CEDR Transnational Road Research Programme
Call 2013: Traffic Management:
Implementation of Innovation
in Traffic Management

funded by Belgium-Flanders, Denmark,
Finland, Norway, UK and Netherlands

METHOD: Managing European Traffic using Human-Oriented Designs

Driving simulator studies evaluating in-vehicle information for traffic management using head-up-displays and gamification techniques

D3.1
February, 2016

TRL	**SWOV**	**TML**	**TNO**	**VTT**
UK	Netherlands	Belgium	Netherlands	Finland

CEDR Call 2013:
Transnational Road Research Programme
METHOD: Managing European Traffic using Human-Oriented Designs

Driving simulator studies evaluating in-vehicle information for traffic management using head-up-displays and gamification techniques

Due date of deliverable: 30/04/2016
Actual submission date: 29/04/2016

Start date of project: 01/06/2014 End date of project: 31/07/2016

Author(s) of this deliverable:
Alistair Weare, TRL, UK
Nick Reed, TRL, UK
Tineke Hof, TNO, The Netherlands
Isabel Moning, TNO, The Netherlands

Table of contents

Executive summary

The report describes two individual, independent studies that investigated how two existing, near future concepts (gamification and a head–up display (HUD)) could be applied to improve drivers' compliance with traffic management guidance. This has the potential to improve overall network efficiency – firstly, through this increased compliance effect but secondly, by giving national roads authorities better and more specific tools for managing the movements of vehicles on their networks.

The UK study (described in Chapter 2) was carried out by TRL and it investigated the use of gamification techniques on driving behaviour. A simple game design was created with the aim of encouraging drivers to adopt behaviours that were safe but also complied with traffic management guidance; the results were mixed but somewhat encouraging. The main positive result was that, on a number of metrics, participants were observed to display significantly better driving behaviour during their drive with the game active than when there was no game present. These effects seemed to apply for male, female, older and younger participants and were observed irrespective of whether the participants liked or disliked the game itself.

The main negative effect of the game was that participants did not like the game and reported that it made driving more difficult and more stressful. On reflection, this is perhaps unsurprising as it is, to some degree, requiring drivers to adopt behaviours that are different from their usual driving habits and it may therefore make them feel uncomfortable. However, it had been expected that enjoyment of the game may offset such feelings. It should be noted however, that although the game was created to the best of the abilities of technical team involved in the project, it would be reasonable to state that there was no specific game mechanics experts that could perhaps have helped to make the game as engaging and enjoyable as possible.

The second study (described in Chapter 3) was carried out in the Netherlands by TNO. It investigated the potential benefits of supplementing (or even replacing) existing driver-oriented traffic management information using spoken warnings or an in-vehicle HUD system. The study found that traffic management information presented by in-vehicle information – either spoken or via a HUD - did not lead to better or more timely compliance than the same information presented via traditional roadside signs. This contrasts with previous studies (e.g. Liu & Wen, 2004) that have found HUDs to improve drivers' reactions – although those results were in a higher workload, urban environment than on the relatively simple highway environment under test.

Participants found the HUD information to be more satisfying but no more useful than the roadside information and the in-vehicle speech interface. It should also be noted that since participants found HUD-presented information no less useful than traditional techniques, then the use of HUDs for providing more specific individual vehicle guidance in traffic management should be explored. This specificity is the real potential benefit of in-vehicle information. As such, it would be of interest to study how individual vehicles given differing traffic management information would behave in a shared motorway environment. It would also be of interest to understand user attitudes to enforcement of traffic management measures implemented by in-vehicle information. This would represent a significant change in policy but in technological terms, it would not be difficult to detect when a vehicle was travelling significantly in excess of its individually prescribed speed limit.

The two simulator studies presented investigated how two different human factors techniques could support future traffic management. The use of simple gamification principles brought about changes in driving behaviour but the approach was not well liked by participants who felt that it made the driving task more difficult. The use of in-vehicle information for presenting traffic management information was found to be no better (but no worse) than traditional techniques but was found to be more satisfying for participants. It can be concluded that each study has therefore provided results that suggest each technique can be explored further. In particular, firstly, it would be of interest to work with game design specialists to create a specific game to encourage safe and compliant driving behaviour that builds upon the results of the UK study. Secondly, the use of in-vehicle information for traffic management was found to be non-detrimental to safe, compliant driving behaviour so further work should focus on how to gain further benefit from the flexibility of traffic management that can be achieved by this approach.

1 Introduction

This document forms Deliverable D3.1 for the METHOD project (Management of European Traffic using Human-Oriented Design), funded under the CEDR Transnational Road Research Programme – Call 2013: Traffic Management: Implementation of Innovation in Traffic Management. The deliverable documents the research and findings of Work Package 3: Simulator studies.

The studies were undertaken to understand how two different human factors techniques can be applied to improve drivers' compliance with traffic management guidance. This has the potential to improve overall network efficiency – firstly, through this increased compliance effect but secondly, by giving national roads authorities better and more specific tools for the managing the movements of vehicles on their networks.

This report documents the two individual studies, carried out independently. The first study (reported in Chapter 2) was carried out in the UK by TRL; it investigates the potential use of gamification as a means for promoting the adoption of driving behaviours that are desirable from a traffic management perspective. The second study (reported in Chapter 3) was carried out in the Netherlands by TNO; it investigates the potential benefits of supplementing (or even replacing) existing driver-oriented traffic management information using spoken warnings or an in-vehicle HUD system.

2 UK Study – The effect of gamification principles for traffic management on driver behaviour

2.1 Introduction

Traffic management[1] is performed with the intention of improving the safety and efficiency of the road network (as well as reducing congestion, improving environmental sustainability and other factors). The effectiveness of traffic management interventions can be improved by increasing the level of driver compliance with those interventions. Additionally, encouraging drivers to adopt desirable driving behaviours more generally can both reduce the need to intervene and improve the ease with which interventions can be planned. Improving driver compliance in line with desirable driving behaviours is therefore beneficial from a traffic management perspective; individual drivers' motivation to comply is a key factor in achieving this.

Gamification is the principle of applying gaming philosophies and mechanics to otherwise non-gaming situations. It is typically applied with the intention of providing a user with a source of motivation to carry out an activity or engage in certain behaviours. This study investigated the effects of a rudimentary game on driving behaviours within a driving simulator. The game was designed to be played and controlled by driving behaviour only. Trial participants were given no direct incentive to perform well or poorly in the game other than for their own personal satisfaction. The primary purpose of the study was therefore to examine whether (and if so, in what way) drivers could be encouraged to improve their driving behaviours and compliance with traffic management interventions simply through being provided with the opportunity to play the game. Secondary to this, the study also sought to understand how response to, and opinions towards, the game differed according to driver demographics.

[1] The definition of traffic management, as defined by the Nordic Road Association, (ITS Terminology (ver. 2008.06.04)) is "Management of traffic flows (people, vehicles and goods) by demand management, traffic information, traffic control and other measures".

2.2 Methodology

2.2.1 Methodology overview

After an initial familiarisation drive, participants completed two drives of a simulated route within TRL's driving simulator. In one drive, participants experienced the normal vehicle interior, in the other, they were also presented with a game displayed on a smartphone. The game presented drivers with images that responded to their driving inputs and overall driving behaviour. It responded with positive feedback to driving behaviours categorised as desirable for safety and/or the traffic network flow, and with negative feedback to driving behaviours categorised as undesirable. The purpose of the game is therefore to promote desirable driving behaviours among users. Effects were identified and assessed by comparing key driving behavioural metrics between the two simulation conditions.

2.2.2 Participants

Forty participants were recruited from TRL's existing participant database of individuals from the local area. Similar numbers of male and female participants were recruited, which were further split into categories of younger (20-35 years) and older (45-60 years) drivers. Thus there were four participant groups of ten participants each: younger male, older male, younger female and older female. All participants were required to have a full UK drivers licence, held for at least two years, and to have normal or corrected-to-normal vision. Typical trial time was around two hours and participants were paid £35 as compensation for their time and travel expenses.

2.2.3 Simulator environment

2.2.3.1 Driving simulator

TRL's driving simulator is based on a manual transmission Honda Civic which is complete with a steering wheel, pedals and other instruments. Images are projected onto four large screens that surround the vehicle; three to the front and one to the rear which enables use of all three of the vehicles mirrors. A sound system and a vibration and motion base give heave, pitch and roll to the car body and add to the sense of 'real-life' driving experience observed by the subject. The system uses MultiGen databases for the 'driving world', which are created by specialist 3D modelling experts. SCANeR II simulation software is used which provides intelligent vehicles that relate their behaviour to that of the simulator vehicle (within the confines of a described behaviour pattern). Surveillance video cameras are mounted in the car and participants can be recorded during their drive. An in-car intercom system allows the experimenter to give instructions and task-related information to the subject from a separate control room.

2.2.3.2 Road environment

Both drives were undertaken in an identical simulated road environment. This was created to represent a generic UK 4-lane motorway (with no hard shoulder) and with light traffic. The scenario contained gantries that displayed variable speed limits and lane closures, as well as bridges to relieve monotony and add context. The gantries were used to show variable speed limits that changed periodically over the length of the route, thus requiring the participant to alter their speed. The gantries were also used to indicate two separate lane closures: firstly of the inside lane (lane 1), then secondly of the two outside lanes (lanes 3 and 4). (See Annex E for clarification of the lane numbering convention used in the two studies.) The route was created to be approximately 20km in length and driveable in approximately 15 minutes if driving at the posted speed limits.

As will be explained further in Sections 2.2.4 and 2.2.5, the route was divided into ten sections for the purposes of the game. The section boundaries were chosen to coincide with stretches that were located away from significant distractions and were therefore of differing length (although broadly similar). Apart from the first section, all had either a change in speed limit or a lane closure at some point within it. Figure 1 shows a schematic diagram of the key features of the test route.

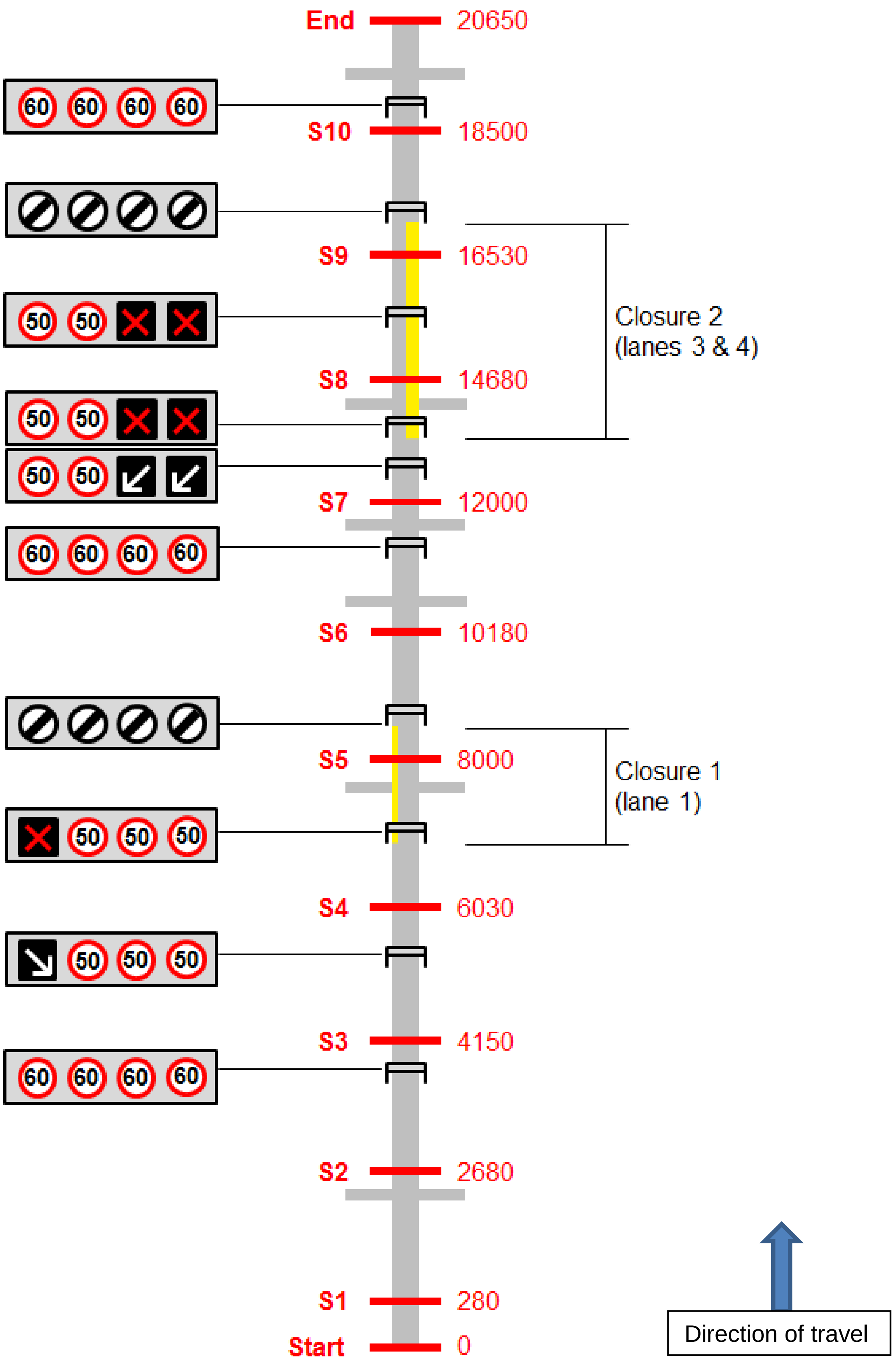


Figure 1: Schematic diagram of the test route (not to scale). The start of each section is shown in metres from the Start.

2.2.4 Game concept

The game was presented on a smartphone, mounted above the dashboard to the left of the driver[2], as might be expected of an aftermarket satellite navigation device. The user was presented with an on-screen avatar that characteristically changed in appearance over the course of the drive depending on the driver's behaviour in the simulator. Ten possible representations were available for the avatar.

A list of desirable and undesirable behaviours was identified from a traffic management and/or safety perspective. Exhibiting desirable behaviours resulted in positive changes (improvements) to the avatar; undesirable behaviours resulted in negative changes (detriments) to the avatar. Table 1 shows the full set of control behaviours.

Table 1: Game input behaviours.

Driving element	Desirable behaviour	Undesirable behaviour
Speed limit compliance	Average speed through section at or below posted limit	Average speed through section above posted limit
Lane closure compliance	Avoidance of restricted lanes	Use of restricted lanes
Lane change behaviour	Smooth changes in advance of closures	Rapid or late changes
Braking behaviour	Smooth decelerations	Harsh decelerations
General lane use	Use of left-hand lane wherever possible	Unnecessary use of middle or right-hand lanes
Vehicle following	Appropriate following distance	Close following

The route was split into ten sections, with avatar changes implemented at the boundaries between sections. The user started with their avatar at level 1 and therefore had the theoretical capacity to reach level 10 by the end of the drive. There was a level 0 so that the driver received some game feedback on their driving if persistent undesirable behaviours were exhibited, but without providing the opportunity for drivers to continue to try to explore negative avatar states and so inadvertently introducing a potential incentive for participants to exhibit undesirable driving behaviours.

When the participant was travelling through each section, an arrow indicated whether they were currently on course to achieve a positive or a negative change at the next section boundary. This provided some level of continuous feedback to avoid unnecessarily distancing the outcome from the input. When positive and negative changes occurred, these were accompanied by a corresponding sound to highlight the change to the participant. Figure 2 shows a basic representation of how this appeared at Level 1, with the green arrow showing an impending positive change. A full set of the game images relating to the eleven game levels (0-10) are presented in Annex A.

[2] Note: for a UK driver, this means the device is located towards the centre of the vehicle.

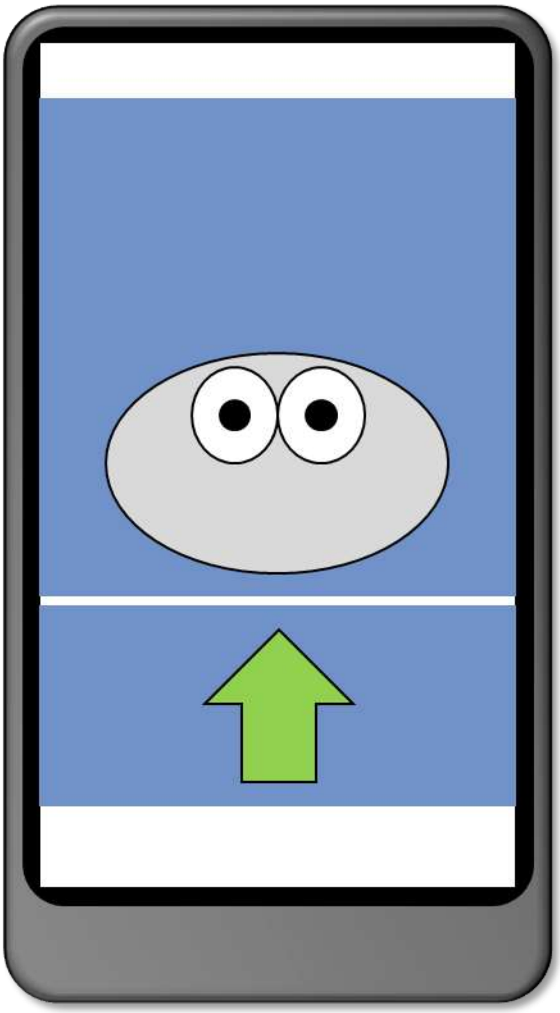


Figure 2: Representation of the ‘Level 1’ game screen as presented to participants at the start of the drive and showing a projected level-up at the next section boundary.

2.2.5 Game mechanics

Each of the ten sections were created to be of roughly similar length, but with section boundaries chosen to be away from locations of higher visual workload (e.g. gantries or bridges). These section boundaries were not represented visually in the road scene and therefore only apparent to the driver through game feedback.

Each section was further subdivided into subsections on the basis of time, defined by the drivers’ distance travelled within each 5-second interval within the section in which they were currently travelling.

At the end of each 5-second subsection, the game calculated the driving performance of the participant, during those five seconds, against the various performance categories (see Table 2). Based on that calculation, each subsection was scored as either positive (+1) or negative (-1).

Table 2: Subsection score calculation criteria.

Driving element	Positive	Negative
Speed limit compliance	Maximum speed at or below posted limit	Maximum speed above posted limit
Lane closure compliance	Avoidance of restricted lanes	Use of restricted lanes
Lane change behaviour	Maximum lateral acceleration <1.25m/s^2	Maximum lateral acceleration >1.25m/s^2
Braking behaviour	Maximum deceleration <7.0m/s^2	Maximum deceleration >7.0m/s^2
General lane use	Use of innermost lane available*	Use of middle or outer lanes when inner lane available*
Vehicle following	Minimum headway to vehicle in front >0.7s	Minimum headway to vehicle in front <0.7s
Final subsection score	No negative scores (+1)	A negative score in any category (-1)

**An inner lane was defined as available if no other vehicle present within this lane within three seconds headway in front of or two seconds behind the participant vehicle.*

At the end of the full section, the avatar state was updated according to the mean score of the subsections in that section. If the participant scored +0.3 or greater as an overall mean, the avatar advanced to the next highest level; if the participant scored -0.3 or less, the avatar regressed to a lower level; and if neither of the above conditions were met, the avatar remained at its current level.

Throughout each drive an indication was presented at the bottom of the screen as to what change was projected to occur to the avatar, based on the driving inputs received so far within the section (as shown in Figure 2). This was updated every five seconds (i.e. at the end of each subsection).

2.2.6 Trial protocol

Potential participants were approached for recruitment through email or telephone and were asked to perform two drives in the simulator of around 20 minutes each. Potential participants were not told about the purpose of the trial.

Trials were conducted with two participants per session. This created natural alternating periods in and out of the simulator that suited the trial and helped to ensure efficient use of the simulator facility. On arrival at TRL, the first participant of each pair was asked to sign a consent form and then performed a 10-minute familiarisation drive in the simulator. This is standard practice for simulator studies and helps to familiarise participants with the feel of the simulator controls and feedback, thus helping to ensure consistency across the two drives.

After completing the familiarisation drive the first participant was given a break whilst the second participant completed their paperwork and undertook their familiarisation drive. During this break, the participant was given a training brief on the nature of the game and how it worked. The training was given verbally, and used visual images of the game screen to aid comprehension. The training took approximately 5 minutes to deliver. After the training, the participant was asked to describe how the game worked, in their own words, to allow the researcher to check their understanding. A copy of the training script used by the researchers is provided in Annex B.

The participant then returned to the simulator to commence the first of their two test drives. The order in which participants experienced the game and no-game drives was counterbalanced, to control for order effects. Each of the four participant groups was further split into two, such that half of each group (five participants) experienced the non-game drive first and the other half experienced the game drive first. Before each drive (regardless of whether or not the game was being presented) participants were instructed to drive as they believe they would in the real world, neither treating the simulator as a game, nor as if they were on their driving test. Participants were not instructed to try to perform well on the game; how they chose to interact with it was their choice. On non-game drives the game was allowed to run within the control room to allow game performance to be recorded without participants receiving the game feedback.

Following completion of the second drive, participants were asked to complete a questionnaire about their perceptions of their own driving behaviour, and their opinions on the design and function of the game. Participants were also asked to complete a simplified version of the standard Driver Behaviour Questionnaire (DBQ), using questions/responses relating to either errors or violations. An example of the questionnaire is provided in Annex C. Once the questionnaire was completed, participants were thanked for their time and paid £35 as compensation for their time and travel expenses.

2.2.7 Simulator data

The following performance data were collected or calculated for the simulator runs:

- Mean speed and standard deviation of speed
- Proportion of time spent over/under the speed limit
- Mean, minimum and standard deviation of headway to vehicle in front
- Mean and maximum accelerator position
- Mean and maximum brake pressure
- Standard deviation of lane position
- Proportion of time spent in lane 1/2/3/4
- Mean time taken to change lanes

Each of the above was calculated within each section and as an overall value for the full drive. In addition, game outcomes from each section were recorded, allowing a record of the game progression and final game score to be logged for each drive.

2.2.8 Research Questions

The research questions for the study were as follows:

Primary research question:

1. Are driving behaviours (represented by the simulator data listed above) seen to differ between the game and no-game drives?

Secondary research questions:

2. Is any difference (i.e. the primary research question) related to driver sex and/or age?
3. Is any difference related to overall performance in the game?
4. Is any difference related to participant scores on the DBQ error and violation scales?
5. Is any difference related to participant subjective ratings of the game?
6. Are participant subjective ratings of the game related to participant score on the DBQ?
7. Are participant subjective ratings of the game related to driver sex and/or age?
8. Are participant subjective ratings of the game related to overall performance in the game?
9. Do subjective ratings of the overall driving experience differ between the game and no-game drives?

2.3 Results

The nature of this research is exploratory, and it aims to identify whether or not the use of gamification principles may hold promise for further development and, if so, to identify any indications of which approaches may prove to be more or less effective than others. The game developed for the trial is not ready for release in the real world, nor would it currently be possible to perform some of the simulated functions in reality. For these reasons, the purpose of the analysis is to identify general trends and possible avenues for further research and not to provide a validation of the game itself. The priority is therefore to minimise the potential for Type II errors[3]. As such, the reader is advised that a large number of exploratory statistical tests have been applied without correction for potential Type I errors. The findings should therefore be viewed with this caveat in mind.

2.3.1 Game score

Final game scores were compared between the ‘game’ and ‘no-game’ conditions (the game was run in the control room for no-game drives, to allow a score to be generated). Figure 3 shows a frequency distribution for the final scores achieved in each drive, split by ‘game’ and ‘no-game’ conditions.

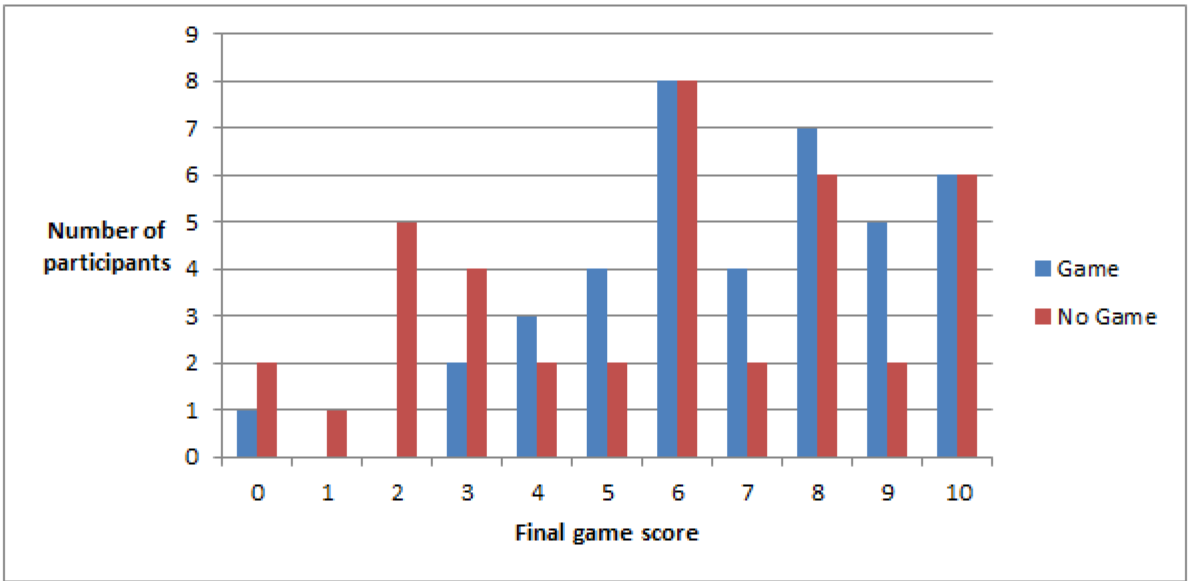


Figure 3: Frequency distribution for final game scores.

A Wilcoxon signed-rank test indicates that participants scored significantly better overall in the ‘game’ condition (mean=6.87, SD=2.33) than the ‘no-game’ (mean=5.73, SD=3.04) condition, p=0.017. Overall, twenty participants scored better in their game drive, nine scored better in their no-game drive, and eleven scored the same in both drives.

On average, participants achieved significantly better scores in the ‘game’ drives than the ‘no-game‘ drives.

[3] Type I errors represent false positives. In other words, detecting an effect that is not really present. Type II errors represent false negatives. In other words, failing to detect an effect that is actually present

2.3.2 Primary effects on driving behaviour

The following analyses are intended to identify any differences in key driving performance metrics between the 'game' and 'no-game' conditions. This relates to the primary research question: "Are driving behaviours seen to differ between the game and no-game drives?" Results summaries are presented in blue text boxes after each subsection.

During our analyses Shapiro-Wilk tests were used to reveal whether the specific data was normally distributed or not. In cases of non-normal distribution a Wilcoxon non-parametric test was performed and in cases where datasets were found to be normally distributed the paired-sample t-test was performed.

2.3.2.1 Driving Speed

Average (mean) speed

The game is designed to encourage drivers to keep below the posted speed limit and therefore it was predicted that the mean driving speed would be slower in the game condition than in the no-game condition.

Mean participant driving speed was calculated across each drive and compared between the game and no-game conditions. The tests revealed that mean driving speed was slightly lower in the game condition, but the result was not statistically significant ($p=0.143$).

Standard deviation of speed

In addition to encouraging compliance with the posted limits, the game is also designed to encourage drivers to drive smoothly and without harsh braking. Both factors lead to a prediction that driving speed would be more consistent in the game condition than in the no-game condition. This would reflect as a smaller standard deviation of driving speed in the game drives.

The standard deviation of driving speed was calculated for each drive and compared between the game and no-game conditions. The tests revealed that the standard deviation of overall driving speed was significantly lower for the 'game' drives (mean=14.37, SD=2.30) compared with the 'no-game' drives (mean=15.12, SD=1.78), $p=0.016$.

Time spent over the speed limit

In addition to calculating overall average speed, the overall proportion of time spent over the posted speed limit was calculated for each drive. A prediction was made that drivers would spend less time over the speed limit in the 'game' drive than the 'no-game drive'.

The proportion of time spent driving above the posted speed limit was calculated for each drive and compared between the 'game' and 'no-game' conditions. The tests revealed that the proportion of time spent driving above the speed limit was lower in the 'game' condition (mean=0.107, SD=0.081) compared with 'no-game' (mean=0.139, SD=0.111), $p=0.074$. The p-value indicates that the result can be considered to be statistically significant only if a 1-tailed prediction is made.

> No significant difference in overall driving speed was detected. However, standard deviation of speed was found to be significantly lower in the 'game' condition than the 'no-game condition'. The percentage of time spent driving above the posted speed limit was found to be significantly lower in the 'game' condition than the 'no-game' condition, provided a 1-tailed significance criterion is adopted.

2.3.2.2 Headway

Three metrics of headway were calculated:

- Average headway (mean headway to the vehicle in front during each drive[4])
- Minimum headway (shortest recorded headway to the vehicle in front during each drive)
- Standard deviation of headway (standard deviation of recorded headways to the vehicle in front during each drive[4])

The game was intended to discourage close following and therefore it was predicted that headway would be longer in the ‘game’ drives than the ‘no-game’ drives. Similarly to the speed metric, it was also predicted that headway would be less variable in the ‘game’ drives due to smoother driving.

Tests showed that the mean and minimum headway were slightly longer in the ‘game’ drives than the ‘no-game’ drives, and standard deviation of headway was slightly lower in the ‘game’ drives. However, none of these differences was statistically significant (p=0.906, 0.563 and 0.422 respectively).

> No significant differences in recorded headway were identified between the ‘game’ and ‘no-game’ conditions.

2.3.2.3 Acceleration

The game did not have any specific feedback mechanism to discourage harsh accelerations, however a prediction was made that participants would accelerate more smoothly in the ‘game’ drives than the ‘no-game’ drives as part of a more general effort to drive smoothly.

Average (mean) accelerator-pedal depression was calculated for each drive, along with the maximum accelerator-depression value (i.e. values from 0-1). The tests revealed that both average and maximum accelerator pedal depression were lower in the ‘game’ condition than the ‘no-game’ condition, but in both cases this difference was not statistically significant (p=0.158 and 0.155 respectively).

> No significant differences in accelerator pedal depression were identified between the ‘game’ and ‘no-game’ conditions.

2.3.2.4 Braking

The game directly penalised harsh braking and therefore a prediction was made that the recorded braking values would be lower in the ‘game’ drives than the ‘no-game’ drives. Two metrics of braking were calculated: mean brake pressure during each drive and the maximum brake pressure value. The tests revealed that both mean and maximum brake pressure were actually slightly higher in the ‘game’ condition than the ‘no-game’ condition, but in both cases this difference was not statistically significant (p=0.446 and 0.440 respectively).

---

[4] Periods when there was no simulated vehicle ahead of the participant and in the same lane were excluded from the calculation

No significant differences in braking pressure were identified between the ‘game’ and ‘no-game’ conditions.

2.3.2.5 Lane use

The game was designed to penalise unnecessary use of outer lanes (lanes 3 and 4) and therefore to encourage drivers to use the innermost lane available. Therefore it was predicted that drivers would spend more time in lane 1 and have an average lane position closer to lane 1 in the ‘game’ drives than the ‘no-game’ drives.

The incentive to use lane 1 did not apply during the lane 1 closure (which occurred in Sections 4 and 5) and therefore lane-use data were excluded from these sections. Data from the remaining sections were used to calculate the proportion of time within each drive that the driver was in lane 1, and the average lane position adopted by the driver over the drive.

The tests on the dataset for mean lane position revealed that drivers adopted a mean position closer to the inside lane in the ‘game’ condition (mean=1.914, SD=0.423) compared with the ‘no-game’ condition (mean=2.010, SD=0.414), p=0.057. This result is statistically significant if a 1-tailed acceptance criterion is adopted.

The tests on the dataset for proportion of time in lane 1 revealed that drivers spent a significantly larger proportion of time in lane 1 in the ‘game’ condition’ (mean=0.344, SD=0.259) compared with the ‘no-game’ condition (mean=0.265, SD=0.261), p=0.010.

Figure 4 shows a graphical representation of the lane use datasets, averaged across all drivers. It can be seen that drivers typically spent a larger proportion of their time in lane 1 in the ‘game’ condition than the ‘no-game’ condition, and that the distribution of lane choice was also skewed towards the inside lane (lane 1).

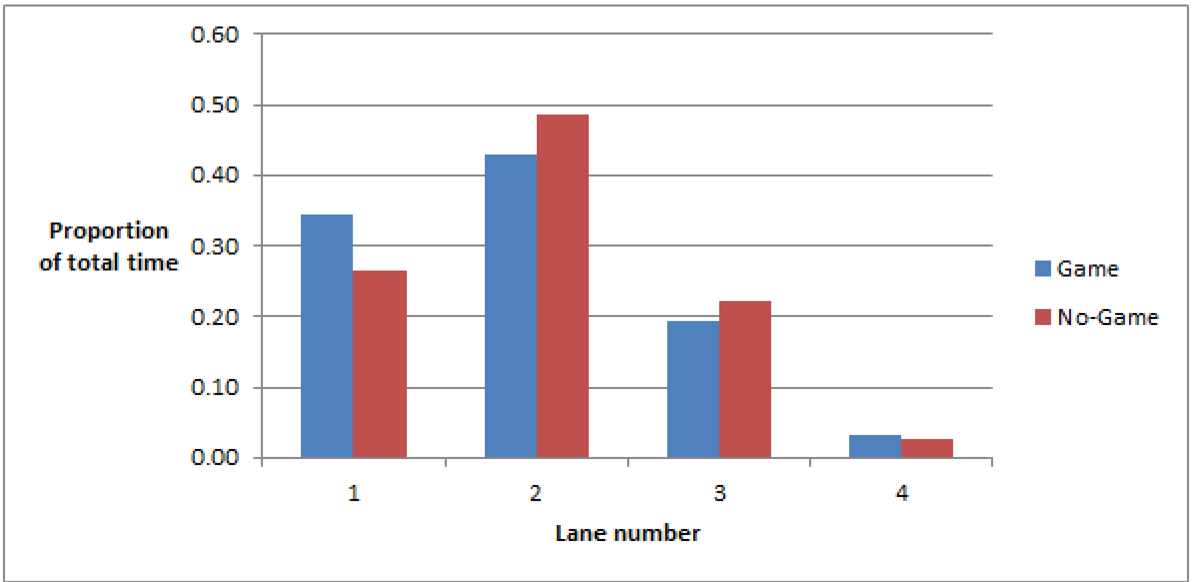


Figure 4: Average proportion of time spent by drivers in each lane between the ‘game’ and ‘no-game’ conditions.

Drivers spent a significantly larger proportion of their time driving in lane 1 in the ‘game’ condition compared with the ‘no-game’ condition. Drivers also adopted mean lane position significantly closer to lane 1 in the ‘game’ condition, assuming a 1-tailed significance criterion is adopted.

2.3.2.6 Lane change behaviour

The game was designed to discourage rapid or erratic lane changing, and therefore a prediction was made that drivers would take longer to change lanes in the 'game' condition than the 'no-game' condition. The period of measurement for each lane change was defined as the time between moving beyond a 1 metre lateral lane offset in one lane, to within a 1 metre lateral lane offset in the corresponding adjacent lane. The mean lane change duration was calculated for each drive.

Counter to the initial prediction, tests showed that the mean time to change lanes was significantly shorter for the 'game' condition (mean=2.58, SD=0.519) than for the 'no-game' condition (mean=2.76, SD=0.568), p=0.017. It is not clear why this would be the case as it would seem to suggest that the game had the opposite effect to that intended. A possible explanation could be that by spending more time in lane 1 in the 'game' condition, participants performed more lane changes between lanes 1 and 2, rather than between the outer lanes, and that these manoeuvres were perceived to be easier and therefore carried out more quickly. There is however no direct evidence to substantiate this theory.

> Drivers performed lane changes, on average, significantly faster in the 'game' condition to the 'no-game' condition, counter to the a-priori prediction.

2.3.2.7 Closed lane compliance

During each drive there were two lane closures. The first was a closure of lane 1; the second was a closure of Lanes 3 and 4. Closures commenced at the point of the first gantry to show a red 'X' in the relevant lane, and end at the next gantry to show the lane open (i.e. a speed limit indicator above that lane).The game directly penalised, with a red flag, any instances of driving within either of these lane closures (i.e. entering lane 1 at any point during the first closure, or entering either lane 3 or 4 in the second closure). In practice, the game markers for the closure boundaries (for the purpose of calculating penalties) were moved upstream 150m from the relevant gantry locations in order to account for the fact that people tend to start their lane-change manoeuvres slightly in advance of the gantries. It was predicted that participants would be less inclined to enter the closed lanes (and therefore spend less time in the closed lanes) in the 'game' condition.

The data were split to assess each of the lane closures separately and, for each drive, the proportion of the distance of the closure in which a driver was in one of the closed lanes was calculated.

The results from the first closure indicated that drivers did drive shorter distances within the closed lane in the 'game' condition compared with the 'no-game' condition, but that this difference was not statistically significant (p=0.569). For the second closure (lanes 3 and 4) participants again drove shorter distances within the closed lanes (as a proportion of the closure length) in the 'game' condition (mean=0.022, SD=0.062) compared with the 'no-game' condition (mean=0.045, SD=0.091), and in this case the difference was statistically significant (p=0.041). Overall, 21/40 participants drove within at least some portion of the closed section (15/40 in the 'game' condition and 17/40 in the 'no-game' condition).

It is not known for certain why there should be an apparent effect for the second closure but not for the first, although anecdotal feedback from participants and the observing researchers indicated that some participants were unaware that the lane was still closed once they had passed the visible obstruction, and therefore did not associate the negative feedback from the game with a need to move back out of lane 1. Those drivers potentially wishing to move into lane 3 or 4 during the second closure appeared to be more aware of the status of the closure and therefore more receptive to game feedback.

> There was no significant difference between the proportion of time spent in the lane 1 closure in the 'game' and 'no-game' conditions; but participants did drive a significantly shorter proportion of the lane 3 and 4 closure within a closed lane in the 'game' condition.

2.3.2.8 Summary of the primary effects of the game on driving behaviour

Table 3 shows a summary of the effects of playing the game on the driving performance metrics.

Table 3: Driving performance metrics influenced by playing the game.

Performance Metric	Mean (SD)		Apparent effect of game*	Statistical significance
	Game	No-Game		
Standard deviation of driving speed (kph)	14.37 (2.30)	15.12 (1.78)	▼	$p<0.05$
Proportion of time above the speed limit	0.107 (0.081)	0.139 (0.111)	▼	$P<0.05$ (1-tailed)
Mean lane position	1.91 (0.423)	2.01 (0.414)	▼	$P<0.05$ (1-tailed)
Proportion of time spent in lane 1	0.344 (0.259)	0.265 (0.261)	▲	$p<0.05$
Time to change lanes (seconds)	2.58 (0.519)	2.76 (0.568)	▼	$p<0.05$
Proportion of driving in Closure 2	0.022 (0.062)	0.045 (0.091)	▼	$p<0.05$

*Green arrow indicates effect in the direction predicted; red arrow indicates effect counter to the direction predicted.

2.3.3 Game effects on driving behaviours related to driver age and gender

This analysis relates to the second research question and seeks to understand whether any of the observed significant effects of the game on driving behaviour (as investigated in Section 2.3.2) were related to the age and/or gender of the participant. All of the significant metrics listed in Table 3 were examined in relation to participant age (participants were categorised as either 'older' or 'younger' for the purposes of the calculation) and participant gender (male or female) using a repeated-measures ANOVA. Of these, two metrics returned results of significant interaction effects ('standard deviation of driving speed' and 'proportion of time spent over the speed limit'). However, the datasets for 'proportion of time over the speed limit' were found to be both significantly non-normal and to have significantly different error variances across the groups, therefore the results of the ANOVA cannot be regarded as reliable.

The data for 'standard deviation of driving speed' were both normal and of similar error variances across the groups and therefore the test results are valid. The test found a significant interaction effect between 'standard deviation of driving speed' and participant age (p=0.021), with older drivers found to be significantly more responsive to the game than younger drivers, and driving with a significantly lower standard deviation of their speed in the 'game' condition (mean=13.03, SD=0.43) than in the 'no-game' condition (mean=14.49, SD=0.37). Figure 5 shows a graphical indication of the interaction effect. It can be seen that older drivers were more responsive to the game than younger drivers. Indeed, the game appeared to have limited effect on younger drivers with regard to their speed variability.

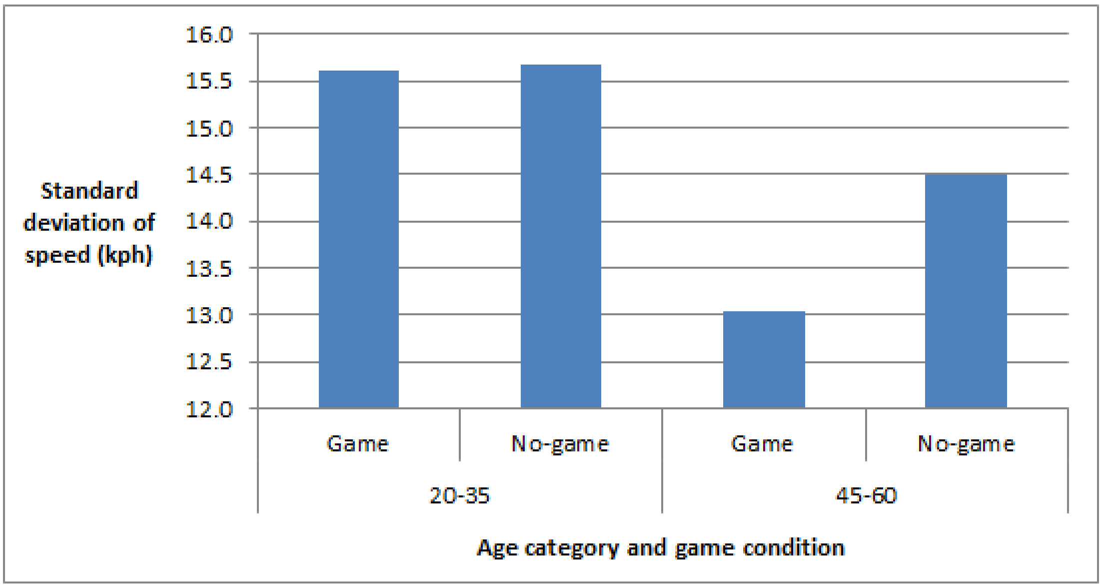


Figure 5: Interaction effect of driver age and game on standard deviation of driving speed.

There was a significant interaction effect between standard deviation of driving speed and participant age category, with older drivers significantly more responsive to the game than younger drivers.

2.3.4 Correlation between changes in driving performance metrics and game score

This analysis relates to the third research question and investigates the relationship between those driving metrics seen to differ significantly between the 'game' and 'no-game' drives and the participants' score in the game. It is essentially a test of whether the game was able to recognise and respond to those changes in driving performance.

The 'game' and 'no-game' datasets for the six significant metrics from Table 3 were transformed to produce a single variable for each metric. This was achieved by subtracting the 'no-game' value from the 'game' value. As a result, each participant had a single data-point for each metric, defining the difference between their scores across their two drives. This was similarly done for the difference in their final game scores across their two drives to allow for a simple correlation test between each driving performance metric and the game score.

A Spearman Rho[5] test was performed for each metric against final game score. Table 4 provides a summary of the results for each of the six correlation tests.

[5] Spearman Rho tests were used in the analysis due to better suitability for ordinal and nominal data than Pearson tests.

Table 4: Correlation test results of driving metrics against game score.

Driving performance metric	Spearman Rho coefficient	Statistical significance
Standard deviation of driving speed (kph)	-0.016	0.923
Proportion of time above the speed limit	-0.376	0.017 ($p<0.05$)
Mean lane position	-0.431	0.006 ($p<0.01$)
Proportion of time spent in lane 1	0.437	0.005 ($p<0.01$)
Time to change lanes (seconds)	0.019	0.907
Proportion of driving in Closure 2	-0.212	0.190

Three of the metrics returned significant results ('proportion of time spent above speed limit', 'mean lane position' and 'proportion of time spent in lane 1'), indicating that the game was able to reflect changes in driving behaviour against these metrics. That the game did not reflect changes in 'standard deviation of speed' and 'time to change lanes' is perhaps not surprising given that the game scoring system did not have any direct mechanism for assessing these behaviours ('time to change lanes' was measured indirectly through lateral acceleration, but the game only registered extreme events, i.e. 'red-flags'). The fact that participants were seen to alter their driving behaviour on these two metrics suggests that these changes were either indirect knock-on effects of changes in other behaviours, or that participants interpreted the game to be sensitive to these behaviours. 'Proportion of driving in closure 2' was clearly a metric to which the game was designed to respond. The lack of a significant correlation between this metric and game score is likely due to the fact that the majority of participants (28/40) did not enter this closure in either drive.

> 'Proportion of time spent above speed limit', 'mean lane position' and 'proportion of time spent in lane 1' were all significantly correlated with game score, indicating the game was able to reflect changes in these behaviours.

2.3.5 Correlation between changes in driving performance metrics and participant scores on the DBQ error and violation scales

As part of the post-trial questionnaire (included in Annex C), participants were asked to complete the Driver Behaviour Questionnaire (DBQ). This contained 16 questions relating to driving errors and eight questions relating to driving violations, all questions being on a 6-point scale, ranging from exhibiting this behaviour "never" to "nearly all the time" Each participant was given a simple score against each scale equal to the mean of their responses.

The six 'changes in driving performance' metrics (see Section 2.3.4) were tested for correlations with participant scores on both aspects of the DBQ, using Spearman Rho tests. However, all tests gave non-significant results, indicating that the effectiveness of the game in inducing changes in driving behaviour does not seem to be significantly related to driving style as measured by the DBQ.

> Changes in driving performance metrics between the game and no-game drives did not appear to be significantly correlated with driving style as measured by the DBQ.

2.3.6 Correlation between changes in driving performance metrics and participant subjective ratings of the game

As part of the post-trial questionnaire, participants were asked to give their subjective ratings of the game against five metrics. The five metrics and the response scales were:

- Rating of their **enjoyment of the game**, out of 10 (10 = extremely enjoyable, 0 = extremely unenjoyable)
- Rating of **how satisfying they found the game play to be**, out of 10 (10 = extremely satisfying, 0 = extremely unsatisfying)
- Rating of the extent to which they felt the game had an overall positive or negative effect on their **driving experience** (asked to mark on a line anchored with the words ‘negative’ and ‘positive’)
- Rating of the extent to which they felt the game had an overall positive or negative effect on the **quality of their driving** (asked to mark on a line anchored with the words ‘negative’ and ‘positive’)
- Rating of **how likely they would be to play the game** if it were available as a free app in the real world, out of 10 (10 = extremely likely, 0 = extremely unlikely).

Summary statistics for participant responses against these five subjective metrics are provided in Table 5.

Table 5: Summary statistics for participant subjective ratings of the game.

Subjective rating metric	Mean score (/10)	Standard deviation
Enjoyment of the game	6.48	2.31
Satisfaction of game play	6.48	2.35
Overall effect on driving experience	6.15	2.10
Effect on quality of driving	6.56	1.77
Likelihood to play the game	4.90	3.39

Spearman Rho correlation tests were performed using the game/no-game ‘difference’ score for each of the six significant driving behaviour metrics (see Section 2.3.4) against each of the five response scales listed above. The results of the tests were non-significant for all of the driving performance metrics save for ‘proportion of time driving in closure 2.’ For this metric, there were significant correlations with ‘enjoyment of the game’, ‘satisfaction of gameplay’ and ‘overall effect of the game on driving experience’. The full test results for ‘proportion of driving in closure 2’ are presented in Table 6.

Table 6: Correlation test results of ‘proportion of driving in closure 2’ against participant subjective ratings.

Subjective rating metric	Spearman Rho coefficient	Statistical significance
Enjoyment of the game	0.335	0.035 ($p<0.05$)
Satisfaction of game play	0.336	0.034 ($p<0.05$)
Overall effect on driving experience	0.434	0.005 ($p<0.01$)
Effect on quality of driving	0.193	0.232
Likelihood to play the game	0.170	0.295

The three significant results all appear to suggest that drivers who spent less time driving in the closure during their ‘game’ drive tended to have lower enjoyment and satisfaction towards the game when compared with their 'no-game' drive. Additionally, these drivers felt it had an increased negative effect on their driving experience It is not clear what may be behind these apparent relationships.

Correlations were identified between opinions towards the game and driving in closure 2, with drivers tending to be less positive about the game if they spent less time in the closure in their game drive than their no-game drive.

2.3.7 Correlation between participant subjective ratings and participant scores on the DBQ error and violation scales

The data for each of the five subjective ratings of the game (see Table 6) were tested for correlation with participant scores on the DBQ error and violation scales using Spearman Rho tests. All tests returned a non-significant result, indicating that driving style (as measured by the DBQ) was not a significant factor in determining participant subjective attitudes towards the game.

Participant scores on the DBQ did not correlate significantly with participant subjective ratings of the game.

2.3.8 Correlation between participant subjective ratings and driver sex and/or age

The data for each of the five subjective ratings of the game (see Table 6) were tested for correlation with participant age using Spearman Rho tests. All five tests returned a non-significant result, indicating that age was not a significant factor in determining participant subjective attitudes towards the game.

Participant age did not correlate significantly with participant subjective ratings of the game.

2.3.9 Correlation between participant subjective ratings and game score

The data for each of the five subjective ratings of the game (see Table 6) were tested for correlation with 'difference in game score' (see Section 2.3.4) using Spearman Rho tests. All five tests returned a non-significant result, indicating that the change in game score (i.e. the effectiveness of the game in inducing driver behavioural change) was not a significant factor in determining participant subjective attitudes towards the game.

Change in game score between the two drives did not correlate significantly with participant subjective ratings of the game.

2.3.10 Effects of the game on subjective ratings of the overall driving experience

As part of the post-trial questionnaire, participants were asked to rate their experience of each drive against five rating scales. Each scale was presented as a 10cm line with a one-word descriptive anchor at each end. Participants were asked to mark on the line the point that best represented their experience of that drive. These ratings were compared between the two drives to determine the effect of playing the game on driving experience. Each comparison was performed using a Wilcoxon signed-rank test and the results of these are presented in Table 7. For each scale, the mean rating is given out of 10, with 10 representing the right-hand anchor of each pairing. The tests were run using data from 39 participants, as one participant did not answer the questions.

Table 7: Results of the subjective ratings of drive experience tests.

Rating scale	**Mean (SD)**		**Apparent effect of game**	**Statistical significance**
	Game	**No-Game**		
Stressful / Relaxing	5.37 (2.36)	6.14 (2.37)	▼	p=0.013
Easy / Difficult	3.39 (2.39)	2.82 (1.99)	▲	p=0.040
Enjoyable / Disagreeable	3.37 (2.32)	3.76 (2.38)	-	p=0.300
Predictable / Unpredictable	4.47 (2.36)	4.00 (2.55)	-	p=0.121
Confusing / Clear	7.26 (2.29)	7.84 (1.71)	▼	p=0.011

The results indicate that across all five metrics, overall participants rated the game to be detrimental to their driving experience, with the effects on three of the five metrics shown to be statistically significant. It is perhaps interesting that participants felt this way, given that objective measures of driving performance suggested that playing the game improved their overall driving. Furthermore, participants were not instructed how to interact with the game, so negative feelings of increased stress and difficulty would appear to be self-imposed.

> Participants tended to think that the game negatively affected their driving experience, rating their game drives to be more stressful, more difficult and more confusing than their no-game drives.

2.4 Summary of results and discussion

The research sought to answer nine research questions, which are examined in this section, followed by a discussion on the implications of the study findings on game design.

Q1. Are driving behaviours seen to differ between the game and no-game drives?
As the primary research question, this research sought to establish whether exposure to the game within the simulator had any significant measureable effects on objective measures of driving behaviour. A number of measures of driving performance were extracted from the simulator, with the following found to differ significantly between participants' 'game' and 'no-game' drives:

- Standard deviation of driving speed
- Proportion of time above the speed limit
- Mean lane position
- Proportion of time spent in lane 1
- Mean duration of lane change manoeuvres
- Proportion of closure 2 driven in

For five of these six measures the effect of the game was apparently beneficial (from the perspective of optimum traffic management). Standard deviation of driving speed reduced, drivers spent less time over the speed limit, drivers adopted a mean lane position closer to lane 1 and spent a greater proportion of time in lane 1, and drove for shorter distances within the second lane closure. Participants did appear to perform lane changes more quickly (which may be viewed as a negative effect) although it is suggested that this may be a side-effect of spending more time in the inner lanes and therefore performing lane changes with potentially fewer distractions. Overall, the data appears to support the idea that the game had an overall positive effect on driving behaviour.

Q2. Is any game-induced difference in driving performance related to driver sex and/or age?
This question was based on the expectation that participant demographics may influence how drivers responded to the game. Of the six performance metrics found to be significantly influenced by the game, only 'standard deviation of driving speed' was found to have any significant interaction effects with driver demographics and it was found that older drivers appeared to be more influenced than younger drivers. In terms of homogeneity of speed, it may indicate that such a game may be more effective in influencing older drivers. Of the other measures, it would appear that drivers of both sexes and age groups responded to the game in a similar manner.

Q3. Is any difference related to overall performance in the game?
This question sought to establish how well the game reflected more objective measures of driving performance and was included as a validation of the game code itself. Of the six performance metrics found to be influenced by the game, three were significantly correlated with game score:

- Proportion of time above the speed limit
- Mean lane position
- Proportion of time spent in lane 1

This indicates that the game was able to at least partially reflect variations in these behaviours and therefore that observed changes in participant driving behaviours were based, at least in part, on genuine feedback on their driving, rather than simply on the perception of playing a game. Exactly how much any effect on driving was influenced by the feedback or simply exposure to the game is not possible to say.

Q4. Is any change related to participant scores on the DBQ error and violation scales?
This question sought to establish whether drivers' self-reporting as having more error-prone and/or less compliant driving styles were more or less responsive to the game. No evidence of any such relationships was found, suggesting that the game may be similarly effective in influencing driving behaviours of people with different driving styles.

Q5. Is any change related to participant subjective ratings of the game?
This question sought to establish whether there was any correlation between the effect of the game and how the game was perceived. It did not necessarily seek to establish causality, i.e. whether any negative attitudes to the game were the result of or cause of a weak game influence, for example. However, no correlations were identified, indicating that subjective opinions were not strongly related to the effectiveness of the game.

Q6. Are participant subjective ratings related to participant score on the DBQ?
This question sought to establish whether there was any correlation between self-reported driving styles and how the game was perceived. The expectation was that those drivers reporting themselves to be more error-prone might be more positive towards a game that gave them feedback on their driving. Likewise, the expectation was that drivers reporting themselves to be more likely to engage in intentional violations might be less positive towards such a game. However, no such links were identified in the data. This suggests that such a game may hold reasonably wide appeal if the gameplay can be refined to be sufficiently enjoyable.

Q7. Are participant subjective ratings related to driver sex and/or age?
This question sought to establish whether demographics might affect the appeal of the game. It was anticipated that younger drivers might be more positive towards the game. However, no significant links were found across the age or sex groupings. As for the results for Q6, this suggests that there is no specific reason to think that such a game may only appeal to limited demographic groups.

Q8. Are participant subjective ratings related to overall performance in the game?
This question sought to establish whether opinions towards the game were dependent on performance. For example, whether positive opinions were associated with those drivers who were more receptive to the game and who showed improvements in their game drive over their no-game drive. However, no such link was found in the data.

Q9. Do subjective ratings of the overall driving experience differ between the game and no-game drives?
This question sought to establish whether the game had any effect on the subjective driving experience. It was anticipated that providing participants with a game to play might make the driving experience more enjoyable. Interestingly, the opposite was found, with participants tending to regard their game drive as being more stressful, more difficult and more confusing. Despite participants receiving no instruction on how to interact with the game, it is possible that the negative opinions relate to feelings of pressure associated with a perceived need to perform well in the trial. It raises some concerns over the practical application of such a game in the real world, especially as such viewpoints would likely not result in a strong uptake of the game amongst the driving population. Anecdotal feedback from participants suggested that some participants found the lag between driving actions and game feedback to be confusing. Equally, the lack of explicit feedback on why a driver may be scoring negatively was reported as confusing. A clear example of this is in those participants who drove in the lane 1 closure (and received red flag penalties accordingly) yet without seeming to realise that they were in a closed lane. The finding suggests that gameplay would need to

be refined in a real-world version of the game, to ensure that drivers do not become frustrated and that negative and positive feedback can be attributed clearly to specific behaviours and therefore allow drivers the opportunity to address those behaviours.

The findings from the study indicate that the game was able to encourage drivers to adopt different driving behaviours in a simulated environment. These changes could be considered as beneficial to the network, from a traffic management perspective. However, the findings from the subjective feedback indicate that the game had a number of potential flaws that would need to be rectified for any version of the game to be played by drivers in the real world.

Distraction: The section boundaries within the game were designed to be away from areas of increased visual workload and thus attempted to minimise the effect of any distraction caused by the game. However, some participants did still raise concerns about distraction. It is not clear whether or not this concern was valid in terms of actual detriments to driving performance or whether it was simply an instinctive aversion to a potential distraction within the vehicle. However, the potential for distraction is clear and any real-world application would need to establish for certain that the potential behavioural benefits outweigh any potential negative effects of distraction. In any case, the design must seek to keep distraction as minimal as possible.

Delay: By providing feedback in terms of the driver's mean score, the game had a level of inherent delay. This delay was small at the start of each section and steadily increased as the section progressed due to the diminishing effect of each subsection score on the overall score. For example, a driver performing consistently poorly through a section would need a sustained period of good performance in order to bring the average score up to a level that would change the feedback he/she is receiving. During this time the driver would continue to receive negative feedback, despite driving well. This disconnection between input and feedback was reported by some drivers to have caused confusion and annoyance and may well be one of the key reasons for some of the negative feedback about the game itself and the 'game' condition drives. Ideally, a future iteration of such a game would seek to minimise any feedback delay.

Specificity: The game measured driving behaviour against six metrics and amalgamated these into a single feedback measure. In many cases drivers were clear about what driving behaviour was causing a negative score, but in others the link was not clear. Two examples were: 1) drivers using the closed lane 1 in the first closure, without realising that the lane was still closed; and 2) drivers travelling only very slightly above the speed limit. With regards to the latter, it was observed that some drivers allow a degree of tolerance in their speed choice (and indeed modern satellite navigation devices have meant that drivers have learned that speedometers in cars often underestimate the actual speed, which leads to the compensatory behaviour of driving slightly faster). The game was designed to start penalising drivers at any speed over the posted limit, which resulted in a number of drivers being penalised whilst appearing to believe that their speed choice was not the factor, or that they were being penalised unfairly. Future iterations of the game should seek to ensure that feedback is specific and so make it clear which behaviour needs to change. In addition, the penalty system should not be viewed by drivers as being so stringent as to put drivers off playing the game.

Plausibility: This was not a finding from the study as it was already known from the outset, but it is worth mentioning that the feedback mechanisms used in the game are currently not practicable in the real world. For example, it is not currently feasible to analyse the positions of surrounding vehicles to determine whether or not the driver is currently driving in the innermost lane available. Even calculating headway would require linking with on-board vehicle systems (such as adaptive cruise control) currently only found on newer and higher-end models. Any future game in the real world would need to address these practicality issues.

3 Dutch Study – The effect of in-vehicle traffic management information on driver behaviour

3.1 Introduction

Traffic management concerns the management of traffic flows by demand management measures, traffic information, traffic control, and other means to keep the transport system available, uncongested, safe and environmentally sustainable (NRA, 2008)[6]. Experiences show that theoretically effective traffic management interventions are not always effective in practice. To get more out of existing and future traffic management interventions taken by the national road administrations in terms of higher throughput and traffic safety, road users' strengths and limitations must be taken into account.

The effectiveness of traffic management interventions can be improved by increasing the level of drivers' perception, comprehension, capability and willingness. Networked and mobile technology offers new ways to (partly) transfer traffic management interventions into the vehicle and, with that, nearer to the driver. In-vehicle communication with the driver may help drivers tonotice and act upon information about changed traffic situations.

It is known from research and practice that in monotonous situations, drivers experience difficulties in perceiving and reacting in a timely manner to unexpected events (Larue, Rakotonirainy, & Pettitt, 2015, Thiffault, & Bergeron, 2003). McBain (1970) considers a monotonous situation when stimuli remain unchanged or when stimuli change in a predictable manner. This study explored – within TNO's driving simulator – how information regarding roadworks and speed could be presented in-vehicle to support drivers in order to enhance perception of and appropriate behaviour towards newly applicable roadworks regulations. Two arrangements of in-vehicle information were investigated. Firstly, spoken forewarnings were used (i.e. spoken messages warning drivers about upcoming changes) that were supplementary to the generally present information on the roadside. Secondly, a Head Up Display (HUD) was used that presented and replaced the information generally present on the roadside, such as speed limits and closed lanes, as well as visual forewarnings about upcoming changes. The primary objective of this research was to examine whether presenting information in the vehicle and in a more direct way to the driver enhanced perception of and compliance with traffic management interventions around roadworks. The nature of this research is exploratory, and aims to investigate whether or not supplementing or replacing roadside information with in-vehicle information is a promising venue for further development.

3.2 Methodology

3.2.1 Methodology overview

After an initial familiarisation drive, participants completed a trip in one of the three conditions: the control condition, the speech condition or the HUD condition. In the control condition, participants did not receive any additional information, only the information that is normally provided through the roadside when driving on highways and encountering roadworks. In the speech condition, information on the roadside was enriched with spoken forewarnings about upcoming changes. In the HUD condition, the information normally presented through the roadside was transferred to the HUD and combined with visual

[6] Nordic Road Association, ITS Terminology (ver. 2008.06.04) http://www.nvfnorden.org/lisalib/getfile.aspx?itemid=5582

forewarnings on the HUD about upcoming changes. It was anticipated that presenting information in the vehicle and nearer to the driver would enhance driver's compliance with (changed) speed limits and closed lanes. This was investigated by comparing relevant driving behaviours between the three conditions.

3.2.2 Participants

Sixty four people were recruited from TNO's existing participant database of individuals. Of these, eight did not appear at TNO at the appointed time. Eleven participants were excluded from the sample during or just after the trial - four encountered simulation sickness and seven were excluded because of technical problems with the simulator and/or data storage. The remaining forty five participants consisted of thirty men and fifteen women, age varied between 20 and 65 years, with a mean of 48.43 years. Participants were paid a fee for their time and travel expenditures. All participants owned a valid driving license and had normal or corrected-to-normal vision.

3.2.3 Simulator environment

3.2.3.1 Driving simulator

The moving base driving simulator at TNO Soesterberg (NL) was used. The driving simulator consisted of a mock-up of a BMW 318i with original controls (e.g. steering wheel, pedals, instruments) with an automatic gear change. The mock-up was controlled by a CCit system. This is a P-104-based Linux system that runs real-time Matlab-Simulink. A sound system produced the sound in and around the mock-up such as other traffic and car noise (i.e. engine, wheels, wind). The sound system is a state machine playing the requested sound files; this can be once or in a loop. The visualization was carried out by a combination of hardware and software components. The total picture (180° front view and 120° back view) was created with multiple channels; a channel consisted of an image generator, warping/blending card and a projector. The update frequency of each channel was 60Hz. The mock-up was mounted on a 6 degrees of freedom motion platform to give additional feedback to the driver. The feedback consisted of onset acceleration in case of accelerating or braking and roll in case of cornering. To generate the feeling of driving, road rumble was set on top of the other movements. The motion had a maximum range of +/- 20cm and +/- 20°, with a bandwidth of approximately 1.5Hz. The motion system was controlled by a motion control computer (50 Hz cycle time). A display in front of the driver, mounted in the dashboard, displayed information about the state of the car, speed and rpm. This display is freely configurable and the update rate was 25Hz.

3.2.3.2 Road environment

The simulated road environment represented a generic Dutch motorway with three lanes, a hard shoulder on the right side, and road signs displaying speed limits. During roadworks situations, arrow wagons, rumble strips, barriers, road workers, and gantries showing variable speed limits and lane closures were presented. Additional signage was used to signal the start and end of the roadworks sections.

In all three conditions, the route was divided into six roadworks sections. In sections 1 to 5, the same roadworks and accompanying information were presented to the participants every time. The aim was for the participants to experience impaired attention due to the monotonous situation. As a result, the scenery during the trip was dull to increase feelings of monotony in the participants and to induce a state of inattention.

In sections 1 to 5, the same information about speed and lane closures was presented each time at the same place and distance. When approaching the roadworks, the speed limit changed from 120 km/h to 90 km/h; the right lane (lane 1) was closed followed by the middle lane (lane 2). (See Annex E for clarification of the lane numbering convention used in the two studies.) Therefore, the roadworks situations in sections 1 to 5 required the same reactions from the participants each time – that is, to move out of the right and middle lane and only driving in the left lane (lane 3) at 90 km/h. The sixth roadworks situation, however, was different from the five previously encountered. In section 6, different information about speed and lane closures was presented and therefore other responses were required from the participants. The speed limit changed from 120 km/h to 90 km/h and then to 70 km/h; the left lane, middle lane and right lane were closed. As a result, participants had to move out of all the lanes and were only allowed to drive on the right shoulder at 70 km/h in this roadworks scenario. It was expected that the information presented through speech and on the HUD would help participants to become aware of and comply with the changed situation that required altered driving behaviour.

3.2.3.3 Head Up Display

In the HUD condition, the roadworks information was presented to the participants only on the HUD. The HUD displayed road signs signalling accessible lanes, speed limits and lane closures. The left side of the HUD continuously presented the current velocity and the mandatory speed limit, whereas the right side of the HUD pre-warned the participant 300 metres in advance about upcoming changes such as lane closures and new speed limits. The pre-warning '300 m' was continuously displayed, and did not count down. The bottom section displayed the distance (in metres) that was still to be covered until a change to the current situation would come into effect. Figure 6 presents an example of the information presented by the HUD.

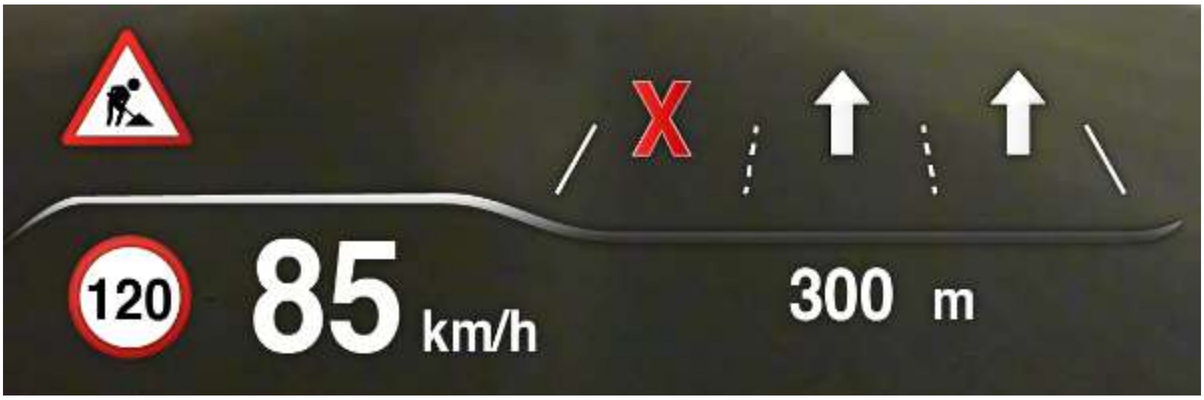


Figure 6. Example of information presented through a Head-Up display.

3.2.3.4 Spoken forewarnings

In the speech condition, participants received spoken messages that warned them about upcoming changes. These messages were played back once to the participants 300 metres before the changes came into effect. For example: "You are approaching roadworks. In 300 meters, the new speed limit will be 90 km/h."

3.2.4 Experimental design

A between-subjects design was used; participants were randomly assigned to either the control condition, the speech condition or the HUD condition. In the control condition, the participants did not receive any additional information, only the information that is normally provided on the roadside when driving on highways and encountering roadworks. In the speech condition, roadside information was enriched with spoken forewarnings about upcoming changes. In the HUD condition, the information normally presented through the roadside was transferred to the HUD and combined with forewarnings on the HUD about upcoming changes. For the purpose of the experiment, each condition entailed a similar number of participants; sixteen participants in the control condition (ten male, six female, with a mean age of 44.5), thirteen participants in the speech condition (eight male, five female, with a mean age of 51.6) and sixteen participants in the HUD condition (twelve male, four female, with a mean age of 49.2).

3.2.5 Trial protocol

Potential participants were approached for recruitment through email or telephone and were told that they would be asked to perform one drive in the simulator for around 40 minutes, but were not told about the exact purpose of the trial.

On arrival at TNO, participants were asked to read a short document about the simulator and the study itself and to sign a consent form. Then, they performed a 5-minute familiarisation drive in the simulator. This is standard practice for simulator studies and helps participants to familiarise themselves with driving in the simulator. After the familiarisation drive, participants were instructed to drive as they normally would do; then their drive in either the control, speech or HUD condition would start. Following completion of the drive, participants were asked to complete a questionnaire assessing acceptance of the varying information presentations (see 3.2.6.2 Questionnaire). The complete questionnaire is provided in Annex D: Questionnaire used in NL trial. Once the questionnaire was complete, participants were thanked for their participation and were paid a fee for their time and travel expenditures.

3.2.6 Data collection

3.2.6.1 Simulator data

The following performance variable data were collected/calculated from the simulator:

- Average speed
- Average time needed to comply with the new speed limits
- Proportion of time spent over the speed limit
- Compliance mistakes with the new speed limit
- Average distance to the arrow wagons when moving out of a closing lane
- Compliance mistakes concerning lane closures

3.2.6.2 Questionnaire

Following completion of the drive, participants were asked to complete a questionnaire consisting of two parts. The first part consisted of questions about gender, age, and driving experience. In the second part, participants were asked to complete the validated acceptance questionnaire that was created by Van der Laan, Heino and De Waard (1997). The exact wording of the statements presented to the participants depended on the experimental condition the participant was in. (For the control condition the statement was “While driving, I perceived the information on the road signs to be …”; for the speech condition participants received an additional statement: “While driving, I perceived the spoken messages to be …”; participants in the HUD condition only rated the statement: “While driving, I perceived the information on the Head Up Display to be ….”). Participants responded to these statements by scoring their responses to nine bipolar adjectives on a five-point Likert scale: useful – useless, pleasant – unpleasant, bad – good, nice – annoying, effective – superfluous, irritating – likeable, assisting – worthless, undesirable – desirable, raising alertness – sleep-inducing.

Twelve statements were constructed and recorded in order to capture whether the participants – during the drive – perceived the presented information to be understandable (understandability scale) and whether they felt alert (driver alertness scale). Six statements had positive wordings, such as 'I felt I was concentrated while driving' and the other six statements had negative wordings, such as 'I felt distracted while driving' (see Annex D). Each time when participants had passed a roadworks section, the statements were played back to the participants while they continued driving. Participants were asked to indicate verbally whether they agreed (yes), agreed slightly (a little bit) or disagreed (no) with the twelve statements that were presented verbally to them. Participants' answers were written down by the researcher. A higher score reflected better understandability and higher alertness levels.

3.2.7 Research questions

The purpose of the present study was to examine whether in-vehicle information about new speed limits and upcoming lane closures enhanced drivers' compliance with new roadworks regulations compared to receiving no in-vehicle information. Passing the roadworks in sections 1 to 5 was meant to induce a state of driver inattention. The research questions and analysis regarding driver behaviour were therefore focused on section 6. On the other hand, understandability of the information and level of driver alertness were assumed to change over time and were therefore analysed per section. Acceptance of the provided information was measured after completion of the trip and was therefore analysed for the trip as a whole.

The main research question was:

1. Is drivers' compliance with new roadworks regulations on the highway found to differ between the control, speech and HUD condition?
    a. Is participants' average driving speed in section 6 lower in the speech condition and HUD condition than in the control condition?
    b. Is participants' time needed to comply with new speed limits in section 6 lower in the speech condition and HUD condition than in the control condition?
    c. Is the proportion of time spent above the speed limit in section 6 less in the speech condition and HUD condition than in the control condition?
    d. Is the percentage of participants who mistakenly continue to drive above the new speed limit in section 6 lower in the speech condition and HUD condition than in the control condition?
    e. Do participants adopt a larger distance to the arrow wagons in section 6 in the speech condition and HUD condition than in the control condition?
    f. Is the percentage of participants who mistakenly continue to drive in a closed lane in section 6 lower in the speech condition and HUD condition than in the control condition?

Secondary research questions were:

2. To which degree do participants understand and accept the information that is presented to them in different ways?
3. To what degree are participants alert during the six roadworks sections?

3.3 Analysis and results

3.3.1 Effects on driving behaviour

The following analyses tested for differences in drivers' compliance with roadworks regulations between the control condition, the speech condition, and the HUD condition.

3.3.1.1 Average speed

The in-vehicle information was expected to have a positive effect on participants' compliance with speed limits, therefore, a prediction was made that the mean driving speed in section 6 would be lower in the speech condition and HUD condition than in the control condition.

Mean participants driving speed was calculated across section 6 and compared between the three conditions. A one-way ANOVA was conducted to test whether there were differences in mean speed in the sixth roadworks section between the control, speed and HUD condition. Results showed that there were no significant differences regarding mean speed in section 6 between the three conditions ($F(2, 42) = 0.48$, n.s.). Participants' mean speed in the control condition ($M = 101.86$ km/h; $SD = 15.79$) did not significantly differ from participants' mean speed in the speech condition ($M = 103.65$ km/h; $SD = 15.04$) and HUD condition ($M = 107.32$ km/h; $SD = 15.94$).

3.3.1.2 Time needed to comply with speed limits

The in-vehicle information was intended to have a positive effect on participants' compliance with speed limits. As a result, a prediction was made that the time to comply with speed limits in section 6 would be lower in the speech condition and HUD condition than in the control condition.

Time needed to comply with the new speed limit of 90 km/h and of 70 km/h (in seconds) was calculated across section 6 and compared between the three conditions. Measurements started 300 m before passing the road sign with the new speed limit and ended when the participant complied with the new speed limit. For the participants that drove at the start of the measurement period over 90 km/h and over 70 km/h respectively, the time needed to comply with the new speed limit was computed and averaged.

With regard to the time need to comply with the new speed limit of 90 km/h in section 6, Levene's test showed that the data violated the assumption of equal variances. Therefore, a Welch's test was conducted to assess whether there were differences in time needed to comply with the new speed limit between the control, speech and HUD condition. Results showed no significant differences in the amount of time taken to comply with the new speed limit of 90 km/h between the three conditions ($F(2, 13.32) = .01$, n.s.). Participants' time needed to comply in the control condition ($M = 8.20$ sec.; $SD = 6.36$) did not significantly differ from participants' time needed to comply in the speech condition ($M = 8.75$ sec.; $SD = 9.56$) or in the HUD condition ($M = 8.29$ sec.; $SD = 6.47$).

With regard to the new speed limit of 70 km/h in section 6, a one-way ANOVA was conducted to test whether there were differences in time needed to comply with the new speed limit between the control, speech and HUD condition. The results showed no significant differences in the time taken to comply with the new speed limit of 70 km/h between the three conditions ($F(2, 29) = .22$, n.s.). Participants' time needed to comply in the control condition ($M = 12.18$ sec.; $SD = 6.35$) did not significantly differ from participants' time needed to comply in the speech condition ($M = .11.30$ sec.; $SD = 8.03$) or in the HUD condition ($M = 10.00$ sec.; $SD = 8.58$).

3.3.1.3 Proportion of time over the speed limit

The in-vehicle information was expected to have a positive effect on the proportion of time a participant exceeded the speed limits. Therefore a prediction was made that the proportion of time spent above the speed limit in section 6 would be less in the speech condition and HUD condition than in the control condition.

Proportion of time spent above the speed limit was calculated across section 6 and compared between the three conditions. Measurements started when participants passed the road sign (either on the roadside or on the HUD) with the new speed limit and ended when the next new speed limit was presented to the driver. A Shapiro-Wilk test showed the data in section 6 to be non-normally distributed and so a Kruskal-Wallis H test was conducted to test whether there were differences in the proportion of time spent above the speed limit in the sixth roadworks section between the control, speed and HUD condition. Results showed that there were no significant differences regarding proportion of time spent driving above the speed limit in section 6 between the control, speech and HUD conditions (($H(2) = .30$, n.s.). Participants' time spent above the speed limit in the control condition (M = 68.95 %.; SD = 43.47) did not significantly differ from participants' proportion of time spent above the speed limit in the speech condition (M = 56.40 %.; SD = 49.32) or the participants in the HUD condition (M = 64.65 %.; SD = 44.97).

3.3.1.4 Compliance mistakes with the new speed limit

The in-vehicle information was expected to reduce mistakes in complying with the new speed limits of 90 km/h and 70 km/h in section 6. Therefore a prediction was made that the percentage of participants who mistakenly continued to drive above the new speed limits would be lower in the speech condition and HUD condition than in the control condition.

A compliance mistake with the new speed limit of 90 km/h was registered when participants mistakenly continued to drive above the new speed limit while driving through the sixth roadworks section. According to SWOV (2012), drivers need one to two seconds to respond to a sudden situation change; we applied that finding to the time needed to start to comply with a new speed limit. With regard to 90 km/h, the predefined period of time included a goodwill allowance of two seconds, meaning that participants had time to comply with the new speed limit 50 metres after they passed the road sign indicating the new speed limit. When participants still drove above the new speed limit at this point (i.e. 50 metres after the road sign) a compliance mistake was registered. With the new 90km/h speed limit in place, participants could therefore make either zero compliance mistakes or one compliance mistake in the sixth roadworks section.

A Chi-square test was conducted to test whether there were differences in numbers of participants that made compliance mistakes while driving through the sixth roadworks section between the control, speech and HUD condition. Results revealed that the number of participants that made a compliance mistake regarding the 90 km/h limit did not significantly differ between the conditions ($\chi(2) = .24$, n.s.). The percentage of participants in the control condition that made a compliance mistake (81.3%) did not significantly differ from the percentage of participants in the speech condition making a compliance mistake (84.6%) or in the HUD condition (87.5%).

With regard to the new speed limit of 70 km/h in the sixth roadworks section, the predefined period of time included a goodwill allowance of two seconds, meaning that participants had time to comply with the new speed limit 39 metres after the road sign displaying the new speed limit. When participants still drove above the new speed limit at this point (i.e. 39 metres after the road sign) a compliance mistake was registered. Therefore, under the new speed limit of 70km/h, participants could make either zero compliance mistakes or one compliance mistake in the sixth roadworks section.

A Chi-square test was conducted to test whether there were differences in the number of participants that made compliance mistakes while driving through the sixth roadworks situation between the control, speech and HUD condition. With the 70 km/h speed limit in force, results revealed that the number of participants that made a compliance mistake did not significantly differ between the conditions ($\chi(2)$ = .89, n.s.). The percentage of participants in the control condition that made a compliance mistake (75%) did not significantly differ from the percentage of participants in the speech condition making a compliance mistake (76.9%) or in the HUD condition (87.5%).

3.3.1.5 Distance to arrow wagons

The mean distance to the three arrow wagons in roadworks section 6 was calculated across the participants who were driving in the lane that was going to be closed; thus excluding participants who drove in other lanes. The distance from the front of the car to the arrow wagon was captured at the moment the participant moved out of the lane that was going to be closed. The distance to the three arrow wagons (in metres) while driving through the sixth roadworks section was computed and compared between the three conditions. The in-vehicle information was expected to have a positive effect on participants' lane changing behaviour, therefore a prediction was made that participants would adopt a larger distance to the arrow wagons in the speech condition and HUD condition than in the control condition.

Levene's test showed that the data regarding distance to the arrow wagons violated the assumption of equal variances. As a result, we conducted a Welch's test to examine whether there were significant differences in distance to the arrow wagons in the sixth roadworks section between the control, speech and HUD condition; results showed no significant differences between the conditions ($F(2, 14.67) = 1.58$, n.s.). Participants' adopted distance to the arrow wagons in the control condition (M = 443.30 metres; SD = 19.71) did not significantly differ from participants' distance in the speech condition (M = 410.15 metres; SD = 90.96) or in the HUD condition (M = 449.23 metres; SD = .39).

3.3.1.6 Compliance mistakes concerning closed lanes

The in-vehicle information was expected to have a positive effect on participants' closed lane compliance. Therefore, a prediction was made that the percentage of participants who mistakenly continued to drive in a closed lane would be lower in the speech condition and HUD condition than in the control condition.

A compliance mistake concerning closed lanes was registered when participants mistakenly continued to drive in the closed lanes while driving through the sixth roadworks section. The predefined period of time included a goodwill allowance of two seconds, meaning that participants had time to move out of the closing lane 50 metres after they passed the road sign. When participants still drove on the closed lane at this point (i.e. 50 metres after the road sign) a compliance mistake was registered. Participants could therefore make either zero compliance mistakes or one compliance mistake in the sixth roadworks section.

Regarding the first lane closure in section 6, only one participant made a compliance mistake. A Chi-square test confirmed that the number of participants that made a compliance mistake regarding the first lane closure did not significantly differ between the conditions ($\chi(2)$ = 1.79, n.s.). The percentage of participants in the control condition that made a compliance mistake (0%) did not significantly differ from the percentage of participants in the speech condition making a compliance mistake (0%) or in the HUD condition (6.3%).

Regarding the second lane closure in section 6, more participants made compliance mistakes. A Chi-square test showed that the number of participants that made a compliance mistake regarding the second lane closure did not significantly differ between the conditions ($\chi(2)$ = 0.59, n.s.). The percentage of participants in the control condition that made a compliance mistake (80%) did not significantly differ from the percentage of participants in the speech condition making a compliance mistake (76.9%) or in the HUD condition (87.5%).

3.3.2 Acceptance and understanding of information and driver alertness

3.3.2.1 Acceptance

After completion of the drive in the simulator, participants' acceptance of the provided information was assessed by the validated acceptance questionnaire (Van der Laan et al., 1997). The exact wording of the statements that were presented to the participants depended on the condition the participant was in (see Annex D). The scores of the bipolar items 3, 6 and 8 were mirrored. The items 1, 3, 5, 7 and 9 represented the usefulness scale with Cronbach's alpha of .765, which indicated an good internal consistency for the scale. The items 2, 4, 6 and 8 reflected the satisfaction scale with a Cronbach's alpha of .902, which indicated an excellent internal consistency for the scale. A lower score reflected a higher level of acceptance.

A one-way ANOVA was conducted to test whether there were differences regarding acceptance of information (i.e. 'usefulness' as well as 'satisfaction' score) between the control, speech and HUD condition. With regard to usefulness, the assumption of homogeneity of variances was violated, so a Welch test was conducted. Results showed that there was no significant difference between participants in the three conditions ($F(2, 26.79) = 81.08$, n.s; $M = 1.96$, $SD = .28$; $M = 1.91$, $SD = .29$; $M = 2.13$, $SD = .49$ respectively).

With regard to satisfaction, the results of the ANOVA showed that there was a significant difference between participants in the three conditions ($F(2, 42) = 9.88$, $p = .001$). Pairwise comparisons with the post hoc Tukey HSD test showed that participants in the HUD condition ($M = 2.89$; $SD = .56$) were more satisfied with the presented information than participants in the speech condition ($M = 3.25$; $SD = .27$; $p = .045$) and in the control condition ($M = 3.50$; $SD = .24$; $p = .001$).

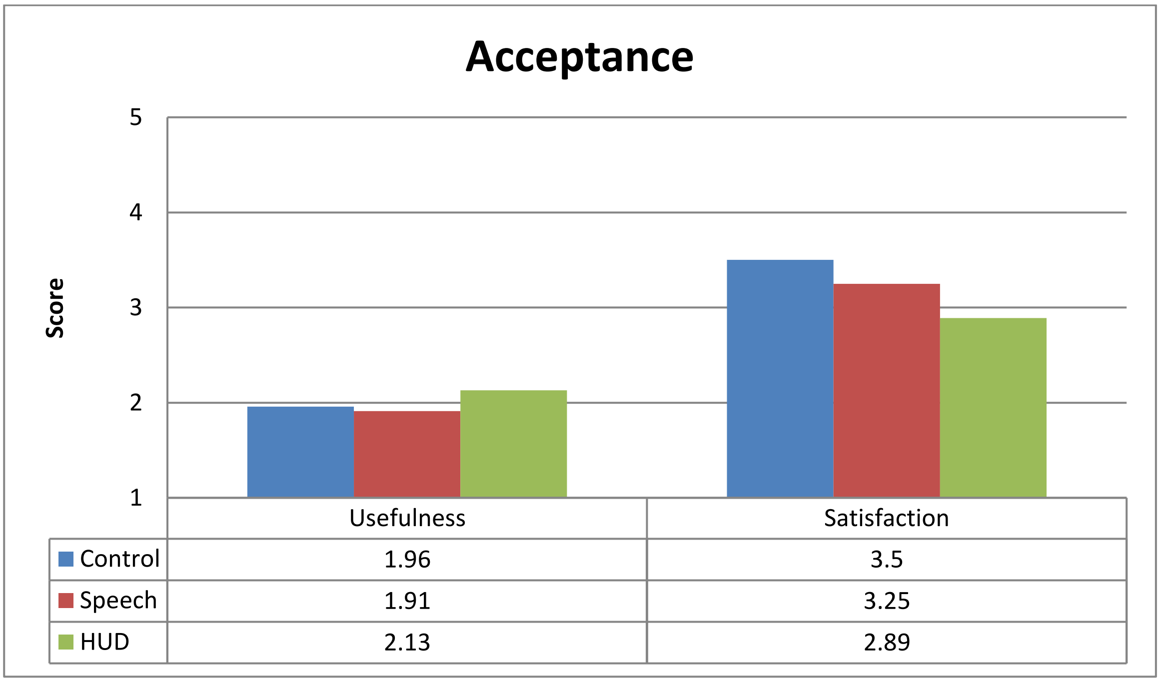


Figure 7 - Perceived usefulness of and satisfaction with information. Note that a lower score reflects higher level of usefulness and satisfaction respectively.

The participants in the speech condition rated acceptance of the roadside information as well as of the spoken forewarnings. Paired samples T tests showed that participants perceived the roadside information ($M = 1.91$; $SD = .29$) to be more useful than the spoken forewarnings ($M = 2.61$; $SD = .81$; $t(12) = -3.43$; $p = 005$). On the other hand, participants were more satisfied with the spoken forewarnings ($M = 2.94$; $SD = .31$) than with the roadside information ($M = 3.25$; $SD = .27$; $t(12) = 2.79$; $p = .016$).

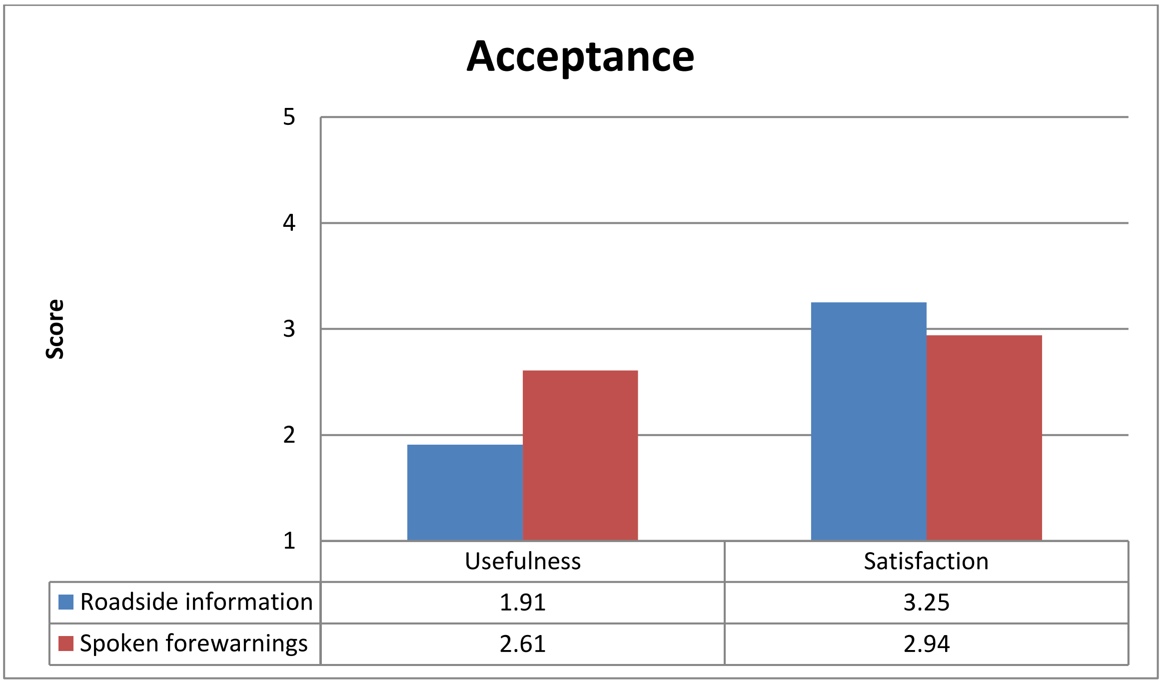


Figure 8 - Perceived usefulness of and satisfaction with roadside information and spoken forewarnings. Note that a lower score reflects higher level of usefulness and satisfaction respectively.

3.3.2.2 Understanding of information

Concerning each roadworks section, participants were asked during the drive to indicate verbally whether they agreed, agreed slightly, or disagreed with six statements assessing understandability of the presented information. The mean scores on these six items for each of the sections were summed and averaged to form the understandability subscale (i.e. items: 2, 3, 7, 8, 10 and 11; see Annex D) with a Cronbach's alpha of .91, which indicated an excellent internal consistency. A higher score reflected better understandability.

A two-way 3 (condition: control, speech or HUD) x 6 (section: 1 to 6) mixed ANOVA with repeated measures on understandability was conducted to test whether there were differences regarding understandability of the information over time (throughout sections 1 till 6) between the control, speech and HUD condition. Mauchly's test showed that the assumption of sphericity was violated slightly (epsilon = .89). Therefore, the Huynh-Feldt correction was used to interpret the results. The results showed that there was no significant interaction effect regarding understandability over time between the three conditions ($F(8.87, 186.17) = 1.11$, n.s.) and no significant main effect of section ($F(4.43, 186.17) = 1.55$, n.s.). There was, however, a significant main effect of condition ($F(2, 20.46) = 9.72$, $p = 001$). A post hoc Games-Howell test (because assumption of homogeneity of variances was violated) showed that participants in the control condition ($M = 2.95$; $SD = .08$) perceived the presented information to be more understandable than participants in the speech condition (M = 2.54; SD = .45; p = .019) and in the HUD condition (M = 2.79; SD = .18; p = 009). There were no significant differences regarding understandability between the speech and the HUD condition.

3.3.2.3 Alertness

Concerning each roadworks section, participants were asked during the drive to indicate verbally whether they agreed, agreed slightly, or disagreed with six statements assessing level of driver alertness. The mean scores on these six items for each of the sections were summed and averaged to form the driver alertness subscale (i.e. items 1, 4, 5, 6, 9 and 12; see Annex D) with a Cronbach's alpha of .94, which indicated an excellent internal consistency. A higher score reflected higher alertness levels.

A two-way 3 (condition: control, speech or HUD) x 6 (section: 1 to 6) mixed ANOVA with repeated measures on level of alertness was conducted to test whether there were differences regarding level of alertness over time (throughout sections 1 till 6) between the control, speech and HUD condition. Mauchly's test showed that the assumption of sphericity was violated considerably (epsilon = .67). Therefore, the Greenhouse-Geisser correction was used to interpret the results. The results showed that there was no significant interaction effect regarding level of alertness over time between the three conditions (F(5.88, 123.53) = .88, n.s.). However, there was a significant main effect of section (F (2.94, 123.53) = 2.77, p = .046). A post hoc Bonferroni test showed that participants were slightly more alert in section 2 (M = 2.85; SD = .23) compared to section 6 (M = 2.77; SD = .30; p = .03). There was no significant main effect of condition (F(2, 24.16) = 2.32, n.s.).

3.4 Summary of results and discussion

The purpose of the study was to examine whether presenting information in the vehicle enhanced compliance with traffic management interventions around roadworks. Spoken forewarnings were provided to the participants in addition to the standard roadside information. A HUD was used to present and replace the information generally present on the roadside, as well as to display visual forewarnings about upcoming changes. As the primary research question, this research sought to establish whether exposure to different ways of presenting information had significant effects on driver's compliance with roadworks regulations. We were especially interested in the situation that drivers were less alert and supposedly had trouble perceiving information about roadworks regulations.

The research sought to answer eight research questions, which are examined in this section.

Q1. Is participants' average driving speed in section 6 lower in the speech condition and HUD condition than in the control condition?

This research question sought to establish whether average driving speed was affected by the different ways of presenting the information. No evidence was found that receiving in-vehicle information – either spoken or via a HUD – supported participants to drive at lower mean speed when they encountered a different roadworks situation on the highway. On the other hand, this finding suggests that relatively new ways of information presentation (by speech and by HUD) do not necessarily worsen drivers' compliance with new roadworks speed limits either.

Q2. Is participants' time needed to comply with new speed limits in section 6 lower in the speech condition and HUD condition than in the control condition?

This research question sought to establish whether participants' time needed to comply with new speed limits was affected by the different ways of presenting the information. No evidence was found that receiving in-vehicle information – either spoken or via a HUD – supported drivers in this when they encountered a different roadworks situation on the highway.

Q3. Is the proportion of time spent above the speed limit in section 6 less in the speech condition and HUD condition than in the control condition?
This research question sought to establish whether participants' proportion of time spent above the speed limit was affected by the different ways of presenting the information. No evidence was found that receiving in-vehicle information – either spoken or via a HUD – reduced the proportion of time above the speed limit when the participants passed a different roadworks situation on the highway. In this driving simulator study, participants drove above the speed limit in section 6 for a large proportion of the time (between 56.4 % and 68.95 %), indicating that in all conditions it was difficult for them to adhere to the speed limit. (The reasons behind this behaviour is outside the scope of this research.)

Q4. Is the percentage of participants who mistakenly continue to drive above the new speed limit in section 6 lower in the speech condition and HUD condition than in the control condition?
This research question sought to establish whether the number of participants who mistakenly continued to drive above the new speed limit was affected by the different ways of presenting the information. No evidence was found that receiving in-vehicle information – either spoken or via a HUD – reduced the number of participants mistakenly driving above the speed limit when they passed a different roadworks situation on the highway. In this driving simulator study, the number of participants that continued to drive above the speed limit in section 6 was relatively large (between 75% and 87.5%), indicating that in all conditions it was difficult for them to adhere to the speed limit. (The reasons behind this behaviour is outside the scope of this research.)

Q5. Do participants adopt a larger distance to the arrow wagons in section 6 in the speech condition and HUD condition than in the control condition?
This research question sought to establish whether participants' distance to the arrow wagons when lanes were closing was affected by the different ways of presenting the information. No evidence was found that receiving in-vehicle information – either spoken or via a HUD – affected participants' adopted distance to the arrow wagons when they encountered a different roadworks situation on the highway. On the other hand, this finding suggests that relatively new ways of information presentation (by speech and by HUD) do not necessarily worsen drivers' compliance with new roadworks lane closures either.

Q6. Is the percentage of participants who mistakenly continue to drive in a closed lane in section 6 lower in the speech condition and HUD condition than in the control condition?
This research question sought to establish whether the number of participants who mistakenly continued to drive in a closed lane was affected by the different ways of presenting the information. No evidence was found that receiving in-vehicle information – either spoken or via a HUD – reduced the number of participants mistakenly driving in a closed lane when they passed a different roadworks situation on the highway. In this driving simulator study, the number of participants that continued to drive in a closed lane in section 6 was relatively large (between 80% and 87.5%), indicating that in all conditions it was difficult for them to move out of the closed lane in time. (The reasons behind this behaviour is outside the scope of this research.)

Q7. To which degree do participants understand and accept the information that is presented to them in different ways?

These research questions sought to establish to which degree participants understood and accepted the information that was presented to them in different ways. Evidence was found that participants in the HUD condition were more satisfied with the presented information than participants in the speech condition and in the control condition. On the other hand, participants in the HUD condition did not perceive the presented information to be more useful than participants in the control or speech condition. Similar results were found for the participants in the speech condition that rated the acceptance of both the standard roadside information and the spoken forewarnings separately. Evidence showed that participants perceived the roadside information to be more useful than the spoken forewarnings. On the other hand, participants were more satisfied with the spoken forewarnings than with the roadside information. Regarding understandability of the presented information, evidence showed that participants in the control condition perceived the presented information to be more understandable than participants in the speech condition and in the HUD condition. This suggests that, although the way of presenting information had no effect on drivers' compliance with speed limits and lane closures, participants apparently understood the information that was presented to them in the normal way (i.e., on the roadside) better than information that was presented in a more unfamiliar way (by speech or on HUD).

Q8. To which degree are participants alert during the six roadworks sections?

This research question sought to establish whether participants' level of alertness over time differed. Evidence showed that participants were slightly more alert in section 2 compared to section 6, suggesting that encountering the same roadworks situation repeatedly decreases participants' perceived level of alertness.

4 Conclusions

The two studies described in this report investigated how two existing, near future concepts (gamification and a head–up display) could be applied to improve driver's compliance with traffic management guidance. This has the potential to improve overall network efficiency – firstly, through this increased compliance effect but secondly, by giving national roads authorities better and more specific tools for managing the movements of vehicles on their networks.

The UK study investigated the use of gamification techniques on driving behaviour. A simple game design was created with the aim of encouraging drivers to adopt behaviours that were safe but also complied with traffic management guidance. The results were mixed but somewhat encouraging. The main positive result was that, on a number of metrics, participants were observed to display significantly better driving behaviour during their drive with the game active than when there was no game present. These effects seemed to apply for male, female, older and younger participants and were observed irrespective of whether the participants liked or disliked the game itself.

The main negative effect of the game was that, contrary to expectation, participants did not like the game and reported that it made driving more difficult and more stressful. On reflection, this is perhaps unsurprising as it is, to some degree, requiring drivers to adopt behaviours that are different from their usual driving habits and it may therefore feel uncomfortable. However, it had been expected that enjoyment of the game may offset such feelings. It should be noted however, that although the game was created to the best of the abilities of the technical team involved in the project, it would be reasonable to state that there was no specific game mechanics experts that could perhaps have helped to make the game as engaging and enjoyable as possible.

It seems that here lies the greatest opportunity from this study – results indicate that gamification can have positive effects on driving behaviours relevant to safety and traffic management. However, the game was not found to be enjoyable with the ultimate consequence that it would be unlikely to be used in the longer term (if at all). A follow-up project in which experts from the mobile gaming industry are brought in to collaborate and create a more attractive and entertaining game could result in an offering that is genuinely useful for traffic management. It would also be necessary to ensure that such a game is not engaging to the extent that it becomes a distraction for drivers, resulting in increased risk of collision. Once satisfied that the game was safe and effective, on-road trials could establish the real world effectiveness of this system and investigate longer term adaptation and usage rates. In that context, it would also be of interest to study broader gamification principles such as the use of competition between users to further encourage safer and more compliant practices.

The Dutch study was interesting in that it found traffic management information presented by in-vehicle information – either spoken or via a HUD - did not lead to better or more timely compliance than the same information presented via traditional roadside signs. This contrasts with previous studies (e.g. Liu & Wen, 2004) that have found HUDs to improve drivers' reactions – although those results were in a higher workload, urban environment than on the relatively simple highway environment under test.

Participants found the HUD information to be more satisfying but no more useful than the roadside information and the in-vehicle speech interface. It should also be noted that since participants found HUD-presented information no *less* useful than traditional techniques, then the use of HUDs for providing more specific individual vehicle guidance in traffic management should be explored. This specificity is the real potential benefit of in-vehicle information. As such, it would be of interest to study how individual vehicles given differing traffic management information would behave in a shared motorway environment. It would also be of interest to understand user attitudes to enforcement of traffic management measures implemented by in-vehicle information. This would represent a significant change in policy but in technological terms, it would not be difficult to detect when a vehicle was travelling significantly in excess of its individually prescribed speed limit.

The two simulator studies presented investigated how two different human factors techniques could support future traffic management. The use of simple gamification principles brought about changes in driving behaviour but the approach was not well-liked by participants and felt to make the driving task more difficult. The use of in-vehicle information for presenting traffic management information was found to be no better (but no worse) than traditional techniques but was found to be more satisfying for participants. It can be concluded that each study has therefore provided results that suggest each technique can be explored further. In particular, firstly, it would be of interest to work with game design specialists to create a specific game to encourage safe and compliant driving behaviour that builds upon the results of the UK study. Secondly, the use of in-vehicle information for traffic management was found to be non-detrimental to safe, compliant driving behaviour so further work should focus on how to gain further benefit from the flexibility of traffic management that can be achieved by this approach.

5 Acknowledgement

The research presented in this report was carried out as part of the CEDR Transnational Road Research Programme Call 2013. The funding for the research was provided by the national road administrations of Belgium-Flanders, Denmark, Finland, Norway, UK and Netherlands.

6 References

Larue, G. S, Rakotonirainy, A., & Pettitt, A. N. (2015) Predicting reduced driver alertness on monotonous highways : a driving simulator study. IEEE Pervasive Computing, 14(2), 78-85.

Liu, Y. C., & Wen, M. H. (2004). Comparison of head-up display (HUD) vs. head-down display (HDD): driving performance of commercial vehicle operators in Taiwan. International Journal of Human-Computer Studies, 61(5), 679-697.

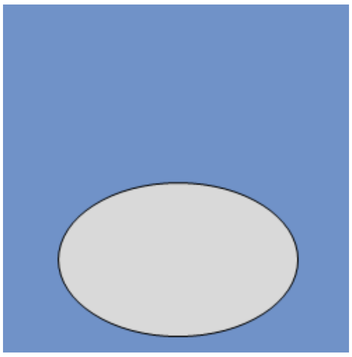
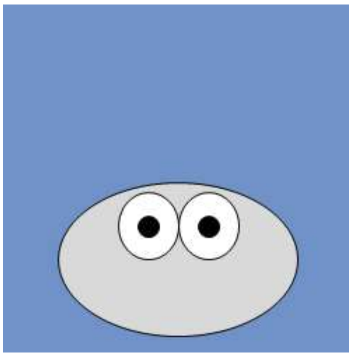
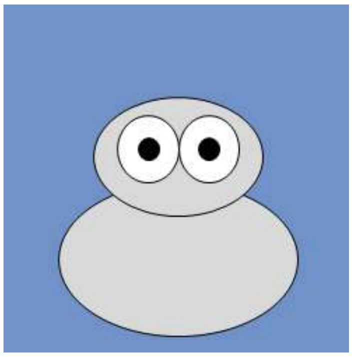
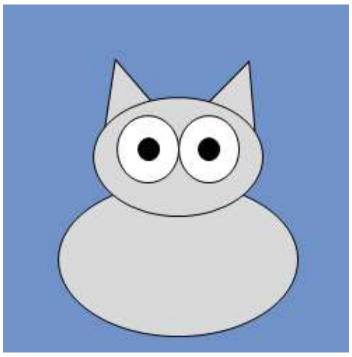
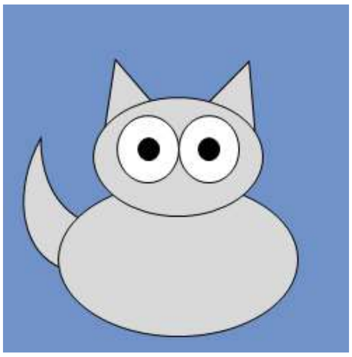
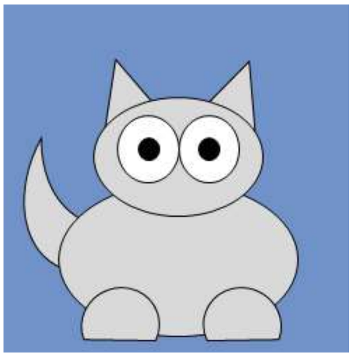
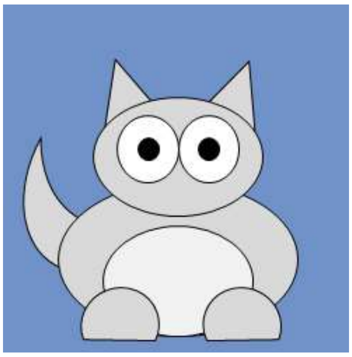
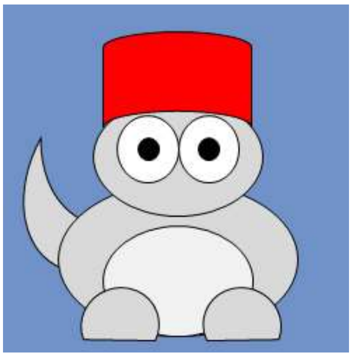
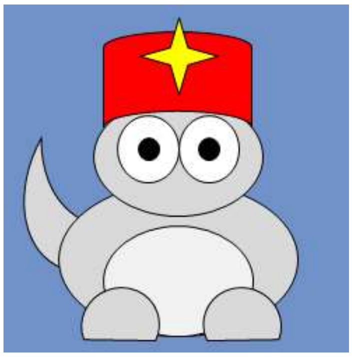
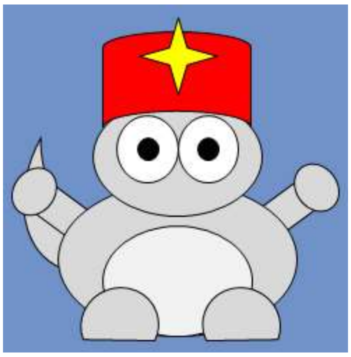
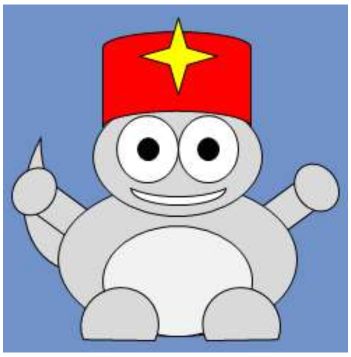
McBain, W. N. (1970). Arousal, monotony, and accidents in line driving. Journal of Applied Psychology, 54(6), 509.

SWOV (2012), SWOV Fact sheet. December. Retrieved from https://www.swov.nl/rapport/Factsheets/NL/Factsheet_Volgtijd.pdf

Thiffault, P., & Bergeron, J. (2003). Monotony of road environment and driver fatigue: a simulator study. Accident Analysis & Prevention, 35(3), 381-391.

Van Der Laan, J. D., Heino, A., & De Waard, D. (1997). A simple procedure for the assessment of acceptance of advanced transport telematics. Transportation Research Part C: Emerging Technologies, 5(1), 1-10.

Annex A: Game screens from UK Trial

0 	**1** 	**2** 
3 	**4** 	**5** 
6 	**7** 	**8** 
9 	**10** 	

Annex B: Game training script used in UK Trial

In one of the two drives you will be doing in the simulator, you will have the chance to play a simple game using a smartphone mounted in the vehicle. You will not need to use your own phone.

At the start of the game you will be shown a simple character on the screen, like a kind or virtual pet. [Show participant the image of the start screen]. The idea of the game is to try to grow and develop this virtual pet during the drive. To do this you don't have to interact with the phone in any way, you do it by the way that you drive.

If you drive well, your pet will advance – growing and developing new features. If you drive poorly, your pet will regress back to its more basic stages.

Good driving means:

- Keeping within the posted speed limits
- Avoiding close-following (or tail-gating) of the vehicle in front
- Using the innermost lane available (for example, avoiding any middle-lane hogging)
- Avoiding any rapid or erratic lane changes
- Avoiding any closed lanes (as indicated by red x's on overhead gantries)
- Avoiding any harsh braking

Overtaking is perfectly acceptable, provided of course that you do not contravene any of the above.

Bad driving means the opposite, which is to say:

- Going above the posted speed limit
- Close-following (or tailgating) the car in front
- Middle-lane hogging if a suitable space is available in the next lane in
- Any rapid or erratic lane changes
- Driving in any closed lanes (as marked by a red x)
- Any harsh braking

The route is split into ten sections. At the end of each section your pet will do one of three things: advance, regress or stay the same, depending on your overall driving behaviour within that section. You will not be able to see the section boundaries, so you will not know exactly when any changes will occur. However, throughout the drive there will be a symbol at the bottom of the game screen indicating what action you are currently on course to achieve. The symbols work as follows:

- If you see a green arrow pointing upwards, this means you are currently averaging a positive score and your pet will go up a level at the end of the section (provided you carry on driving well) [show green arrow screen]
- If you see a red arrow pointing downwards, this means you are currently averaging a negative score and your pet will go down a level at the end of the section (unless you change your driving behaviours) [show red arrow screen]
- If you see a white horizontal dash, this means you are somewhere in the middle and are currently projected to stay on the same level at the end of the section. [show white dash screen]

In addition to the visual indicators, you will also hear a sound whenever your pet either goes up or down a level. If it goes up you will hear a 'Ding' sound, and if it goes down you will hear a 'NaNaNa' sound.

Some driving behaviours will act as a sort of red flag and mean that you will automatically go down a level at the end of the section. These are:

- Going more than 10 mph over the speed limit
- Harsh braking
- A rapid or erratic lane change

At the start of the next section the red flag is removed and your general driving will be scored again.

Theoretically it is possible to reach level 10 of the game by the end of your drive.

During both drives we would like you to drive as closely as possible to how you would in the real world. How you interact with the game is up to you.

Do you have any questions?
To help check your understanding, please give me a quick explanation of how the game works, in your own words.

Annex C: Questionnaire used in UK Trial

Section 1 – Feedback on simulator drives

In this section we'd like to gather your views and opinions on your experiences in the simulator today.

Drive 1

1. Was your first drive (not including the familiarisation drive) with or without the in-car game? [please circle]

   With / Without

2. The following table contains a number of scales with opposing descriptions. For each scale, please mark on the line the point that best describes your experience of the **first** drive.

Stressful		Relaxing
Easy		Difficult
Enjoyable		Disagreeable
Predictable		Unpredictable
Confusing		Clear

Drive 2

3. Was your second drive (not including the familiarisation drive) with or without the in-car game? [please circle]

   With / Without

4. The following table contains a number of scales with opposing descriptions. For each scale, please mark on the line the point that best describes your experience of the **second** drive.

Stressful		Relaxing
Easy		Difficult
Enjoyable		Disagreeable
Predictable		Unpredictable
Confusing		Clear

Section 2 - Feedback on the in-car game itself

1. Please rate your enjoyment of the game out of 10 (10 = extremely enjoyable, 0 = extremely unenjoyable).

0	1	2	3	4	5	6	7	8	9	10

2. Please rate how satisfying you found the game play to be out of 10 (10 = extremely satisfying, 0 = extremely unsatisfying).

0	1	2	3	4	5	6	7	8	9	10

3. Please mark on the line below the extent to which you think the game had an overall positive or negative effect on your driving experience (If you think it had no effect, please mark in the middle).

Negative		Positive

4. Please mark on the line below the extent to which you think the game had an overall positive or negative effect on the quality of your driving (If you think it had no effect, please mark in the middle).

Negative		Positive

5. If such a game were available as a free app in the real world, how likely do you think you would be to play it? (10 = extremely likely, 0 = extremely unlikely)

0	1	2	3	4	5	6	7	8	9	10

6. Please provide any additional feedback you may have on the game design, including any ways you think it could be improved.

Section 3 – Driver Profile

Please fill in the following details:

Age at your last birthday	
Sex (M/F)	
Number of years driving	
Approximate annual mileage	
Primary vehicle (make and model)	
Primary mobile phone (make and model)	

Section 4 – Driving style

How often do you do the following? For each item, please mark a box in the appropriate column:

		Never	Hardly ever	Occasionally	Quite often	Frequently	Nearly all the time
1	Attempt to overtake someone that you hadn't noticed to be signalling a right turn						
2	Get into the wrong lane when approaching a roundabout or a junction						
3	Miss 'Stop' or 'Give Way' signs and narrowly avoid colliding with traffic having right of way						
4	Misread the signs and take the wrong exit from a roundabout						
5	Fail to notice that pedestrians are crossing when turning into a side street from a main road						
6	Drive especially close to the car in front as a signal to its driver to go faster or get out of the way						
7	Forget where you left your car in the car park						
8	Queuing to turn left onto a main road, you pay such close attention to the mainstream of traffic that you nearly hit the car in front						
9	Hit something when reversing that you had not previously seen						
10	Cross a junction knowing that the traffic lights have already turned against you						

		Never	Hardly ever	Occasionally	Quite often	Frequently	Nearly all the time
11	On turning left nearly hit a cyclist coming up on your inside						
12	Disregard the speed limits late at night or very early in the morning						
13	Attempt to drive away from the traffic lights in third gear						
14	Fail to check your rear-view mirror before pulling out, changing lanes, etc.						
15	Have an aversion to a particular class of road user, and indicate your hostility by whatever means you can						
16	Become impatient with a slow driver in the outer lane and overtake on the inside						
17	Underestimate the speed of an oncoming vehicle when overtaking						
18	Switch on one thing, such as the headlights, when you meant to switch on something else, such as the wipers						
19	Brake too quickly on a slippery road, or steer the wrong way in a skid						
20	Intending to drive to destination A, you ‘wake up’ to find yourself on the road to destination B, perhaps because the latter is your more usual destination						
21	Drive even though you realise you may be over the legal blood-alcohol limit						
22	Get involved in unofficial ‘races’ with other drivers						
23	Realise that you have no clear recollection of the road along which you have just been travelling						
24	Angered by another driver’s behaviour, you give chase with the intention of giving him/her a piece of your mind						

That is the end of the Questionnaire. Thank-you for your time

Annex D: Questionnaire used in NL trial

Questionnaire - Acceptance

You just made a trip in the driving simulator.

[control condition] You received information through road signs along the highway.
[speech condition] You received information through road signs along the highway and through spoken messages.
[HUD condition] You received information through the Head-Up display.

There are always 5 answer possibilities. Please tick a box on every line.

While driving I found the *information on the road signs/ through spoken messages/ through the Head-up display …*

Useful						Useless
Pleasant						Unpleasant
Bad						Good
Nice						Annoying
Effective						Superfluous
Irritating						Likeable
Assisting						Worthless
Undesirable						Desirable
Raising alertness						Sleep inducing

Questions - demographics

What is your age? ____ years old

What is your gender? M F

For how long do you have your driving license? ______ years

How many days per week do you drive on average?______ days

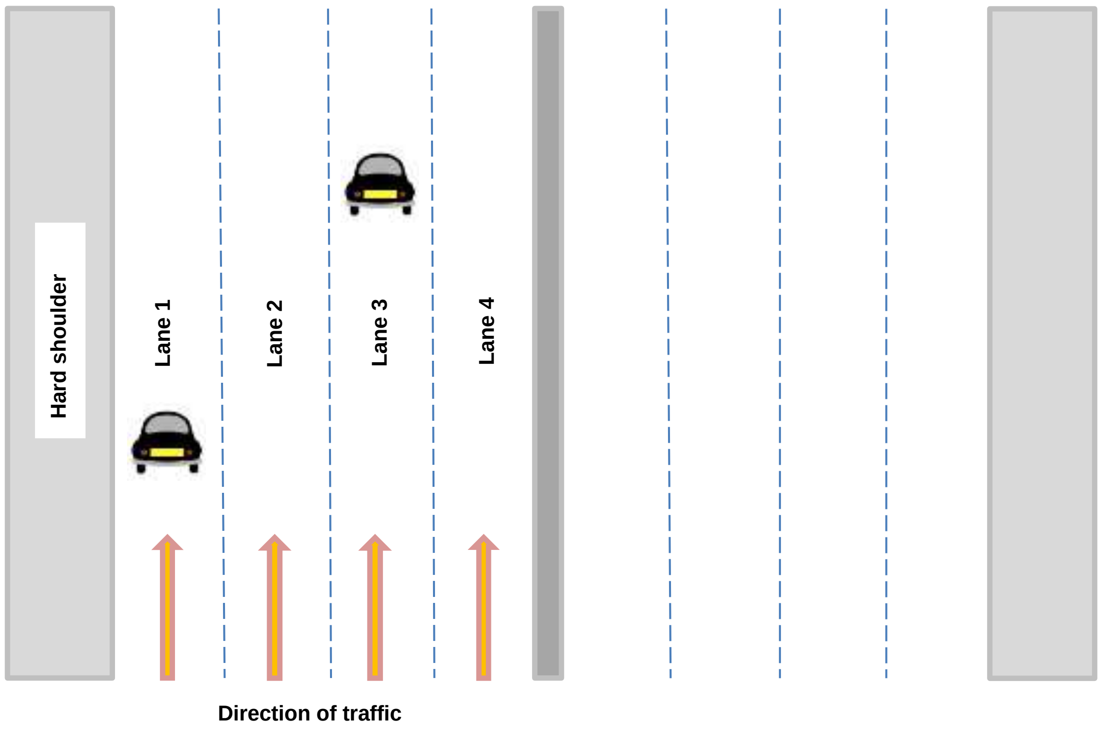
How many kilometres do you drive on average per week? ______ km

Statements – presented acoustically to the participants while driving. Please respond with yes, no, or a little.

1. I felt comfortable while driving
2. I found the presented information/ spoken messages easy to follow
3. I found the presented information/ spoken messages redundant
4. I felt concentrated while driving
5. I felt sleepy while driving
6. I felt distracted while driving
7. I found the presented information/ spoken messages irritating
8. I found the presented information/ spoken messages distracting
9. I felt stressed while driving
10. I found the presented information/ spoken messages sufficient
11. I found the presented information/ spoken messages easy to understand
12. I felt alert while driving

Annex E: Lane numbering convention

UK Study



Dutch Study