection reports.

Checklists, interview protocols, survey forms will need to be prepared for the collection of data from site visits and observations at stations (i.e. before and after the provision of the lighting) and from staff at stations on the likely effects of the lighting.

How do you collect these different types of data?

A checklist for observations at stations could prompt for the following:

For observations of where people stand (in relation to the lighting)

- annotation of diagrams of the station platforms, including details of counts of people in different areas of the platform.

For how long

- recording of dwell times in selected locations (e.g. near to lighting, in locations away from the lighting)

Whether their attention appears to be drawn to the lighting

- notes on the gaze of passengers, changes in gaze, fascination with the lighting.

Whether their behaviour appears to be influenced by the lighting

- descriptions of movements towards the source of the lighting or away from the source of the lighting, distance from critical aspects of the infrastructure (e.g. platform edge), numbers of people in secluded areas that are illuminated by the lighting (or planned to be illuminated by the lighting).

Real time observations could be carried out directly from station platforms, bridges or other locations. Alternatively, it is possible to use cameras to record the movements and behaviours of people at the station, enabling in-depth analysis of the recordings at a subsequent time. Consideration may need to be given to ethical issues involved in observing people in this type of location.

For interviews with staff, questions could be drawn from those reported in Appendix 3. Examples include the following: Since the implementation has there been any change in behaviour of people at the station? What types of behaviour have changed? How have passengers responded to the

Global
Top 100
University

The University of
Nottingham

UNITED KINGDOM · CHINA · MALAYSIA

lights? Do passengers behave differently near blue lights in comparison to other platforms or in areas without the blue lights?

Ethical issues for these interviews may also need to be considered.

It is important that there is a process in place for formally recording progress with the implementation of the lighting against the early intentions for the blue lighting programme of work. This should record relevant time periods and critical dates for completion of different phases of the programme (e.g. completion of the lighting, full operation of the lighting) to enable interpretation of outcome data (i.e. enabling analysis of the potential impacts of the lighting on the numbers of incidents). This should also record details of any deviations from the plans for provision of the lighting at the station or any part of the station.

3.1.5. Analysis of data

What do you need to do with the data that you collect? How do you deal with different types of data (e.g. quantitative data, descriptive / qualitative data)?

In this example, this can include descriptive accounts and supporting summary statistics, plans and annotated diagrams that are indicative of any changes in behaviour. These would include summaries of findings over the range of different time periods that have been studied. Some positive results could include reductions in the numbers of people in high risk areas (e.g. close to the platform edge), fewer people and shorter dwell times in secluded areas of stations that are illuminated by the lighting.

Descriptive accounts can also be produced from the analysis of the data from questioning of staff. Positive results from this part of the analysis could include low numbers of reported negative reactions or complaints to staff, or reports of positive comments relating to the visual environment at stations. Any indications of changes in antisocial behaviour or other incidents should ideally be supported by evidence to support these types of claims.

3.1.6. Concluding and reporting

How do you draw appropriate conclusions from the data that you have collected and analysed?

In this example a report should include conclusions that are linked to the study objectives. This will therefore include conclusions on whether the lighting may have had impacts (sustained over a period of time) on the behaviours of people at the station (including the nature of these changes) and conclusions on how people appear to have responded to the introduction of the lighting. It is important to explain the extent to which the programme has been implemented as it was intended. For example, if there have been deviations in the programme it may be these deviations that have contributed to an effect or the fact that there is no observable effect.

Global
Top 100
University

The University of
Nottingham

UNITED KINGDOM · CHINA · MALAYSIA

3.2. Example 2- Evaluation of a Smart CCTV camera

In order to extend this approach, the simple structure of the framework has also been applied to an example of another technology intervention (a Smart CCTV camera). This is presented as a means of seeking to test the ability to adapt this framework to a new context. The examples of data types and commentary are provided as an example and not based on any consultation with staff in the industry so far.

Once again, it is hard in the early stages of development of this idea to establish the extent to which this intervention could reduce the number of fatalities at a station. However, it is possible to investigate other aspects of performance and use of the technology in the railway environment.

3.2.1. Objectives

How does the preventative measure work?

We know that there are certain behaviours that can be identified (post incident) in people who commit suicide on the railway. There are also many interventions that take place on the railway, where staff and others identify someone acting suspiciously and intervene to prevent an incident. Unfortunately, it can be difficult to respond in time and there may not be sufficient numbers of staff or others with the necessary knowledge to be able to identify people at the right time. CCTV technology may offer greater coverage of the railway, especially if smart algorithms can be used to help with identification of those at risk.

What exactly are you trying to show in your study?

A long term objective could be - To reduce railway suicide through application of a SMART camera system for the timely identification of suspicious behaviours at stations.

There are many intermediate objectives that are important to consider in this type of project. Two examples are listed for the purpose of this example

-        to distinguish a set of target behaviours from other behaviours in the railway setting

-        to examine how security users interact with and respond to notifications from the Smart camera system

Early stages of the project would collect relevant information about the design and specification of the equipment and determine how this could be applied in the complex context of the railway station.

Global
Top 100
University

The University of
Nottingham
UNITED KINGDOM · CHINA · MALAYSIA

3.2.2. Identifying evidence that is needed to show that the preventative measure works as it should

What evidence is needed to fulfil your study objectives (including the effectiveness, quality and value of the intervention)?

For the purpose of narrowing the scope for this example only one of the objectives has been used *i.e. distinguishing a set of target behaviours from other behaviours in the railway setting.* This measure might not be expected to decrease in effectiveness over time.

The types of evidence needed include findings from study of the correct identification of target behaviours in relevant parts of the station environment.

When is the right time to collect the evidence that you need?

There is need for evidence of achieving this in relevant conditions (e.g. external - in sunlight or poor weather; under canopies; in different lighting conditions; in different densities of people / crowding). Therefore, it would be necessary to test the camera for periods of time in a range of locations. The choice of these locations would be dependent on circumstances at the station, but might include the following:

At platforms ends, where there is no shading of the sun from canopies, or protection from rain. Observation periods / data collection periods should be planned to take account of different lighting levels (bright sunlight and potential for glare, dawn / dusk, night time with use of typical station lighting)

Under canopies, with potentially lower lighting levels. Observation periods / data collection periods should be planned to take account of different lighting levels (daylight and no station lighting, night time with typical station lighting, effects of shadows from the station infrastructure in daylight and night time.

Observations at selected areas of the station (e.g. near the platform edge, near stairwells, platform ends or other isolated locations, in periods with different levels of occupation at the station (dense crowds at peak times, moderate crowds / small groups of people, few people on the station platform).

Evidence to determine the quality of the intervention

It is also important to determine if there is evidence that there was good support in the literature for the performance of this type of technology, and evidence from other studies in the industry in GB or worldwide.

Evidence to determine the value of the intervention

Qualitative data could be collected on the extent to which this meets a need that can not be satisfied by other means. Costs of supply and maintenance of the equipment can be calculated.

Global
Top 100
University

The University of
Nottingham
UNITED KINGDOM · CHINA · MALAYSIA

3.2.3. Selecting locations and people to study

Where should the study take place?

Decisions about the numbers of locations for inclusion in the study will be dependent on the stage of development and readiness of the technology. Early trials may be carried out in a small number of locations, which provide appropriate variations in the operating environment and local context, to meet the objectives of this specific study.

At a later stage of development, to investigate additional study objectives, consideration would need to be given to implementation of the technology in a wider range of railway stations. Issues of sampling would need to be considered.

In this example, a camera(s) could be used at various locations at a single rail station that is capable of producing all environmental / population based test conditions that have been mentioned above.

Who should be involved?

Consideration would need to be given to the individuals and groups that would be observed. In the early stages of development, it may be difficult to plan studies where real-life, at-risk behaviours can be observed in the rail environment. Therefore, there would be need for elements of simulation in this type of research study. Actors could be used to simulate target behaviours in varying locations and population densities at the station. This would need to be planned in conjunction with the developers of the equipment, who would be aware of the technical capabilities of the equipment (e.g. whether it can identify whole body movements / postures, single limb movements etc.). It would then be possible to design and conduct observations to attempt to pick out the target behaviours against the background of other behaviours at the stations, in the selected environmental and population based conditions.

Ethical issues associated with the study would also need to be considered.

A clear plan for implementation and evaluation is needed. This could build upon successful elements of preliminary studies or recommended methodologies in the literature. This would include engagement of other stakeholders (e.g. project managers, design, safety and support staff, train operating companies, technology suppliers, station manager / staff, security / CCTV operators and analysts). It is therefore important to collect information from relevant groups of staff about their practical experiences of use of the technology. In this case the cameras may not have a direct effect on influencing behaviours and responses of people at the station (cameras are a common feature at the station). For this type of intervention the important issues are likely to be: how people are expected to respond to warnings or alerts from the camera system; how this impacts on their current job (e.g. how they are expected to monitor displays?). This type of consultation could be achieved with a series of short interviews with relevant staff who have roles in interacting with the technology.

Global
Top 100
University

The University of
Nottingham

UNITED KINGDOM · CHINA · MALAYSIA

3.2.4. Collection of data

What are the best methods to collect the data that you need?

Observational data would be collected via the Smart CCTV system. It would be necessary to have an additional source of information on the actual numbers of suspicious behaviours to test the success of the new camera system (e.g. a script for an actor).

Surveys, questionnaires, diaries and interview studies could be developed to collect information from CCTV operators about their responses to notifications and their perceptions of use of the technology in different contexts.

How do you collect these different types of data?

It is important to study the use of the camera in the real railway environment, so site survey data, environmental measurements and descriptive data could also be collected on the types of railway contexts that pose problems (e.g. lighting levels, locations, densities of people, locations of cameras, viewing distances, viewing angles).   These types of descriptive data are important in determining the extent to which the intervention is implemented in this trial in the way in which it was intended.

3.2.5. Analysis of data

What do you need to do with the data that you collect?

One part of the analysis of data could be focused around the frequencies of correct / incorrect identification of target behaviours (e.g. numbers of times where there was positive identification of a target behaviour, numbers of times where there was failure to identify target behaviour, linked to different contexts).

How do you deal with different types of data (e.g. quantitative data, descriptive / qualitative data)?

The statistical data can be supported by descriptive data on how operators interpret notifications and their responses to different types of notifications. Analyses of qualitative data would include classification and description to identify the circumstances in which the camera is effective or ineffective. For example, this could reveal the types of behaviours that are likely to produce errors in identification, because they are hard to identify or distinguish from other behaviours. This would also produce knowledge of the circumstances in which it is hard to correctly identify behaviours (i.e. what types of environmental conditions or levels of crowding).

3.2.6. Concluding and reporting

How do you draw appropriate conclusions from the data that you have collected and analysed?

Conclusions will need to be drawn on the likely success of the camera in supporting the identification of different postures / movements in different operating contexts. The report in this case is likely to be written to inform a project team that is developing the Smart cameras for use in the railway environment. The report will highlight the original purpose of the study (i.e. distinguishing a set of target behaviours from other behaviours in the railway setting) and provide

Global
Top 100
University

The University of
Nottingham
UNITED KINGDOM · CHINA · MALAYSIA

sufficient details of the method, results and related interpretations, to enable the team to make decisions on the next steps in the project development process.

3.3. Concluding thoughts on the development and use of the examples

The simple framework has been applied to two different examples of technology interventions. Both have different objectives. Efforts have been made to explain the types of issues that are important when considering how to conduct evaluations for these situations, also giving some insight to the decisions that have to be taken (e.g. in relation to what data to collect, where to carry out the study and who to involve) when planning the evaluation study. Setting up an effective study can be difficult without sufficient experience of scientific methods. This type of expertise may be available in a project of this kind by building a team of industry staff who have sufficient experience of scientific methods for study design and data collection and analysis (e.g. including qualitative and quantitative methods). There may be situations where specialist support is needed to ensure that reliable data can be collected.

4. Conclusions

Findings from the earlier review of the implementation of blue lighting at stations in GB have been used to develop of a framework that can help to support industry staff in planning and evaluating similar interventions. Initial tests of the utility of this framework have been carried out by applying this to describe the research methods and data requirements for two studies involving a blue light scenario and another example of new technology (Smart CCTV).

Designing and evaluating technologies in this type of context is not a straightforward activity. Successful design requires a good understanding of scientific methods, such as experimental / study design, data collection and analysis. Understanding of qualitative and quantitative methodologies is likely to be important. Unfortunately, there is no single methodological approach that can be used to study what could amount to many variations of preventative measures that are employed within a wide range of operational circumstances on the railway. Efforts have been made within the proposed framework to provide guidance through the provision of a series of prompt questions and examples, based around the blue light and CCTV preventative measures. The extent to which this framework can be adopted in practice will need testing. Further development of this type of framework will benefit from further consultation with potential users.

It is proposed that those with an interest in applying this type of framework test the use of this with a wider set of examples. As a first step, it would be possible to explore this within a facilitated workshop setting, where a group of stakeholders could have the opportunity to work on designing an evaluation study for a new safety intervention. This could be used to examine the utility of the framework and collect feedback on additional amendments that might be needed before this is made more widely available to other stakeholders in the industry.

Global
Top 100
University

The University of
Nottingham
UNITED KINGDOM · CHINA · MALAYSIA

5. References

Fink, A. 2015, Evaluation Fundamentals, Insights into Program Effectiveness, Quality and Value, third edition, Sage, USA.

Hills, D,. Junge, K. (2010) Guidance for transport impact evaluations. The Tavistock Institute, London. Available from http://www.roadsafetyevaluation.com/evaluationguides/index.html.

Ryan, B., Philippou, A. (2015), Evaluation of rail safety interventions – a review of use of blue lighting at two stations, Report for Network Rail

Appendix 1 – Using the structure of the framework to summarise findings from the Blue Light studies in progress at Network Rail

Step in the evaluation process	Commentary on examples of evidence and data types from the blue light study that are relevant to this framework
1. Define the **objectives** of the study of the technology	It is not clear that the objectives of the studies were articulated precisely. This may have contributed to some unintended shift in the overall direction of the projects, as a result of stakeholder pressure. Currently there are no apparent targets that can be used to demonstrate success of the projects. A number of potential mechanisms (for how the lighting may work) have been identified in the preliminary research for the blue lighting intervention. One or more of the mechanisms in this case are plausible, but these are not really linked to the schemes that are operating. It is therefore difficult to determine whether the measures are working in the way that it is thought that they are working, or if they will continue to work over time. There are good efforts to consider the potentially harmful effects (health and performance) of the new lighting, as well as the potential benefits, though the source of the evidence could be clearer to enable verification.
2. Identify the **evidence that is needed** to show that the technology works as it should	There are limited examples of data so far that demonstrate the effectiveness of the blue light intervention. Much more can be done to observe behaviours and ask questions to collect accounts from staff at stations on user opinions, to understand the likely impacts of the lighting. Efforts have been made to determine the quality of the lighting solution prior to implementation, through some consideration of the theory around the mechanism of the effect. More systematic studies of behavioural responses to different types of lighting could perhaps have been considered prior implementation. The blue lighting is implemented with the intention of meeting a need that is not satisfied by other prevention measures at the moment, though more evidence is needed to demonstrate the value of this as an intervention. There has been some discussion of the relevance of cost in decision making on suicide prevention measures, though some differences in opinion are reported. There is evidence to suggest considerable planning and design activities in the development of each of the blue lighting initiatives. This included dialogue between organisations (Network Rail and train operating companies), though it appears that there have been differences of opinion that, whilst some compromises and agreements have been reached, may have had implications for what could be achieved within these projects. It does not seem that there has been much exchange of ideas been the different schemes in this early part of the development of the blue lighting solutions. There seems to have been little focus on the evaluation of the intervention. The projects at these stations are taking place early in the development of the use of blue lighting in the GB rail industry and experiences from these trials will help to inform the development of subsequent blue lighting initiatives.
3. Select the **locations** for the	For understandable reasons the two studies have been considered so far as separate, single station case studies. There is some evidence to explain the reasoning for inclusion of each of the two stations in the study (i.e. relating to

Step in the evaluation process	Commentary on examples of evidence and data types from the blue light study that are relevant to this framework
study and **people / groups** that should be involved	questions of need – there has been evidence of a problem at the stations, other solutions are not practicable, there has been an opportunity to implement and test the lighting). It is important to be aware of statistical problems associated with regression to the mean when selecting stations for inclusion in an intervention study (i.e. an intervention following a number of recent incidents may appear to be effective – any change in numbers of incidents may result from normal fluctuations in the data rather than as a result of the intervention). This problem can be minimised by collecting data from a control station which is matched on relevant criteria to the station that is being studied. Identification of a suitable control station (e.g. matched on size, function, location (rural, urban, inter-city), manned / unmanned, footfall) could help with analysis and interpretation of the outcome and other data from the studies.
4. **Collection of** data	There is evidence of good initial work in collecting data on the proposed lighting from research publications, industry sources and specifications, and site visits and inspection reports. Accounts within the interviews have helped to identify some of the organisational / inter-organisational factors that have impacted on progress with the lighting programme. There is limited use of outcome data (i.e. formal recording of the numbers of incidents, threats, reports of trespass) to evaluate the likely success of the lighting and no specific efforts to collect data from staff or users at stations on the likely effects of the lighting. There has been limited formal recording of progress with the implementation of the lighting against the early intentions for the blue lighting programme of work. There has been no collection of data at control stations.
5. **Analysis** of the data	Limited evidence has been presented to date on the collection and analysis of data from the intervention.
6. **Concluding** and **reporting** on results	A post intervention report was available for one of the stations. Other documents described relevant features of the lighting programme and some provisional results at the other station. Generally, reports on the outcome were anecdotal in nature, reporting on the perceived success of the lighting. There may be some inconsistencies with other data sources.

Appendix 2 – An example of a simplified logic map of short term medium term and longer term outcomes for a blue light safety intervention

Short term outcomes   *Medium term outcomes*   *Long term outcomes*

Achieve consensus between organisations for implementation of preventative measures

Change behaviours at stations (e.g. less risky behaviours)

Reduction of suicides at stations where blue lights are fitted

Identify appropriate locations for the installation of blue lights

Improve the perceptions / mood of people at stations

Maintain the reduction of suicides at stations where blue lights are fitted

Install lights in conjunction with the intended design proposal

Global
Top 100
University

The University of
Nottingham
UNITED KINGDOM · CHINA · MALAYSIA

Appendix 3 – Interview questions that were used in the review that is reported in Ryan and Philippou (2015)

Category	Question
Implementation - Role	• Could you explain your role in the implementation of blue lights at your station? o What decisions were you part of/responsible for?

Implementation - Design	• How did you select the specifications of the lights?
	   o e.g. blub Wattage
	   o type of bulb
	   o the shade of blue

Implementation - Problems	• Were there any difficulties involved with implementation?
	   o Is the final outcome as planned?
	   o Is it for ease of maintenance e.g. attached to existing pole/replaced existing lights with blue?

Implementation – Variations	• When was the lighting fitted (date, at least month if precise details not known, start and end of any programme)
	• Are blue lights fitted on all platforms?
	   o How did you decide where to fit the lighting?
	   o Are there any differences in the lighting of different places of the station?
	   o What types of differences?
	   o Why?

Environmental Factors - Visibility	• When are the blue lights visible?
	   o e.g. day time, night time, dusk, dawn
	   o Are they visible in sunlight?
	   o What times are the lights turned on/off?

Evaluation – Data Types	• Are you collecting any data about the potential effects of the lighting? … before … after fitting of the lighting • Can you tell me about the types of data that you collect? • Is it possible to get a copy of the data? • What types of data would be useful?
History of Incidents	• Have you witnessed personally, or are aware of any suicidal events, a fatality or near miss? • Was this before or after installation of blue lights?

Observations - Before	• Why were the specific locations chosen?
	   o Were they seen as a high risk area?
	   o Why?
	   o Was there any suicidal behaviour noticed at the exact position?

Observations – Passenger responses	• Since the implementation has there been any change in behaviour of people at the station?
	   o What types of behaviour have changed?
	   o How have passengers responded to the lights?
	   o Do passengers behave differently near blue lights in

	comparison to other platforms or in areas without the blue lights?
	o Has there been any unusual behaviour since the blue lights have been fitted?
	o Is it possible to say whether passengers are calmer?
	o Have they appeared to be unfamiliar or uneasy with the lighting?
	o Have there been any complaints or questions about the lights?
Observations – Staff responses	• How have staff responded to the lights? o Has there been a change in staff performance since the implementation? o Absenteeism? o Health changes? o Any complaints e.g. distraction?
Observations – Variation in response	• Have there been any differences in the effects of the lights on different platforms? If so, what? And why do you think this these differences exist?
Other influencing factors	• Were there any changes at the station around the same time as the blue lights installation? o any other prevention tactics employed? o any changes in staff/staff protocol? o any maintenance work? o any changes in train schedules?
Additional Information	• Is there anything else you would like to tell me today?

Table A1Topics for Interview for Programme staff / senior managers