Delivery Report for

MeBeSafe

Measures for behaving safely in traffic

Deliverable Title	Results of field trials
Deliverable	D5.4
WP	WP5 Field evaluation
Task	Task 5.4 Data collection Task 5.6 Data analysis

This project (MeBeSafe) has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 723430.

Disclaimer

Copyright

Grant Agreement No.	723430
Project Start Date	01/05/2017
Project End Date	31/10/2020
Duration of the Project	42 months
Deliverable Number	D5.4
Deliverable Leader (according to GA)	Volvo Cars
WP Leader	Mikael **Ljung Aust**, Volvo Cars
Deliverable Leader(s)/ (Editor(s))	Mikael **Ljung Aust**, Volvo Cars
Dissemination Level (Confidentiality)	PU: Public
Nature	Report: PU Other: PU
Status	Final
Due Date	M39 (July 2020) -> M40 (August 2020)
Main Author(s)	Mikael **Ljung Aust**, Volvo Cars
Other Contributing Author(s) *alphabetical order by last name	Niccolò **Baldanzini**, University Firenze Bram **Bakker**, Cygnify Moritz **Berghaus**, ISAC/ RWTH Aachen Saskia **de Craen**, SWOV / Shell Marianne **Dyer**, Shell Pär **Gustavsson**, Volvo Cars Jordanka **Kovaceva**, SAFER /Chalmers Anna-Lena **Köhler**, ika/ RWTH Aachen Matin **Nabavi Niaki**, SWOV Norah **Neuhuber**, Virtual Vehicle Olaf **Op den Camp,** TNO Alberto **Perticone**, University Firenze Elizabeth **Uduwa-Vidanalage**, Shell Vincent **de Waal**, Heijmans Pontus **Wallgren**, SAFER/ Chalmers

	Marijke **van Weperen** TNO Anders af **Wåhlberg**, Cranfield University
Reviewer(s)	MariAnne **Karlsson**, SAFER/ Chalmers Olaf **Op den Camp,** TNO Anna-Lena **Köhler**, ika/ RWTH Aachen Stefan **Ladwig**, ika/ RWTH Aachen Saskia **de Craen**, SWOV / Shell
Formal review and upload by Coordinator	Institute for Automotive Engineering (ika) – RWTH Aachen University

Please refer to this deliverable as follows:

Ljung Aust, M., et al. (2020). Results of field trials (Deliverable 5.4). Retrieved from MeBeSafe website: https://www.mebesafe.eu/results/.

Abstract

The main objective of WP5 is to run a set of field trials with naïve users (i.e. not experts involved in the development of the measures) for all nudging and coaching measures developed in WP2-4. Field trials with naïve users are necessary in order to validate the estimated effectiveness of each measure.

The field trials were set up in as realistic settings as possible, given the possibilities to implement/distribute each measure. This deliverable gives a short description of the field trial setup for each measure, and then reports the effects of the nudge on road user behaviour.

Version	Date	Comment
1	7 May 2020	First Draft
2	15 June 2020	Input from Field Trials
2	25 June - 02 July	I Review loop within Steering Committee
3	02–12 July 2020	II Review loop within Steering Committee
4	12 – 25 July 2020	III Review loop within Steering Committee
5	31 July 2020	Second Draft
6	03–17 Aug 2020	Loop Volvo Cars & ika/ RWTH
7	17 Aug 2020	Document ready for formal review
8	17-24 Aug 2020	Loop ika/ RWTH
9	24-28 Aug 2020	Formal review ika/ RWTH
10	15 Oct 2020	Correction
11	29 Oct 2020	Deliverable released for revision by the EC on request of consortium
12	30 Oct 2020	Final version after release for revision by the EC

table of document history

Table of Contents

List of Figures 5

List of Tables 8

Acronyms 9

1 Executive Summary 10

2 Contribution by each Partner 14

3 Introduction 15

4 Field trial results for O1: Driver alertness feedback (Volvo Cars) 16

4.1 Field trial setup 16

4.1.1 Participants 16

4.1.2 Materials, procedure and test design 16

4.1.3 Nudging measure 17

4.1.4 Data collection 18

4.1.5 Dependent variables 18

4.2 Results 19

5 Field trial results for O2: Usage of safety ADAS to prevent close following (Volvo Cars) 20

5.1 Field trial setup 20

5.1.1 Participants 20

5.1.2 Materials, procedure and test design 20

5.1.3 Nudging measures 21

5.1.4 Data collection 23

5.1.5 Dependent variables 24

5.2 Results on Ambient display nudge 24

5.3 Results on the Competitive Leaderboard nudge 25

5.4 Comparing the effects of Ambient Display and Competitive Leaderboard nudges 26

6 Field trial results for O3: Attention to potential hazards in intersections (TNO) 28

6.1 Field trial setup 28

6.2 Results 30

7 Field trial results for O4: Behavioural change through online private driver coaching (Volvo Cars) 35

7.1 Field trial setup 35

7.2 Field trial cancelled 36

8 Field trial results for O5: HGV driver behavioural change through online coaching (Shell) 38

8.1 Field trial setup 39

8.1.1 Groups 39

8.1.2 Timeline 39

8.1.3 Data processing and storage 40

8.1.4 Intervention 40

8.1.5 Dependent variables 41

8.1.6 Analysis 41

8.2 Results 42

8.2.1 Number of drivers and trips 42

8.2.2 Driver behaviour change 42

8.2.3 Covid-19 43

8.2.4 Summary and outlook 44

9 Field trial results for O6 and O7: Safe speed/trajectory on inter-urban roads (ika/ RWTH Aachen)....46

9.1 Field trial set-up....46

9.2 Results....49

9.2.1 Overall Results....49

9.2.2 Behavioural Results of Fastest Drivers....51

9.2.3 Results of On-Site Survey....54

9.2.4 Results of Online Resident Survey....55

9.2.5 Results of PTW Analysis....55

10 Field trial results for O8: Cyclists' speed reduction – Sweden (SAFER/ Chalmers University)....61

10.1 Field trial setup....61

10.1.1 Description of the test sites....61

10.1.2 Installed equipment....63

10.2 Results....64

10.2.1 Measurements of effect on speed....64

10.2.2 On site interviews with cyclists....67

10.2.3 Instrumented bikes study....68

11 Field trial results for O8: Cyclists' speed reduction – Netherlands (TNO/SWOV) 69

11.1 Field trial setup....69

11.2 Nudge design....69

11.3 Location description....69

11.4 Video data collection and analysis....71

11.5 Results 73

12 Summary and conclusions 75

References 76

List of Figures

Figure 4.1: Screen shots from the Driver Alert Control Nudge app 18

Figure 5.1: Screen shots from Ambient Display nudging concept 22

Figure 5.2: Screen shots from Competitive Leaderboard nudging concept 23

Figure 5.3: The relative change in ACC usage between baseline and treatment for each driver participating in the field trial when exposed to the Ambient Display nudge. Each vertical bar represents one driver. 25

Figure 5.4: The relative change in ACC usage between baseline and treatment for each driver participating in the field trial when exposed to the Leaderboard nudge. Each vertical bar represents one driver. 26

Figure 5.5: Relative change in ACC usage for all drivers under the Ambient Display Nudge and the Leaderboard nudge respectively. 27

Figure 6.1: Different stages of the HMI escalation. In 30 km/h zones, the small green figure is shown from 50 m before the intersection. The figure evolves into a larger orange cross (30 m before the intersection) to the largest red cross at the intersection. The indent into the cross denotes the direction from which the potential hazard might be expected, in this case from the right. 28

Figure 6.2: The HMI integrated as a HUD in between the steering wheel and the windshield. 29

Figure 6.3: Results regarding the gaze in the correct direction of the potential hazard at different distances from the intersection aggregated for all participants and all intersections. In red, the percentage of time is given without HMI activated, in green, the results with nudging HMI...31

Figure 6.4. Velocity distributions aggregated for all drivers and intersections in 30 km/h zone as a function of the distance to the intersection...33

Figure 6.5: Velocity distributions aggregated for all drivers and intersections in 50 km/h zone as a function of the distance to the intersection....33

Figure 9.1: On the left: Set-up of the field test with thermal cameras and roadside marking lights. The beginning of the lights is the beginning of the exit lane. On the right: Nudging threshold based on speed over the course of the exit lane....47

Figure 9.2: On the left: Ratio of vehicles faster than 85 % quantile of speeds (V85) in baseline scenario. On the right: Mean speed of vehicles that fulfil nudging criteria at x = 50....51

Figure 9.3: Results for the velocity development of fast drivers (drivers faster than 87.01 km/h at x = 50) of testing phase 1. The x-axis shows the within-factor position, the y-axis shows the velocity in km/h, and the lines represent the different scenarios (between-subjects factor) of testing phase 1. Error bars depict standard deviations.... 53

Figure 9.4: On the left side: mean peak velocity below each camera for drivers above Q3 quartile; Blank x-axis due to different peak velocity positions under each camera. On the right side: instant mean velocity for drivers above Q3 quartile at x = [50, 140, 230].... 58

Figure 9.5: On the left side, trajectories under each camera, superimposed on real map. On the right side, comparison on lateral displacement for bikers above Q3 threshold.... 59

Figure 10.1: The nudge at site 1.... 62

Figure 10.2: Site 1 - Nobelplatsen (57°42'49.6"N 12°00'23.4"E). The bike lane is unidirectional, and the bikes comes from the low right corner of the picture moving towards the top left.... 62

Figure 10.3: Site 2, note the Viscando Otis measuring equipment mounted on the light pole nearest in the picture.... 63

Figure 10.4: Site 2 - The Götaälvbron site (57°43'12.4"N 11°57'45.5"E), the bike lane is the lane most to the right in picture. ... 63

Figure 10.5: The equipment installed at one of the locations (Götaälvbron). ... 64

Figure 10.6: Distribution of cyclist speed at different positions and conditions for location 1 (left) and location 2 (right). The numbers on the top report the sample size for each boxplot. ... 65

Figure 11.1: Schematics of the transverse line nudge reducing in distance ... 69

Figure 11.2: Location of study indicating the cyclists traveling through the underpass that will turn right (yellow arrow), the through cyclists (blue arrow), and the location of the nudge (orange) ... 70

Figure 11.3: Implemented nudge along the through traveling cycling facility, showing two of the three implemented cameras (red boxes) ... 71

Figure 11.4: Camera view 1 indicating the through traveling cyclists where the nudge will be implemented ... 72

Figure 11.5: Camera view 2 at the intersection ... 72

Figure 11.6: Camera view 3 showing the intersection from an angle where the right turning cyclist approach is more visible ... 73

List of Tables

Table 6.1: Participants level of (gaze) attention towards the potential hazard as a result of the HMI activation (for 18 participants, the analysis of gaze direction is available). 32

Table 6.2: Number of participants that decrease speed, increase speed or do not change speed as a result of the HMI escalation, compared to the baseline. 34

Table 8.1: Mean values of smoothness, harsh braking and acceleration events, before and after coaching started for the drivers who passed beyond the onboarding stage. Lower values indicate better driving 43

Table 8.2: Mean values of smoothness, harsh braking and acceleration events, compared for drivers who did and did not pass beyond onboarding. Cohen's d values were computed for the differences 43

Table 9.1: Scenarios of testing phase 1. 48

Table 9.2: ANOVAs for PTWs above Q3 60

Table 10.1: Field experiment design. 64

Table 10.2: Summary of ANOVA results for site 1 and site 2. 66

Acronyms

ACC	Adaptive Cruise Control
ADAS	Advanced Driver Assistance Systems
DAC	Deriver-Alert Control
HGV	Heavy Goods Vehicles
HMI	Human Machine Interface
HUD	Head-up Display

1 Executive Summary

The main objective of WP5 is to run a set of field trials with naïve users (i.e. not experts involved in the development of the measures) for all nudging and coaching measures developed in WP2-4. This deliverable describes the first results of the Field Trials that were run to empirically establish the impact of each measure.

All these activities have taken place in Tasks 5.4 (Data collection) and 5.6 (Data analysis). The deliverable gives a first overview over the overall results. More detailed results and descriptions will be given in deliverable D5.5.

For Objective 1 - driver alertness feedback, a fleet of N = 49 drivers were provided with an additional incentive to stop and take a break (a gift card type of reward) when the Driver Alert Control system indicated that a break would be beneficial (i.e. when a high level of drowsiness had been detected). The offer was displayed on an additionally installed in-vehicle screen whenever Driver Alert Control would trigger.

The results showed a clear positive effect on driver behaviour. The number of drivers who stopped within 20 minutes after getting a drowsiness warning doubled in the treatment phase, i.e. when offered an incentive to stop in combination with the drowsiness warning. For drivers who received DAC warnings in both baseline and treatment, the average stopping time after receiving the warning was reduced with 8 minutes in the treatment phase.

For Objective 2 - usage of safety ADAS to prevent close following, a fleet of N = 49 drivers were provided with nudging that consisted of different types of visual in-vehicle feedback on the extent to which they were using Adaptive Cruise Control (ACC) while driving. Two types of visual feedback were tried: A) Ambient Display concept and B) Competitive Leaderboard concept.

Both concepts had significant effects on driver behaviour. For the ambient display nudge, the average ACC use was 14.24 % in baseline and 20.82 % in treatment. This means that drivers on average increased their ACC usage level with about 46 % when being nudged with the Ambient Display concept. For the Competitive Leaderboard nudge, the average ACC use was 14.48 % in baseline and 30.67 % in treatment. Drivers thus on average increased their baseline ACC usage level with 118 % when being nudged with the Competitive Leaderboard concept.

For Objective 3 - Attention to potential hazards (i.e. to improve timely attention to a potential hazard in intersections), the field trial involved a total of *N* = 22 naïve drivers who twice drove a prescribed 1-hour route through central Eindhoven (NL). Each driver received a nudge at unsignalized intersections, to direct their attention towards areas of the intersection where view obstructions would hide a possibly approaching bicyclist.

With the nudging HMI to direct driver attention, drivers spent on average 20% more time looking in the direction of a potential hazard at a distance of 20-30 m before entering the intersection. Out of *n* = 18 participants, *n* = 10 increased their gaze in the direction of the possible hazard when the HMI was activated. Additionally, *n* = 13 and 14 out of *N* = 22 participants decrease their speed while approaching an intersection in respectively the 30 km/h and 50 km/h zone.

For Objective 4 - behavioural change through online private driver coaching, it was determined that ACC oriented coaching would have its largest impact not on drivers who are already using ACC, but rather on drivers who do not use ACC at all. Since nudging toward increased ACC usage only can be applied on drivers who already use the function, non-users must first become users before nudging can be applied.

A key assumption for the field trial setup, based on previously collected driving data, was that 20 to 30 % of the ACC nudging field trial participants (see Objective 2) would be determined non-ACC users, who would not respond to the ACC nudging concepts. These non-users would therefore provide great test group for coaching.

As it turned out, this assumption did not hold. Data from the ACC nudging concept trial showed that all participants were nudged into some level of ACC usage by the nudging concepts. While a positive finding in the sense that the Objective 2 field trial was more successful than expected, this also meant that there was no-one left to coach for the Objective 4 field trial. The latter therefore had to be cancelled.

For Objective 5 - HGV driver behavioural change through online coaching, two fleets of company drivers were recruited, one in Norway and one in the UK. However, due to delays in the development of the coaching app, the field trial start was delayed until late February 2020. Unfortunately, this placed the start of the field trial right at the start of the corona pandemic. Over the course of the world wide spread of the pandemic it has severely affected both the companies involved in the field trial and the traffic environment in which they normally drive. This places restrictions upon the possible interpretations of the field trial outcome. While the data indicates that the app was well received and used by the drivers, and also that peer-to-peer coaching was a viable approach, today it is not possible to conclude whether coaching does change HGV driver behaviour or not.

For Objectives 06 and 07 - Safe speed/trajectory on inter-urban roads, the field trial took place on an exit lane in Eindhoven, Netherlands, where roadside marking lights were installed in such a way that drivers who entered the exit lane at speeds above a predefined threshold could be exposed to systematically varying light patterns along the lane.

The results indicate that vehicles do slow down significantly when being nudged by the light patterns. This is especially true for drivers exceeding 87 km/h at stimulus onset at x = 50 encountering static lights, while lights moving towards the driver revealed ambiguous results and need further investigation (see D5.5). For Powered Two-Wheelers however, no systematic effect of the nudge could be found in the data.

For Objective 8 - Cyclists' speed reduction the field trials involved a random sample of cyclists passing two test sites implemented in Gothenburg, Sweden, and another random sample of cyclists who passed a test site implemented in Eindhoven, the Netherlands.

In both instances, passing cyclists were visually nudged by transverse lines on the bicycle lane that got closer to each other as the distance to the respective intersection decreased.

Both trials showed positive effects on cyclist behaviour. In the Gothenburg trial, 9-17% more cyclists reduced their speed in treatment depending on location and other factors. In the Eindhoven trial, cyclist speeds were reduced, and deceleration rates were also higher during treatment.

2 Contribution by each Partner

This deliverable was written by Mikael Ljung Aust, with contributions by Marijke van Weperen (chapter 6, 11), Olaf Op den Camp (chapter 6, 11), Pontus Wallgren (chapter 10), Jordanka Kovaceva (chapter 10), Anders af Wåhlberg (chapter 8), Bram Bakker (chapter 8), Saskia de Craen (chapter 5), Marianne Dyer (chapter 8), Matin Nabavi Niaki (chapter 8), Norah Neuhuber (chapter 8), Elizabeth Uduwa-Vidanalage (chapter 8), Anna-Lena Köhler (chapter 9), Moritz Berghaus (chapter 9), Vincent de Waal (chapter 9), Niccolò Baldanzini (chapter 9.2.5), and Alberto Perticone (chapter 9.2.5).

All other partners in WP5 contributed by giving their feedback on this deliverable. All partners have fulfilled their tasks in time and with satisfactory quality.

3 Introduction

The objective of tasks 5.4 (data collection) and 5.6 (data analysis) is to empirically ascertain the impact on road user behaviour of each nudging and coaching measure developed in WP2-4. This report describes the field trial implementations, and the resulting findings regarding behavioural impact of the measures on road users, on a per measure basis.

Typically, a baseline data set (no measure in place) is compared to a treatment data set (measure in place and active), and significant differences between baseline and treatment is what is being reported in the following. Assuming adequate consideration of potential confounding factors (i.e. factors outside the nudge that may have influenced field trial outcomes), these differences represent the safety effects of the nudges applied, and thus form the basis for the EU wide safety impact calculations for the measures that are reported in D5.5.

4 Field trial results for O1: Driver alertness feedback (Volvo Cars)

For driver alertness feedback, the nudging concept consists of providing the driver with an incentive to stop and take a break when the Driver Alert system indicates that a break would be beneficial (i.e. when a high level of drowsiness has been detected). The details and results of this field trial are described below.

4.1 Field trial setup

4.1.1 Participants

The field trial test fleet had *N* = 49 participants. All were Volvo Cars employees driving Volvo XC60 MY 2020 company cars. A gender- and aged-balanced test population was targeted in recruitment. The final fleet consisted of n = 26 women and 23 men in the age span 39 - 62 (*M* = 50.4, *SD* = 6.07). Driving experience ranged from 20 to 44 years (*M* = 32.0, *SD* = 6.25).

Note that these are the same participants that were nudged to increase their usage of ACC more, as described in chapter 5 below. While it was not predicted that there would be any interaction effects between the two nudging types (i.e. receiving an incentive to stop when drowsy versus being nudged to engage ACC more often) as they address very different events and mechanisms, this still deserved to be explicitly mentioned.

4.1.2 Materials, procedure and test design

All participants were given the same written information stating that the purpose of the test was to examine a new platform for driving feedback. Participants were informed that their company car would be fitted with an additional screen (in the form of an iPhone 6 or 7 where they would receive visual feedback related to their driving and use of vehicle systems. Participants were asked to keep the phone “alive” and visible at all times when driving and report any problems that occurred.

Note that DAC was not explicitly mentioned in the information drivers received, to avoid influencing driver behavior in any way except by the nudge itself. The field trial used a within-group design, i.e. baseline and treatment data were collected from the same 49 cars in a sequential manner.

4.1.3 Nudging measure

In order to nudge drowsy drivers to take a break by providing an incentive to do so, a *Driver Alert Control nudge app* was implemented. When a driver received a Driver Alert Control warning from the car, the app informed the driver that s/he would receive a surprise gift if s/he took a break within 20 minutes. A timer then started to count down. If the driver did not stop within that time, the app would tell them that they missed their chance to receive a gift.

If the driver did stop, the app would reveal what the surprise gift was and inform that it would be delivered by email. The gifts were vouchers valid in different online/physical stores, restaurants or recreational attractions, at values between 30 and 90 €. A driver could not receive more than one voucher per 24 hours. Furthermore, a driver could not get the same voucher type twice. The different views of the app are shown below.

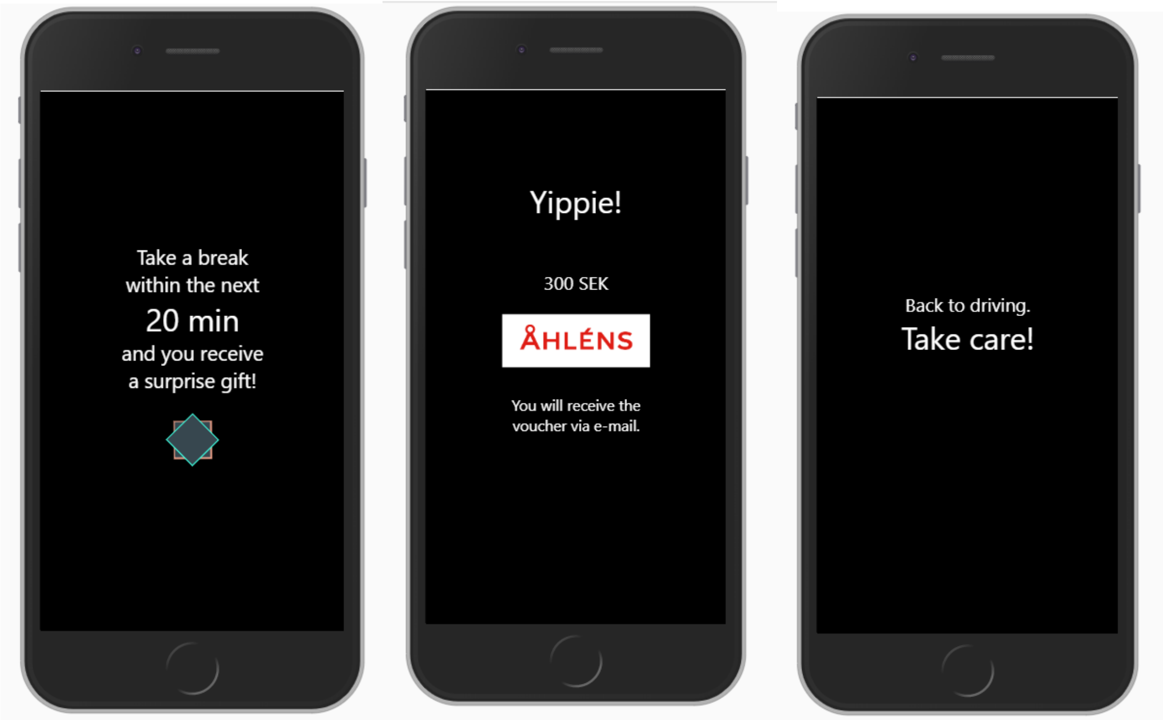

Figure 4.1: Screen shots from the Driver Alert Control Nudge app

4.1.4 Data collection

The participants' cars were equipped with remote data acquisition units set up to record vehicle data including engine status, vehicle speed, DAC status and standalone GPS-data. Data recording was triggered at every engine start and continued until engine shutdown. The general baseline data collection lasted from October 2019 to mid-April 2020. Treatment data was collected from mid-April to August 2020.

4.1.5 Dependent variables

DAC stopping time was calculated for each trip by calculating the duration from DAC warning status until the vehicle was at standstill (i.e. vehicle speed = 0). For trips including several DAC warnings, the first warning was used as the time reference point.

4.2 Results

During baseline, *N* = 23 drivers received at least one DAC warning and there was a total of 59 trips which included at least one DAC warning. In 44 % of the trips where a driver received a warning they stopped within 20 minutes. During treatment, *N* = 11 drivers received at least one DAC warning and there was a total of 15 trips which included at least one DAC warning and an incentive offer to the driver. In 87 % of the trips where a driver received a warning and the incentive was offered, the drivers stopped within 20 minutes and received the incentive. In other words, the proportion of drivers who stopped wihin 20 minutes of a DAC warning almost doubled when drivers were offered an additional incentive.

Looking at a within driver comparison, there were *N* = 9 drivers who received at least one DAC warning in both baseline and treatment. For these *N* = 9 drivers, their stopping time was on average reduced with 8 minutes. The highest average decrease in stopping time between baseline and treatment was 32 minutes.

5 Field trial results for O2: Usage of safety ADAS to prevent close following (Volvo Cars)

5.1 Field trial setup

For usage of safety ADAS to prevent close following, the nudging consisted of providing the drivers with different types of visual feedback on the extent to which they were using Adaptive Cruise Control (ACC) while driving. The details and results of this field trial are described below.

5.1.1 Participants

All *N* = 49 participants were Volvo Cars employees driving Volvo XC60 MY 2020 company cars. To the extent possible, a gender- and aged-balanced test population was targeted. The number of females were 26 and males 23 aged between 39 and 62 (*M* = 50.4, *SD* = 6.07). Driving experience ranged from 20 to 44 years (*M* = 32.0, *SD* = 6.25).

Note that these are the same participants that were given an incentive to stop if receiving a DAC warning, as described in chapter 4 above. While it was not predicted that there would be any interaction effects between the two nudging types as they address very different events and mechanisms (i.e. receiving an incentive to stop when drowsy versus being nudged to engage ACC more often) this still deserved to be explicitly mentioned.

5.1.2 Materials, procedure and test design

All test participants were given the same written participant information stating that the purpose of the test was to examine a new platform for driving feedback and driver behaviour. The participants were informed that their company car would be fitted with an additional screen (in the form of an iPhone 6 or 7) to which software would be remotely downloaded that would give them visual feedback related to their driving and their use of the car's systems. Furthermore, participants were asked to keep the phone alive and visible at all times when driving and report any problems that occurred.

The terms *Adaptive Cruise Control* (ACC) or *Driver Alert Control* (DAC) were not explicitly stated anywhere in the information the drivers received, in order to avoid influencing drivers in other ways than through the app design.

The field trial used a within-group design, where baseline and treatment (i.e. driving with the app) data were collected from the same 49 cars. One driver participated in baseline and the ambient design treatment but not in the competitive leaderboard treatment.

5.1.3 Nudging measures

ACC nudging was tried in two different versions, one using an ambient design concept and the other a Competitive Leaderboard concept. The *ACC ambient design nudge* was designed based on the assumption that many humans prefer order over chaos in their lives. Thus, the design aimed to nudge drivers into using ACC by continuously transforming the visuals from a chaotic to an orderly pattern, with the transformation continuing as long as they drove with ACC engaged. The concept provided drivers with a daily goal of 10 minutes of ACC use. The app informed the drivers whether ACC was available or not, as well as indicating if the function was active or inactive.

The structure of the ambient display nudge is as follows: the start view of the app (when the engine is turned off) shows a yellow ACC symbol and the text "Adaptive Cruise Control" and "Not started". When the engine is turned on, the symbol changes to grey and the text changes to "Not available" as long as vehicle speed is below 15 km/h. When ACC is available (speed exceeds 15 km/h), ten grey dots start to move with random speeds on the screen and the text states "Available".

When the driver activates ACC, the dots lower their speeds and turn white while the text changes to "Active". For each minute of ACC driving one dot will move into the centre of the screen, turn yellow and slowly circulate. If the driver temporarily deactivate ACC, the yellow dots will stay in the centre while the others will behave as they did before ACC was activated and the text states "Paused". When the drivers have driven with ACC for ten minutes all the dots will centre and slowly circulate together in what is perceived as harmony and the text "Goal reached" is shown. Following 10 minutes of ACC driving the only visual difference is the colour of the ACC symbol (yellow when ACC is active and grey otherwise). The corresponding screen views of the app are shown below in Figure 5.1.

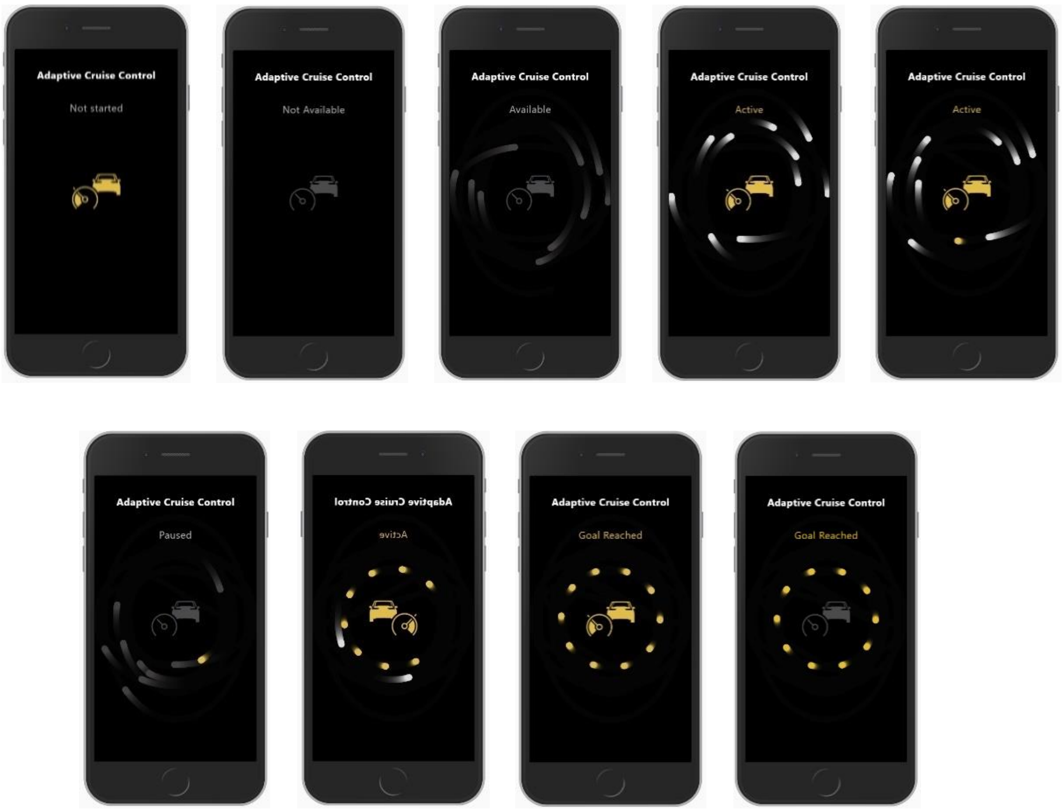

Figure 5.1: Screen shots from Ambient Display nudging concept

The *ACC Competitive Leaderboard Nudge* was designed to test another nudging approach – social comparison and competition. This app presented the drivers with a leaderboard ranking all the participants by weekly ACC minutes.

The structure of the Competitive Leaderboard Nudge is as follows: during driving the visuals show a Volvo XC60 either at standstill (when ACC is not activated) or driving (when ACC is active) and the minutes of ACC use today. Whenever vehicle speed is zero or the engine is turned off the app displays the leaderboard. The leaderboard shows your rank, your weekly ACC minutes, your daily ACC minutes and trend (position change since the leaderboard was last shown). In addition to this, the leader and his/her minutes as well as other drivers around your rank is shown. All participants were assigned a fake name that was shown in the app. Every Sunday night the leaderboard was reset, and the participants received an email with their weekly rank and ACC minutes as well as the name of the weekly winner. The different views of the Leaderboard Nudge are shown below.

Figure 5.2: Screen shots from Competitive Leaderboard nudging concept

5.1.4 Data collection

The participants' cars were equipped with data acquisition units set up to record vehicle data including engine status, vehicle speed, ACC status, DAC status as well as standalone GPS-data. Data recording was triggered at every engine start and continued until engine shutdown.

The general baseline data collection lasted between October and November 2019 and the general treatment data collection between December 2019 and July 2020. The baseline data included in the analysis below consists of N = 16,604 trips, adding up to a total of 4,342 hours of driving. The treatment data consists of *N* = 41,012 trips, with 30,189 trips (7,399 hours of driving) for the ambient ACC concept and 10,823 trips (2,529 hours of driving for the competitive ACC concept. Data files not including any driving data (i.e. only ignition on and off without any speed increase) were filtered out.

5.1.5 Dependent variables

ACC use percentage was calculated for each trip by dividing the duration of ACC status ON with the trip duration (the time between engine start and stop). Furthermore, average ACC use on an individual level was calculated by summing all ACC status ON duration and divide that by total trip duration.

To calculate the effect of the respective nudge on ACC usage, the percentage of ACC usage over total trip time was first calculated for each driver in the baseline (no nudge) and treatment (nudge active) phase, and then summed on a group level. This was done for both nudging concepts.

5.2 Results on Ambient display nudge

For the ambient display nudge, the average ACC use was 14.21 % in baseline and 20.82 % in treatment. This means that drivers on average increased their normal level of ACC usage with about 46 % when being nudged with the Ambient Display concept. A paired t-test showed that this increase was significant ($t(48) = 5.25$, $p < .001$).

In Figure 5.3 below, the difference between ACC use in baseline and when being nudged with the Ambient Display concept for all drivers are visualised, ranked from largest increase to lowest.

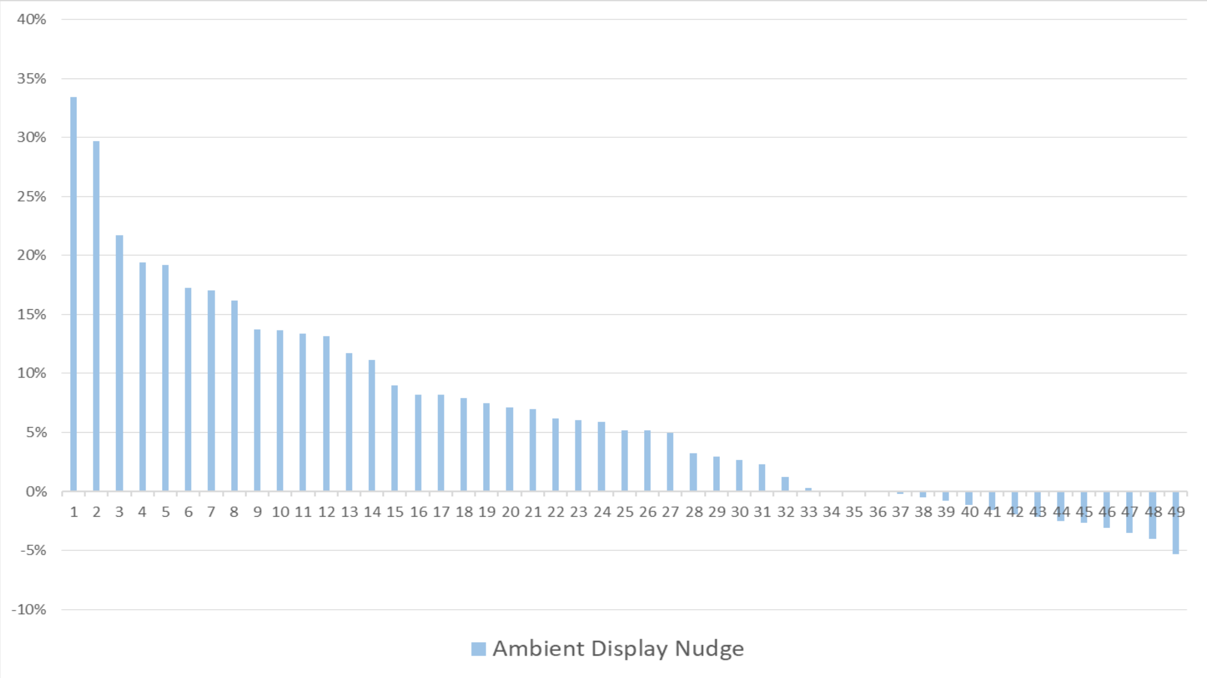

Figure 5.3: The relative change in ACC usage between baseline and treatment for each driver participating in the field trial when exposed to the Ambient Display nudge. Each vertical bar represents one driver.

As can be seen, when using the *Ambient Display nudge*, *N* = 26/49 drivers increased their ACC use by at least 5 %. However, it should also be noted that there were some drivers for whom ACC usage decreased in treatment. This illustrates the need to keep track of nudging impacts in real time if possible, so one can remove the nudge, or switch to a different paradigm, if drivers are negatively affected.

5.3 Results on the Competitive Leaderboard nudge

For the Competitive Leaderboard nudge, the average ACC use was 14.48 % in baseline and 30.67 % in treatment. This means that drivers on average increased their normal level of ACC usage with about 118 % when being nudged with the Leaderboard concept. A paired t-test showed that this increase was significant ($t(47) = 6.64$, $p < .001$).

In Figure 5.4 below, the difference between ACC use in baseline and treatment II for all drivers are visualised, ranked from largest increase to lowest.

Figure 5.4: The relative change in ACC usage between baseline and treatment for each driver participating in the field trial when exposed to the Leaderboard nudge. Each vertical bar represents one driver.

5.4 Comparing the effects of Ambient Display and Competitive Leaderboard nudges

An obvious question to ask when deploying two different nudging concepts targeting the same population is whether drivers were affected similarly or differently by the two concepts. In Figure 5.5 below, the relative change in ACC use is shown for both the Ambient Display concept and the Competitive Leader board concept, on a per driver basis. As can be seen, the answer to the question seems to be that the two different nudging concepts have affected most drivers differently, i.e. there are very few instances where the two bars per driver are of exactly the same height. Some drivers have responded better to the Ambient Display concept, but most drivers seem to have responded the best to the Competitive Leaderboard nudge. This provides interesting learnings for the future, in the sense that if one wants to create a particular type of change in a large driver population, quite a bit of experimentation will need to be applied to find the right concepts, and the final result is likely to include more than one type of nudge if the outcome is to be robust across the whole population.

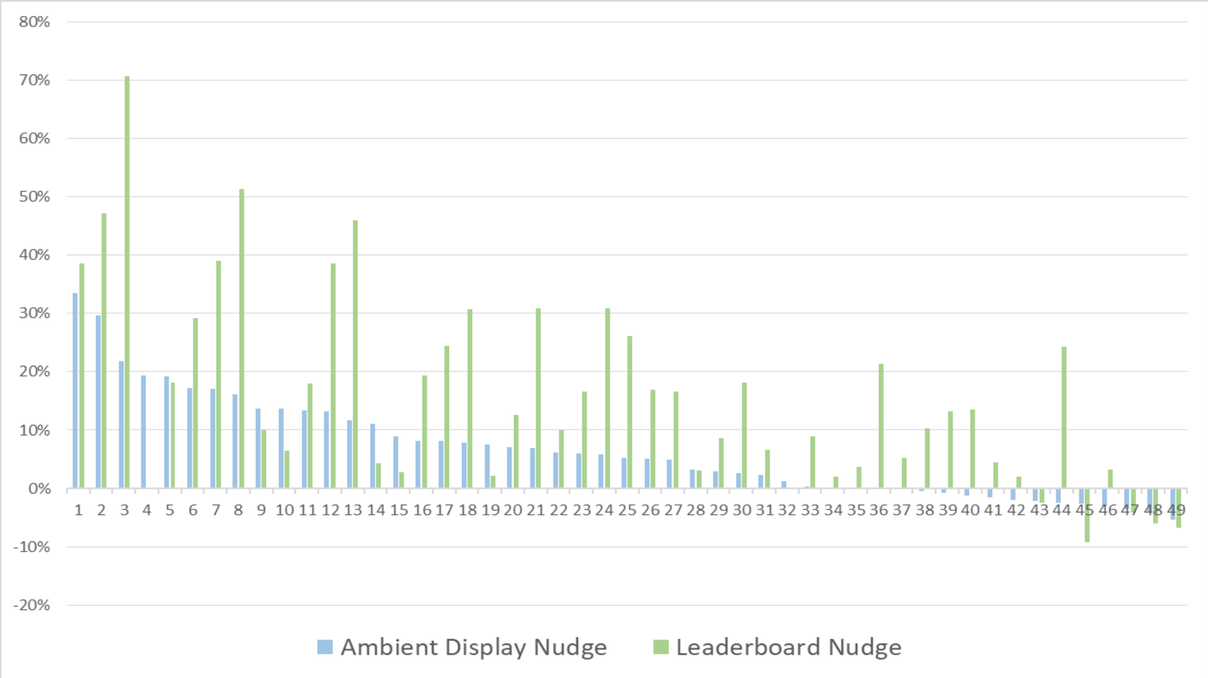

Figure 5.5: Relative change in ACC usage for all drivers under the Ambient Display Nudge and the Leaderboard nudge respectively.

6 Field trial results for O3: Attention to potential hazards in intersections (TNO)

6.1 Field trial setup

For Objective 3 - Attention to potential hazards (i.e. to improve timely attention to a potential hazard in intersections), the field trial has consisted of a driving experiment, where a total of $N = 22$ naïve drivers drove twice over a prescribed route of about 1 hour through the centre of Eindhoven (NL). Each driver received a nudge at unsignalized intersections, to direct their attention towards areas of the intersection where view obstructions would hide a possibly approaching bicyclist.

The study is conducted using a TNO laboratory vehicle (VW Jetta), equipped with:

- A nudging Human Machine Interface (HMI), developed by OFFIS in MeBeSafe, which escalates in approach of each of the qualified intersections in the test according to Figure 6.1. The HMI is integrated as a Head-up Display (HUD) in front of the driver (see Figure 6.2).

Figure 6.1: Different stages of the HMI escalation. In 30 km/h zones, the small green figure is shown from 50 m before the intersection. The figure evolves into a larger orange cross (30 m before the intersection) to the largest red cross at the intersection. The indent into the cross denotes the direction from which the potential hazard might be expected, in this case from the right.

Figure 6.2: The HMI integrated as a HUD in between the steering wheel and the windshield.

- A data collection system to measure the response of the drivers (speed of the vehicle, gaze direction of the driver derived by Cygnify from cameras directed towards the driver), the output of the HMI and the accurate position, speed and acceleration of the vehicle with time.
- Context cameras, to provide a view from the vehicle towards the environment and the local traffic situation.

Participants received short instructions before entering the car for the sensor calibration procedure and the subsequent two-hour test drive. They were informed about the objective of the experiment in general terms. The instructions included showing the HMI symbols and the escalation thereof, but only limited information about what the signs meant (i.e. that the size and colour of the cross indicate the level and direction of a potential hazard and that the arrow points in the direction from which the hazard is expected to be largest). Additional questions about the meaning of these signs were not answered, because this could influence the nudging effect, but general questions regarding the trial were answered as long as answers would not bias the subjects.

Each participant in the test has driven the same route twice, one time as a baseline without the nudging HMI activated, and one time with the HMI activated. Half of the drivers started with the baseline, the other half started with the HMI switched on, in order to minimize the bias possibly caused in the order of baseline and treatment.

The route passes 74 intersections that meet the criteria for the test (non-signalized urban crossing) of which 21 are located in a 50 km/h zone (test vehicle has priority) and 53 are located in a 30 km/h zone (priority for traffic from the right).

6.2 Results

To get insight into the effect of the nudging HMI on the direction of attention of drivers approaching an intersection, two specific metrics have been used. One is the gaze direction of the driver (gaze direction measurements were available for 18 of the 22 participants in the test). The other is the response of the driver represented in a change of speed of the vehicle.

The first metric calculates the percentage of time the driver gazes in the direction of the potential hazard. At each qualified intersection on the route, the hazard (e.g. a cyclist hidden by a view-blocking obstruction) possibly approaches from the near or the far side of the test vehicle. It is possible to determine the percentage of the total time looking into the correct[1] direction, from an analysis of the gaze direction of the driver, taking the layout of each intersection -which is known since a fixed route is driven- into account. In Figure 6.3Figure 6.3, the percentage of time the driver directs the attention in the correct direction is given at different distances from the intersection (from 40 m in 5 m intervals down to 0 m at the intersection).

[1] The “correct” direction is defined by a narrow view angle corridor as a function of the distance to the intersection. A differentiation is made between the corridor for the near side and for the far side.

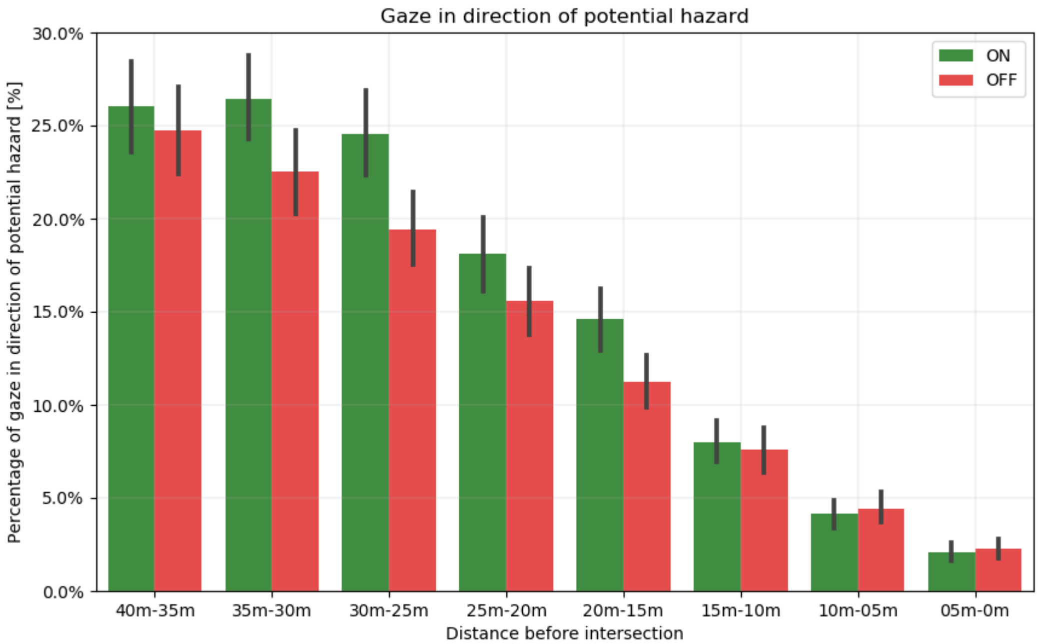

Figure 6.3: Results regarding the gaze in the correct direction of the potential hazard at different distances from the intersection aggregated for all participants and all intersections. In red, the percentage of time is given without HMI activated, in green, the results with nudging HMI.

Figure 6.3 shows the aggregated results for all test participants and all intersections, with the red bars showing the results for the baseline (no nudging) and the green bars indicate the results with activated nudging HMI. In the range of 25 to 30 m before the intersection, the nudging HMI leads to an average increase of 26 % in time that the driver looks in the direction of the potential hazard. This is interpreted as an increase in attention to the direction of the potential hazard.

Not all participants are similarly affected by the HMI. The gaze direction for the individual participants has also been studied (following the approach as shown in Figure 6.3). This leads to the following result:

Effect of HMI on gaze direction	Number of participants
More attention to potential hazard	10
No clear effect	3
Less attention to potential hazard	5

Table 6.1: Participants level of (gaze) attention towards the potential hazard as a result of the HMI activation (for 18 participants, the analysis of gaze direction is available).

From *N* = 18 participants, the HMI activation leads to 10 participants to pay more attention to the direction of the potential hazard, which represents 56% of the drivers. A percentage of 28% of the drivers pays less attention to the potential hazard, when considering gaze direction.

As a second metric, the response of drivers in the approach has been determined by analysing the change in speed during the approach of the intersections on the route, with and without HMI activated. A distinction is made between the zones with a speed limit of 30 km/h (priority for traffic from the right) and the zones with a speed limit of 50 km/h (priority for the test vehicle with the nudging measure integrated).

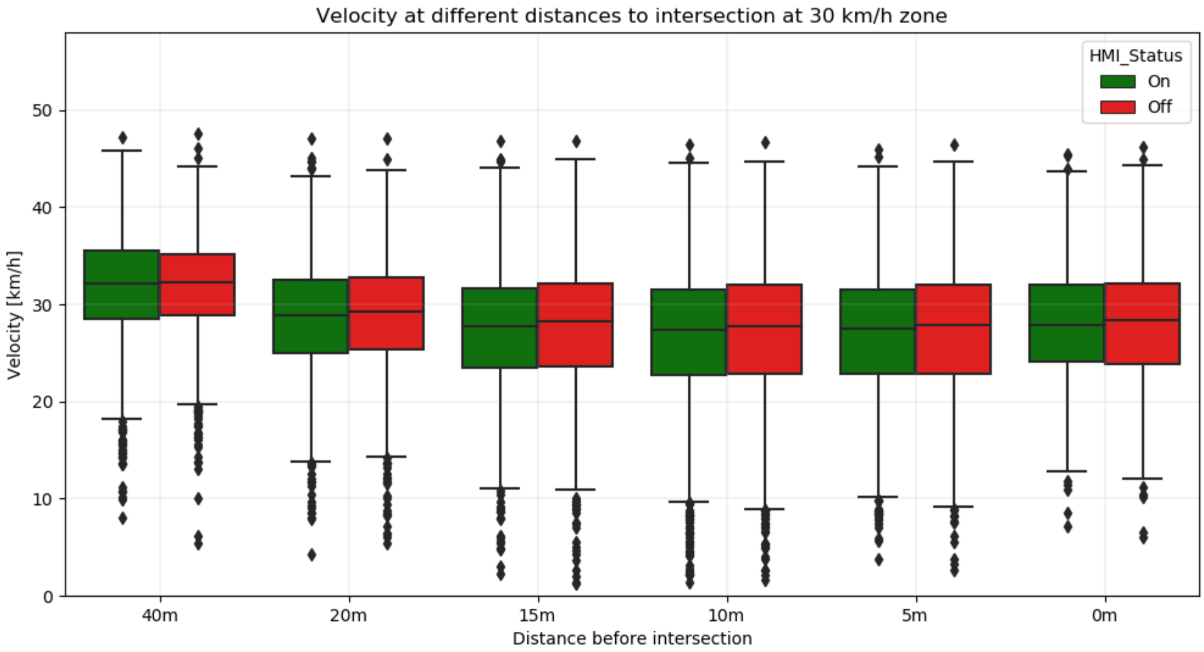

Figure 6.4. Velocity distributions aggregated for all drivers and intersections ***in 30 km/h zone*** *as a function of the distance to the intersection.*

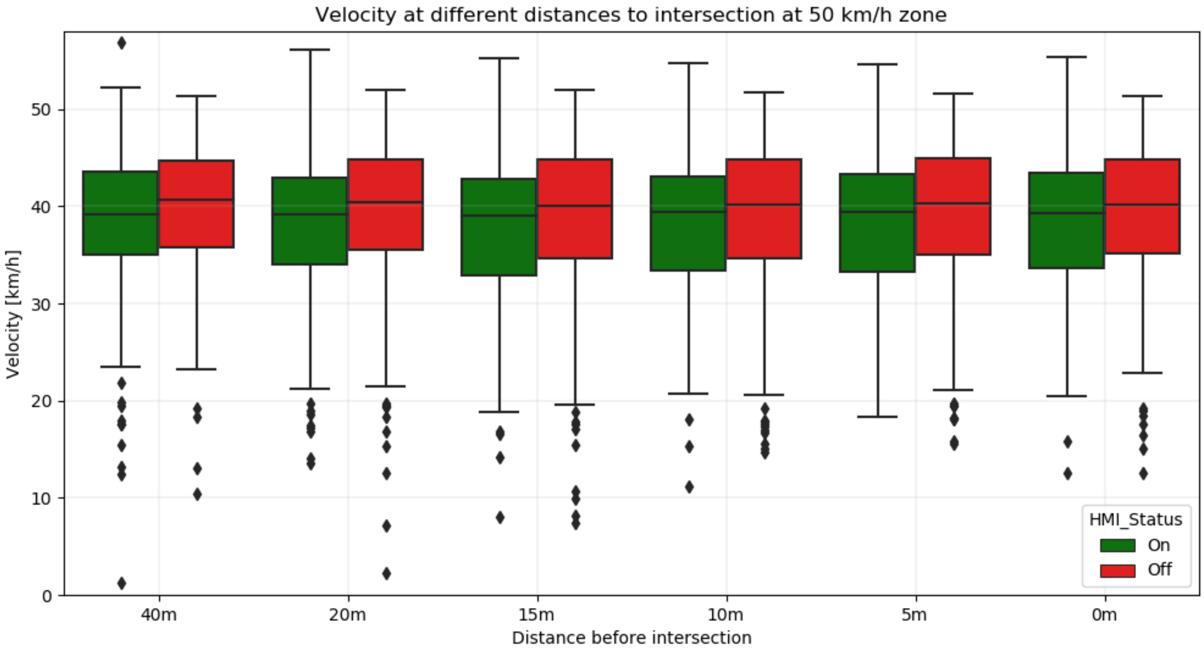

Figure 6.5: Velocity distributions aggregated for all drivers and intersections ***in 50 km/h zone*** *as a function of the distance to the intersection.*

In all cases, the speed during the approach is lower (though only slightly) for the situations in which the nudging HMI is activated compared to the baseline without HMI escalation. Note that it was not the objective of the nudging HMI to accomplish a speed reduction; the intention has always been to increase the level of attention of the drivers in an approach of an intersection with a potential hazard appearing from behind a view-blocking obstruction. A speed reduction might be interpreted though as an increased level of attention. This does not mean that drivers that increase speed have a decreased level of attention.

Not all participants are similarly affected by the HMI. The change in speed for the individual participants in approach of the intersections has also been studied. The results are summarized in Table 6.2.

	Number of participants	
Effect of HMI on speed	**30 km/h**	**50 km/h**
Decrease speed	13	14
No clear effect on speed	6	1
Increase speed	3	7

Table 6.2: Number of participants that decrease speed, increase speed or do not change speed as a result of the HMI escalation, compared to the baseline.

Though the HMI seems to effectively nudge the majority of the drivers to (slightly) reduce speed, the influence on the gaze direction by the nudging HMI is considered a stronger indicator for the conclusion that the proposed in-vehicle nudge is effective in increasing the level of attention of drivers in an approach of an intersection.

7 Field trial results for O4: Behavioural change through online private driver coaching (Volvo Cars)

7.1 Field trial setup

As previously detailed in WP4 (see Deliverable 4.3), when it comes to ACC usage, the overall picture of ACC usage levels indicated that ACC users could be grouped into three types; the intensive users, the modest users and the non-users, where the last group does not use ACC at all. Furthermore, a clear difference in mind-set could be identified between users and non-users. Both the intensive and modest users were well aware of how ACC operates and comfortable with using it while driving. Drivers in the non-user group on the other hand were afraid of activating ACC, because they did not trust it to be capable of actually regulating speed and the distance to lead vehicles.

From that analysis, it was determined that ACC oriented coaching would have its largest impact not on drivers who are already using ACC, but rather on drivers who do not use ACC at all. In principle, since nudging toward increased ACC usage only can be applied on drivers who are already function users, non-users must first become users before nudging can be applied. It was thus decided that coaching would be applied primarily toward non-users, with the goal of turning them into users, and hence become available subjects for nudging efforts.

Ways of coaching non-users using an in-vehicle app was developed and tried out in WP4 (see Deliverable 4.5). In total, three development studies were performed, one in Sweden ($N = 30$ test persons), one in the US ($N = 10$ test persons) and one in England ($N = 6$ test persons). These studies lead to three important insights.

First, the app as developed was not robust and natural enough in its speech interaction, especially for users with limited interest in technology (i.e. the target group for coaching). To reach this level, a natural speech-based app with performance much closer to common speech recognition systems like Siri©, Alexa©

etc. and a high level of dialogue localization in the driving support domain would have to be developed.

Second, for the results of the field trial to be clear, it is important to avoid possible confounders. One MeBeSafe research question is whether non-users of ACC can be turned into users through coaching. In that perspective, it would be unfortunate if technical activation difficulties were to interfere with the effects of coaching.

Third, while the coaching toward ACC usage was successful in development pilots, it cannot be ruled out that the presence of a test leader in the vehicle might have had an increased influence. In other words, even if the App was perfectly build, some drivers who now activated functions may have refrained from doing so in absence of a test leader in the vehicle.

Given these conclusions, it became quite clear that the best way forward was to employ a Wizard of Oz approach in the field trial. Wizard of Oz testing is commonly used to understand interaction patterns for functionality which is not yet fully developed. The test participant is led to believe to be interacting with a computer-based function of some type (such as a self-driving car), while in reality an experimenter (the "wizard") is simulating the behaviour of the application (in the case of self-driving cars, a hidden back seat driver is controlling the vehicle).

7.2 Field trial cancelled

A key foundation of this field trial setup was the assumption (based on previously collected data) that among the test participants recruited to the ACC nudging field trial, 20-30 % would be determined non-ACC users who would not be affected by the ACC nudging concepts, and who therefore could benefit from coaching.

This assumption did not hold. When processing the data from the ACC nudging concepts, it turns out that all participants in the field trial were nudged into some level of ACC usage by the nudging concepts. In other words, there is no-one left to coach.

There is also no immediate way to remedy the problem. At a minimum, a new ACC nudging field trial with a significantly larger test population would have to be carried out, where one monitors test participants actively during the treatment phase to make sure there is a sufficient group of non-users left to coach at the end of the trial. For certainty reasons, the field trial should likely apply some form of staggered release design, with groups of $N = 15\text{-}20$ drivers entering a treatment with 3-4 weeks delay between groups, and the field trial is kept going until a set target level of non-nudged users will be reached. This is not feasible within MeBeSafe, so the assessment of the potential for coaching non-users will not come further than the work carried out in WP4.

On the positive side regarding the ACC nudging, there was enough time and competence around to roll out an additional in-vehicle ACC nudging concept (reported above), so the field trial on ACC nudging actually was doubled in size compared to what was initially planned.

8 Field trial results for O5: HGV driver behavioural change through online coaching (Shell)

In WP4 of MeBeSafe, a coaching system for truck drivers was developed, as well as a first version (V1) of the DriveMate app, which was to measure driver behaviour by in-phone sensors and deliver feedback and coaching material. V1 was field tested at the company Litra in Norway by four drivers starting in December 2018 and found to exhibit a high number of bugs. After the shift of funds within the project, V2 of Drivemate was developed in WP5, starting in November 2019. In February 2020, field tests were started at Litra (*N* = 13 drivers) and at Bertschi in the UK (*N* = 20 drivers). This deliverable reports on preliminary results of this field trial.

Principles of the MeBeSafe coaching system

DriveMate and the coaching to be delivered were designed to address issues which had been identified as associated with similar systems in WP1. This, however, also lead to the system differing from others, and the guiding principles will be therefore briefly described here.

- The drivers are anonymous. The account of a driver is associated with a phone ID, but the owner of the phone is not known to the researchers.
- The data is not shared with the company unless aggregated over all drivers.
- There is no real control as to whether the drivers actually read the material, or what they do in the coaching sessions. This is due to the principle of supplying the drivers with a support tool, and not to introduce further surveillance.
- All data is gathered by the phone from its internal sensors, and it is thus not connected to the vehicle CAN-bus.
- Very little information is displayed by the app during driving, due to the risk of distraction.
- Coaching is based mainly upon cognitive-behavioural principles.

- Coaching is peer-to-peer, i.e. drivers are paired and instructed in coaching and deliver this themselves.

8.1 Field trial setup

8.1.1 Groups

Two companies supplied drivers who had volunteered for the project; Litra (*N* = 13 drivers, Norway) and Bertschi (*N* = 20 drivers, UK). There was no control group, as those companies approached to participate in the project declined their participation. However, as there was a baseline period of measurement of at least 18 days when no intervention was delivered (no feedback and no coaching), and drivers would proceed at different paces through this period, a sort of staggered design was used.

The Litra drivers were issued with new Android© phones specifically for DriveMate, while the Bertschi drivers installed the app on their company phones.

8.1.2 Timeline

The introduction to DriveMate and coaching was held on the 27th of February 2020 for the Litra drivers in Bergen (Norway) and on the 7th of March 2020 for Bertschi in Middlesbrough (UK). The timeline thereafter became individual for each driver, as it was dependent upon how fast the driver undertook the sessions. After eighteen sessions of onboarding (coaching techniques material), the actual coaching was started.

Due to the developmental level of the app, the onboarding was not delivered once a day (or rather 22 hours after the completion of the previous session) as planned, and the onboarding was therefore delayed beyond the expected three to four weeks. To speed up the pace, on 4th of June 2020 the onboarding setting was changed so that a new session could be delivered once a minute.

8.1.3 Data processing and storage

In V2 of DriveMate, raw data was gathered by the app and sent to the Cygnify server for processing of feedback values on the three parameters of smoothness, harsh braking and harsh acceleration. These calculated values were stored in a database hosted by Shell, along with the raw data files for each trip.

Summary values for each trip were calculated by Cygnify, which were then rendered as coloured bars in the app and added to average values for the driver and the company (also shown as bars).

8.1.4 Intervention

The intervention has two distinct parts; the DriveMate app and the coaching material. The app has a simple setup where the drivers start DriveMate when they are driving their truck, and then save the trip afterwards. The app uses GPS and time to calculate smoothness of driving (the average of all speed changes when moving), and harsh acceleration and braking (further described below).

Time wise, there are also two distinct phases to the intervention. First, drivers only receive written instructions in the app about how to do peer-to-peer coaching (onboarding), with a time period of at least 22 hours in between sessions. After 18 sessions, during which driving data is gathered but not displayed (baseline data), the drivers move to the second phase, and coaching is started including feedback in the form of driving data after each trip. Drivers pair up and meet for discussions when the app indicates this to be due (every two weeks at the beginning). Discussion subjects are suggested by the app, including summaries of the user's driving behaviour since the last session. These driving data are compared to previous behaviour and that of all drivers of the same company.

Also, there are events which have been recorded by the app and saved for coaching by the driver, safety topics and videos of truck driving events (gathered from the web). Some coaching alerts invite the drivers to take a survey about a road safety topic (e.g. speeding, dangerous overtaking, distraction, etc.). If the results of a survey indicate that a driver lacks awareness or competence, the topic is suggested for discussion in a coaching session.

8.1.5 Dependent variables

The project used the three parameters of smoothness of driving, harsh braking and harsh acceleration, both as feedback to the drivers and as outcome variables. The goal was to reduce these values (with zero as the absolute minimum value for all variables).

Smoothness was calculated as the average of all absolute acceleration values during movement (given by GPS position and time). This variable has been found to be associated with crash involvement for car (Lajunen & Summala, 1997; Quimby, Maycock, Palmer & Grayson, 1999) and bus drivers (Khorram, af Wåhlberg & Tavakoli, 2020; af Wåhlberg, 2006; 2007; 2008).

Harsh braking events have no single physical definition and goes by many different names in research (e.g. Duarte, Gonçalves & Farias, 2013; Klauer et al., 2009; Tapp, Pressley, Baugh & White, 2013). In *MeBeSafe*, preliminary analyses on the UDRIVE database had indicated that there existed differences in what could be considered harsh braking at different speeds; at low speeds, most strong braking was found to be due to traffic lights turning red, i.e. not a situation of some kind of risk. Therefore, two different criteria for harsh braking events were implemented; 1.4 m/s^2 when speed was <40 km/h and 0.9 m/s^2 when speed was >40 km/h. Acceleration events were calculated in a similar manner.

8.1.6 Analysis

Due to the quasi-experimental setup of the trial, and the individual delivery of the intervention, the arrangement of the data became complex. As each driver pair would start the intervention at different times, and record different numbers and lengths of trips, a separate arrangement was needed for each driver pair.

8.2 Results

8.2.1 Number of drivers and trips

Litra: A total of *N* = 705 trips had been recorded by the Litra drivers, of which *N* = 668 (93.4%) were error-free and could be used. One driver reported that his MeBeSafe phone had stopped working because it could not be charged. Due to corona, this problem could not be resolved. Three drivers did not record any trips. This left nine drivers who recorded trips during the whole intervention period. However, due to corona, Litra experienced a strong setback in business and many drivers were laid off, while the remaining worked less than their usual hours. Also, the onboarding did not proceed as expected, due to technical problems the onboarding sessions were not presented to the drivers as planned. By early June 2020, only one driver had proceeded to the coaching stage.

Bertschi: This company continued business rather much as before corona. By early June 2020, *n* = 4 drivers, out of *N* = 20, had proceeded to the coaching stage. These drivers recorded 1615 trips, of which 1398 (86.6%) were error-free.

None of the drivers had reached a coaching session that included a driver-competence survey.

8.2.2 Driver behaviour change

The effect of the intervention was calculated as the differences in means between before and after the coaching intervention started for each driver. This means that differing numbers of trips were used for each driver. The length of the trips could also differ strongly.

Table 8.1 and Table 8.2 show the results by 2020-06-10 (means and standard deviations on the outcome variables). For both companies, the number of drivers in coaching are too small for further analysis to be meaningful.

	Litra, N=1			Bertschi, N=4		
	Before	After	d	Before	After	d
	Mean/std	Mean/std	-	Mean/std	Mean/std	-
Smoothness	0.264/0	0.273/0	-	0.226/0.070	0.203/0.087	-
Harsh braking	0.524/0	0.584/0	-	0.439/0.407	0.449/0.596	-
Harsh acceleration	0.397/0	0.446/0	-	0.339/0.399	0.365/0.578	-
Number of trips	172	24	-	96	21/14	-

Table 8.1: Mean values of smoothness, harsh braking and acceleration events, before and after coaching started for the drivers who passed beyond the onboarding stage. Lower values indicate better driving.

	Litra			Bertschi		
	Coaching, N=1	No coaching, N=8	d	Coaching, N=4	No coaching, N=16	d
	Mean/std	Mean/std	-	Mean/std	Mean/std	-
Smoothness	0.273/0	0.254/0.031	-	0.203/0.087	0.356/0.161	-
Harsh braking	0.584/0	0.400/0.087	-	0.449/0.596	0.410/0.010	-
Harsh acceleration	0.446/0	0.286/0.074	-	0.365/0.578	0.338/0.022	-
Number of trips	24/0	59/72	-	21/14	58/37	-

Table 8.2: Mean values of smoothness, harsh braking and acceleration events, compared for drivers who did and did not pass beyond onboarding. Cohen's d values were computed for the differences

8.2.3Covid-19

The field trial was started at the very moment when the corona crisis was acknowledged in most European countries, and lockdown and other measures restricting movement where put in place. This had the direct effect upon the coaching trial that all drivers in Norway had their driving strongly reduced. Also, the focal point manager was put on part time and could no longer support MeBeSafe or the drivers.

Furthermore, the driver behaviour measurements of MeBeSafe assumes an unchanging driving environment, a condition which has been violated by the changes in traffic due to the corona crisis. There are also seasonal changes in driving environment which were planned to be handled by statistical controls in these data. The change due to corona, however, is currently not possible to estimate in the areas where the field trial is taking place. Any changes in measured truck driver behaviour in the MeBeSafe project can therefore be due to different factors, which can currently not be disentangled. Although drivers who did not reach the coaching stage of the intervention could in principle be used as a control group for the same time period as the drivers who started coaching, the amount of data was deemed too small to be used for this end. Also, these drivers were probably to some degree self-selected, and any difference therefore not really reliable due to the intervention.

Also, feedback from the drivers concerning the use of the app was limited. An explanation could be that the corona crisis might have had an impact in the sense that drivers would see the project as less important than many other factors in their lives, and therefore have abstained from responding to surveys and queries.

8.2.4 Summary and outlook

The limited test period and the corona pandemic places restrictions upon the possible interpretations of the field trial results. Conclusions on whether coaching changes driver behaviour can therefore not be drawn at this stage. However, conclusions can to some degree be drawn about the feasibility of delivering peer-to-peer coaching in trucking companies, the technical standard of the DriveMate app, as well as whether the users are satisfied with the app. For the latter point, we may conclude that most drivers where probably reasonably satisfied, as they continued to record trips during the whole field trial time period, despite technical problems and the corona crisis.

Supporting that hypothesis, whenever we have had meetings with truck drivers about the coaching and the app, they have been very positive about the concepts behind the app and this approach to coaching. Thus, our careful conclusion is that there is potential for the approach and technology we have developed, even though due to the limitations and issues we have encountered, and which are described elsewhere, at this point in time it is too early to say whether it leads to statistically significant benefits in the KPIs we have identified.

9 Field trial results for O6 and O7: Safe speed/trajectory on inter-urban roads (ika/ RWTH Aachen)

9.1 Field trial set-up

For Objectives 6 and 7 - Safe speed/trajectory on inter-urban roads, the field trial took place on an exit lane in Eindhoven, The Netherlands, where roadside marking lights were installed (40 LED road studs on each side of the exit lane for a total length of 240 m, see Figure 9.1 on the left) in such a way that drivers who entered the exit lane at speeds above a predefined threshold (see Figure 9.1 on the right) could be exposed to a set of systematic light patterns along the lane. We measured the vehicle's speed using thermal cameras along with computer vision algorithms. An intelligent decision control logic identified those vehicles that fulfilled the nudging criteria (exceeding the speed threshold as shown in figure 9.1 on the right and a minimum distance of 90m between two nudged vehicles) in order to display the light pattern only to relevant vehicles at the relevant position, thereby avoiding distraction of other drivers. More details about this set-up are described in deliverables D3.3 and D5.3 (chapter 7). For details on trial design, please see D5.1.

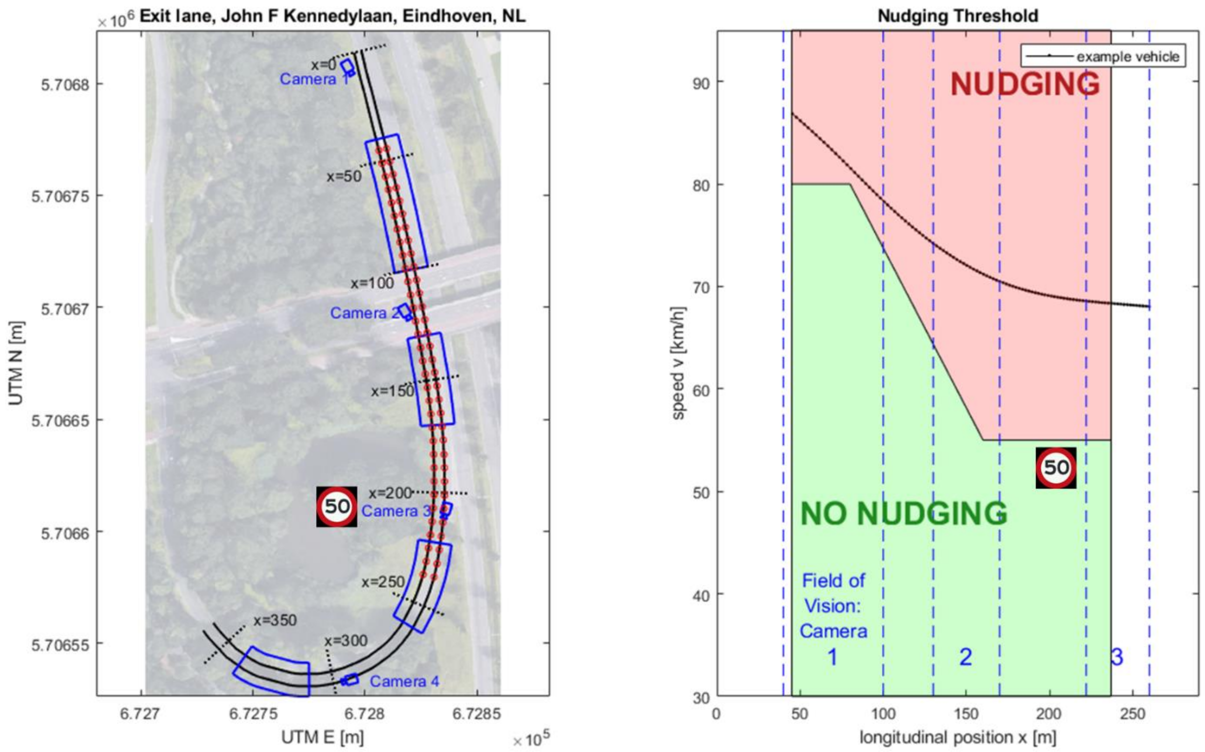

Figure 9.1: On the left: Set-up of the field test with thermal cameras and roadside marking lights. The beginning of the lights is the beginning of the exit lane. On the right: Nudging threshold based on speed over the course of the exit lane

Nine different scenarios were tested based on the results of the driving simulator studies (see Deliverable 3.2) including systematic variations of light pattern, spacing between activated lights, brightness levels, as well as light movement speed. Each scenario was usually tested for one week (Monday to Monday), some shorter in case of technical issues. The scenarios were divided into four different testing phases. Testing phase 1 compared a baseline with static lights and with lights moving towards the driver. Both nudging conditions were tested in two different variations of spacing between active lights, resulting in four scenarios plus baseline (see Table 9.1). Subsequent testing phases compared different movement speed of lights (phase 2) and conditions with different spacing between sets of activated lights (phase 3). Between testing phase 2 and testing phase 3, we conducted an intermediate baseline for control purposes.

Furthermore, we tested a speed indicator device displaying a smiley sign placed right in front of the 50 km/h-sign ("☺" for drivers below the threshold, "☹" for drivers above this threshold) while the nudging system itself was turned off (testing phase 4). These additional testing phases and a more in-depth analyses of testing phase 1 will be evaluated in D5.5. This deliverable D5.4 focuses on the main results of the four scenarios plus baseline evaluated in testing phase 1; Table 9.1 illustrates testing phase 1.

No.	Scenario			Number of all vehicles on the exit in the testing week
	Colour	Movement	Spacing between lights	
0	No nudge – baseline			N = 19,030
1	red	Moving towards the driver at 50 km/h		N = 10,059
2	red	Static lights		N = 18,458
3	red	Moving towards the driver at 50 km/h		N = 19,417
4	red	Static lights		N = 19,211

Table 9.1: Scenarios of testing phase 1.

Within the first test phase, the factors *movement of lights* as well as *spacing between the lights* were varied and tested against a baseline with no lights. The colour of lights was held constant across nudging conditions. In correspondence to the results of the simulator testing (see D3.2) and since humans are conditioned to respond to red lights with caution (Donald, 1988; Edworthy & Adams, 1996) we expected drivers to reduce their driving speed more in static light conditions compared to a baseline (**H1.1**). Since the speed perception in conditions with lights moving towards the driver is altered (e.g. Gibson, 1950; Manser & Hancock, 2007; Gates, Qin, & Noyce, 2008), we expected drivers to reduce their speed more in conditions with lights moving towards them compared to a baseline (**H1.2**) and compared to static lights (**H1.3**). The spacing of the lights was not expected to have an effect on driving speed between static light conditions (**H1.4**) and conditions with moving lights (**H1.5**). This was tested exploratory.

9.2 Results

The effectiveness of the Infra Driver Nudge was evaluated based on the speeds recorded by the thermal imaging cameras (overall descriptive results: chapter 9.2.1, driving behaviour of fastest drivers: chapter 9.2.2), as well as via an on-site survey (chapter 9.2.3) and a resident survey (chapter 9.2.4). UNIFI analyses the potential effectiveness of the system on PTWs (chapter 9.2.5). This deliverable gives only an overview over the results. Detailed hypotheses, methods, results, and discussion will be provided in D5.5.

9.2.1 Overall Results

During the first testing phase (21st Oct 2019 to 09th Dec 2019), the trajectories of *N* = 881,087 vehicles (≈17,300 per day) were gathered. This is not equivalent to the actual number of vehicles on the road as vehicles on the left lane were ignored since they do probably not change to the exit lane. *N* = 283,673 (31.8 %) vehicles have used the exit lane, on average ≈5,600 per day. The trajectory data were thoroughly checked for plausibility so that *N* = 143,797 (50.7 %) of the vehicles were used for the analysis. *N* = 86,175 (59.9 %) of these vehicles took the exit while a scenario was active (including the baseline scenario). *N* = 56,253 (65.3 %) of them fulfilled the criteria for nudging (speed above threshold, headway large enough to show the light pattern) for at least one moment. This number is not equal to the number of actually nudged vehicles for two reasons: in the baseline scenario (0), vehicles were not nudged even if they fulfilled the nudging criteria, and some vehicles were nudged but excluded from the analysis due to implausible trajectories.

For the analysis of the effect of nudging, the mean speed is not preferable for two reasons: (1) the mean speed at x = 50 varies in the different scenarios, presumably due to varying weather and traffic conditions and (2) the mean speed includes also those vehicles that do not fulfil the nudging criteria, i.e. their speed is below the threshold and/or they cannot choose their speed freely due to a vehicle in front of them. Therefore, we compared only those vehicles within the same speed range at x = 50 and selected only those vehicles that were uninfluenced by another vehicle. Vehicles with a speed between 80 and 85 km/h at x = 50 reduced their speed by up to 3.0 km/h (4.9 %) more than in the baseline scenario, depending on the position and the scenario. Vehicles with a higher speed at x = 50 were slowed down even more (e.g. 4.6 km/h between 95 and 100 km/h). In road design, the 85 % quantile of speeds "V85" is often used for safety analyses (Lippold, 1999). The percentage of vehicles faster than V85 of the Baseline scenario is reduced by up to 39 % at x = 200 depending on the scenario. Further analyses will be reported in D5.5. Figure 9.2 (on the left) illustrates the ratio of vehicles faster than the 85% quantile of speeds (V85) in the baseline scenario, Figure 9.2 (on the right) shows the mean speed of vehicles that fulfil the nudging criteria at x = 50 (speed above 80 km/h, headway above 90 m).

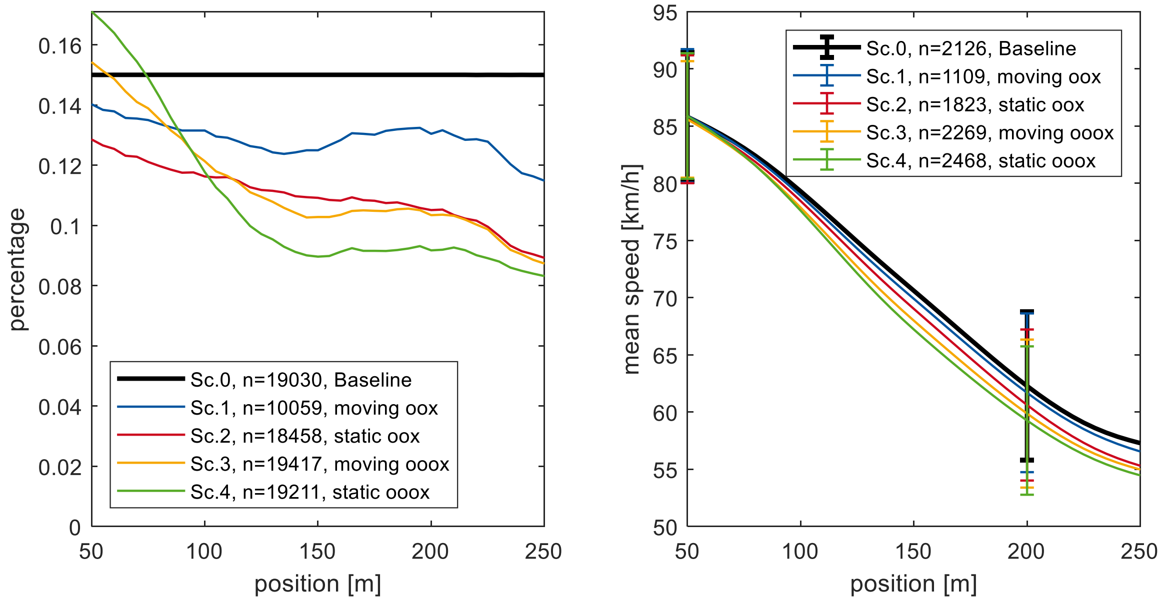

Figure 9.2: On the left: Ratio of vehicles faster than 85 % quantile of speeds (V85) in baseline scenario. On the right: Mean speed of vehicles that fulfil nudging criteria at x = 50.

9.2.2Behavioural Results of Fastest Drivers

The effect of the nudging measure on driving speed is estimated by evaluating the speed reduction between the light onset at x = 50 until the start of the curve of the exit at x =205. As the nudging system targets mainly fast drivers, we narrowed the sample as described in 9.2.1 to the fastest drivers. For this, we included only fast drivers for further analysis. Therefore, we included only those drivers, whose driving speed was two standard deviations ($SD_{all\ drivers}$ = 9.468 km/h) above the mean speed of the baseline ($M_{all\ drivers}$ = 68.076 km/h). Therefore, all drivers in the sample for the analysis of fast drivers exceeded 87.01 km/h at stimulus onset at x = 50. This is a common means to determine the cut-off value based on sample statistics.

The test design in test phase 1 was a mixed design with repeated measures on the within-subjects factor *position on the curve* (2; x = 50 and x = 205) and the between-subjects factor *nudging scenario*, The nudging scenario had five levels, being either a) baseline (scenario 0), b) lights moving towards the driver with a narrow spacing of lights (oox, scenario 1), c) static lights with a narrow spacing of lights (oox, scenario 2), d) lights moving towards the driver with a wide spacing of lights (ooox, scenario 3), or e) static lights with a wide spacing of lights (ooox, scenario 4). The dependent variable was the driving speed in km/h.

We conducted a mixed ANOVA with the between-subjects factor *nudging scenario* (5; baseline, static_oox, static_ooox, towards_oox, and towards_ooox) and repeated measures on the within-subjects factor *position on the curve* (2; x = 50 and x = 205) for the fast drivers. The ANOVA revealed a significant main effect for position ($F[1, 1582] = 26284.30$, $p < .001$, $\eta_p^2 = .94$) and a significant main effect for scenario ($F[4, 1582] = 10.17$, $p < .001$, $\eta_p^2 = .03$). In addition, the results also showed a significant interaction between the factors position and condition ($F[4, 1582] = 8.70$, $p < .001$, $\eta_p^2 = .02$). Figure 9.3 illustrates the results.

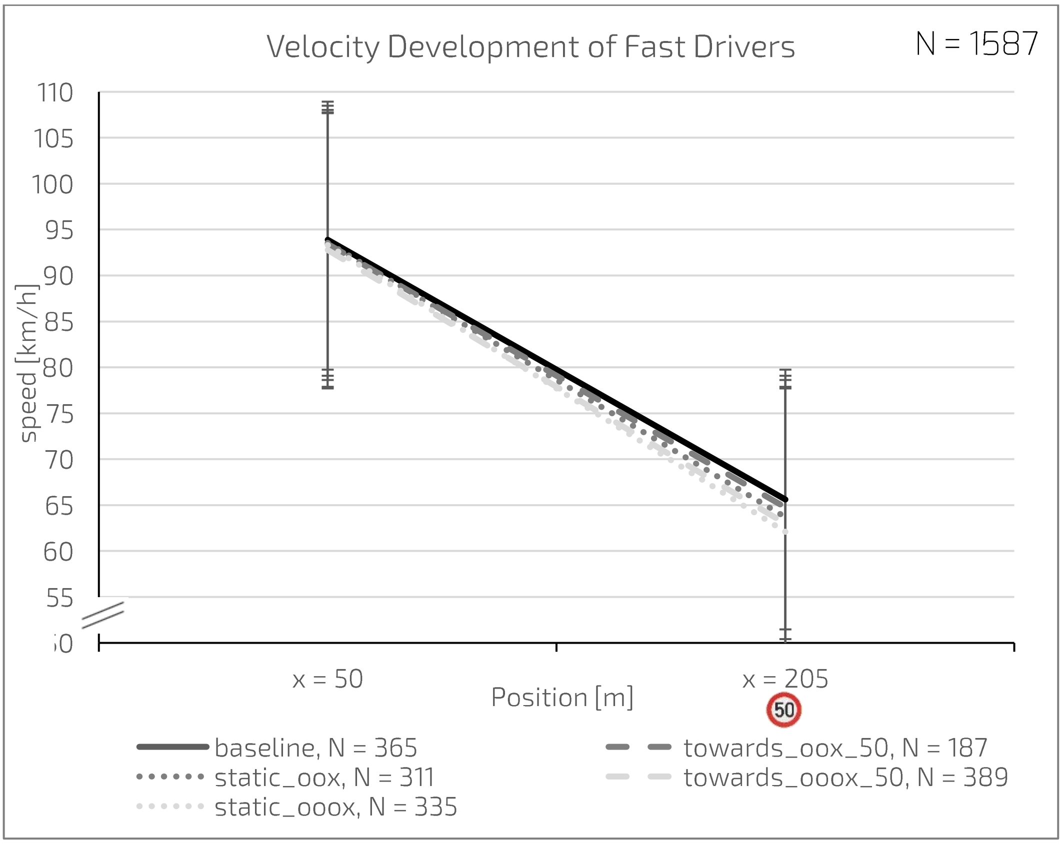

Figure 9.3: Results for the velocity development of fast drivers (drivers faster than 87.01 km/h at x = 50) of testing phase 1. The x-axis shows the within-factor position, the y-axis shows the velocity in km/h, and the lines represent the different scenarios (between-subjects factor) of testing phase 1. Error bars depict standard deviations.

Drivers slowed down through the exit between x = 50 and x = 205. Due to the main effect of the factor scenario and the significant interaction, we can conclude that the five conditions tested in phase 1 differ significantly. As can be seen in table 9.3, drivers are generally fastest when they are not nudged. Post-hoc tests were conducted to determine the differences between scenarios. The results led to the acceptance of H1.1 and to the rejection of H1.4 and H1.5 for fast drivers (see D5.5). We found ambiguous findings for H1.2 and H1.3, which need further investigation and discussion. The detailed results and discussions will be part of D5.5.

9.2.3Results of On-Site Survey

An on-site survey with *N* = 20 participants was conducted in October 2019 and consisted of test drives with three experimental conditions: a baseline with no lights, static lights and lights moving towards the driver. Aim of the study was to gain insights into how the light conditions are perceived by drivers in a real-life testing environment. In order to get direct feedback right after the recruited drivers experienced the light conditions in randomized order, we conducted a semi-structured on-site survey.

All participants stated that they saw the lights on the roadside and both light conditions were thought to aim at reducing speed and raising attention. Further, participants stated that they thought the moving lights operate as a warning signal. For both light conditions, participants stated that the lights influenced their driving behaviour. The static lights were perceived as increasing participants' attention and the moving lights were mostly perceived as making the participants reduce their speed and drive more cautiously.

A Wilcoxon signed-rank test revealed that the lights moving towards the driver were rated significantly better than the static lights for the statement *"The light system made me decrease my driving speed"* ($Z = -2.25$, $p < .05$). A one-tailed Wilcoxon signed-rank test showed that the participants regarded the lights moving towards them as significantly more "useful" ($Z = -2,89$, $p = .002$), "nice" ($Z = -2,60$, $p = .005$) and "effective" ($Z = -1,96$, $p = .025$) compared to the static lights, rated on the Van der Laan Acceptance Scale (Van der Laan, Heino, & De Waard, 1997). On the same rating scale, none of the light conditions was perceived as "useless", "unpleasant" or "annoying", and both light conditions received mostly positive evaluations, especially the moving ones.

9.2.4 Results of Online Resident Survey

An online survey with *N* = 287 participants was used to evaluate how people perceive the infrastructure nudge in November 2019. As participants could answer the survey anytime within a four-week timeframe, we were not able to determine which scenario participants experienced. Therefore, we asked for their experience with the system as a whole.

Responding on a 4-point Likert-scale (1 = "completely agree" to 4 = "completely disagree", and "I don't know"-option) they rather agreed that they felt safe when the light system was on, "rather agreed" that the system guided their driven trajectory, felt that the lights system made them aware of a hazardous situation. Further, they replied ambiguously whether they felt as if they decreased their driving speed due to the light system, disagreed when asked if the light system made them nervous, did not feel as if the light system irritated them, and disagreed considerably when asked if the system made them increase their speed.

The majority of participants evaluated the lights concerning their general acceptance and experience on the Van der Laan Acceptance Scale (Van der Laan et al., 1997). Participants rated the light patterns as helpful, pleasant, good, fun, effective, nice, supportive, desirable and vigilant rather than pointless, unpleasant, bad, boring, unnecessary, annoying, useless, undesirable and sleep inducing. Detailed statistical results and will be reported in D5.5.

9.2.5 Results of PTW Analysis

Powered Two-Wheeler (PTW) raw data were recorded in video format by ISAC in the periods October 8-11 and June 21-23 2020 and were analysed at the University of Florence. The field trial set-up was the same as in Figure 9.1. Since different light scenarios were used in those days (see D5.5 chapter 9.5), we did not distinguish each scenario, but only considered the difference between baseline (light off) and treatment (light on). The evaluation of the nudging system was conducted based on PTW trajectories and velocities. As in the previous sections, the vehicle speeds were measured processing videos from thermal cameras, but vehicles in the videos were identified (and the kinematic parameters determined) exploiting image contrasts, generated by the different temperatures of moving objects on the background. The main advantage of the method based on image contrast, compared to an artificial intelligence (AI) technique, is that it detects the vehicles after a few frames of observation. The drawbacks of the method are the necessity of a manual pre-selection of the video sections with PTWs, since the method is not capable of categorizing PTWs separately from other vehicles, and of a clear view without large overlaps between vehicles. More details about the method for video processing will be given in D5.5, to explain better how it works and how it copes light differences to detect PTWs.

Eighty-one hours of video recordings were available and processed: *N* = 470 motorcyclists were identified in all the lanes (i.e. in the main highway lanes and in the exit one). We decided to consider the whole road width, as PTWs riders tend to change the lane more often compared to car drivers. This is different to the analysis of car drivers. In the dataset only *N* = 54 riders (12 %) of them decided to take the exit lane and their data were processed to assess the nudging effect. The riders who took the exit lane were divided into two groups: 'baseline' and 'treatment'. The former includes *N* = 35 riders (65 %), who used the exit lane when the system was off; whereas the latter contains *N* = 19 riders (35 %), who were detected by the video cameras while the nudging system was active. In the latter group, *N* = 17 out of *N* = 19 riders activated the nudging system at least in one section.

Furthermore, more recent video recordings, related to the period from the 21st to the 23rd June 2020, were examined: *N* = 36 PTWs were detected. Since the light system was previously dismantled, these bikers were considered as 'Post Treatment' and a potential residual effect of nudging system on them was investigated. In the definition of the baseline, we assumed that there was no permanent effect on the riders from the preliminary tests of the nudging system. In fact, preliminary tests were run for approximately four weeks before the video recordings. This assumption was made since there were no velocity data before September 2019 (i.e. previous to the preliminary tests of the nudging system).

As the nudging system targets fast drivers, we excluded 75% of acquisitions from the preliminary analysis, thus considering only the fastest riders, defined as the highest speed quartile (Q3) for each camera. The 'V85' design criterion was not considered because it requires vehicles to travel in free-flowing conditions, i.e. when the preceding vehicle has at least 4 seconds headway, and above the 85th percentile. Because of the limited number of cases, the criterion was too restrictive and thus not appropriate for the analysis of this dataset of PTW data.

As already highlighted in 9.2.1, the mean velocity is likely biased by many factors. In addition to the previously cited factors, since we took advantage from a different computer technology for acquisitions, different tracking length is another factor that could affect the mean value. Therefore, it was assumed that the peak speed was better suited to the purpose of this comparison. On the left side of Figure 9.4, the mean peak velocity for riders above Q3 was plotted for baseline and treatment data. A blank x-axis is plotted, since the peak speed value occurs at a different position under each camera for each rider. On the right side of the figure, we plotted the mean instant velocities for riders above Q3 at three sections (x = 50, x = 140 and x = 230), located approximately at the beginning of the field of view for each camera. In the data analysis, no interpolation was done between cameras, and thus the set of riders is defined independently for each camera (approximately *N* = 9 riders were considered for the baseline group and *N* = 3 riders for the treatment group). By examining riders that are speeding in the specific road section, we eliminate any confounding effect (i.e. differently for riders over the threshold for the first camera, but under the threshold for the other two cameras would have still been considered speeders).

In Figure 9.4 on the left side, it's possible to see that after an initial gap of 3.0 km/h between baseline and treatment groups, the difference decreases to 1.4 km/h, underneath the third camera. On the right side, the difference between baseline and treatment are 1.0 km/h at x = 50, 2.9 km/h at x = 140, and 2.0 km/h at x = 230.

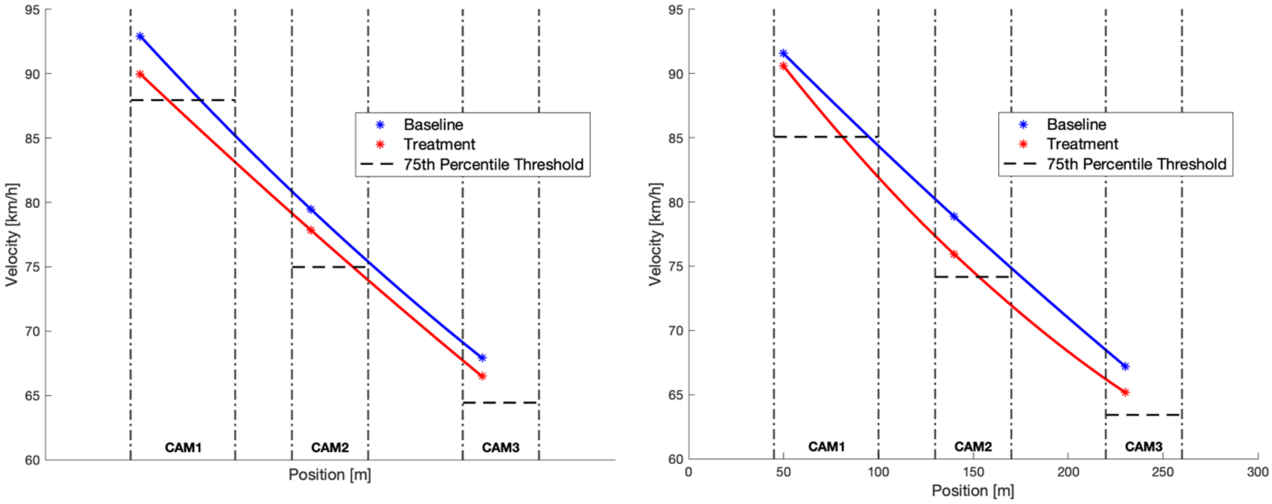

Figure 9.4: On the left side: mean peak velocity below each camera for drivers above Q3 quartile; Blank x-axis due to different peak velocity positions under each camera. On the right side: instant mean velocity for drivers above Q3 quartile at x = [50, 140, 230].

The peak velocity gap under CAM1 was analysed in relation to the trajectory (Figure 9.5): a change in velocity could not be ascribed to the nudging system. In fact, in the first road section, most of the riders moved from the second and third lane to the exit lane and, therefore, they were not affected by the light patterns (Figure 9.5, right side). At x = 50 in the baseline group, *N* = 5 out of *N* = 8 riders (63 %) were outside the exit lane; the same for the *N* = 2 riders (100 %) of treatment group. This rejects the hypothesis that the initial velocity gap was an effect of nudging, but more consistently it might be due to weather and traffic condition.

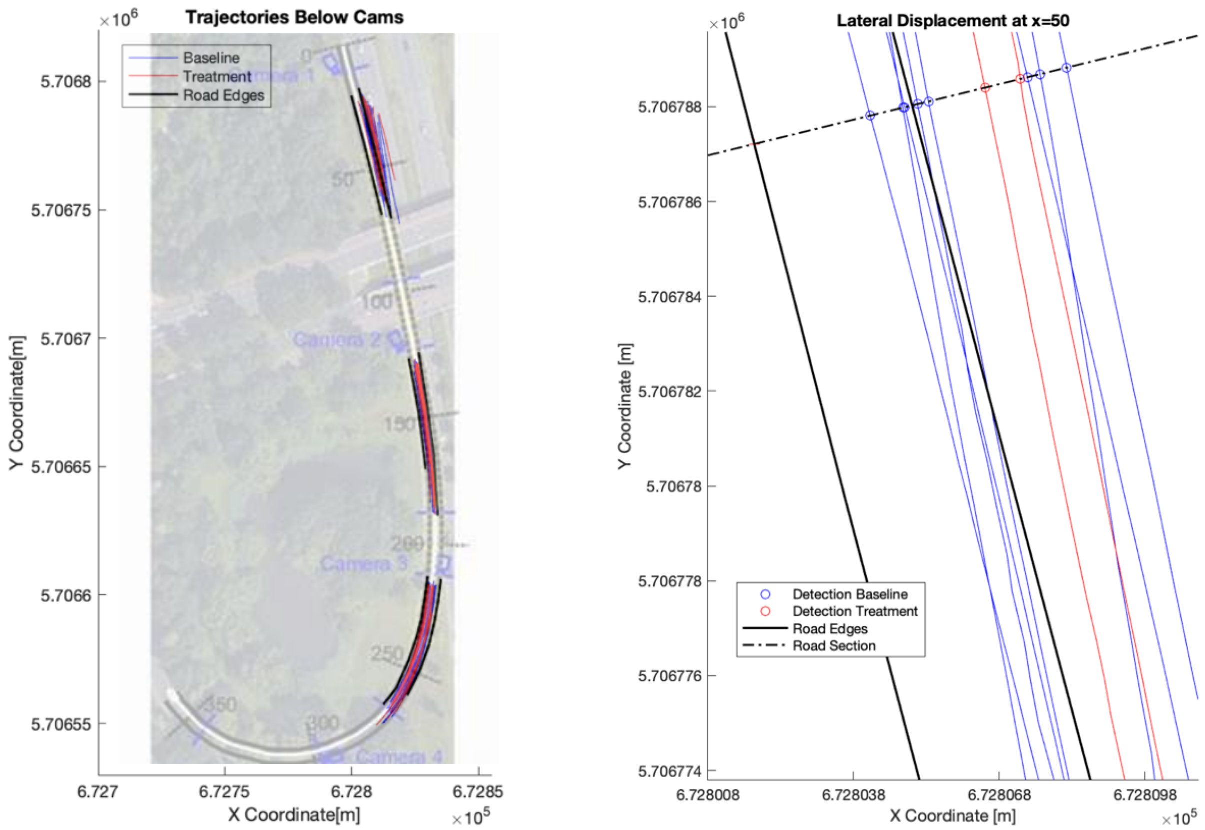

Figure 9.5: On the left side, trajectories under each camera, superimposed on real map. On the right side, comparison on lateral displacement for bikers above Q3 threshold.

In addition, an ANOVA was performed: a two-way ANOVA for unbalanced design, since we had different acquisition for the baseline pattern and the treatment one. Three effects were investigated. The main effect on velocity of (1) scenario, that is 'baseline' or 'treatment' ($F[1, 35] = 1.6$ for peak velocity and $F[1, 28] = 0.75$ for velocity at x) ; the position along the road (2) ($F[2, 35] = 90.31$ for peak velocity and $F[2, 28] = 46.77$ for velocity at x); eventually, the interaction between pattern and position ($F[2, 35] = 0.65$ for peak velocity and $F[2, 28] = 0.01$ for velocity at x). In Table 9.2, we show the results of the two-way ANOVA for the *between-groups* variation.

Criterion	p-values			Number of PTWs above Q3 in the ANOVA	
	Longitudinal Position	Scenario	Interaction		
Peak Velocity	< .001	0.214	0.527	N = 9	N = 4
Velocity at X	< .001	0.393	0.994	N = 9	N = 3

Table 9.2: ANOVAs for PTWs above Q3.

P-values show that longitudinal position has a significant effect on velocity and that the results we collected are statistically significant. The scenario p-values indicate that none of the data sets is significantly different from each other, so the nudging system is not effective on speed for PTWs in this field trial. The p-values for the interaction show that there are no significant interactions between scenario and position; that is, we have an *additive model*.

In this deliverable, we reported some preliminary results for the baseline and treatment sets. In D5.5, we'll thoroughly analyse velocities and lateral position in more sections along road. In addition, the 'post-treatment' group will be included in the data analysis and the comparison with 'baseline' will be supplied in order to investigate the existence of any residual nudging effect on riders.

10 Field trial results for O8: Cyclists' speed reduction – Sweden (SAFER/ Chalmers University)

10.1Field trial setup

For Objective 8 - Cyclists' speed reduction as carried out in Sweden, the field trial involved a random sample of cyclists who passed the two test sites implemented in Gothenburg. In addition to the collected bicycle speed data, a number of passing bicyclists were interviewed on the two sites. Furthermore, N = 17 of the passing cyclists were equipped with instrumented bicycles and then interviewed while watching their trips documented on video.

10.1.1Description of the test sites

The nudges were installed at two locations in Gothenburg, the sites chosen on the basis that they are intersections where there have been previous serious accidents involving bicycles and cars (and their respective drivers). The Swedish national traffic accident data base, STRADA, as well as City of Gothenburg data have been used to confirm that these are indeed spots where slowing down bicyclists may have a positive impact on safety. The chosen spots are Nobelplatsen and Götaälvbron.

Figure 10.1: The nudge at site 1.

Figure 10.2: Site 1 - Nobelplatsen (57°42'49.6"N 12°00'23.4"E). The bike lane is unidirectional, and the bikes comes from the low right corner of the picture moving towards the top left.

At site 1, Nobelplatsen, shown in Figure 10.2, the bike lane is single lane, going in one direction (width 150cm). The bike lane is separate from the street, but on the same level as the pedestrian pathway. It has on street parking for cars on the left, and shops and restaurants on the right. The bike lane has a slight downwards slope, adding some extra speed to the bicyclists. The nudge (see Figure 10.1) was placed just before an intersection with car and bus traffic, where there has been a number of accidents and incidents historically.

The second site, Götaälvbron (or Göta river bridge), shown in Figure 10.4, has a two-way bike lane (width 120 cm per lane) with a pedestrian pathway to the right. It has a downwards slope but in addition the slope down from the bridge will give the bicyclists quite a bit extra speed. The nudge was installed just before an intersection, see Figure 10.3, which is considered one of the more dangerous in Gothenburg where there has been quite a lot of accidents and incidents over the years.

Figure 10.3: Site 2, note the Viscando Otis measuring equipment mounted on the light pole nearest in the picture.

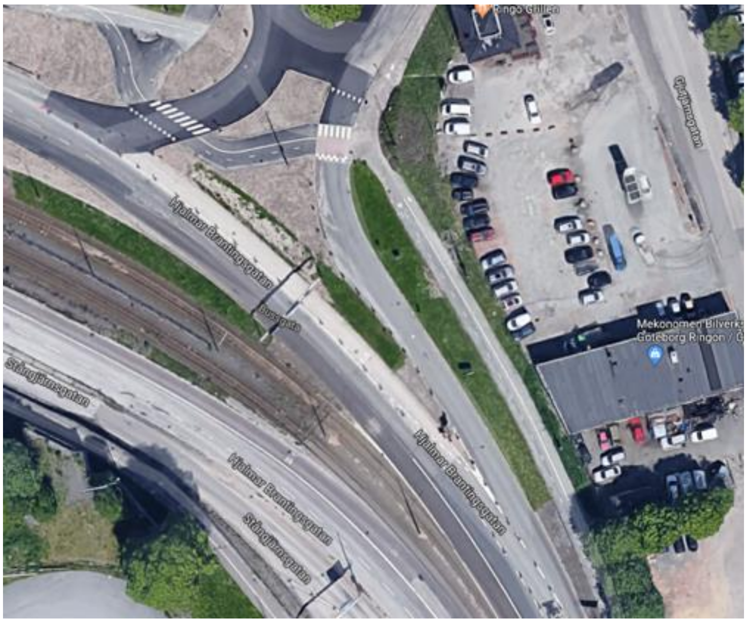

Figure 10.4: Site 2 - The Götaälvbron site (57°43'12.4"N 11°57'45.5"E), the bike lane is the lane most to the right in picture.

10.1.2 Installed equipment

The measuring equipment installed at the locations consists of two Viscando **OTUS3D FLEX** units to cover the area of interest (viscando.com, see Figure 10.5). Each unit has a pair of cameras producing a stereo image that is interpreted by the device resulting in tracks of individual bikes (as well as pedestrians, cars, buses, etc.) that is transferred wirelessly to Viscando. As the data is generated in real-time, no images are sent or stored, which makes the units GDPR compliant. The data enabled an analysis of speed, as well as trajectory, and control for factors such as cyclist type, weather, and wind. The baseline and treatment periods for the two sites are shown in Table 10.1.

Figure 10.5: The equipment installed at one of the locations (Götaälvbron).

Location	Comparison situation	Month
1) Nobelplatsen	Baseline (3 days)	September
	Treatment 1 (4 days)	September
	Treatment 2 (4 days)	October
2) Götaälvbron	Baseline (4 days)	April
	Treatment (4 days)	March

Table 10.1: Field experiment design.

10.2 Results

10.2.1 Measurements of effect on speed

The number of cyclist trajectories that satisfied the selection criteria for baseline, treatment 1 and treatment 2 for location 1 were 740, 1151 and 995, respectively, while the numbers for location 2 were $N = 1301$ and $N = 1292$. The speed distribution at three positions for baseline and treatment for both locations is shown in Figure 10.6. The effect of the nudge on average speed at the three positions was different across sites and treatment repetition. The main effects of condition (baseline or treatment), position (start, middle, and end of the nudge) and the interaction effect were all significant for both locations, see Table 10.2. For location 1, the speed at baseline was higher than at treatment 1, $p < .001$, which means that the cyclist speed decreased with the nudge. On the other hand, for location 1, the speed at treatment 2 (one month after the first treatment) increased in comparison to baseline, $p < .001$. For location 2, the speeds at treatment were higher than the baseline, $p < .001$.

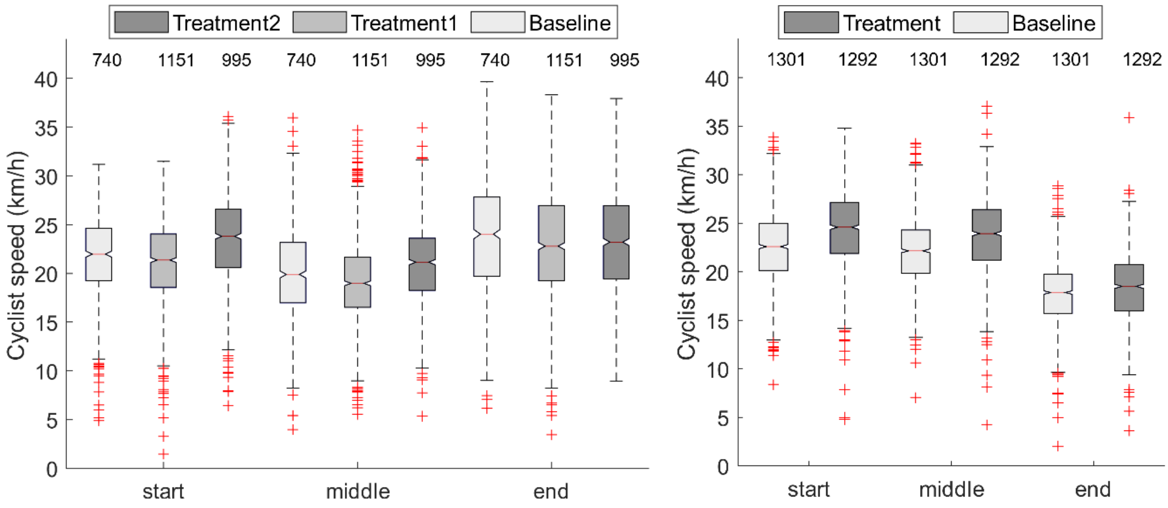

Figure 10.6: Distribution of cyclist speed at different positions and conditions for location 1 (left) and location 2 (right). The numbers on the top report the sample size for each boxplot.

	Site 1		*Site 2*	
	F	*p*	*F*	*p*
Condition	82.04	< .001	267.30	< .001
Position	297.89	< .001	1802.47	< .001
Condition x Position	17.12	< .001	18.85	< .001

Table 10.2: Summary of ANOVA results for site 1 and site 2.

Leisure cyclists (cycling on weekends between 12-18:00) showed lower speeds in treatment than in baseline for both sites, while commuters (cycling on weekdays in between 7-8:00 and 16-17:00) were less affected by the nudge than leisure cyclists. The main effects of condition and cyclist type were significant for site 1, $F(2,1472) = 3.34$, $p = 0.03$ and $F(1,1472) = 26.51$, $p < 0.001$. At site 2, the main effect of condition was not significant $F(1,1296) = 0.01$, $p = 0.98$, while cyclist type and the interaction between condition and cyclist type were significant, $F(1,1296) = 16.71$, $p < 0.001$ and $F(1,1296) = 26.34$, $p < 0.001$, respectively.

When taking into consideration the wind, the apparently contradicting results across sites were not evident or statistically significant any longer. The main effect of wind was significant, both for site 1, $F(2,2877) = 3.66$, $p = 0.03$ and site 2, $F(2,2587) = 16.42$, $p < 0.001$, confirming that tailwind increases speed while headwind reduces it.

The percent of cyclist that arrived at the end of the nudge at speeds greater than 25 km/h was in general lower in treatment than in baseline for location 1. More specific, the proportion of cyclists that ride above 25 km/h, at the end of the nudge, was lower in treatment 1 than in baseline $\chi^2(1) = 6.44$, p = 0.01, while this effect was not significant for treatment 2 in comparison to baseline, $\chi^2(1) = 2.32$, p = 0.13. However, once wind was considered, the nudge did not slow down faster bikers more than slower ones.

Individual cyclists decreased their speed when cycling through the nudge more in treatment than in baseline. For site 1, the proportion of cyclist decreasing their speed more than 10% was larger in treatment 2 in comparison to baseline, $\chi^2(1) = 22.05$, $p < 0.001$. For location 2, the proportion of cyclist decreasing their speed more than 20 % ($\chi^2(1) = 77.98$, $p < 0.001$) and 30 % ($\chi^2(1) = 76.22$, $p < 0.001$) was larger in treatment than in baseline.

In summary, the average speed was significantly lower by 0.7 km/h in treatment 1 than in baseline on site 1, while it was significantly higher by 1.3 km/h on site 2. On the individual level, the percent of cyclists that reduced their speed, from the start to the end of the nudge, was between 12-55% for baseline and 21-72% in treatment, depending on the test site, treatment repetition and speed thresholds. In other words, between 9-17% more cyclists reduced their speed through the nudge in treatment than in baseline when approaching urban intersections, depending on the test site, treatment repetition, speed thresholds and wind condition.

10.2.2 On site interviews with cyclists

At both test sites, short interviews were held with passing bicyclists; in total $N = 54$ interviews, of which $N = 31$ were conducted site 1 and $N = 23$ at site 2.

Out of all interviewees, $N = 30$ were women and $N = 24$ men. The reason the gender distribution was not entirely balanced was that more women than men stopped for the interview. Forty-four of the participants reported to have noted or recognized the nudge when shown a picture of it. Thirty-nine of the participants interpreted the nudge as something that intended to slow down bicyclists and/or warn for a dangerous intersection. Half of the participants thought that the nudge had affected their behaviour. An important finding was that 48 of the 54 participants said they were OK with this type of markings in bicycle lanes, while the six who were not stated that they did not understand the purpose. Thus, acceptance for this type of nudge was high among the cyclists in the study.

10.2.3 Instrumented bikes study

The result shows that although the participants show quite different behaviour in terms of how and where they choose to bike, they all have a clear rational for their choices. The study also shows that small differences in the infrastructure could act as nudges, changing the users' behaviour. In terms of the implemented nudges, all participants were positive in that they thought it is relevant to affect speed in dangerous intersections and believed that the implemented nudges would have this effect.

11 Field trial results for O8: Cyclists' speed reduction – Netherlands (TNO/ SWOV)

11.1 Field trial setup

For Objective 8 - Cyclists' speed reduction as carried out in the Netherlands, the field trial involved a random sample of cyclists who passed the test site located at the cyclist path at the Kruisstraattunnel in Eindhoven during a 14-day period (to cover both baseline and treatment at the same days of the week).

11.2Nudge design

The nudge design is similar to the one adopted in Sweden as shown in Figure 11.1. In this design, transverse lines are implemented getting closer to each other as the distance to the intersection decreases. This type of nudge has been shown to have a positive effect on speed reduction of motor vehicles (Denton,1973; Godley, 2000; Gates and Qin, 2008) therefore, it is implemented as a nudge to evaluate the speed and safety changes at this location.

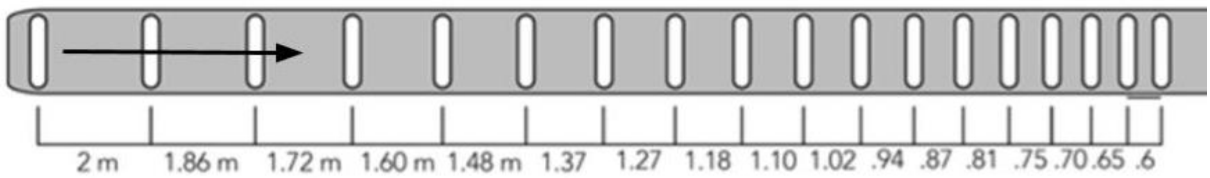

Figure 11.1: Schematics of the transverse line nudge reducing in distance

11.3Location description

This location was selected as there are many cyclist-cyclist interactions in the Netherlands which is a safety concern. The location is a T-intersection of two dedicated cycling facilities with no car traffic at the intersection of the Kruisstraattunnel and Fellenoord in Eindhoven. It is the main cycling facility leading to the train station and therefore has a high cyclist traffic volume especially during rush hour, with up to almost 11.000 cyclists travelling through this intersection per day (Dufec/Sweco, 2018).

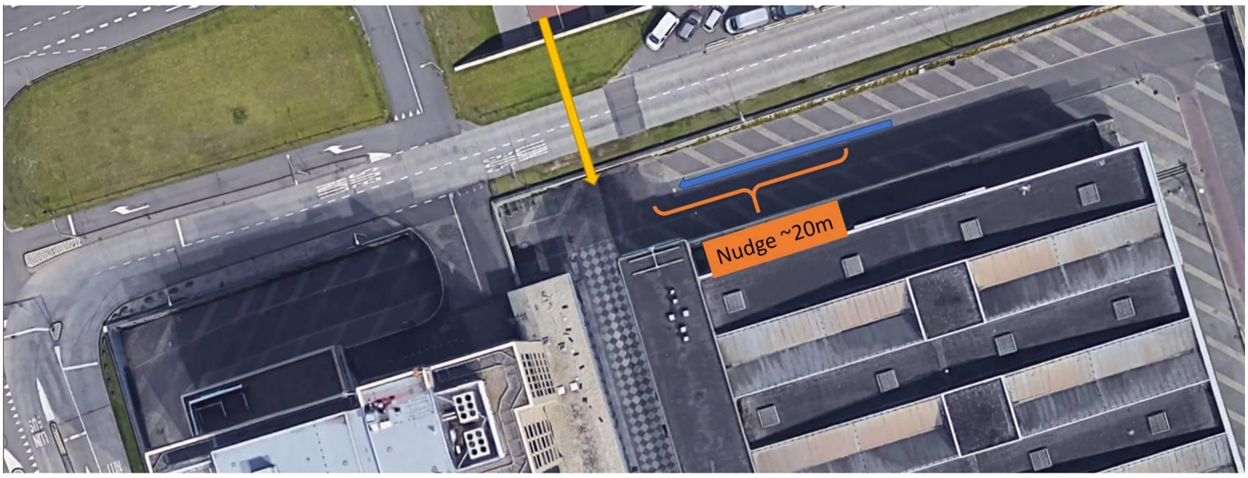

Figure 11.2: Location of study indicating the cyclists traveling through the underpass that will turn right (yellow arrow), the through cyclists (blue arrow), and the location of the nudge (orange)

One leg of the T-intersection is an underpass as indicated with the yellow arrow in Figure 11.2. As a result, cyclists travelling along the blue arrow with high speeds do not have a clear view of the cyclists coming from the right. This leads to conflicts between the through travelling and right turning cyclists at this location, especially since through traveling cyclist must give way to the right turning cyclists, which is not the case in practice. The implementation of the nudge along the through-traveling direction aims to reduce the speed of cyclists approaching the intersection in order to reduce their speeds, providing them with more time to see and give way to the right turning cyclists. The location and length of the nudge is shown in orange in Figure 11.2.

The width of the cyclist lane is 2.25 m and the width of the applied transverse lane marking nudge is 2.45 m. Figure 11.3 shows the implemented nudge and a through cyclist traveling along the nudge approaching the intersection with the underpass cycling facility on the right.

Figure 11.3: Implemented nudge along the through traveling cycling facility, showing two of the three implemented cameras (red boxes)

11.4 Video data collection and analysis

Video data collection was performed by an external company "Connection Systems" (www.connectionsystems.nl). They installed three temporary cameras and collected video data from different angles for two weeks. The cameras were setup on existing poles overlooking the location as shown in Figure 11.3. Figure 11.4, Figure 11.5 and Figure 11.6 show the camera views.

Connection systems provides road user trajectory files using an object detection method and no video files are transferred to comply with the GDPR rules for privacy. Road users are detected and classified, in our case into cyclists, motorcyclists (mopeds and scooters), and pedestrians. From the provided trajectories, we are able to analyse potential speed and safety changes between the before and after nudge scenarios.

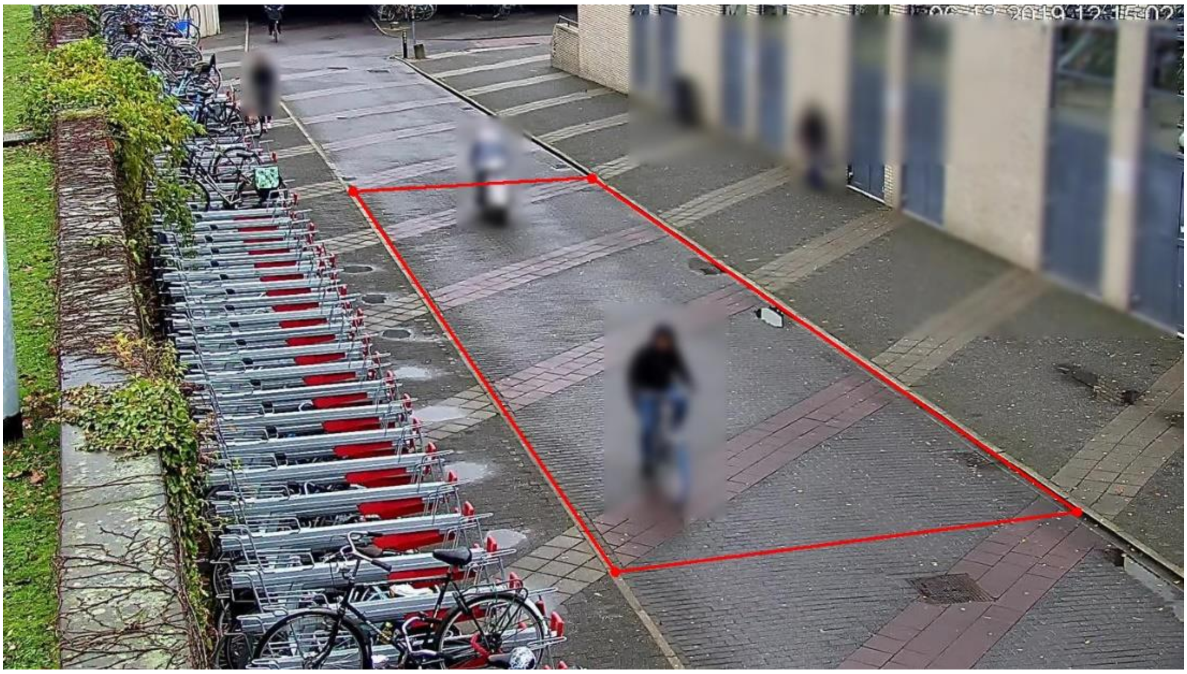

Figure 11.4: Camera view 1 indicating the through traveling cyclists where the nudge will be implemented

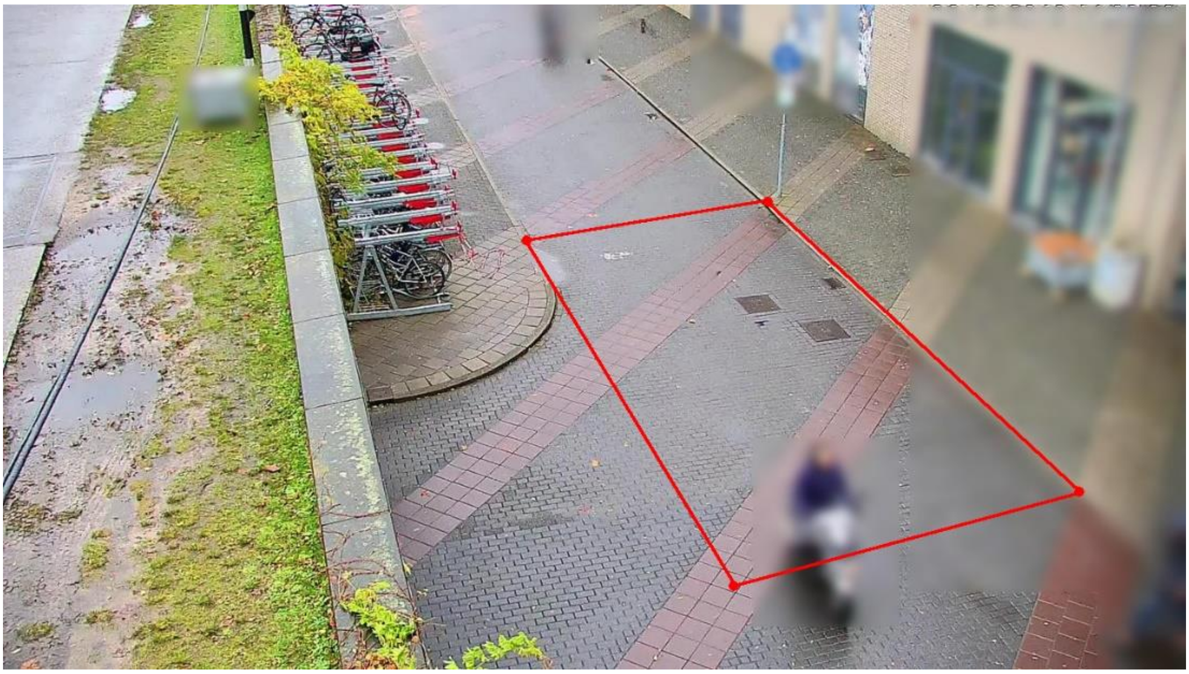

Figure 11.5: Camera view 2 at the intersection

Figure 11.6: Camera view 3 showing the intersection from an angle where the right turning cyclist approach is more visible

The data collection period for the nudge scenario is from December 1st, 2019 at 12:00AM until December 3rd at 10:00AM, and for the base scenario without the nudge, the same weekdays (Sunday through Tuesday), were selected on December 8th at 12:00AM until December 10th at 10:00AM. The trajectory processing includes evaluation of the location where cyclists are riding along both approaches and the change in speed of cyclists going through the nudge. The through cyclist speeds are computed from camera 1 (Figure 11.4) to focus only on cyclists traveling along the nudge.

11.5Results

First quick analysis results indicate a positive effect on reducing the average speed of through-travelling cyclists that pass the nudge. However, more detailed analyses show that this effect is not significant, based on the analysis of the average speed of through-travelling cyclist over the nudge area. The speed of the through-travelling cyclists is influenced not only by the nudge but also by the manoeuvres of other cyclists and pedestrians in the area. It seems difficult to discriminate the different influence factors from the data. More detailed analyses are foreseen in an attempt to isolate the effect of the nudging solution on the speed of the through-travelling cyclists (to be reported in Deliverable D5.5).

12 Summary and conclusions

As has been described above, most nudges developed in MeBeSafe overall show positive effects, i.e. they have influenced road user behaviour significantly and in the expected way when being applied in their respective contexts.

This is a very encouraging result that hopefully can motivate other interested parties to start exploring different nudging approaches as tools toward improving road safety. More details on both the nudges themselves, and on what future implementations might look like, will be reported in deliverable D5.5.

References

Berghaus, M. et al (2019). *Infrastructure measures (Deliverable 3.3)*. Retrieved from MeBeSafe website: https://www.mebesafe.eu/results/.

Choudhary, V., Shunko, M., Netessine, S. & Koo, S. (2020). Nudging Drivers to Safety: Evidence from a Field Experiment. *INSEAD Working Paper No. 2020/28/TOM. Available at SSRN: https://ssrn.com/abstract=3491302 or http://dx.doi.org/10.2139/ssrn.3491302*

De Craen, Saskia, et al (2019). *Report on effective feedback (Deliverable 4.5)*. Retrieved from MeBeSafe website: https://www.mebesafe.eu/results/.

Denton, G. (1973). The influence of visual pattern on perceived speed at newbridge. Laboratory Report LR531. *Crowthorne: TRL Limited*, *1*(1).

Donald, N. (1988). *The design of everyday things*. New York: Basic Books.

Duarte, G. O., Gonçalves, G. A., & Farias, T. L. (2013). Vehicle monitoring for driver training in bus companies - Application in two case studies in Portugal. *Transportation Research Part D*, *18*, 103-109.

Dufec/Sweco. (2018). *Telgegevens 2018 fietsers Kruisstraattunnel*. For the Municipality of Eindhoven, June 2018.

Edworthy, J. & Adams, A. (1996). *Warning design: an integrative approach to warnings research*. London: Taylor & Francis.

Gates, T. J., Qin, X. K., & Noyce, D. A. (2008). Effectiveness of Experimental Transverse-Bar Pavement Marking as Speed-Reduction Treatment on Freeway Curves. *Transportation Research Record*. *2056*(1), 95-103.

Gibson, J. J. (1950). *Perception of the Visual World*. Boston, MA: Houghton Mifflin.

Godley, S. T. (1999). *A driving simulator investigation of perceptual countermeasures to speeding*. Melbourne,, Australia: Monash University.

Khorram, B., af Wåhlberg, A. E., & Tavakoli, A. (2020). Longitudinal jerk and celeration as measures of safety in bus rapid transport drivers in Tehran. *Theoretical Issues in Ergonomics Science*. https://doi.org/10.1080/1463922X.2020.1719228

Klauer, S. G., Dingus, T. A., Neale, V. L., Sudweeks, J. D., & Ramsey, D. J. (2009). *Comparing Real-World Behaviors of Drivers with High Versus Low Rates of Crashes and Near-Crashes*. Report DOT HS 811 091. Washington: National Highway Traffic Safety Administration.

Köhler, A.-L., et al. (2019). *Report Infrastructure Measures (Deliverable 3.2)*. Retrieved from MeBeSafe website: https://www.mebesafe.eu/results/.

Lajunen, T., & Summala, H. (1997). Effects of driving experience, personality, driver's skill and safety orientation on speed regulation and accidents. In T. Rothengatter & E. Carbonell Vaya, *Traffic and Transport Psychology: Theory and Application*, pp. 283-294. Amsterdam: Pergamon.

Lippold, C. 1999. Zur Geschwindigkeit V85 als Projektierungsgröße im Straßenentwurf. *Straßenverkehrstechnik 43*(1)

Ljung Aust, M., et al. (2019). *Trial Design (Deliverable 5.1)*. Retrieved from MeBeSafe website: https://www.mebesafe.eu/results/.

Ljung Aust, M., et al. (2019). *Locations ready for field trials (Deliverable 5.3)*. Retrieved from MeBeSafe website: https://www.mebesafe.eu/results/.

Ljung Aust, M., et al. (2020). *Final Measures (Deliverable 5.5)*. Retrieved from MeBeSafe website: https://www.mebesafe.eu/results/.

Manser, M. P. & Hancock, P. A. (2007). The influence of perceptual speed regulation on speed perception, choice, and control: Tunnel wall characteristics and influences. *Accident Analysis & Prevention, 39*(1), 69-78.

Op den Camp, Olaf, et al. (2019). *Report on test vehicles (Deliverable 2.3).* Retrieved from MeBeSafe website: https://www.mebesafe.eu/results/.

Pereira, M., Beggiato, M. & Petzoldt, T. (2015). Use of adaptive cruise control functions on motorways and urban roads: Changes over time in an on-road study. *Applied Ergonomics 50*, 105-112.

Quimby, A., Maycock, G., Palmer, C., & Grayson, G. B. (1999). *Drivers' Speed Choice: An In-depth Study*. TRL Report 326. Crowthorne: Transport Research Laboratory.

Tapp, A., Pressley, A., Baugh, M., & White, P. (2013). Wheels, Skills and Thrills: A social marketing trial to reduce aggressive driving from young men in deprived areas. *Accident Analysis and Prevention, 58*, 148-157.

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

af Wåhlberg, A. E. (2006). Speed choice versus celeration behavior as traffic accident predictor. *Journal of Safety Research, 37*, 43-51.

af Wåhlberg, A. E. (2007). Long-term prediction of traffic accident record from bus driver celeration behavior. *International Journal of Occupational Safety and Ergonomics, 13*, 159-171.

af Wåhlberg, A. E. (2008). The relation of non-culpable traffic incidents to bus drivers' celeration behavior. *Journal of Safety Research, 39*, 41-46.

af Wåhlberg, A. E., et al. (2019). *Final coaching scheme (Deliverable 4.3).* Retrieved from MeBeSafe website: https://www.mebesafe.eu/results/.