Department for Transport

TRL THE FUTURE OF TRANSPORT

PUBLISHED PROJECT REPORT PPR2009

Driver2020 - an evaluation of interventions designed to improve safety in the first year of driving

Report D1: Effectiveness of interventions delivered to novice drivers

J Weekley, S Helman, S Chowdhury, J Hammond, J Hutton

DRIVER 2020

Report details

Report prepared for:	Department for Transport
Project/customer reference:	1-170
Copyright:	© TRL Limited
Report date:	April 2024
Report status/version:	6
Quality approval:	Lucy Fuller **(Project Manager)** Graham Grayson **(Technical Reviewer)**

Disclaimer

This report has been produced by TRL Limited (TRL) under a contract with the Department for Transport. Any views expressed in this report are not necessarily those of the Department for Transport.

The information contained herein is the property of the Department for Transport. Whilst every effort has been made to ensure that the matter presented in this report is relevant, accurate and up-to-date, TRL Limited cannot accept any liability for any error or omission, or reliance on part or all of the content in another context.

Acknowledgements

Firstly, TRL would like to thank the Department for Transport for funding and commissioning this research. Secondly, we would like to thank the delivery partners who assisted in the design of the interventions, and who delivered them during the Driver2020 trial; these were The Royal Society for the Prevention of Accidents, The Driving Instructors Association, Trak Global, DriverMetrics, and Agilysis. Further thanks are due to the following individuals who assisted in the design of the trial: Dr Graham Grayson, Professor Frank McKenna, Professor Steven Stradling, Dr Lisa Dorn, Professor Nicola Christie, Professor Jane Hutton, Professor Johnathon Ehsani, Professor Alan Tapp, Professor Mark Horswill, and Dr Willem Vlakveld. Finally, the following organisations served on the stakeholder panel that contributed insights around how the trial might best support policy: The Welsh Government, Transport Scotland, BRAKE the road safety charity, National Highways, The Parliamentary Advisory Council for Transport Safety, The National Fire Chiefs Council, National Police Chiefs Council, RAC Foundation, Road Safety GB, The Driver and Vehicle Standards Agency, The Driver and Vehicle Licensing Agency, Transport for London, and The Association of British Insurers.

Table of Contents

Executive summary

Background

It is a known challenge in road safety that young and novice drivers are at a greater risk of involvement in collisions than more experienced drivers, for reasons associated with their younger age, and their relative lack of on-road experience. A large evidence base confirms this (see Appendices A.1 and A.2).

Historically there have been two broad approaches to reducing collision risk for young and novice drivers. The first is to use the licensing system to directly target age and experience factors, using legislation to set a minimum age for licensure or minimum periods of time for people to spend in the supervised learning phase, and to set restrictions on higher risk driving situations after solo driving begins. There is a strong evidence base showing that this approach can be effective at reducing risk in other jurisdictions (see Appendix A.3). One post-test measure in the Great Britain system is the two year probationary period during which new drivers on their first full licence who accumulate six or more penalty points have their licence revoked and are required to go through the learning process again. The second broad approach is to use non-legislative approaches such as driver training and education to try to equip drivers with better skills and knowledge to keep safe. There is less evidence that this approach is effective (see Appendix A.4).

The Department for Transport (DfT) commissioned TRL to conduct the Driver2020 project, which sought to build evidence for educational, training and technology-based non-legislative approaches by evaluating the most promising of these in a randomised controlled trial reflecting a potential real-world voluntary rollout of such measures.

The Driver2020 project

In a previous DfT-funded review of evidence, including a consultation with stakeholders, Pressley et al. (2016) identified five interventions that showed promise, and were the most suitable for trialling in a research programme committed to in the 2015 Road Safety Statement. Two of these were designed to be delivered during the post-test period, when people have passed their practical driving test and have begun solo driving. These interventions were:

- **Mentoring agreements:** a set of materials for use by novice drivers and mentors (for example parents) in voluntarily setting restrictions on early post-test driving (for example driving with peer-age passengers, or driving at night)
- **Telematics:** an app-based intervention designed to provide feedback to novice drivers on their driving style, with modest incentives (weekly treats such as coffee vouchers, and a monthly entry into a prize draw to win a month’s free car insurance) provided for safer driving.

This report presents findings from the novice driver arm of the DfT-funded Driver2020 study, in which the effectiveness of these two interventions was evaluated. Three other reports present: findings on three additional interventions delivered to drivers in the learner stage, findings from qualitative research with users and providers to understand engagement with the interventions, and a summary report for the whole project (Weekley et al.2024b; Hitchings et al., 2024; Helman & Weekley, 2024).

The two research questions for this element of the evaluation were (for each novice intervention):

- How effective is the intervention at reducing collisions in the first 12 months of post-test driving?
- Does engaging with the intervention lead to other changes in relevant surrogate measures? (Surrogate measures are outcomes used as substitutes for collisions either because they are known to be related to collision risk – for example speeding – or they reflect things that are targeted by an intervention).

Method

The novice driver arm of the study used a randomised encouragement design. This meant that participants who registered their interest in taking part were randomly assigned to one of three groups, with two providing the opportunity to engage with an intervention (mentoring agreement or telematics), and one being a control group that was offered no intervention. One specific requirement for the project was that participants should not be provided with overly generous incentives to engage with the intervention to which they were assigned; incentives were required to be commensurate with those that might be expected to be sustainable in any wider roll-out of a given intervention, meaning in practice that they were very modest.

The novice arm of the study was advertised to people using two main methods. The first was an email sent to randomised samples of test passers (who had previously opted into research communications) by the Driver and Vehicle Standards Agency (DVSA). The second was a leaflet shown to test passers at the 100 busiest driving test centres across Great Britain. Most (84%) of the 12,307 participants who registered their interest in the trial were recruited via the former method, with the remainder finding out through the latter, or through other means.

In order to be eligible for the study, participants needed to be aged between 17 and 24 years, have passed their practical driving test no more than four weeks before registering, and either have, or expect to have, access to a vehicle for driving for the following 12 months. If assigned to an intervention group, participants were put in touch with their intervention provider following registration. All participants were also asked to complete a survey which asked them about their experience learning to drive, and further surveys 3, 6 and 12 months after their test pass date. These later surveys asked about their experience driving post-test, about collisions they had been involved in, and about several other surrogate measures of safety, some of which were related specifically to outcomes targeted by the interventions. Participants were given a £5 retail voucher (or equivalent charity donation if they preferred) for each survey they completed and were entered into a prize draw if they wished, to win a year’s car insurance and other prizes such as retail vouchers and tablet computers.

Recruitment for participants began in October 2019 and continued until early January 2021. Therefore, there were a number of ways in which recruitment, learning to drive, and post-test driving experience were impacted by the COVID-19 pandemic and associated lockdowns and restrictions. These factors were considered and controlled where possible in the analysis.

Analysis

When all participants had completed their 12 months post-test period, two analyses were run to understand the effectiveness of the interventions.

First, the study investigated how effective each intervention was at reducing collisions in the 12 months post-test period; this was assessed by comparing the average number of self-reported collisions in each intervention group with that of the control group. This analysis is sometimes referred to as an 'intention to treat' analysis; it included all participants (for whom there were survey data) in each group, regardless of whether they engaged with the intervention. Such an approach aims to understand what would happen at the population level in real-world conditions, rather than looking specifically at people who follow the treatment protocol as intended.

The second analysis is known as a 'per-protocol analysis' – so-called as it includes only those people who followed the treatment protocol as intended – so looking only at the impact of the intervention on those people who were known to engage to some degree. This analysis used the surrogate measures as outcomes, since the sample sizes were much too small to permit any robust analysis of collision numbers. Because the numbers of people who engaged were low, the study compared relevant surrogate measures – which matched the intended outcomes of the interventions – between the intervention and control groups.

For the mentoring agreement intervention, engagement was defined as setting at least one agreement; for the telematics intervention, engagement was defined as downloading and using the app to log at least one journey.

Limitations

Several limitations to the study design should be considered when reading the findings and conclusions.

- The data were gathered during the COVID-19 pandemic, which would have affected learning and post-test driving. Generalisability to future circumstances in which there is no global pandemic may be affected.
- The sample is biased towards females and is an opportunity sample, which limits generalisability.
- The self-selecting nature of the sample (both in terms of registering for the trial and in terms of providing survey data) means that the findings cannot necessarily be generalised to all novice drivers in Great Britain. However, the design (randomised allocation to groups) does mean that self-selection bias should not play any role in group differences.

- Most data are self-reported meaning some social-desirability effects may be present, and data reflect the interpretation of participants to survey questions.
- The engagement with interventions was at a very low level, meaning the findings might not reflect what would happen with greater engagement.
- The findings related to surrogate measures provides us with evidence of the behaviour change that results from the interventions as these analyses only include those who engaged; however, the very low numbers engaging with some interventions mean that some very small effects may have been missed.
- The findings only cover the first year of driving and cannot be extended beyond this period (for example we cannot say whether behaviour would be impacted beyond this period, or whether any changes seen would persist).

Findings

The findings from the study are reported below, first for the effect of the interventions on collisions, and then on surrogate measures. It should be noted that the COVID-19 pandemic affected either learning to drive, post-test driving, or both during the study for the majority of participants. Thus, some caution should be exercised in how generalisable the findings are to future circumstances.

Collisions

- Greater mileage was associated with more collisions.
- As novice drivers accumulated on-road experience the increased risk associated with increased mileage reduced; in other words, although mileage overall increased risk, the increase was not linear.
- The COVID-19 pandemic also led to changes in collision risk (after controlling for mileage). Those who passed their test before the pandemic were more likely to be involved in collisions than those who passed after. Those who reported that their learning was impacted more by the pandemic were also more likely to be collision-involved.

The findings from the intention to treat analysis were:

- There was no statistically significant difference between the mentoring agreement (N=2,356), telematics (N=2,234), and control (N=2,258) groups on the main outcome measure of self-reported collisions in the first 12 months of post-test driving.
- This was true after controlling for mileage driven and for the impacts of the COVID-19 pandemic, age, and gender.

The lack of an effect on collisions is not surprising given the low engagement rates, since even if the interventions were extremely effective for those who engaged, they would only be having an effect on a small proportion of the intervention group; an analogy is that if a medicine exists that cures a disease with 100% effectiveness, and only 1% of people take the medicine, then the real world effectiveness of the medicine will be 1%. Unfortunately, it is not possible to examine differences in collisions solely in the people who engaged with the interventions, as the samples sizes were insufficient to support such an analysis; this is why surrogate measures were used for the per-protocol analysis (see below).

Surrogate measures

The lack of differences in collisions between groups lends greater importance to the per-protocol analyses, which look only at those people known to have engaged to some degree with the interventions, on a range of surrogate measures targeted for change by those interventions. It needs to be noted that the engagement 'dose' or 'amount' of intervention exposure was low even in those people who did engage, and it may be that the effects seen were limited by this.

The per-protocol analyses showed the following effects for those who engaged with the interventions (all findings relative to the control group):

For the mentoring agreement intervention:

- A statistically significantly higher proportion of engagers reported that they set voluntary agreements limiting their post-test driving with peer-age passengers in their first six months of post-test driving. Proportions reporting limits were 10% and 12% at three-months (N=145) and six-months (N=138) post-test for engagers, compared with 5% and 3% for control participants (N=2,908 and N=2,669 respectively). (Would likely improve safety)
- A statistically significantly higher proportion of engagers reported that they set voluntary agreements limiting their post-test driving in the dark in their first six months of post-test driving. Proportions reporting limits were 13% and 15% at three-months (N=145) and six-months (N=138) post-test for engagers, compared with 7% and 5% for control participants (N=2,908 and N=2,669 respectively). (Would likely improve safety)
- Engagers reported a statistically significantly lower proportion of their driving as being in the dark in months 4-6 post-test (26% for engagers, N=138, versus 30% for the control participants, N=2,669). (Would likely improve safety)

For the telematics intervention:

- Engagers (N=669) reported a driving style that was statistically significantly more 'inattentive, careless, irresponsible, and risky' than control participants (N=2,447). (Would likely reduce safety)
- Engagers reported a statistically significantly lower proportion of their driving as being in the dark at months 7-12 post-test (36% for engagers, N=514, versus 38% for control participants, N=1,752). (Would likely improve safety)

Conclusions

The novice driver arm of the Driver2020 project aimed to test the real-world effectiveness of two interventions on both collision involvement, and a number of surrogate safety measures, in the first 12 months of post-test driving.

Neither the mentoring agreement intervention, nor the telematics intervention, led to any change in self-reported collision involvement in the Driver2020 novice driver sample. Therefore the findings from this sample suggest that if either of these interventions were offered on a similar voluntary basis as in this study to novice drivers aged 17-24 in Great Britain, they would be unlikely to lead to any measurable reduction in collision risk.

An important finding that helps in the interpretation of this main conclusion is that when such interventions were offered under conditions of voluntary engagement, levels of engagement were extremely low.

When only participants who had engaged with the interventions were considered, both interventions were associated with statistically significant changes in relevant surrogate measures, relative to the control group.

Further work with the Driver2020 dataset could help further elucidate the effects of the COVID-19 pandemic on learning to drive and on early post-test driving in young and novice drivers. It also provides a recent dataset that can be used to examine this group, and the long-understood road safety challenge they present.

1 Introduction

1.1 The challenge of young and novice drivers

This section summarises the known road safety challenge presented by young and novice drivers, and a short description of research with this group including the programme of applied work delivered in Great Britain throughout the 1990s and 2000s, that laid the foundations for the approach taken in the Driver2020 project. It is a summary of the more detailed background provided (with references) in Appendix A.

The conclusion that emerges from this evidence is that young and novice drivers have an elevated collision risk for reasons associated with their youthfulness and their inexperience. In short, younger drivers and less-experienced drivers are at greater risk. Successfully reducing risk requires interventions that act on one or both areas.

Approaches that bring about an increase in the age at which someone becomes licensed have reduced collisions. The same is true of approaches that provide drivers with greatly increased levels of on-road experience before licensure. Finally, approaches that limit exposure to the riskiest situations in early driving after licensure – allowing experience to build up in lower risk situations before access to higher risk situations is granted – are effective in reducing collisions. In Great Britain, the Road Traffic (New Drivers) Act 1995 mandates a two year probationary period during which new drivers on their first full licence who accumulate six or more penalty points have their licence revoked and are required to go through the learning process again. Two years after passing their first driving test, or on passing again after having their licence revoked under the Road Traffic (New Drivers) Act 1995, drivers enter the full licence stage. Note that driver licensing is devolved in Northern Ireland and so this project covers Great Britain only.

Many interventions based on education, technology and training have not been developed and delivered in an evidence-led way. Currently, there are no standards or guidance for road safety education, so interventions can vary hugely in quality, delivery and content. They have delivered poor results because they either target things that may not be relevant for safety, or they target things that are relevant, but do so inadequately. An example of the former is the traditional 'skills' approach to driver training. In the past this has failed to bring about safety improvements due to a focus on specific vehicle control skills such as 'recovering from skids' on the assumption that improving such skills will reduce collision risk. An example of potentially inadequate targeting can be found in attempts to provide drivers with knowledge about risk on the assumption that this knowledge will lower drivers' collision risk by bringing about changes to their driving behaviour. The assumptions underlying these approaches have turned out to be incomplete. Vehicle control skills do not appear to be adequate to ensure safety, and changing behaviour requires more than just a provision of information about risk. Approaches that have fared better include hazard perception training, which focuses on the skill of anticipating potential hazards on the road ahead, although the literature on such approaches is still relatively immature in so far as being able to demonstrate a direct link from the training to collision outcomes.

Some training and education approaches have even been associated with increased crash risk. There are plausible mechanisms that can explain how this can happen. For example, it is believed that some skid training courses have led to some drivers taking additional risks, as they assume that their new training will keep them safe; in reality the skills fade quickly without practice, and therefore this confidence is unfounded. Some education approaches have resulted in people gaining access to driving earlier (and therefore at a younger age) than would otherwise have been the case, putting them at more risk as a result.

In light of the relative lack of success of many existing interventions based on education, technology and training, even in graduated licensing systems, research attention is now revisiting such approaches such that they can support other policy approaches. The hope is that if interventions can be based on sound behavioural science, and can be focused on the right mechanisms, they can add value.

1.2 Origins of the project

In the 2015 Road Safety Statement (DfT, 2015) the Government committed to “Undertaking a £2 million research programme to identify the best possible interventions for learner and novice drivers” (p8). The Statement also set the context for this research programme by stating the following in paragraphs 1.11 and 1.12:

> “1.11: Ten years ago, there were fewer options for reducing the elevated collision risk within the young driver population. Many foreign governments placed legislative ‘graduated driver licencing’ restrictions on their young people. These options include restricting driving to the hours of daylight or not allowing the carriage of passengers, for months or even years after passing tests.
>
> “1.12: Technology is one of the ways that we can help young drivers be safer. Technology is now emerging that can manage novice driver risk in a more bespoke way without restricting the freedoms of all of our young people. In short, there are modern and sophisticated non-legislative alternatives that treat each young driver as an individual with their own distinct risk profile.”

The work that underpinned the commitment to the research programme that became the Driver2020 project was therefore focused on finding non-legislative interventions that were best suited for trialling.

1.3 New approaches to reducing collision risk

The year after the publication of the Road Safety Statement, the Department for Transport (DfT) commissioned a review of interventions for young and novice drivers, focusing specifically on identifying interventions based on education, training and technology that showed promise either theoretically (for example being based on accepted models of learning or behaviour change) or based on evaluation data on some relevant surrogate measure of risk (for example driving behaviour) (Pressley et al., 2016). The intention was that these interventions could then be evaluated in a controlled trial in Great Britain, looking specifically at their effectiveness in reducing collisions.

The brief for the Pressley et al. project specifically excluded approaches that are legislative in nature and focused on non-legislative options.

In this way the Pressley et al. review was the work that laid the ground for the Driver2020 project; the intention was that the candidate interventions identified would subsequently be tested in a large-scale controlled trial, using collisions as the main outcome measure.

Pressley et al. concluded that there were seven approaches that showed promise. Three of these involved parental engagement in the learning and early driving of young and novice drivers, three involved the use of telematics or app-based approaches to monitor and manage driving risk, and one was hazard perception training. All were found to show promise either based on existing evidence of changes in relevant surrogate measures or based on linking to a known risk factor for young and novice drivers in a theoretically coherent and plausible way.

1.4 The Driver2020 project

Following a stakeholder workshop on the feasibility of the interventions identified in Pressley et al. (2016), four were recommended for trialling.

Two of these were designed to be applied in early post-test driving (“novice driver interventions”):

- **Mentoring agreements:** a set of materials for use by novice drivers and mentors (for example parents) in voluntarily setting restrictions on early post-test driving (for example driving with peer-age passengers, or driving at night)
- **Telematics:** an app-based intervention designed to provide feedback to novice drivers on their driving style, with various incentives provided for safer driving.

Two were designed to be applied during the learning stage (“learner driver interventions”):

- **Logbook:** an app designed to encourage more on-road practice, covering a broader range of driving conditions and road types, during the learning period
- **Hazard perception training:** a set of three e-learning modules designed to improve hazard perception skill.

Following the recommendations from Pressley et al., a decision was also taken by the Department for Transport to include an education-based intervention for trialling (to be applied during the learning period). The reasoning for this was that, despite there being limited evidence uncovered in the Pressley et al. review, delivery capacity for this type of intervention exists in Great Britain, and there was already widespread delivery. Driver2020 provided an opportunity to develop an education intervention informed by behaviour change techniques and targeted at learner drivers. It was therefore decided that if strong content, informed by behaviour change theory, could be designed, then evaluating this alongside the other approaches would be of value:

- **Classroom-based education:** a whole-day classroom-based intervention in which attendees take part in several activities designed to equip them with knowledge and skills, and ongoing self-monitoring strategies, to make them safer as drivers.

The Driver2020 project was therefore a national trial to evaluate the effectiveness of these five interventions. The project was procured in 2017 and was registered at the ISRCTN registry (https://www.isrctn.com/ISRCTN16646122).

1.5 Research questions

The main research question for each intervention was as follows:

- How effective is the intervention at reducing collisions in the first 12 months of post-test driving?

A secondary research question for each intervention was:

- Does engaging with the intervention lead to other changes in relevant surrogate measures?

For all interventions, due to the fact that participants were free to choose to engage, the following research question was also asked:

- What were the factors that led people to engage with the intervention, or the barriers that stopped them engaging?

1.6 This report

This report presents the findings from the novice driver arm of Driver2020 project. Four reports cover the whole project:

- D1 (this report) – Effectiveness of interventions delivered to novice drivers (Weekley et al., 2024a). Presents analysis and findings from the quantitative evaluation of the novice driver interventions.
- D2 – Effectiveness of interventions delivered to learner drivers (Weekley et al., 2024b). Presents analysis and findings from the quantitative evaluation of the learner interventions.
- D3 – Delivery of interventions and engagement by novice and learner drivers (Hitchings et al., 2024). Presents analysis and findings from the qualitative evaluation using interviews with participants and delivery partners for both novice and learner interventions.
- D4 – Summary of findings (Helman & Weekley, 2024). Overall project report summarising the key findings from each part of the trial.

Please note that these four references are cross-referenced in the reports (including this one) when it is useful to do so; however, they have been published at the same time as each other as part of the Driver2020 project, rather than being part of the wider existing literature. A supplementary appendix containing the data collection surveys and intervention logic models is also available (Weekley & Helman, 2024).

This report is structured as follows:

Section 1 has described the background to the Driver2020 project – covering its origins within policy, the aim of the project and the research questions. Section 2 describes the overall method for the study in terms of its high-level design and approach. The impact of the COVID-19 pandemic on the trial is also discussed in this section. Note that much of Sections 1 and 2 are included in all project reports, with adjustments relating to the focus of the report (for example in this document – COVID-19 impact is described mainly in terms of its impact on the novice arm).

Sections 3, 4 and 5 then discuss the novice driver interventions only. Section 3 describes the method used in the novice driver arm of the study, including the detailed design, the interventions, materials and recruitment. Section 4 describes the findings. Section 5 discusses these findings and draws conclusions.

2 Method – overall study

2.1 Design

The trial had two arms. One arm tested the effectiveness of the novice driver interventions (mentoring agreements, telematics). The other tested the effectiveness of the learner driver interventions (logbook, hazard perception training, classroom-based education). See section 1.3 for high level descriptions of these interventions and section 3.3 for fuller descriptions. Figure 2-1 and Figure 2-2 show (again at a high level) the route participants took through the study. Participants were only able to sign up for one arm of the trial.

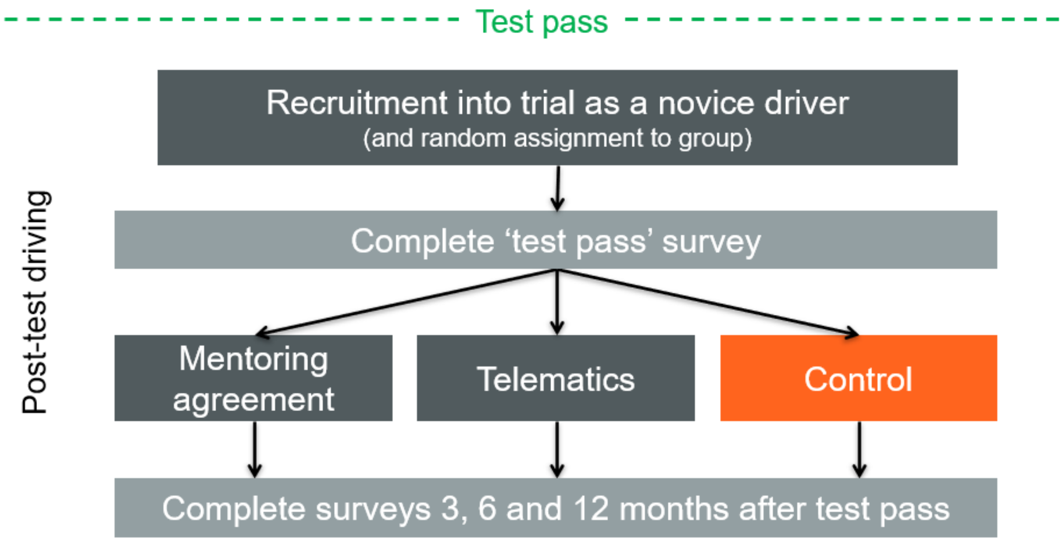


Figure 2-1: Design – novice driver arm

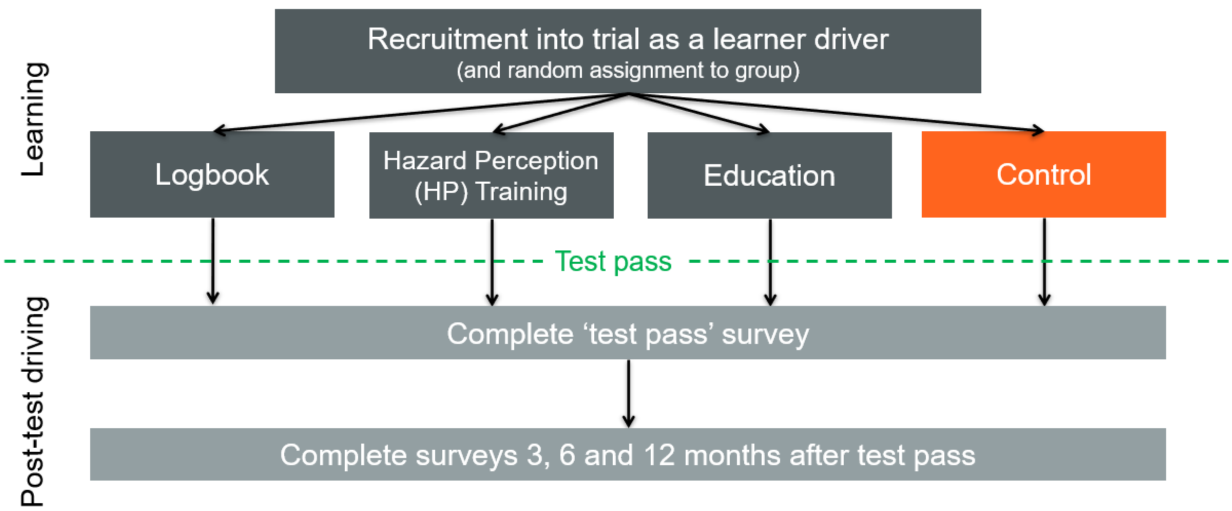


Figure 2-2: Design – learner driver arm

Because participants were assigned randomly to treatment or control groups in their respective arm of the study, the Driver2020 project is a type of randomised controlled trial (RCT). Incentives were required to reflect those that might be expected to be sustainable in any wider roll-out of a given intervention, meaning in practice that they were very modest.

The study could thus be better described as using a ‘randomised encouragement’ design (West et al., 2008). Unlike in classic RCTs in which participants are expected to adhere fully to a treatment protocol (and are given incentives to ensure this), in this design participants were randomly assigned only to an opportunity or ‘encouragement’ to receive an intervention. Crucially, the participants chose whether or not to engage with the intervention – and if they did engage, to what extent. This approach is useful when the incentives required to guarantee engagement would be unrealistic, undermining a generalisation of the findings to real-world roll-out. It is also useful when testing a voluntary (rather than legislatively-enforced) behaviour, as it allows exploration of engagement.

Thus, in simple terms Figure 2-1 shows that:

- Novice drivers were recruited into the trial shortly after passing their practical driving test, and then randomly assigned to either the mentoring agreement, telematics, or control group. They decided whether and how much to engage with any intervention they were offered during the next 12 months.
- They were immediately asked to complete their test pass survey on being recruited (because they had already passed their practical driving test)
- They were then asked to complete further surveys 3, 6 and 12 months after the date they passed their practical driving test.

Figure 2-2 shows that:

- Learner drivers were recruited into the trial and were randomly assigned to either the logbook, hazard perception training, classroom education, or control group. They decided whether and how much to engage with any intervention they were offered during their learning period (which varied in length, depending on the participant).
- They were asked to complete their test pass survey shortly after passing their practical driving test (if they passed within the timescales of the study) .
- They were then asked to complete further surveys 3, 6 and 12 months after passing their practical driving test.

2.2 Approach to research questions

2.2.1 Research questions related to impact of the interventions

The main research question for each intervention was as follows:

- How effective is the intervention at reducing collisions in the first 12 months of post-test driving?

This was assessed by comparing the average number of self-reported collisions (controlling for other factors such as mileage) in the first 12 months of post-test driving in each intervention group with that of the corresponding control group.

A secondary research question for each intervention was:

- Does engaging with the intervention lead to other changes in relevant surrogate measures?

This was assessed by comparing the relevant surrogate measures for each intervention group (these differed by group and were largely self-reported) with the same in the corresponding control group. Surrogate measures included things like driving style (targeted by the telematics intervention), and the setting of limits on driving in some post-test contexts (targeted by the mentoring agreements).

In Great Britain it is not feasible to identify individual drivers in official casualty data (STATS19). This was the main reason why self-reported survey measures were used to collect data. The reliance on self-reported measures in answering surveys is also a cost-effective approach with such large samples – in total around 27,000 people registered into the study, though not all completed data collection. This approach may be subject to some measurement error and bias (for example people may respond in a socially-desirable way) and these limitations need to be taken into account when interpreting the findings; however, while not perfect, this approach has offered many insights in previous studies of this kind (e.g. Wells et al. 2008; Forsyth et al. 1995; Maycock et al. 1991).

2.2.2 Research question relating to engagement with the interventions

For all interventions, due to the fact that participants were free to choose to engage, the following research question was also asked:

- What were the factors that led people to engage with the intervention, or the barriers that stopped them engaging?

This was assessed through undertaking interviews with participants in each of the intervention groups who were known – usually through system data from each intervention provider – to have engaged to varying degrees (including some who had not engaged at all). Findings from these interviews – and from interviews with delivery partners – are reported in Hitchings et al. (2024).

The key focus of the engagement research was to complement the effectiveness evaluation, so that any interventions found to be effective could be rolled out with the best chance of appealing to or being encountered by young and novice drivers.

2.3 Timeline of the study

Figure 2-3 shows the relative timings of the recruitment and delivery of the learner and novice arms, including in relation to the COVID-19 pandemic.

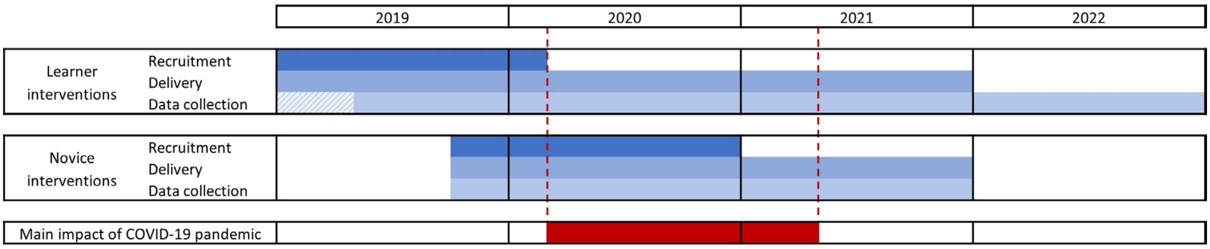


Figure 2-3: Timings for recruitment and delivery of the learner and novice interventions in relation to the COVID-19 pandemic

Recruitment for learners began in January 2019 and continued until early March 2020 (prior to the start of the COVID-19 pandemic). The delivery period for the learner interventions began at the same time and continued until December 2021, at which point delivery to those learner participants who had not yet passed their test was stopped. Data collection continued until December 2022 to allow 12 months post-test driving for those participants who passed their test close to this date.

Recruitment for novices began in October 2019 and continued until early January 2021. Delivery of the novice interventions began immediately and continued until December 2021 when the final novice participant completed 12 months post-test driving (and associated data collection).

Figure 2-3 also shows the timing of the main impact of the COVID-19 pandemic in relation to the recruitment and delivery of the interventions; this was from the beginning of the first lockdown in March 2020 until the end of the third lockdown in early May 2021. However, it should be noted that the impact of the pandemic, particularly on the waiting time for practical driving tests and therefore the pass rate of learner participants, continued beyond this point and was still ongoing at the end of the trial.

Details of the impact of the pandemic on the novice arm of the study and how these were addressed is discussed in the next section. The impact on the learner arm is discussed Weekley et al. (2024b) which reports the results of the learner intervention trial.

2.4 The impact of the COVID-19 pandemic on the novice driver arm

2.4.1 Practical tests, theory tests and formal driving lessons (with Approved Driving Instructors (ADIs))

A key impact of the pandemic on the novice arm was that practical driving tests were suspended (for all except key workers) during the periods of lockdown. These dates varied by country (and by region) and are summarised below in Table 2-1.

Table 2-1: Dates of practical driving test suspension in England, Scotland, and Wales during the COVID-19 pandemic

Practical (car) driving test	England	Scotland	Wales
1st lockdown			
Suspended	20 March 2020	20 March 2020	20 March 2020
Started	22 July 2020	14 September 2020	17 August 2020
2nd lockdown			
Suspended	5 November 2020	Not suspended	24 October 2020
Started	2 December 2020	Not suspended	9 November 2020
3rd lockdown			
Suspended	5 January 2021	26 December 2020	20 December 2020
Started	22 April 2021	6 May 2021	22 April 2021

Before and after the 2nd lockdown, all three countries were subject to localised restrictions known as tiers (England), protection levels (Scotland) and alert levels (Wales). In addition to the nationwide lockdowns listed above, practical driving tests were also suspended when an area was in Tier 4 / Protection Level 4 / Alert Level 4, which happened at various times.

Theory tests were also suspended (for all except key workers); the dates for these were broadly similar to the dates for practical tests.

Formal driving lessons were also cancelled over the same period. The advice for most driving school instructors and independent instructors was that lessons should not go ahead except for key workers who were preparing for an essential driving test.

Suspension of the practical driving test had a significant impact on novice recruitment (and an impact on the participants in the learner arm – see Weekley et al., 2024b). As participants were recruited shortly after they passed their test, recruitment was effectively on hold each time the practical tests were suspended. However, this did not have an impact on individual novice participants, as once recruited they were unaffected by the suspensions.

Depending on their recruitment date however, their learning to drive process may have been affected; all those who were recruited after March 2020 may have experienced an impact – the cancellation of their tests (if booked), cancellation of lessons, and the inability to book and take the test when ready.

In March 2020, additional questions were added to the test pass survey to ask all participants who passed from that point on about the impact of the pandemic on their learning to drive process (for survey details, see the supplementary appendix document (Weekley & Helman, 2024)). Questions were also added to the topic guides for the interviews in the qualitative evaluation, reported in Hitchings et al. (2024).

2.4.2 *Reduced levels of driving*

As well as formal suspensions, novice participants may have been affected by reduced levels of driving more generally. For example, during lockdowns, only essential journeys were permitted by restrictions.

For almost all participants this will likely have had an impact on the amount of driving experience they were able to get during this period (unless for example they were key workers) – either in their post-test driving, in their learning to drive process, or both.

Impact on the learning- to-drive process has been identified through additional questions in the test pass survey as mentioned above. Impact on post-test driving has been estimated using self-reported measures of mileage during the survey reporting periods.

2.4.3 *Implications*

None of these impacts should introduce any bias into the study results as all groups (control and intervention) were subject to the same circumstances. Nonetheless the analysis is more complicated to account for changes at different times for the participants (in exposure to driving, for example) and this is discussed further in later sections of this report and in Weekley et al. (2024b) which covers the learner driver interventions.

Another implication of the impacts was on the original intention in the study to achieve an approximate alignment in time of the post-test driving periods in the learner and novice arms. COVID-19 disrupted recruitment (and test-passing) substantially, and this means that the alignment was much less than anticipated.

The generalisability of the findings to future circumstances (where there is no pandemic) may be affected; the pandemic will have significantly altered the learning to drive process and post-test driving of most of our trial participants. For example, the time to obtain a licence for this cohort is likely to have been significantly increased compared with what has been observed before. This provides both challenges for the current study and an opportunity to gain further insight into the effect of age and experience on young drivers' collision risk.

3 Method – novice driver arm

3.1 Participants and recruitment

The eligibility criteria for novices were that they were aged between 17 and 24 years, had passed their practical driving test less than four weeks ago, and reported having or planning to have some access to a vehicle for driving for the following 12 months. The various recruitment methods directed participants to the registration survey, which assessed whether participants met these criteria.

The target sample size was 2,036 per group (this achieving 80% power to detect a 20% reduction in collision involvement from an expected baseline of 15%; so a change from 15% of drivers having a collision within their first 12 months of driving, to 12%).

The main recruitment mechanism was through an email invitation sent by the Driver and Vehicles Standards Agency (DVSA). This email invitation (see Appendix A) was sent to multiple random samples of between approximately 6,000 and 12,000 17-24-year-olds who had recently passed their test (and indicated that they were open to taking part in research). This was sent to a new sample approximately every two weeks throughout the recruitment period (excluding March to September 2020 when driving tests were suspended). The sample size was increased where possible following the pandemic in order to boost recruitment numbers. A second recruitment method was through leaflets distributed to the 100 busiest test centres in Great Britain. Test centre managers and driving examiners were also given digital versions of the leaflet. Digital and physical versions of the leaflet were shown to novice drivers after they had passed their test, by the examiner.

Novice participants were asked in the registration survey how they had heard about the study; 84% of registrants indicated that they had been invited via the email from DVSA, while 9% stated that they had been shown a leaflet on passing their practical test; the remaining 7% of participants stated that they had heard about the trial through an alternative mechanism, for example a friend, driving instructor, social media, or other. Figure 3-1 shows the number of participants registered per month, clearly showing the impact of the pandemic with almost no registrations between April and August 2020.

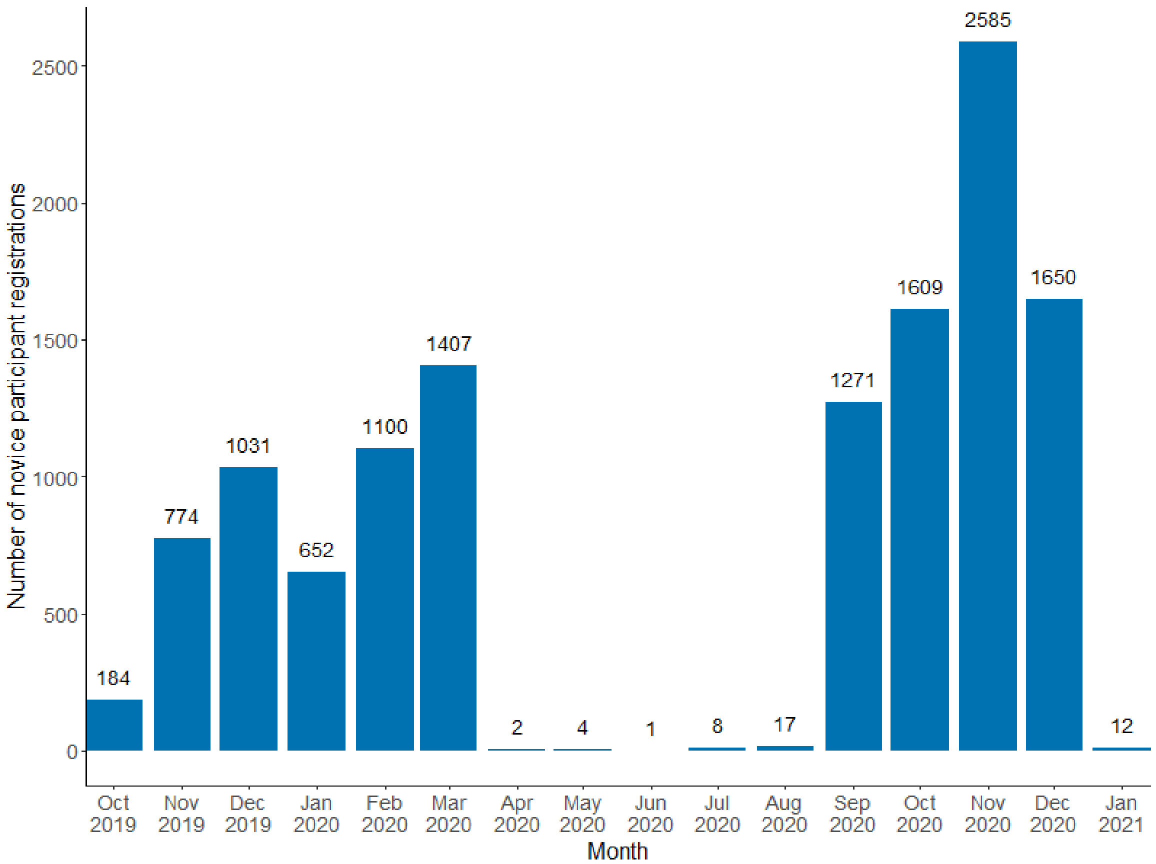


Figure 3-1: Number of novice participant registrations per month during the recruitment period (Oct 2019 to Jan 2021)

The total number of novice participants recruited into the novice driver arm of the study was 12,307 across Great Britain. The gender distribution and mean age for each group of recruited participants is shown in Table 3-1.

Table 3-1: Mean age and gender distribution for the total number of novice participants recruited

Group	Total recruited	Female	Male	Mean age
Mentoring agreement	4183	60%	40%	18.84
Telematics	4173	60%	40%	18.89
Control	3951	61%	39%	18.87
Total	12,307	60%	40%	18.86

Please note that this is all participants recruited, not those who provided data through surveys. For demographics of participants who provided data, and detail on those who were included in the final trial sample for analysis, see Section 4.

Participants were provided with a £5 retail voucher (or equivalent donation to charity) for each of the four data collection surveys they completed (detail on the surveys is included in section 3.3.3). In addition, all those who completed the 12-month survey were automatically entered into a prize draw to win one of four top prizes (a year's free car insurance) or other prizes such as iPads and retail vouchers. These incentives were designed to encourage completion of the data collection surveys and minimise drop out of the participants from engaging with this element of the trial. The incentive structure was informed by previous studies including Helman et al. (2017). In line with UK law on prize draws, people were also informed that they could send their contact details to be entered into the prize draw, even if they did not take part in the trial data collection.

3.2 Design

3.2.1 Logic models

The logic models created for the mentoring agreement and telematics interventions are available in the supplementary appendix document for the project (Weekley & Helman, 2024). A logic model is a diagram that shows how an intervention is expected to produce its effects – how it is meant to work (O'Cathain et al., 2019; Smith, Li & Rafferty, 2020).

3.2.2 Group assignment

The design of the study was a randomised encouragement design, meaning that participants in the intervention groups were given the opportunity to engage with their intervention, but were not required to do so. Registered novice participants were randomly assigned to the mentoring agreement, telematics, or control group. The assignment was weighted slightly more heavily towards the mentoring agreement and telematics group than towards the control group, as it was judged that the drop-out rate from these groups may be higher than in the control group (although this did not turn out to be the case). The randomisation therefore proceeded as follows:

- The database assigned a given participant to either the mentoring agreement, telematics, or control group by assigning a study ID number randomly (weighted according to how many 'vacancies' – compared with the target number – in each group remained)
- For every 2,700 control group assignments, 2,850 assignments were made to the mentoring agreement group, and 2,850 were made to the telematics group

Appendix C shows the data flows and way in which novice participants were handled when brought into the study (see also section 3.4).

3.2.3 Analyses

Two analyses were undertaken. First, an ‘intention-to-treat’ analysis compared the number of self-reported collisions in the first 12 months of post-test driving between the three groups, including every participant for whom there was survey data (even if they did not engage with the intervention to which they had been given access). The purpose of this approach is to understand the real-world effectiveness of an intervention, including any effects of non-engagement (Montori & Guyatt, 2001). Generalised linear modelling was used in this analysis. First, a base model was constructed to understand the role of exposure, experience, age, and gender on collisions. This base model was then used to check for the effect of the COVID-19 pandemic, and the interventions. See section 4.4 for a detailed description of the modelling approach.

The two-person team of statisticians undertaking the intention-to-treat analysis was blinded to group identity until after the full analysis had been run. Any variables which might give away group membership (for example the question in the 12-month survey on which interventions the participant had heard of) were removed at this stage of the analysis.

The second analysis was a ‘per-protocol’ analysis looking only at the impact of the intervention on those people who were known to engage to some degree. The purpose of such an analysis is to examine the potential efficacy of the interventions in individuals who choose to engage in the event that the numbers of people who engaged were low. In this analysis some of the non-collision measures in the surveys were used to assess efficacy; some of these were general surrogate measures for safety, and some where those that matched the intended outcomes of the interventions.

The following were used as general surrogate measures for safety:

- Self-reported near misses
- Self-reported Driving Events scale score (items on hazard involvement)
- Self-reported frequency of exceeding the speed limit

For the mentoring agreement intervention per-protocol analysis the following surrogate measures were compared between the mentoring agreement and control groups:

- Self-reported limits set on post-test driving
- Self-reported proportion of mileage driven in the dark, and with peer-age passengers
- Self-reported attitudes to enforcement and to limits for new drivers

For the telematics intervention the following surrogate measures were compared between the telematics and control groups:

- Self-reported proportion of mileage driven in the dark
- Self-reported driving style scores
- Self-reported attitudes to enforcement and to limits for new drivers

It should be noted that surrogate measures were used for the ‘per-protocol’ analysis instead of collisions, as the samples of engagers were far too small to permit a reliable analysis of the collision measure. The sample sizes for engagers were sufficient to support analysis of the surrogate measures, as these do not suffer from the very low baselines/high variability associated with the collision variable (which consequently requires much greater sample sizes for analysis).

Prior to these main analyses, that were focused on the research questions, other comparisons were made between groups to ensure that they were matched on variables such as age, gender and personality variables; this was especially relevant for the per-protocol analysis to check that those who engaged with the interventions were matched with appropriate participants in the control group.

Statistical tests were used throughout this study to test for differences between groups (for instance, control, telematics, or mentoring agreement groups, or between pre-pandemic and post-pandemic groups).

In all cases where statistical tests were used to compare data, the convention from the behavioural sciences of reporting *p*-values was adopted. The probability value, *p*-value, is used to determine statistical significance. However, statistical significance alone can be misleading in trials with very large samples because it is influenced by the sample size. Therefore, this study presents the effect sizes where the results were statistically significant. Effect sizes help understand the magnitude of differences found. Both are essential to understand the full impact of the interventions. Findings are statistically significant only when explicitly stated.

The meaning of an effect size varies on the type of statistical test that is being used; the same effect size value has different meanings for different statistical tests. To give an indication of real-world impact, values of a given effect size statistic are described as 'negligible', 'very small', 'small', 'medium' or 'large'. For more information on statistical significance, effect sizes and details on the specific statistical tests used for each comparison, see Appendix D.

Throughout the analyses, appropriate graphical outputs have been used to present the data and give a quick visual representation of the distribution of values. This includes tables and charts such as histograms, bar and line plots, boxplots and violin plots. Histograms are used to show the frequency of data in particular groups. (Please note that, due to rounding, the sum of the data labels on some histograms does not equal exactly 100%.) Box plots are a visual representation of the inter-quartile ranges (25th and 75th percentiles) and median (50th percentile or 'average') for a continuous variable (values outside of the 95% confidence interval are shown as individual data points). Violin plots are similar to boxplots but also show the density plot (where the width of each curve corresponds with the frequency of data points in each region) of a continuous variable.

3.3 Materials

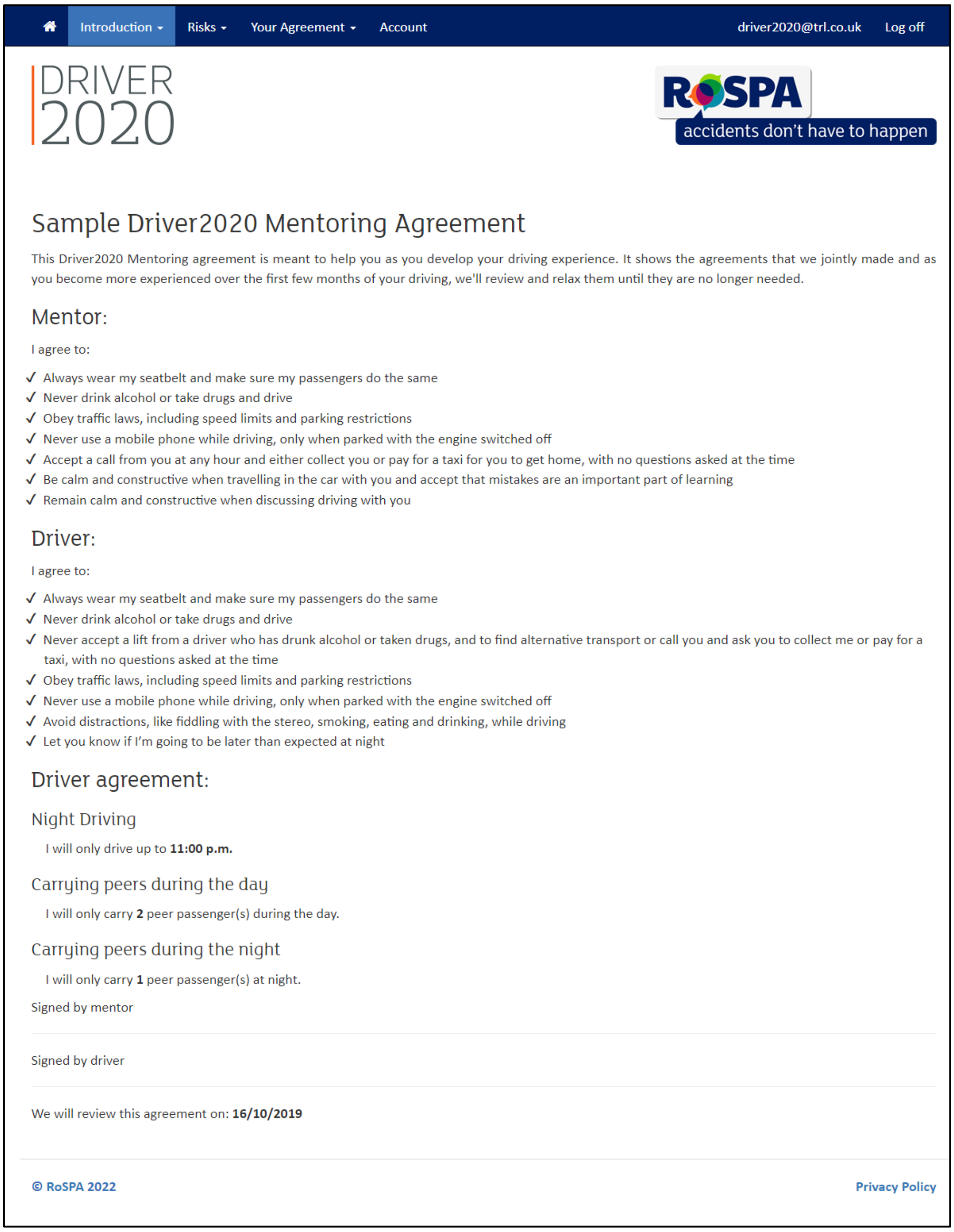
3.3.1 Intervention: Mentoring agreement

The mentoring agreement intervention (delivered by the Royal Society for the Prevention of Accidents - RoSPA) was in the form of a website: www.drivermentoringagreement.com. Participants in this intervention group were invited to login to this website.

The website provided general guidance on reducing risks for new drivers and the role that mentors can play in this; it also provided specific information and guidance on the following risks: night driving; driving with peer passengers; urban roads; rural roads; motorways; distractions; driving in poor weather; drink and drugs; seat belts; and driver fatigue. The intervention was based on the 'Checkpoints' intervention used in the US, described in Pressley et al. (2016).

It also provided detailed information on how the Driver2020 mentoring agreement process worked, and an example agreement and additional guidance and checklists for those completing the process without a mentor. An overview of the process is given below:

1. After signing up, the participant was prompted to complete the first agreement. This involved both the mentor (usually a parent) and driver agreeing to seven pre-set driving conditions, such as obeying speed limits, not drinking (or taking drugs) and driving, and wearing their seatbelts.
2. The mentor and new driver were then expected to discuss some options for conditions that the new driver must drive within for the first agreement period. These were based on the key risks faced by new drivers – driving at night and with peer age passengers (Clarke, Ward & Truman, 2002; Ouimet et al., 2015).
    - Time conditions for night driving (e.g., only driving up to 10pm)
    - Conditions on number of peer passengers carried during the day
    - Conditions on number of peer passengers carried at night.
3. Around a month later, they received an automated text and email reminding them to review their agreement. At this point (or later), they were expected to revisit the agreement and discuss how it went with their mentor, and what experience they had gained. The aim was that the conditions could be relaxed as the driver gains experience.
4. They were then advised to set up a second driving agreement with new driving conditions and set a date to review it around 4 weeks later. As with the first agreement, they were prompted when it was time to review and also encouraged to set the third agreement – again, the timing could be extended if preferred.
5. Once the third agreement was reviewed, the programme was completed.



Introduction Risks Your Agreement Account driver2020@trl.co.uk Log off

DRIVER 2020

RoSPA accidents don't have to happen

Sample Driver2020 Mentoring Agreement

This Driver2020 Mentoring agreement is meant to help you as you develop your driving experience. It shows the agreements that we jointly made and as you become more experienced over the first few months of your driving, we'll review and relax them until they are no longer needed.

Mentor:

I agree to:

- ✓ Always wear my seatbelt and make sure my passengers do the same
- ✓ Never drink alcohol or take drugs and drive
- ✓ Obey traffic laws, including speed limits and parking restrictions
- ✓ Never use a mobile phone while driving, only when parked with the engine switched off
- ✓ Accept a call from you at any hour and either collect you or pay for a taxi for you to get home, with no questions asked at the time
- ✓ Be calm and constructive when travelling in the car with you and accept that mistakes are an important part of learning
- ✓ Remain calm and constructive when discussing driving with you

Driver:

I agree to:

- ✓ Always wear my seatbelt and make sure my passengers do the same
- ✓ Never drink alcohol or take drugs and drive
- ✓ Never accept a lift from a driver who has drunk alcohol or taken drugs, and to find alternative transport or call you and ask you to collect me or pay for a taxi, with no questions asked at the time
- ✓ Obey traffic laws, including speed limits and parking restrictions
- ✓ Never use a mobile phone while driving, only when parked with the engine switched off
- ✓ Avoid distractions, like fiddling with the stereo, smoking, eating and drinking, while driving
- ✓ Let you know if I'm going to be later than expected at night

Driver agreement:

Night Driving

I will only drive up to **11:00 p.m.**

Carrying peers during the day

I will only carry **2** peer passenger(s) during the day.

Carrying peers during the night

I will only carry **1** peer passenger(s) at night.

Signed by mentor

Signed by driver

We will review this agreement on: **16/10/2019**

© RoSPA 2022 Privacy Policy

Figure 3-2: Sample mentoring agreement provided on the intervention website

3.3.2 Intervention: Telematics

For the telematics intervention (delivered by Trak Global), a new participant was invited to download the Driver2020 telematics app and connect with the Bluetooth in their car. This meant that each time a journey was taken in that car the journey would be recorded by the app. Those participants who did not have Bluetooth in their vehicle could order a (free) Wingman – a small battery-powered device to be affixed to the windscreen that connects to the app via Bluetooth.

Using the technology available within modern smartphones (e.g., accelerometers, GPS), the app recorded telematics data. A custom-built driving behaviour algorithm used these data to provide the driver with a score based on their driving behaviour. This algorithm took into account the following variables when calculating risk (numbers in brackets reflect weighting assigned to the measure by the algorithm in its calculation of risk for feedback): speed vs road speed limit (25%); speed vs average speed for the section of road travelled (where the dataset permits – 25%); braking severity (17.5%); cornering (course over ground/direction change calculation – 7.5%); time of day (day/night, peak/off-peak – 20%); duration of journey (5%).

The participant was provided with feedback about their driving via the app dashboard (Figure 3-3) in the form of a weekly driving score out of 10 and rating, either red, amber, or green.

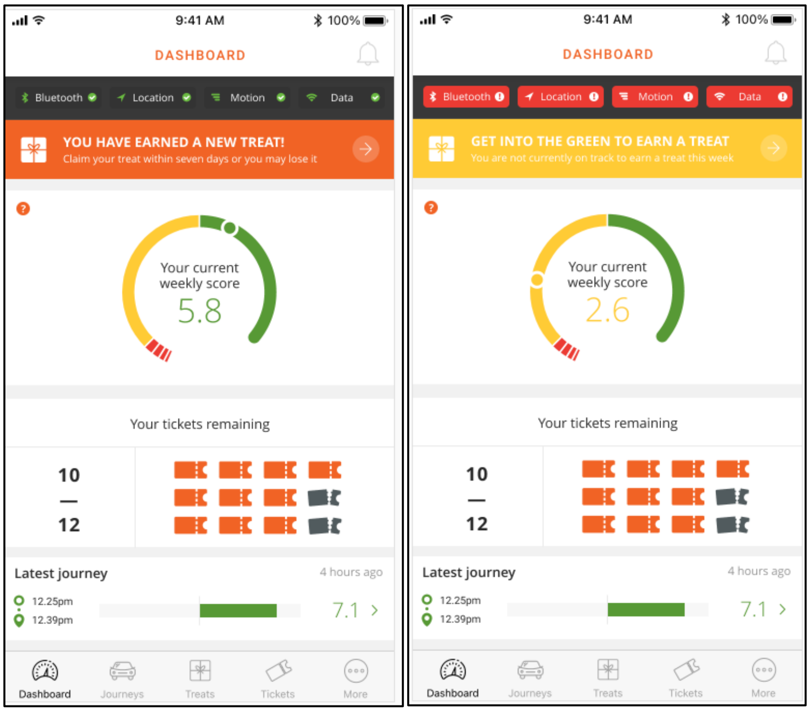


Figure 3-3: Example dashboards from the telematics app

The weekly score was made up from individual scores for speed, smoothness and usage for each journey, and this journey-level information was also available through the app; nudge notifications such as that shown in Figure 3-4 reminded participants of when improvements to their behaviour were needed.

Figure 3-4: Example notification from the telematics app

Participants were incentivised by earning weekly 'treat rewards' (e.g., vouchers for coffee) if the weekly score was 'green', and by retaining their monthly 'ticket' into a prize draw to win a year's free car insurance. Poor driving behaviour would result in a low 'red' weekly driving score; getting two 'red scores' during any month would lead to losing the prize draw ticket for the month, even though it was still possible to improve performance to obtain weekly rewards in the other weeks (Figure 3-5). (Note that the prize draw mentioned here was separate from the main study prize draw for survey completion, which is open to all trial participants, including the control groups.)

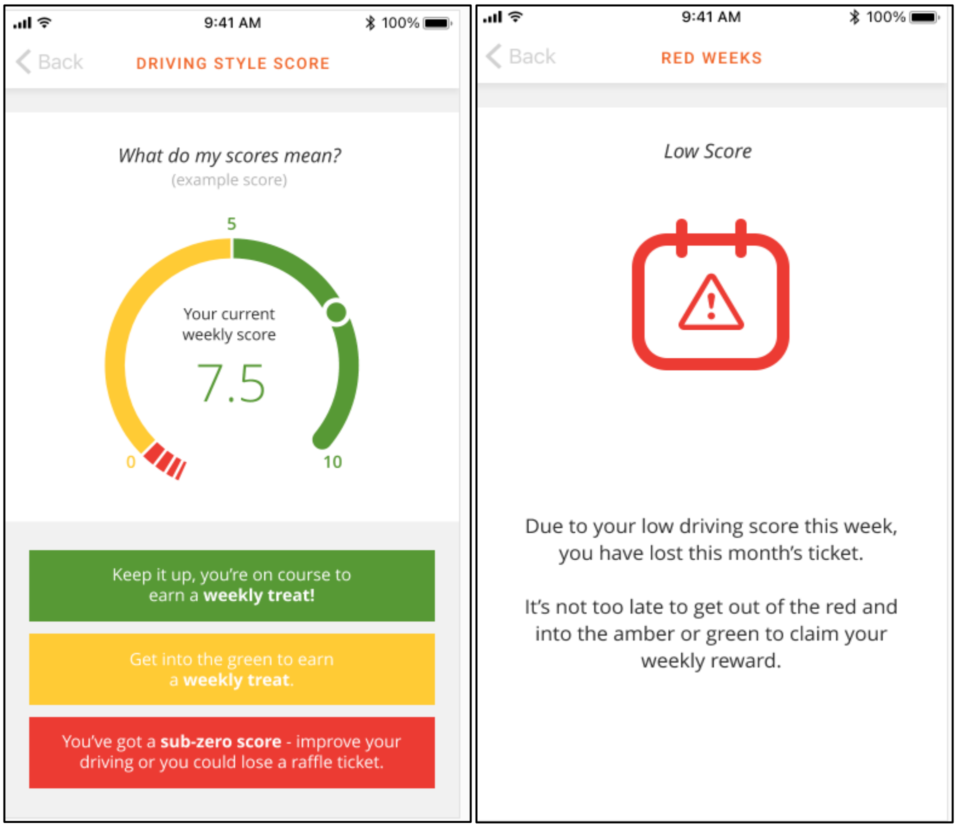


Figure 3-5: Feedback on score from the telematics app

3.3.3 Surveys

Five surveys were used to collect data from all participants, in all intervention and control groups. These are described below and are included in full in the supplementary appendix for the project (Weekley & Helman, 2024).

3.3.3.1 Registration survey

The registration survey (10-15 minutes, online) first collected information about participants to check they were eligible to register with the study (see section 3.1 for criteria).

For those participants who were eligible, the remainder of the survey served as an informed consent form including relevant data protection information, and collected contact details, driver licence number, information on education and employment, postcode, how the participant had heard about the study, and finally a 30-item form of the Big-Five Personality Inventory (Soto & John, 2017) to be used as a matching variable (this was one of a number of measures that was used to control for potential bias in engagers, relative to control participants – see section 4.5.2).

3.3.3.2 Test pass survey

The test pass survey first collected name and participant ID for the purposes of data matching. The remainder of the survey asked a range of questions about the participant's experience learning to drive (for example the types of roads they practised on, the amount of practice they had with instructors and with other supervising drivers), various items on attitudes to post-test restrictions and more general enforcement of road laws, information on education and employment, and items around access to a vehicle.

3.3.3.3 3-month, 6-month and 12-month post-test surveys

The 3-month, 6-month and 12-month post-test surveys began with questions collecting name and participant ID (for the purposes of data matching). The remainder of the survey asked a range of questions about the participant's experience driving post-test. The main items in the 3-, 6- and 12-month post-test surveys were the same; the difference in the surveys was that they asked about different periods of time (the first three months post-test, the fourth to sixth months, and the seventh to twelfth months respectively). The items included were:

- The number of collisions in which people had been involved during the period of interest, if any (and some further questions about when, where, and how each collision occurred)
- Mileage driven
- Proportion of mileage driven on different road types
- Proportion of mileage driven with a peer-age passenger, driven in the dark, driven on wet roads, and driven for work
- Near misses

- Six of the seven items from the Hazard Involvement/Driving Events scale from Quimby et al. (1999)
- Frequency of driving while tired, and driving over the speed limit
- Any limits or restrictions on their driving during the period of interest, for example less access to a car than expected, parents imposing restrictions, impacts of the COVID-19 pandemic or other changes in circumstances.
- The driving style scales from Guppy, Wilson and Perry (1990)
- Confidence in driving ability
- Attitudes towards enforcement of various road safety laws, and restrictions and limits for learner and new drivers
- Access to a vehicle and details of insurance policy type
- Levels of education and employment details

A final question in the 12-month survey asked participants if they had come into contact with any of the other interventions (a check for spill-over effects – for example if someone had come into contact with one of the interventions through contact with a friend in the study – see section 4.3.3).

3.4 Procedure

This section describes the procedure for each participant on entry into the study and the subsequent communication that took place. This procedure and the related data flows are illustrated in Appendix C.

Each participant who completed the registration survey and was eligible for the study was assigned to a group on the basis of the process described in section 3.2.2 and sent a welcome email. All participants were told that TRL would send a link to the first survey within the next few days; those in one of the intervention groups were also told that the relevant intervention provider would be in touch (for the mentoring agreement, RoSPA; for the telematics, Trak Global). The participant details were then passed to the DVSA to confirm that their test pass date was within the past four weeks. Once confirmed, the participant was invited to complete the test pass survey and the details of those assigned to one of the intervention groups were passed on to the relevant intervention provider.

After this initial contact, there were two ways in which participants could subsequently have contact with the study. All participants had contact with TRL regarding the data collection surveys. Those in the intervention groups also had contact with the intervention provider. All communications were branded as Driver2020 to minimise confusion.

After they received the contact details of the participants in their specific intervention group, providers contacted them to invite them to register with the specific intervention. For RoSPA (mentoring agreements), new participants were processed weekly. Each participant was invited to sign up via email; participants were also emailed to confirm each agreement set / process completed. Reminders (up to three) were sent for people who did not register, and for setting each agreement stage in the process. Text messages were also sent if the participant provided their mobile number at sign up. For Trak Global (telematics), new participants were processed weekly – each new participant was invited to download and pair with the Driver2020 telematics app. If the app was not paired in one week a reminder email was sent; a second reminder was sent if there was no action in two weeks, after which no more contact was made with the participant. If the pairing was made, further communication with the participant was made through the app, depending on their actions and engagement.

All participants were sent emails by TRL at 3, 6 and 12 months after their test pass date providing links to the relevant data collection surveys. If a survey was not completed, the participant received an email reminder with the survey link every week, up to a maximum of six reminders. Text message reminders (also including the survey link) were sent at the same time as the third and fourth emails. These reminders were also used for the test pass survey. For some participants who had not responded to reminders for the 12-month survey, the TRL research team also telephoned them to encourage them to complete the survey.

3.5 Limitations

There are a number of limitations of the research that need to be considered when drawing conclusions. These are mentioned throughout the report and are listed here for clarity.

1. The data were gathered during the COVID-19 pandemic, which would have affected learning and post-test driving. Generalisability to future circumstances in which there is no global pandemic may be affected.
2. The sample is biased towards females and is an opportunity sample. Therefore, any generalisation of the findings to the population of interest (novice drivers aged between 17 and 24 years of age in Great Britain) needs to be done with caution.
3. The self-selecting nature of the sample (both in terms of registering for the trial and in terms of providing survey data) means that the findings cannot necessarily be generalised to all novice drivers in Great Britain. However, the design (randomised allocation to groups) does mean that self-selection bias should not play any role in group differences.
4. The self-reported nature of most of the data means that conclusions again require caution; it is possible that the reported data are biased to some degree with social desirability; although this is not critical for the main comparisons between groups, it may mean that reported levels of behaviour are different to what would actually be observed. Self-reported data also reflect the interpretation of participants to survey questions.
5. Given the very low numbers of participants engaging with the interventions, and the low ‘dosages’ or ‘amounts’ of contact with interventions in those participants who did engage, the findings on collision reduction reflect only one potential set of roll-out conditions for the interventions.
6. The findings related to surrogate measures provide us with evidence of the behaviour change that results from the interventions as these analyses only include

those who engaged; however, the very low numbers engaging with some interventions mean that some very small effects may have been missed.

7. The findings only cover the first year of driving and cannot be extended beyond this period (for example we cannot say whether behaviour would be impacted beyond this period, or whether any changes seen would persist).

4 Results – novice driver arm

4.1 Introduction

This section presents the results of the learner driver arm of the Driver2020 study. The focus is on the intention-to-treat analysis – sections 4.2 to 4.4 cover the trial sample used, exploratory analysis and statistical modelling. Section 4.5 presents the per-protocol analysis.

4.2 Trial sample

The final dataset to be used for the novice driver 'intention-to-treat' analysis was collated on 8th January 2022. Any survey responses received after 23:59 on 7th January 2022 have not been included in analysis.

Table 4-1 shows the completion rate for each of the four data collection surveys for novice participants, along with the total number recruited, for each group.

Table 4-1: Number and percentage of novice participants who completed each survey

	Mentoring agreement (Recruited: 4,183)		Telematics (Recruited: 4,173)		Control (Recruited: 3,951)		Total (Recruited: 12,307)	
Test Pass	3,615	***86%***	3,526	***84%***	3,373	***85%***	10,514	***85%***
3-month	3,029	***72%***	2,916	***70%***	2,908	***74%***	8,853	***72%***
6-month	2,761	***66%***	2,641	***63%***	2,669	***68%***	8,071	***66%***
12-month	2,578	***62%***	2,468	***59%***	2,468	***62%***	7,514	***61%***

Note that, for each survey, the number and percentage shown represents the completion for that survey point, independent of whether previous surveys had been completed; not all the participants who completed the 12-month survey had also completed all previous ones.

Table 4-2 shows the number of participants in each group who completed all four surveys and that number as a percentage of those who were eligible to complete all four surveys (for novices, this was all participants recruited).

Table 4-2: Number of participants in each group who completed all four surveys.

Group	Final sample – number of participants with all four surveys completed (as % of total eligible
Mentoring agreement	2,356 (56%)
Telematics	2,234 (54%)
Control	2,258 (57%)
Total	6,848 (56%)

The original sample size calculations for the trial resulted in a target of 2,036 participants per group; since the number of participants completing all four surveys exceeded this target in all three groups, these participants were used as the final sample. This final sample had complete data for all surveys and thus missing data did not need to be imputed for modelling purposes. An attrition analysis (see section 4.3.7) was completed to determine if there were any differences between the final sample and those recruited.

4.3 Exploratory analysis

This section presents the exploratory analysis of the trial sample, their learning to drive experience, and their experiences during their first year of driving. The aim of this was to explore whether characteristics differed between the three groups: control, mentoring agreement intervention, and telematics intervention to assess the comparability of the groups for further analysis.

The demographics of the sample are explored in section 4.3.1, and these are then compared with the demographics of test passers nationally in section 4.3.2 to examine how representative the trial sample is of the national population. This is for information only, as the study was always targeted at younger novice drivers rather than the national population of test passers. In addition, in studies of this kind females are typically over-represented (for example see Wells et al., 2008; Helman et al., 2017) as was the case in this study.

Section 4.3.4 presents analysis of the conditions in which participants in each group learned to drive and drove during their first post-test year, such as different road types, weather, and light conditions. The impact of the COVID-19 pandemic on participants and their learning to drive and post-test driving experiences is explored in section 4.3.5.

A summary of the number of collisions reported by participants is presented in section 4.3.6 and compared between different demographics, post-test mileage levels, and intervention groups. Finally, a brief summary of the attrition analysis carried out is described in section 4.3.7.

The key insights from the exploratory analysis of the novice driver sample are as follows:

- The sample was biased towards females (66% of the sample was female) meaning caution is needed in generalising findings.
- Around 67% of the sample was in full time education when registering to take part.
- There was no evidence of spill-over or contamination effects (people in one group being exposed to interventions from the others).
- Participants who passed their test after the COVID-19 pandemic reported a number of impacts from the pandemic on their learning to drive, mainly not being able to book practical tests as soon as they would like, tests being postponed or cancelled, and having less practice with a driving instructor.
- Participants who passed their test before the COVID-19 pandemic reported lower annual mileage than those who passed after.

- There were no significant differences between the groups for any of the characteristics explored, meaning that the groups would be suitable for comparison in the intention-to-treat analysis.
- There were changes to the characteristics of the sample over time as some types of people were more likely to drop out of the study than others, but these changes were consistent between the groups.

4.3.1 Demographics

This section presents a summary of the demographics of the trial sample. The distributions of age, gender, Social Deprivation Index (SDI), and education status were compared between the three groups (control, mentoring agreement and telematics) but statistical tests found no differences between groups.

The participants included in the analysis presented in this section are those who answered all surveys (6,848 participants).

4.3.1.1 Age

Figure 4-1 shows the proportion of participants in each group by age at registration. The majority of trial participants were 17 or 18 years old when they registered for the trial. These two age categories account for 41% and 20% of all trial participants respectively. Differences in age distribution between groups were not statistically significant. The mean age (at registration) of a participant in the trial was 18.9. This is over a year older than in the learner arm, which is due to the fact that learners were recruited when they began learning, while novices were recruited when they had passed their practical test.

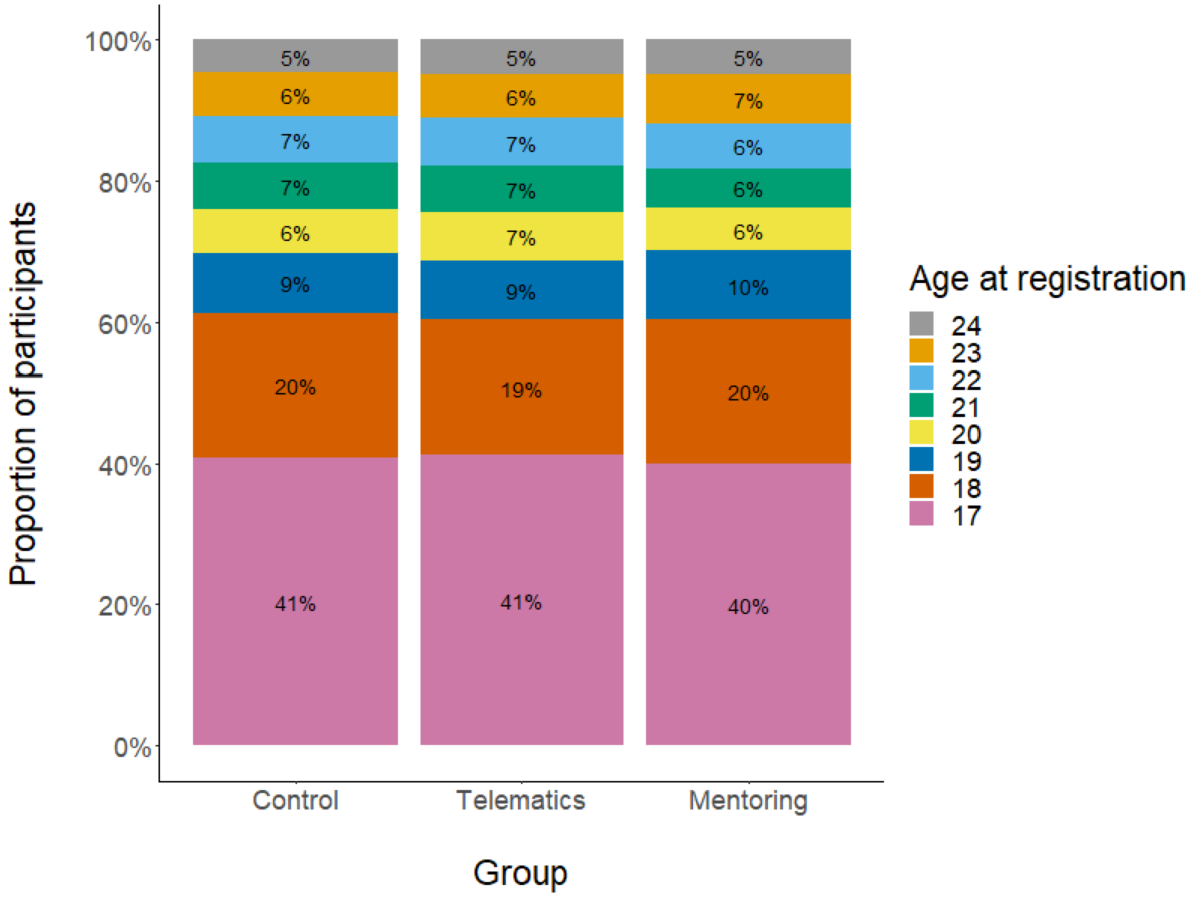


Figure 4-1: Proportion of trial participants in each group by age at registration (N=6,848)

4.3.1.2 Gender

The proportion of male and female participants in each group are shown in Figure 4-2. The trial sample was approximately two-thirds female (66%) and one-third male (34%); as noted previously, in studies of this kind females are typically over-represented (for example, see Wells et al., 2008; Helman et al., 2017). This means that findings cannot be generalised to the population of young and novice drivers in Great Britain without some caution. There was no difference in gender distribution between groups.

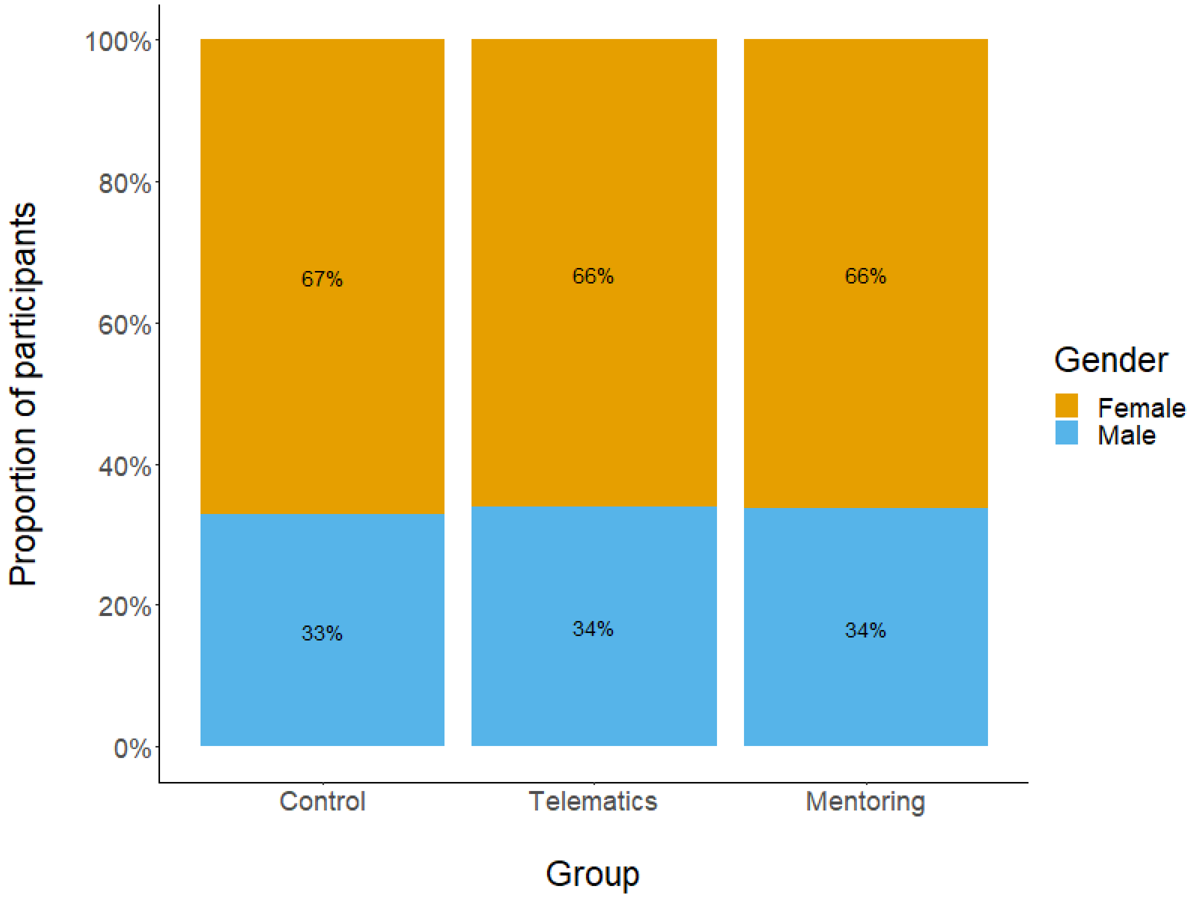


Figure 4-2: Proportion of trial participants in each group by gender (N=6,848)

4.3.1.3 Social Deprivation Index

Index of Multiple Deprivation (or social deprivation index) is the official measure of relative deprivation for small areas (or neighbourhoods) in the UK. Data from England, Wales, and Scotland (gov.uk, 2019; gov.scot, 2020; gov.wales, 2019) were downloaded and combined to create one dataset. This was matched to the participant data by location information (postcode at time of registration) and the social deprivation ranks were then assigned to four quartiles from most deprived to least deprived.

Figure 4-3 shows the proportion of participants in each group by Social Deprivation Index quartile. Participants in all three groups were similarly distributed between the four quartiles, albeit with slightly lower proportions in the lower quartiles (more deprived) than in the higher quartiles (less deprived). Q1 (most deprived) accounted for the smallest proportion of participants across all groups (19%), while 22% and 26% of participants were in Q2 and Q3 respectively, and the largest proportion (33%) was in Q4 (least deprived). Differences between groups were not statistically significant.

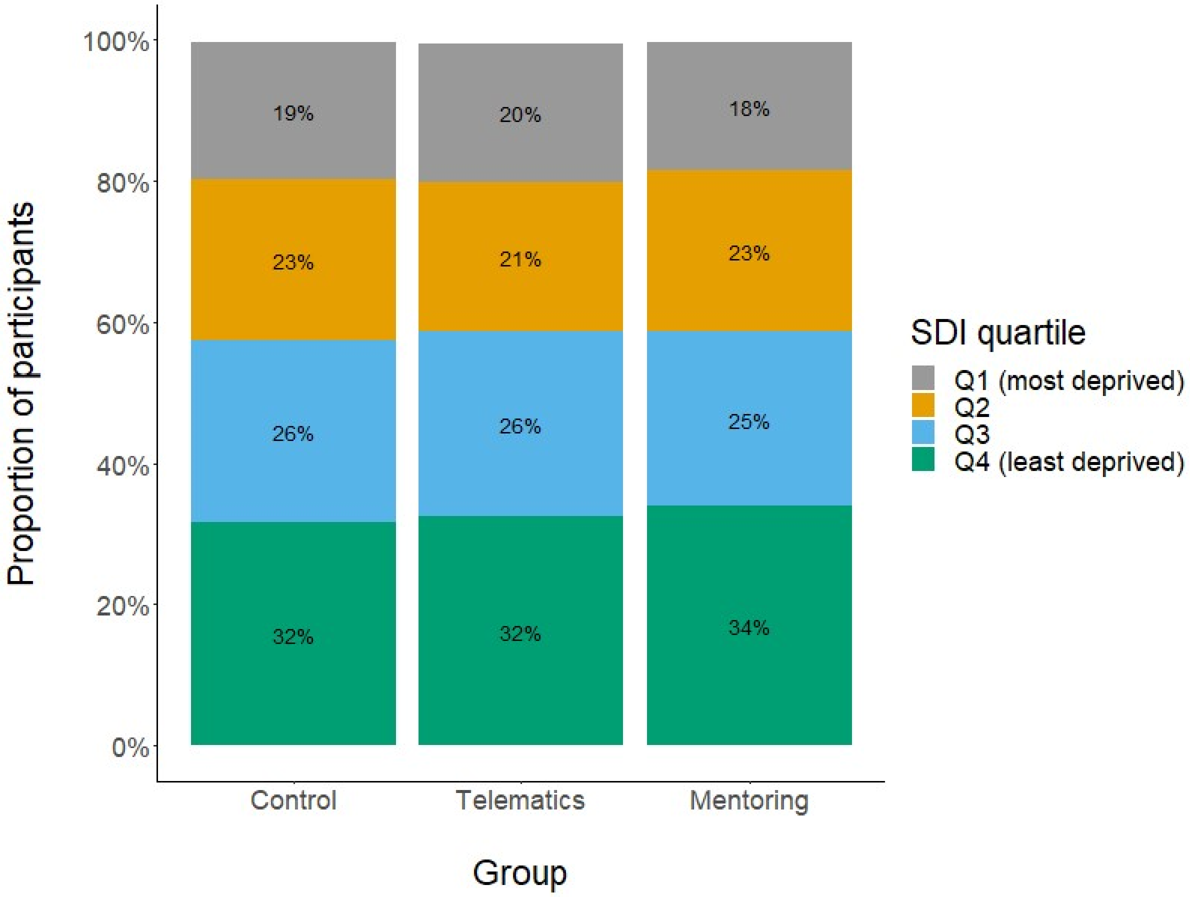


Figure 4-3: Proportion of trial participants in each group by Social Deprivation Index quartile (N=6,848)

4.3.1.4 Education status

Figure 4-4 shows the proportion of participants in each group who were in full-time education when they registered for the trial. The majority of trial participants (67%) in the novice arm were in full-time education when they registered, and this proportion was the same within each group.

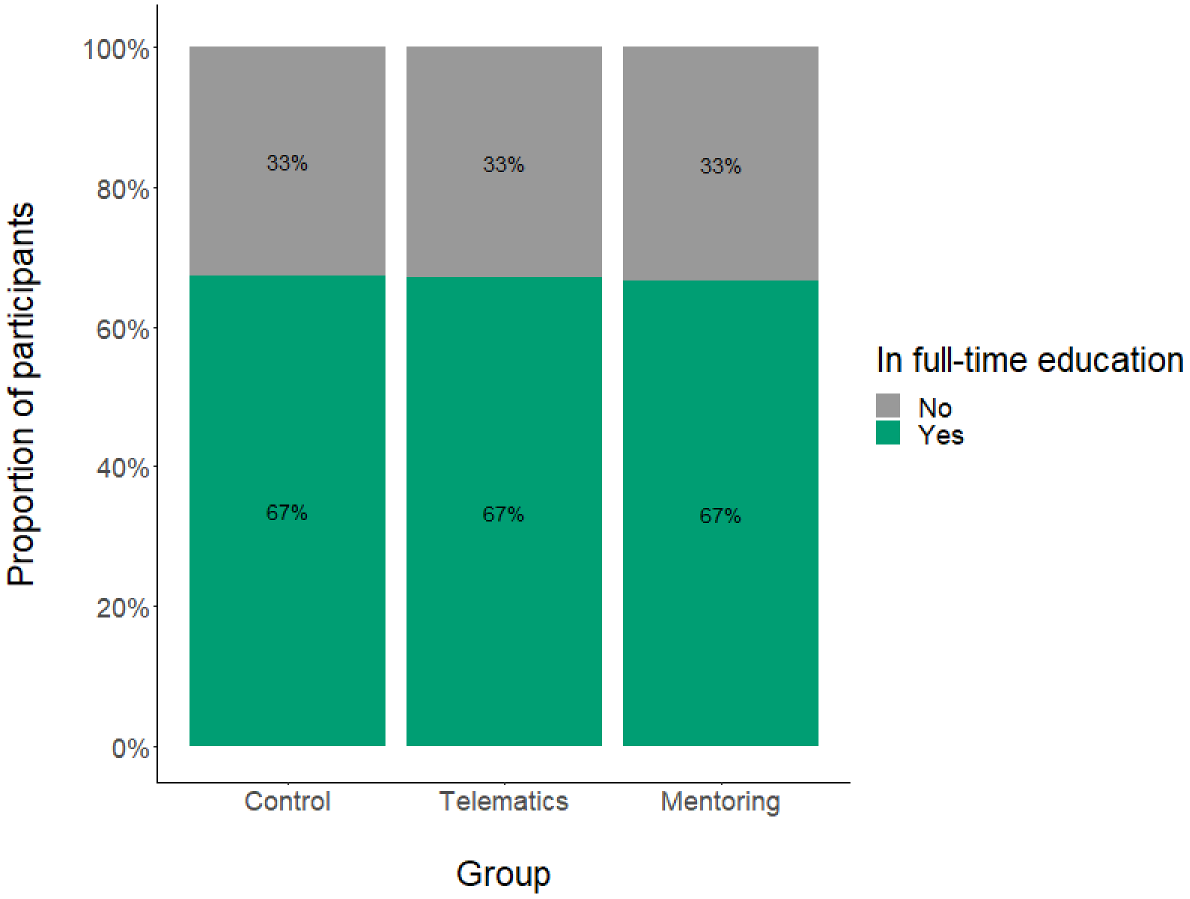


Figure 4-4: Proportion of trial participants in each group by education status (N=6,848)

4.3.2 Comparison with other data sources

The age and gender distributions of the trial participants were compared with test passers in the whole of Great Britain during the period April 2019 to March 2021 as published by DfT in table DRT0203 (DfT, 2021). These comparisons were done to explore how similar those registered in the trial were to the population of test passers nationally.

The results of these comparisons are presented in Figure 4-5 (age) and Figure 4-6 (gender), and showed that the trial sample includes higher proportions of younger people and females than seen across all test passers in Great Britain during a similar period.

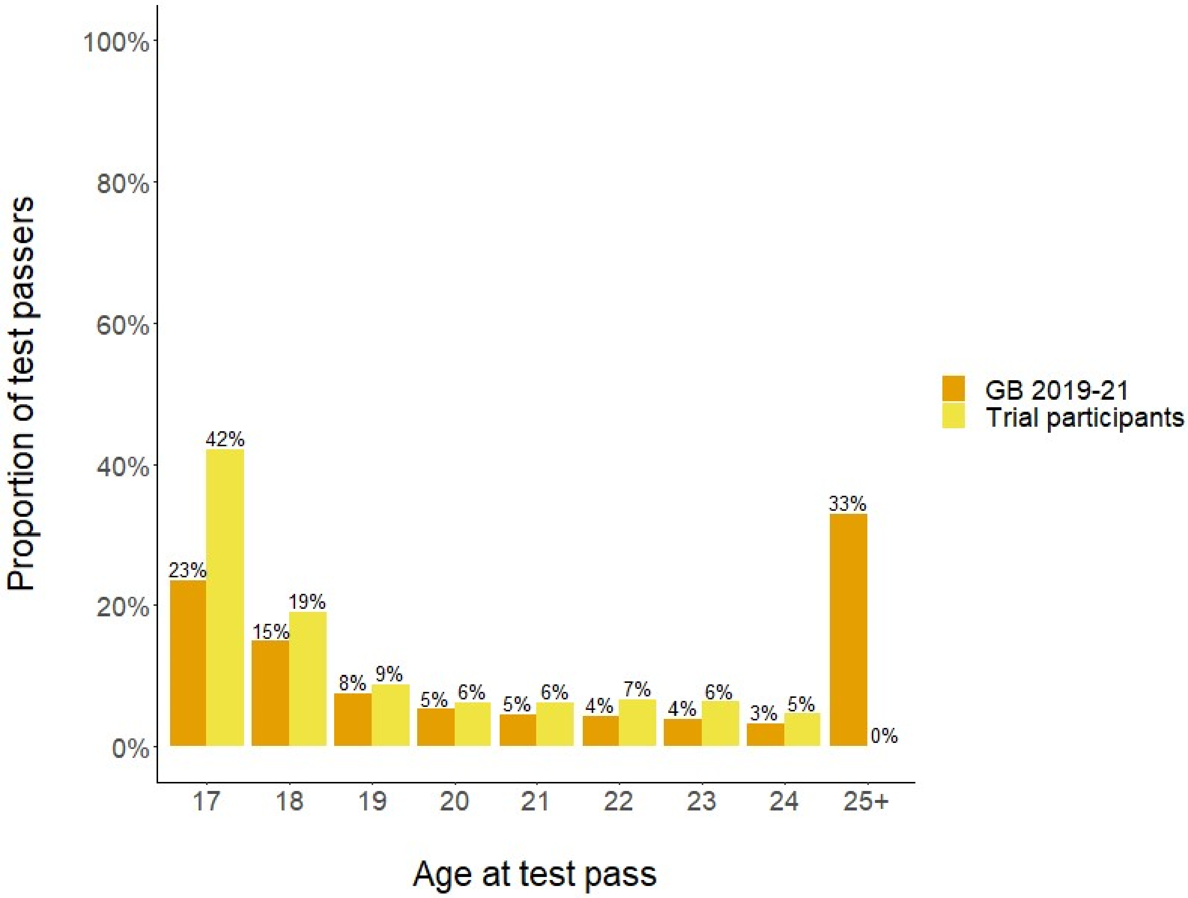


Figure 4-5: Proportion of test passers in Great Britain (2019-21) and trial participants by age at test pass (N=6,848)

There was a much higher proportion of 17-year-olds (at test pass) among the trial participants (42%) than seen across all test passers in Great Britain (23%). There were also higher proportions of drivers of every age from 18 to 24 in the trial sample compared with all test passers. One third (33%) of all test passers in Great Britain were aged 25 or over.

These differences were expected given that the trial sample only included participants aged 17-24 at test pass. It means that findings should only be generalised to this age group.

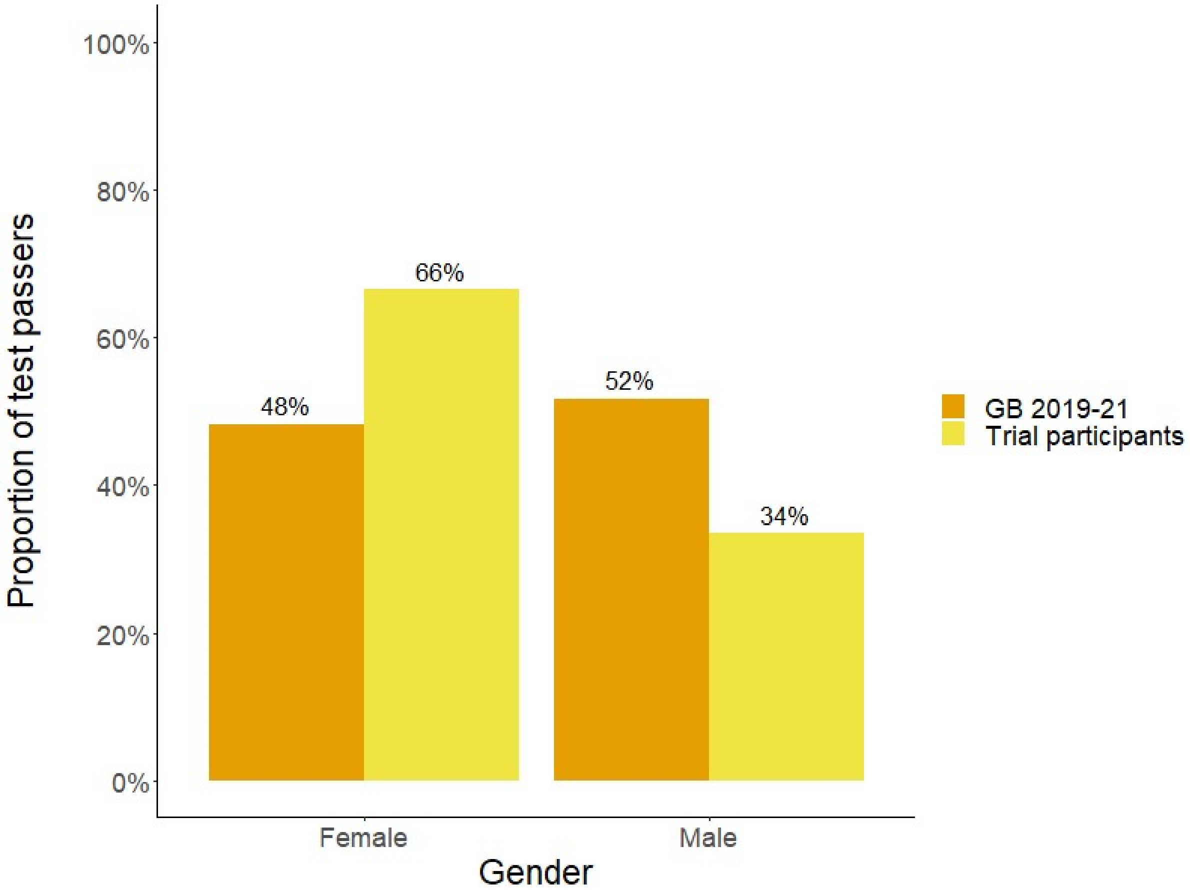


Figure 4-6: Proportion of test passers in Great Britain (2019-21) and trial participants by gender (N=6,848)

Figure 4-6 shows the difference in gender distributions between the trial sample and test passers nationally. The trial sample includes a higher proportion of females (66%) compared with test passers across Great Britain (48%). This is typical for studies of this type (for example Wells et al. (2008) had 63% and 37% respectively); the study was not reliant on achieving a representative gender balance, meaning that findings cannot be generalised to the population of young and novice drivers in Great Britain without some caution.

No comparison was done between the trial sample and test passers nationally for SDI or education status; this is because these data are not readily available for the national test passer population.

4.3.3 Spill-over and contamination checks

Checks were undertaken to ensure that participants in each group had not been affected by the interventions applied to any other groups (spill-over) or other known external factors (contamination) – in this case whether or not they had a telematics-based insurance policy.

The original intention was that checks for spillover would need to be carried out through a question in the 12 months post-test survey, therefore participants were asked whether they had heard of, or engaged with, any of a list of interventions (those in this trial). During the development of the interventions however, it became apparent that access to the interventions could be controlled much more closely than originally expected. From the control over access, and supported by the answers to the question, spill-over is extremely unlikely to have been an issue.

Participants were also asked at the 12-month time point whether they had a telematics insurance policy (this is similar to the telematics intervention but is also common among young drivers) in order to be able to control for any differences between the groups. The proportion of participants with this type of insurance policy was similar for all three groups (36% for control group and telematics group and 35% for mentoring agreement group). This indicates that it is unlikely that any bias was introduced into the trial from contamination through telematics insurance policies.

4.3.4 Learning and driving conditions

This section explores the conditions in which trial participants learned to drive and the types of driving they did in the first year after passing their test. There were no notable differences identified between groups in either learning or post-test driving conditions.

4.3.4.1 Time spent learning

In the test pass survey participants were asked to estimate how many hours they had spent driving with an Approved Driving Instructor (ADI) and with another supervising driver (e.g., friends or family) before they passed their test. Table 4-3 presents the mean number of hours participants in each group spent learning with an ADI, with another supervising driver, and in total.

Trial participants reported spending an average of 73.5 hours driving before passing their test; 43.8 hours learning with an ADI and 29.7 hours practising with another supervising driver.

This was similar across all groups, which would be expected, given that neither of the interventions was designed to change preparation for the test, as people did not sign up until test pass.

Table 4-3: Mean (and standard deviation) of hours spent learning to drive with an instructor and with a supervising driver, by group (N=6,848)

		Group			All participants
		Control	Telematics	Mentoring agreement	
Mean (and standard deviation) hours spent learning to drive…	With ADI	44.0 (32.1)	42.8 (31.2)	44.7 (32.8)	43.8 (32.1)
	With other supervising driver	29.2 (37.5)	30.6 (38.3)	29.3 (37.9)	29.7 (37.9)
	In total	73.2 (46.9)	73.4 (47.5)	74.0 (49.1)	73.5 (47.9)

4.3.4.2 Learning conditions

In addition to answering questions about the number of hours spent learning, participants were also asked to estimate the proportion of their learning mileage they spent driving in certain conditions including on different road types, with additional passengers (not including the supervising driver), in the dark, and in wet conditions. The proportion spent learning in each condition has been averaged across the participants in each group.

Figure 4-7 shows the proportion of learning hours participants in each group spent driving in different conditions. All three groups show very similar results, which again was expected due to the interventions being delivered post-test.

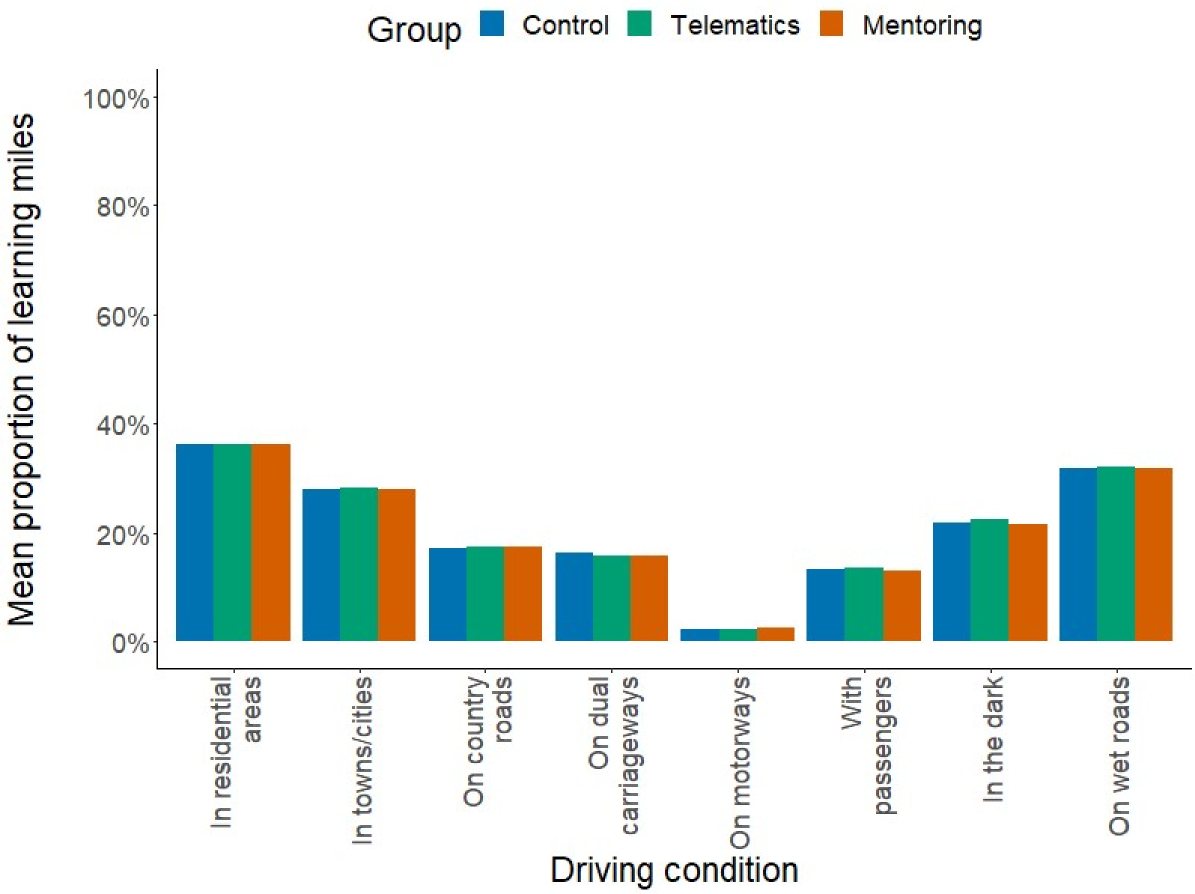


Figure 4-7: Proportion of learning miles spent driving in different conditions by group (N=6,848)

Of the different road types asked about, participants said they spent the most time learning in residential areas (36% of learning hours) and the least on motorways (2% of learning hours). This is to be expected, because residential areas are typically quieter and more suitable for learning and practising driving, especially at the beginning, and in particular for learning manoeuvres. Note also that motorway driving is not included in the practical test.

Participants reported spending 13% of their learning hours driving with passengers additional to their supervising driver. The proportions of learning hours spent driving in the dark and on wet roads were 22% and 32% respectively.

4.3.4.3 *Mileage since test pass*

Table 4-4 shows the mean number of miles driven in each reporting period and the total driven during the first twelve months since passing the driving test, by group.

Participants stated that they drove an average of 5,813 miles in the first twelve months after passing their test. Those in the mentoring agreement group had the highest mean annual mileage (5,918 miles) and those in the control group had the lowest (5,635 miles). Statistical testing showed a significant difference in the distributions of mileage between the three groups (p=0.01), albeit with a very small effect size of 0.0009 (see Appendix D for a description of p-values and effect sizes); the effect size suggests the significant result could be attributed to the large sample sizes rather than genuine differences between groups.

Table 4-4: Mean (and standard deviation) of total miles driven by reporting period and in first twelve months since test pass by group (N=6,848)

		Group			
		Control	Telematics	Mentoring agreement	All participants
Month 1-3	Mean (and standard deviation)	1,289 (1,490)	1,291 (1,423)	1,341 (1,580)	1,308 (1,500)
Month 4-6	Mean (and standard deviation)	1,474 (1,590)	1,522 (1,482)	1,528 (1,593)	1,508 (1,557)
Month 7-12	Mean (and standard deviation)	2,871 (2,410)	3,069 (2,404)	3,049 (2,520)	2,997 (2,448)
Miles driven in 12 months since test pass	Mean (and standard deviation)	5,635 (4,397)	5,881 (4,208)	5,918 (4,586)	5,813 (4,404)

4.3.4.4 *Post-test driving conditions*

At the end of each post-test reporting period (months 1-3 post-test, months 4-6 post-test, and months 7-12 post-test), participants were asked about the conditions in which they had driven during that reporting period. The questions were the same as those asked about learning conditions, but also included a question about the proportion of mileage spent driving for work (not including commuting).

Figure 4-8 shows the proportion of miles reported as driven in different conditions since test pass for participants in each group. All three groups show very similar results. The proportions reported at each survey point have been weighted based on reported mileage for each person and then averaged across the participants in each group.

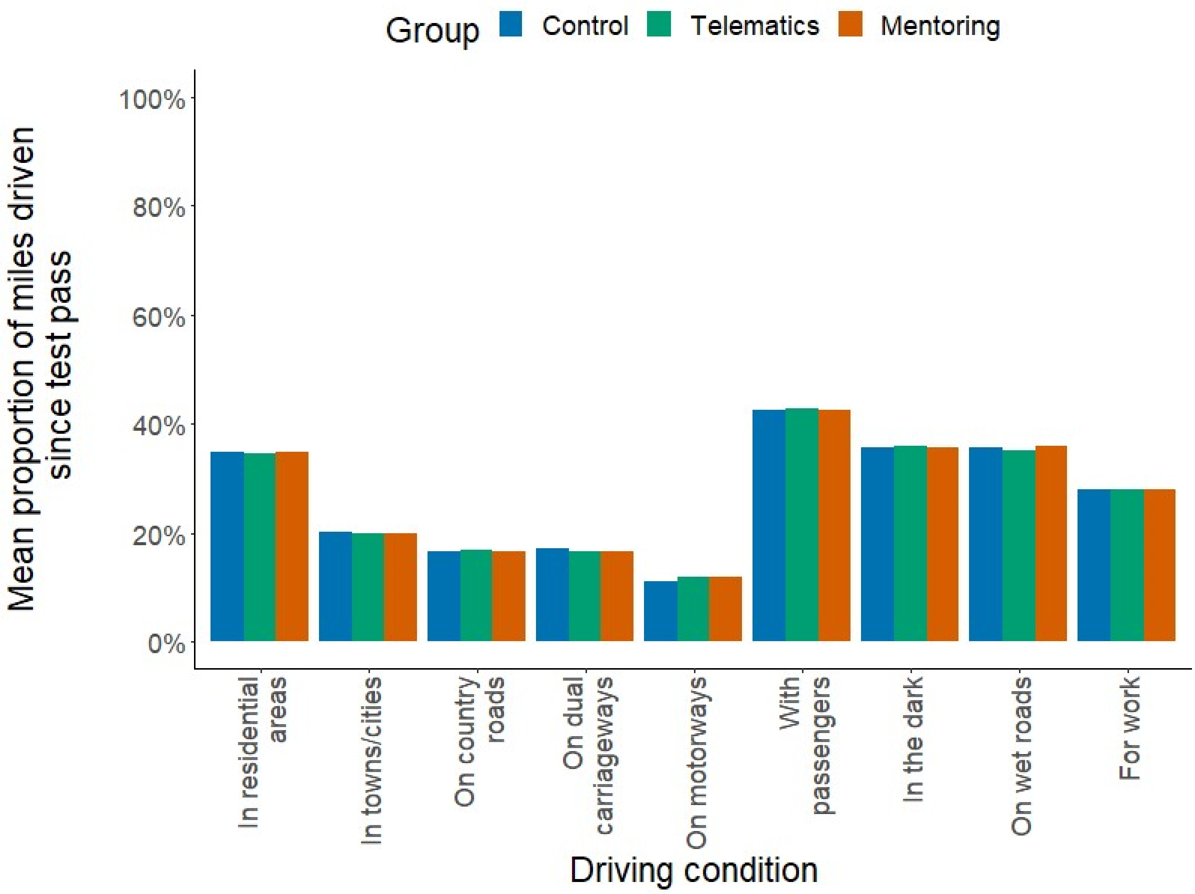


Figure 4-8: Proportion of total miles driven in 12 months since test pass in different conditions by group (N=6,848)

Figure 4-8 shows that, out of the different road types, participants most commonly reported that they drove in residential areas after passing their test, and motorways accounted for the smallest proportion of post-test mileage. This pattern is similar to that reported by participants whilst they were learning to drive (Figure 4-7) but the difference in proportions between motorways and other road types is smaller than for the learning phase.

The proportion of reported post-test mileage driven with passengers (43% across all groups) was much higher than seen pre-test (13%, see section 4.3.4.2). The proportion of reported driving done in the dark after participants passed their test (36% across all groups) was also higher than before they passed (22%) but the proportion of reported driving on wet roads was more similar before and after passing (32% and 36% respectively).

Figure 4-9 and Figure 4-10 show how the proportions of reported post-test mileage in different road types and conditions changed across the three reporting periods. Again, all three groups showed similar results.

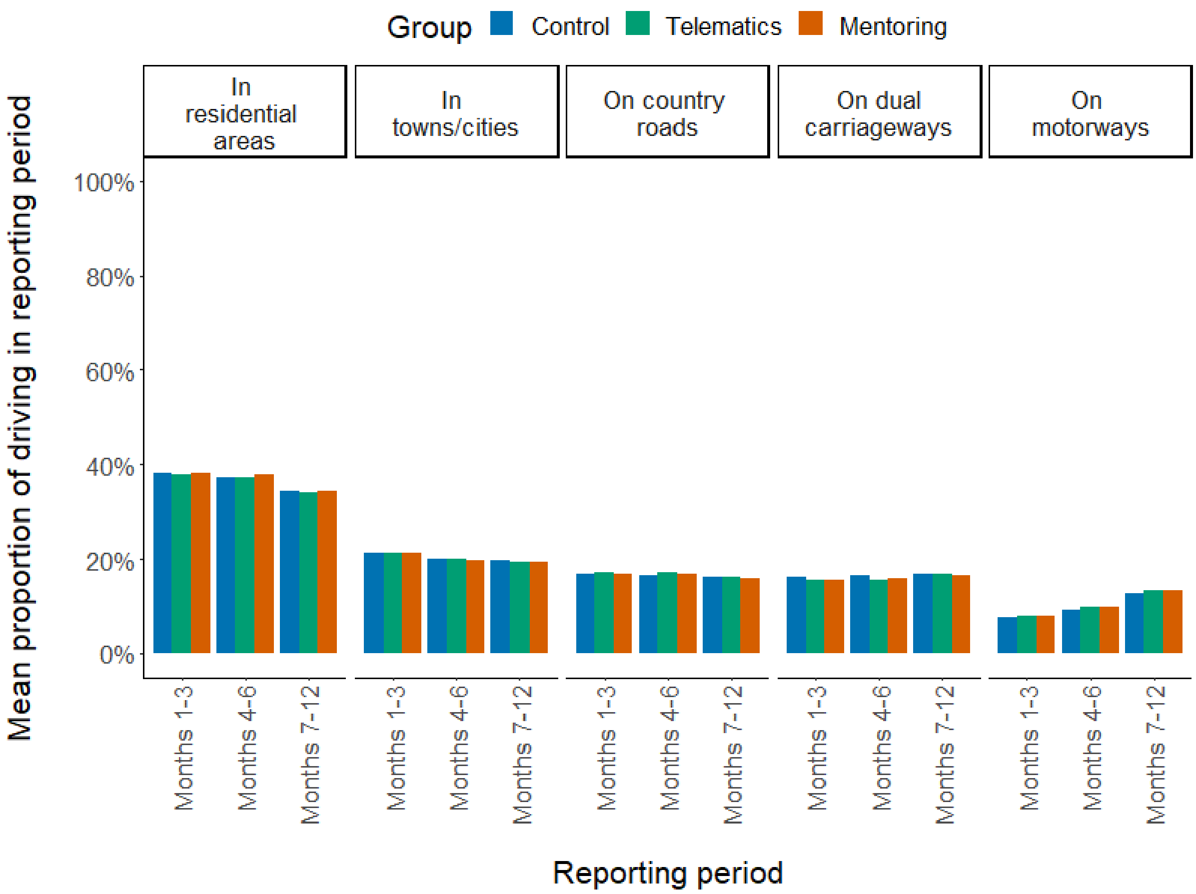


Figure 4-9: Proportion of miles driven in each reporting period on different road types by group (N=6,848)

Figure 4-9 shows that the proportion of reported post-test mileage driven by participants on different road type stayed broadly similar across the three reporting periods in their first year of driving, with the exception of the proportion of time spent driving on motorways. As time since test pass increased, participants reported an increase in the proportion of their driving that they did on motorways from 8% during the first three months post-test to 13% in the final six months of their first year of driving (proportions across all groups).

There were also small decreases over the three reporting periods in the proportions of reported mileage driven in residential areas and in towns and cities, although this does not necessarily mean that participants drove less in these locations during the reporting periods later in their post-test driving. Rather, it is likely that the increased proportion of motorway driving in later reporting periods meant that driving on residential and urban roads correspondingly accounted for a slightly smaller proportion of total mileage.

There were no noticeable differences between groups in the proportion of mileage driven on the different road types. Statistical tests showed no significant differences.

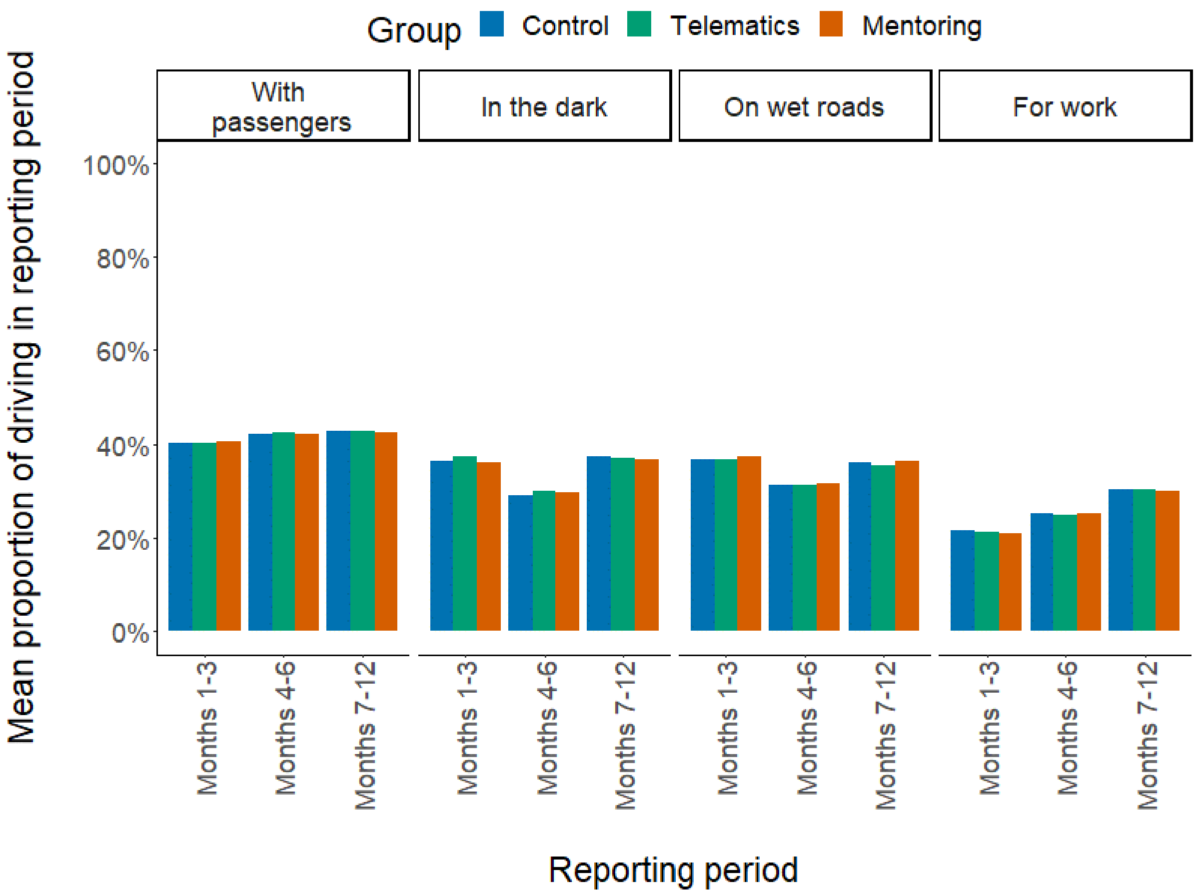


Figure 4-10: Proportion of miles driven in each reporting period with passengers, in the dark, on wet roads, and for work; split by group (N=6,848)

Figure 4-10 shows that the proportion of reported mileage driven with passengers was broadly similar across the three reporting periods, but the proportions of reported post-test mileage driven in the dark, on wet roads, and for work, differed more across reporting periods.

Participants reported an increase in the proportion of their driving which was done for work (any driving for work purposes excluding commuting to and from their usual place of work) across the three reporting periods; 21% of mileage (across all participants in all groups) was driven as part of work in participants' first three months after passing their test and this increased to 30% in the final six months of their first post-test year.

Within each reporting period, there were no significant differences between the groups.

Both the proportions of reported mileage driven in the dark and on wet roads showed a noticeable decrease during the months 4-6 reporting period when compared with the months 1-3 and the months 7-12 periods. The proportion of mileage driven by participants (all groups) in the dark was 29% during the months 4-6 reporting period compared with 36-37% in the months 1-3 and months 7-12 periods. Participants reported that driving on wet roads accounted for 36-37% of their driving during the months 1-3 and months 7-12 reporting periods respectively but only 31% in the months 4-6 period.

It is likely that these differences in the months 4-6 reporting period are linked with the times of year during which this period fell. The COVID-19 pandemic meant that the rate of participants passing their test was not consistent throughout the period covered by the trial; a large number of participants answered the six-month survey in late summer 2020 or late spring 2021. Hence, the months 4-6 reporting period for these participants was during spring or summer when the weather conditions and daylight hours meant that less driving in the dark and on wet roads would be expected.

4.3.5 Impact of COVID-19 pandemic

This section explores the impact of the COVID-19 pandemic on participants' learning to drive process and post-test mileage; exploratory analysis showed that the pandemic had an impact on both.

4.3.5.1 Impact of COVID-19 pandemic on test pass date

Of the 6,848 participants who answered all surveys, 2,960 (43%) passed their test before 20th March 2020 (the start of the first lockdown) and 3,888 (57%) passed their test after this date. Henceforth, the period prior to 20th March 2020 is labelled as "Pre-COVID-19" and the period after this date is labelled as "Post-COVID-19".

Figure 4-11 shows the distribution of participants by date of passing their practical driving test.

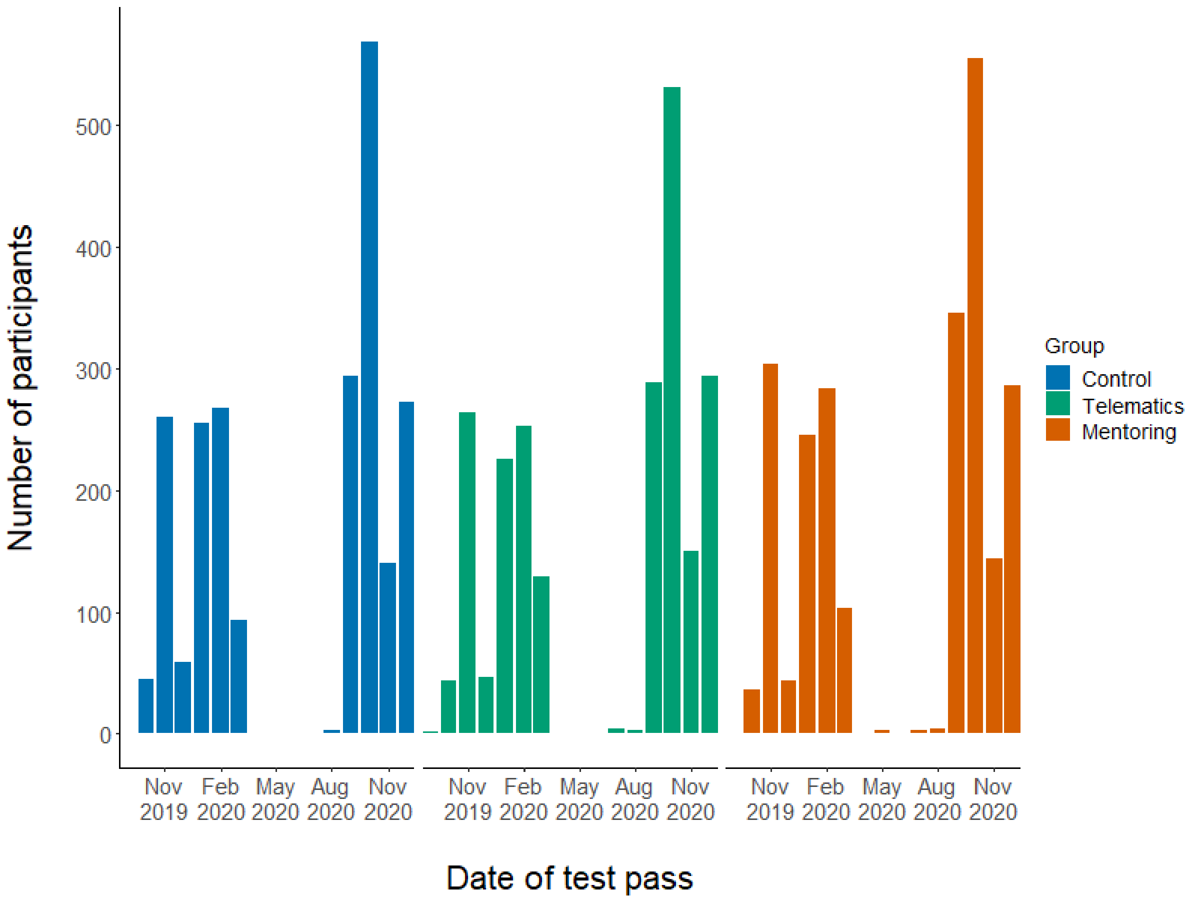


Figure 4-11: Distribution of participants by date of test pass (N=6,848)

There was no significant difference in the distribution of date of test pass between groups. Just under half of the participants in each group passed their test prior to the start of the pandemic.

4.3.5.2 Impact of COVID-19 pandemic on learning to drive

In the test pass survey, participants were asked if they had experienced any impacts on their learning to drive process due to the COVID-19 pandemic. These included impacts such as postponement or cancellation of their theory or practical tests, being unable to book their tests when they were ready to, having more or less driving experience with their instructors or parents and having the same amount of experience but taking a longer time due to the pandemic.

The results are shown in Figure 4-12. Participants were allowed to select all options that applied to them and therefore, the proportions do not add up to 100%.

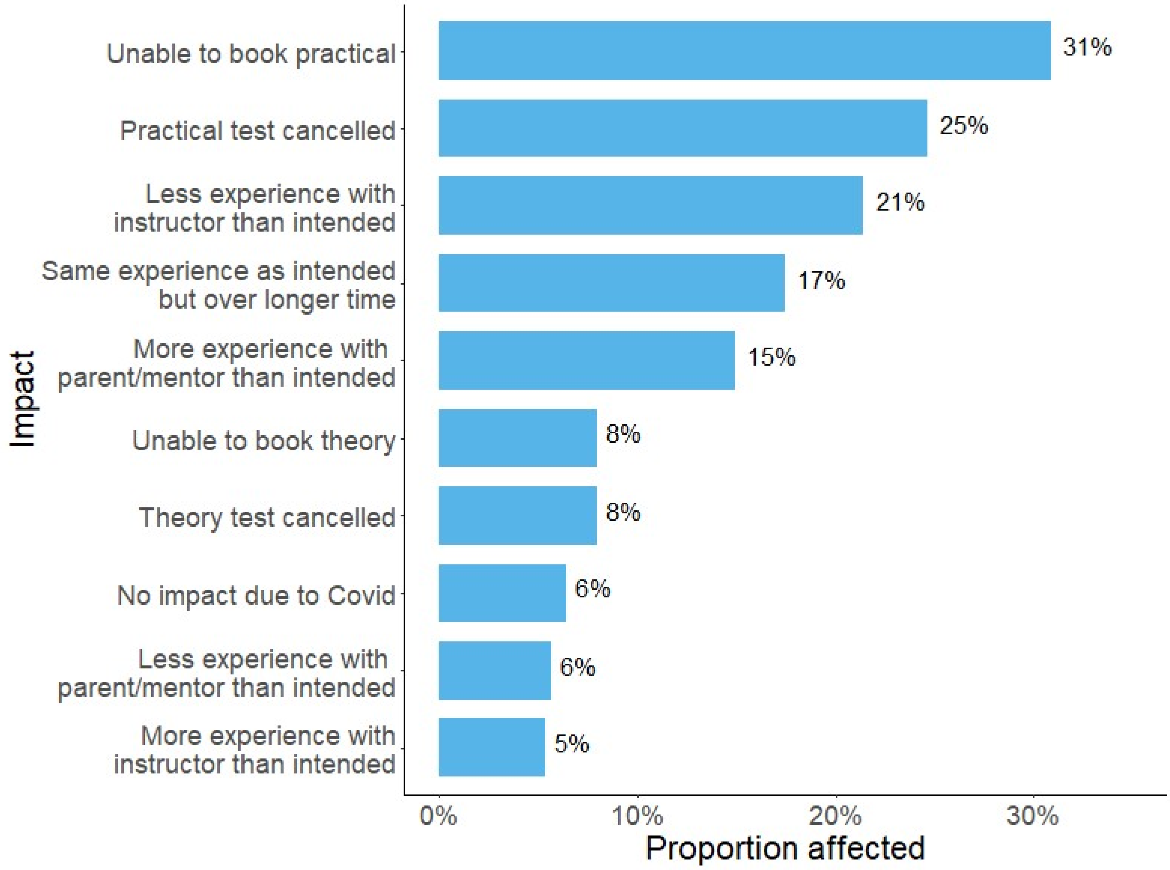


Figure 4-12: Impact on the learning to drive process due to the pandemic (N=3,888)

The top three impacts reported by participants were not being able to book their practical driving test when they were ready to, practical driving test being postponed or cancelled, and having less experience with their driving instructor. The results did not vary by group.

The above responses were converted into a numerical score for easier comparison during the intention to treat (ITT) modelling (see section 4.4). A score of 0 was given if participants said that the pandemic had no impact on their learning to drive process. The impact factors are shown in Table 4-5. Factors such as being unable to book their theory or practical tests, or tests being postponed or cancelled, increased the participant's score by 1 as it impacted their learning to drive process in a negative manner. However, if participants reported having more driving experience with their instructor or parent then this reduced their score by 1, on the grounds that this impact was positive in nature.

Table 4-5: Calculating a score for the impact of COVID-19 pandemic on learning to drive

Factor	COVID-19 pandemic impact on learning to drive score modifier
No impact on learning to drive process due to the pandemic	0
Practical test booked which was postponed or cancelled	1
Theory test booked which was postponed or cancelled	1
Unable to book practical test when ready to	1
Unable to book theory test when ready to	1
Less experience with instructor than intended	1
Less experience with parent than intended	1
More experience with instructor than intended	-1
More experience with parent than intended	-1
Took longer to get the amount of experience intended	1

This resulted in a COVID-19 Impact Score for each participant, where a higher value denotes a greater negative impact of the pandemic on the learning to drive process and a lower value denotes a lower negative impact of the pandemic on the learning to drive process. A negative value suggests that the pandemic actually had a positive impact on the participants' learning to drive process (for example by helping them get more practice).

As this question did not apply to those who passed their test prior to the pandemic, all participants in this group were given a score of 0. Participants who passed their test after the start of the pandemic had a score ranging between -2 and 7, with the average score for the group being around 1.6 (standard deviation of 1.2), based on the scoring system for impact factors noted above.

4.3.5.3 Impact of COVID-19 pandemic on mileage after passing their test

This section explores the impact of the pandemic on participants' mileage after they passed their practical driving test. Figure 4-13 and Table 4-6 show the summary of annual mileage reported by participants pre and post the start of the pandemic.

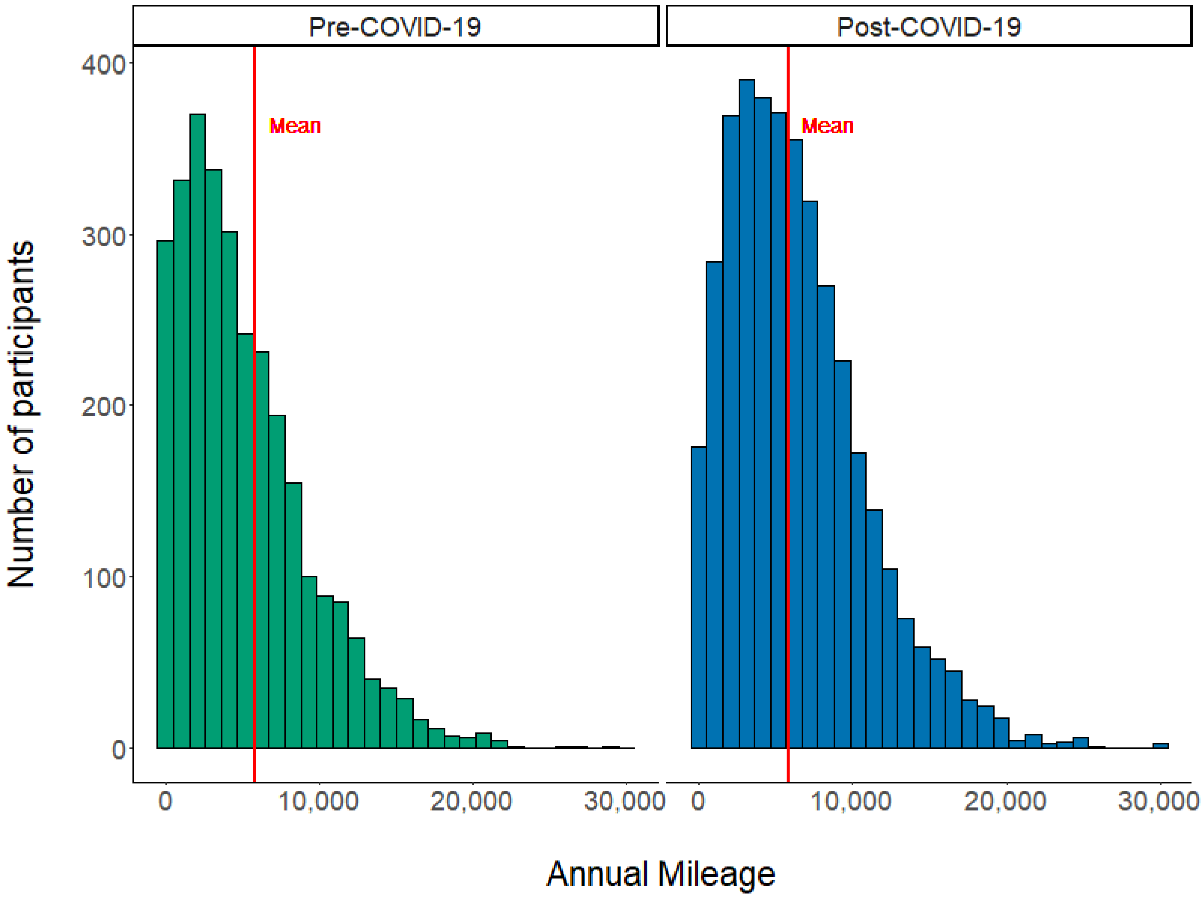


Figure 4-13: Distribution of annual mileage by test pass date (pre or post pandemic) (Pre-pandemic N=2,960, Post-pandemic N=3,888)

Table 4-6: Summary of annual mileage by test pass date (pre or post pandemic) (N=6,848)

		Pre-COVID-19	Post-COVID-19	All participants
Miles driven in 12 months since test pass	Mean (Standard deviation)	5,087 (4,215)	6,365 (4,465)	5,813 (4,404)
Total number of participants		2,960	3,888	6,848

Participants who passed their driving test prior to the start of the pandemic reported lower annual mileage (5,087 miles) than those who passed their test after the pandemic began (6,365 miles). This difference was statistically significant (p<0.01) with a small effect size of 0.18. This result is expected as those who passed their test prior to the start of the pandemic would have been impacted by the various lockdown periods after test pass to a greater extent.

4.3.5.4 Reasons for not driving

Some participants did not drive in each reporting period. This is summarised in Table 4-7. Participants who did not drive at all in the first twelve months (157 participants) were excluded from the statistical modelling (conducted in section 4.4).

Table 4-7: Number of participants who did not drive in each reporting period

Reporting period	Number of participants who did not drive	Proportion of participants who did not drive
Months 1-3	497	7%
Months 4-6	408	6%
Months 7-12	279	4%
First 12 months	157	2%

At each reporting period, participants were asked for specific reasons that may have limited their driving in the reporting period. This included reasons such as not having access to a vehicle, not needing to drive due to change in where participants lived or worked, not needing to drive due to health conditions or to save money, being banned from driving or having lost their licence and due to the COVID-19 pandemic. Figure 4-14 presents the trends of the top three reasons for limitations on their driving over time, for participants who did not drive in each reporting period. As participants would select multiple reasons, the proportions do not add to 100%.

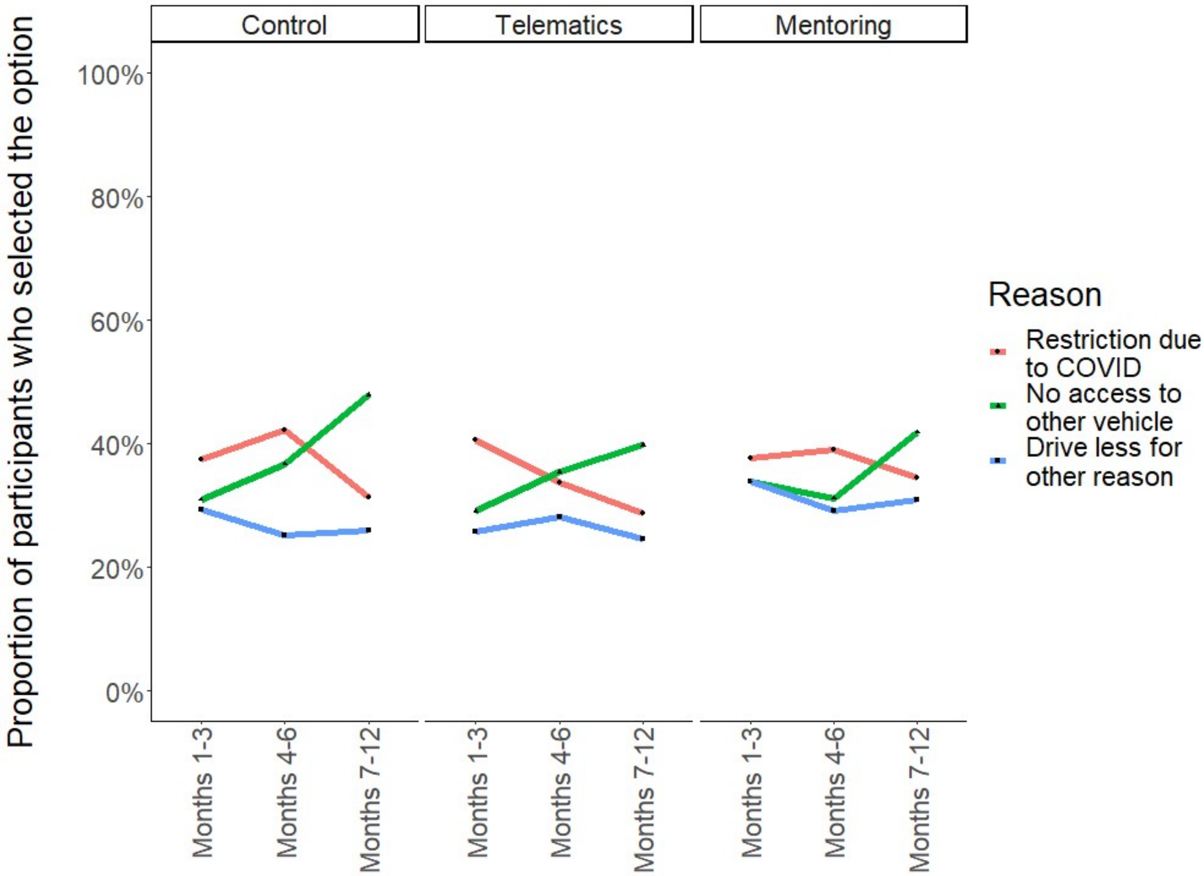


Figure 4-14: Top three reasons for not driving by reporting period

Restrictions due to the COVID-19 pandemic was the most common reason for not driving in the first reporting period (months 1-3 after test pass) shown by the red line in the figure. This was closely followed by not having access to other vehicles (shown by the green line). Over time, the impact of restrictions due to the pandemic typically reduced. By the third reporting period (months 7-12), a higher proportion of participants reported not having access to another vehicle as the main reason for not driving (shown by the green line). This suggests that the effect of the pandemic on driving reduced over the twelve months since test pass; there was no significant difference between the control and intervention groups for this variable.

4.3.6 Exploratory analysis of collisions

At each reporting period (months 1-3, months 4-6 and months 7-12) participants were asked how many collisions they had been involved in over the duration of that period. The responses from each reporting period were combined to obtain the total number of collisions involved in at 12-months since test pass. The descriptive analysis shown in this section excludes participants who did not drive at all in 12 months and aligns with data used to develop the statistical model in section 4.4.

4.3.6.1 *Exploring collisions by age, gender and mileage*

Previous studies (for example, Maycock et al., 1991; Forsyth et al., 1995; Wells et al. 2008) examining the collision rate of young and novice drivers found that the number of reported collisions was related to age, gender, and exposure. More recent studies (Helman et al., 2017) found no clear association between collisions and gender, suggesting that the gender differences have reduced over time when using this kind of measure (self-reported collisions, which tend to be dominated by 'damage only' collisions). Figure 4-15 shows the average number of self-reported collisions per person by age and gender for the 12 months post-test driving. There is no clear pattern by age or gender.

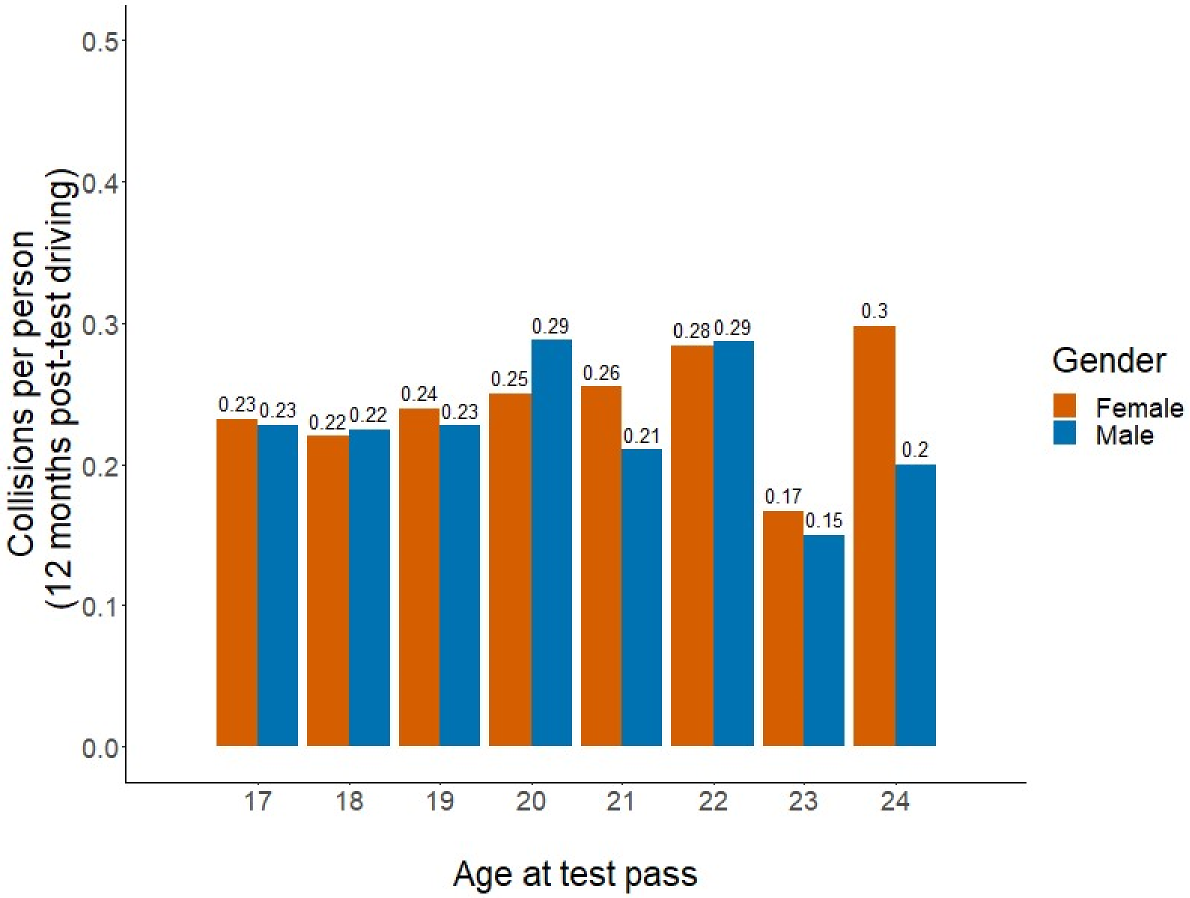


Figure 4-15: Average number of self-reported collisions per person by age and gender (excluding those who did not drive in 12 months, N=6,691)

The lack of difference by gender (average collisions per person for males=0.23 and females=0.24) appears to support the 'closing' of gender differences over time in studies in Great Britain. The lack of differences by age (average collisions per person for 17-year-olds=0.23, 20-year-olds= 0.26, and 24-year-olds=0.27) could be due to the impact of the pandemic typically increasing the age at test pass due to delays in practical driving tests, or a reduced variation in collisions due to the more focused age range (17-to-24-year-olds) recruited in this study. It must be noted that previous studies (e.g., Wells et al., 2008; Helman et al., 2017) included older age groups (17-60+ in both cases) which increased the variation of age in the data.

Figure 4-16 shows the distribution of mileage by the number of collisions reported by participants. This has been represented by a violin plot (shown by the outer shaded region) and boxplot. The violin plot shows the density of the data at different values and the boxplot shows the inter-quartile ranges (25th and 75th percentile) and the median (50th percentile or 'average') value. For instance, the violin plot shows that participants reporting no collisions generally reported a lower annual mileage with the most around 2,500; for participants reporting three or more collisions, the distribution of mileage reported was more spread.

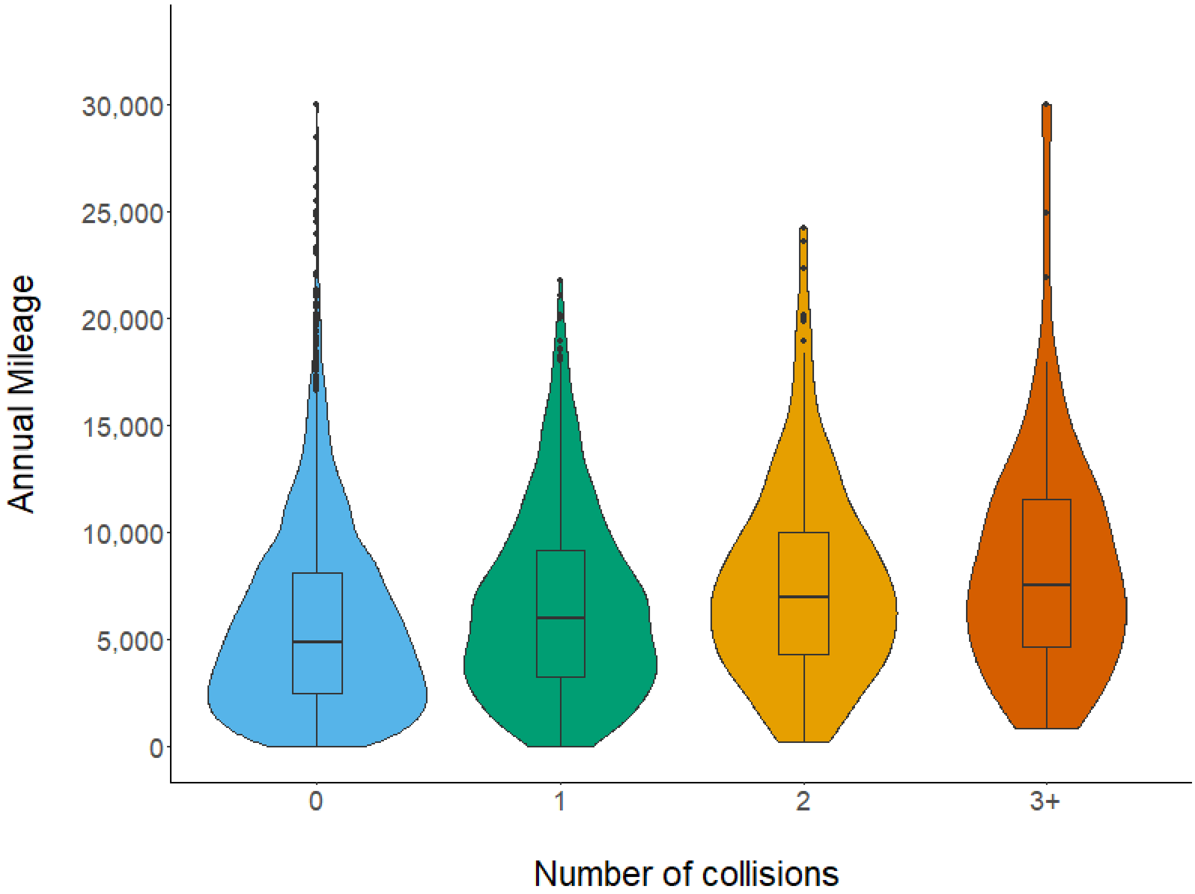


Figure 4-16: Distribution of mileage by the number of collisions (excluding those who did not drive in 12 months, N=6,691).

Participants who had higher annual mileage reported a higher number of collisions on average. The average annual mileage for those who reported having three or more collisions (8,667 miles) was significantly different (p<0.01) compared with those who reported having no collisions (5,746 miles). As a result, mileage was considered an important factor to account for in the collision modelling in section 4.4.

4.3.6.2 Exploring collisions by intervention group

Figure 4-17 shows the proportion of collisions reported by group. The majority (around 80%) of the participants across all groups reported that they had not been involved in any collisions. Roughly 10% in each group reported being involved in one collision. There was no significant difference (*p*=0.49) in the number of people reporting at least one collision between groups.

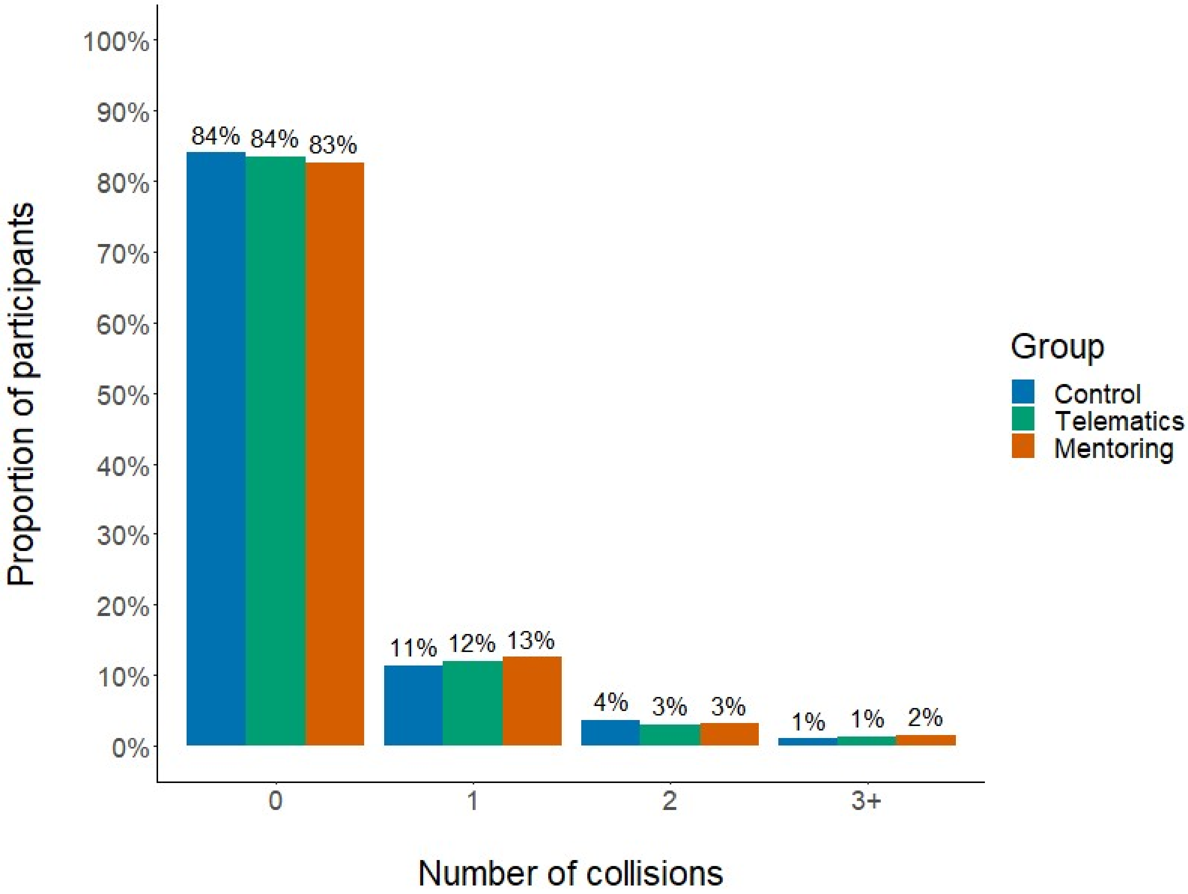


Figure 4-17: Number of self-reported collisions by group (excluding those who did not drive in 12 months, N=6,691)

Table 4-8 below shows collision rate per thousand miles driven by each group. The results show there was no apparent difference between the groups. No statistical testing was undertaken on these rates. Statistical modelling of the effect of the interventions on collisions, accounting for various potential confounding variables in addition to mileage, is presented in section 4.4.

Table 4-8: Collision rate per thousand miles of driving by group

	Control	Mentoring agreement	Telematics	All participants
Total number of participants	2,205	2,294	2,192	6,691
Total number of collisions reported in 12 months	493	568	497	1,558
Total miles driven in 12 months	12,723,068	13,943,613	13,138,773	39,805,454
Average mileage driven per person	5,770	6,078	5,993	5,949
Collision rate per thousand miles	**0.039**	**0.041**	**0.038**	**0.039**

4.3.7 Attrition analysis

Section 3.2.2 described how participants were randomly assigned to groups (with a slight weighting towards the mentoring agreement intervention group); this ensured that the characteristics of the sample across the groups were similar.

As participants dropped out of the study, the samples for the groups changed. This meant that the characteristics of the groups could change in comparison with the characteristics of the registration sample. An attrition analysis was therefore undertaken to investigate how the groups changed over time as participants dropped out, and whether the way the characteristics of the samples changed were consistent over time across the groups. In other words, the analysis checked for any bias introduced by different types of people dropping out of different groups.

Overall, there were changes in the characteristics of the sample over time as participants dropped out. However, these changes were consistent across the groups. There were no noticeable differences in the behaviour of the characteristics for the groups. Therefore, there was no need to correct samples for differences between them.

4.4 Intention to Treat (ITT) analysis

The intention to treat (ITT) analysis analysed the effect of the interventions on collisions (relative to the control group) at the population level. It included participants who did, and did not, engage with the interventions.

This section first examines the relationship between self-reported collision numbers and key driver variables, namely age, gender, mileage, and the impact of the COVID-19 pandemic in greater detail. Participants who did not drive at all in the first twelve months (157 participants) were excluded from the statistical modelling.

This model, independent of any intervention effect, forms the base model (shown in section 4.4.4) and can be used to hold certain factors constant when undertaking other analyses. Next, the intervention groups (control, telematics, and mentoring agreement) are added to the base model forming the ITT analysis (results in section 4.4.5). This tests whether the group to which a participant was assigned (mentoring agreement, telematics, control) impacted on their collision risk at twelve months post-test, independently of their level of engagement with the intervention on offer. In addition to the models above, the effect of driver experience (effect of months since test pass) on the number of reported collisions was modelled.

The key insights from this analysis are as follows:

- Higher mileage, passing the test before the COVID-19 pandemic, and a higher self-reported impact of COVID-19 on learning to drive were all associated with greater collision risk.
- Neither of the interventions was associated with any change in collision risk.

4.4.1 Background

The hypotheses for the two interventions were as follows:

Mentoring agreement: A web-based 'mentoring agreement' provided to novice drivers aged 17-24 in Great Britain (for them to use if they wish) will result in a lower proportion of such drivers who have a crash in their first 12 months of post-test driving, relative to a control group of novice drivers who do not receive the opportunity to use the web-based parent-teen mentoring agreement.

Telematics: A mobile-phone-based app which provides feedback (including rewards and penalties) on driving style provided to novice drivers aged 17-24 in GB (for them to use if they wish) will result in a lower proportion of such drivers who have a crash in their first 12 months of post-test driving, relative to a control group of novice drivers who do not receive the opportunity to use the app.

Based on previous studies of this type (Maycock, Lockwood & Lester, 1991; Forsyth et al., 1995; Wells et al., 2008; Helman et al. 2017) one would expect to see some or all of the following have independent impacts on collision risk in the first year of post-test driving:

- Exposure: drivers who drive more would be expected to have higher collision risk
- Age: younger drivers would be expected to have higher collision risk
- Gender: in some of the previous studies males have had a higher collision risk than females (although this difference was reducing throughout the 1990s and 2000s, and was not present in the most recent analysis reported in Helman et al., 2017)
- Experience: drivers would be expected to have a higher collision risk in their early months of driving post-test than in their later months of driving.

There may also be other variables that need to be accounted for in the current dataset – specifically there are potential impacts of the COVID-19 pandemic to consider; these were explored.

4.4.2 Modelling method

The modelling method applied to the data in this study was a multivariate regression technique known as Generalised Linear Modelling (GLM). Multivariate regression techniques are designed to explore the relationship between the response variable (in this case the number of self-reported collisions at 12 months since test pass) and a number of exploratory variables (such as age, gender, etc.) on which the response variable is assumed to depend. In this study, the response variable, number of collisions at 12 months since test pass, is assumed to follow a negative binomial distribution. The decision of using a negative binomial distribution over a Poisson distribution was determined through likelihood ratio and overdispersion tests. This distribution is frequently used to model count data such as collisions and has been used in previous studies of similar nature (Wells et al., 2008; Helman et al., 2017).

The base model was developed by individually adding key driver variables to establish whether the addition has a statistically significant effect on the response variable (number of collisions). The key driver variables were identified from previous studies (section 4.4.1) on young and novice drivers. In addition to these variables, certain impacts of the COVID-19 pandemic were included in the model to explore the degree to which it may have impacted the number of collisions.

4.4.3 Impact of experience

Previous studies have shown that experience has a significant impact on the likelihood of a collision. Wells et al. (2008) modelled experience as the number of years the participant has been driving and the same process was followed in this study (using months rather than years). In order to model this, a repeated measures design was applied, and experience was coded as 0.25 for those who reported driving in the first six months and 0.75 for those who reported driving in the second six months since test pass.

Generalised Linear Models (GLM) assume that the data modelled are independently distributed (the response from one participant is independent of the other). However, in order to develop a model which includes experience over time as an explanatory variable, an approach which considers that some of the data are related is needed (as each participant responds to the surveys at multiple time points). Therefore, a mixed effects model was used to examine the relationship between driver experience and number of collisions. Mixed effects models are applied in settings where repeated measurements are made on the same participant, and they can also account for any missing values. It was assumed that the number of collisions followed a negative binomial distribution (similar to the base model using GLM).

The variables included in the model were mileage, age at test pass and gender along with the additional variable, experience. The model showed that experience was statistically significant ($p<0.01$) with a negative coefficient of 0.40 suggesting that the likelihood of a collision reduces as experience since test pass increases. This is what would be expected based on previous findings from studies of this kind.

4.4.4 Base model

The base model for the reported number of collisions in the first 12 months after passing their practical driving test was of the form:

log (Collisions) = $\beta_0 + \beta_1$ log(Mileage) + β_2 Gender + β_3 Age + β_4 TestPassDate + β_5 COVIDImpact

where β_n are the coefficients to be estimated. The variables were included one by one in the following form:

- Mileage: This is number of miles driven in the 12 months since test pass. This variable was included as a natural logarithm to account for the fact that the likelihood of a collision increases as mileage increases, but not in a linear manner.
- Gender: This was included as a categorical variable, where 0 was coded as female and 1 was coded as male.
- Age: This is age at the time the participant passed the practical driving test. A variety of functional forms were tested (logarithmic, exponential, inverse and categorical) but a continuous variable was found to be the most appropriate fit for the data.
- TestPassDate (pre- or post-pandemic): This was determined based on the date of test pass (before or after the start of lockdown in March 2020) as a surrogate for the impact of the COVID-19 pandemic on passing the test. Post-COVID-19 pandemic was coded as 0 and Pre-COVID-19 pandemic was coded as 1.
- COVIDImpact: This variable was calculated as a continuous score based on participants' response around how the pandemic affected their learning to drive process. The scores are explained in greater detail in section 4.3.5.2.

No term for experience was included as the main analysis considered the first 12 months of driving as a whole; GLMs are not able to handle temporal variables well, and a decision was made to exclude experience from the base model on this basis. The coefficients for the base model are shown in Table 4-9.

Table 4-9: Coefficients for the base model (* denotes statistically significant coefficients)

Variable	Coefficient	Standard Error	*p*-value
Intercept	-4.935	0.426	<0.01(*)
Log(Mileage)	0.362	0.037	<0.01(*)
Gender: Male	-0.088	0.067	0.19
Age at test pass	0.011	0.014	0.40
Test pass: Pre-COVID-19 pandemic	0.235	0.085	<0.01(*)
COVID-19 pandemic impact on learning to drive	0.073	0.031	0.02(*)

The base model showed that:

- The log of mileage coefficient was statistically significant ($p<0.01$) with a positive coefficient of 0.36. This suggests that as mileage increases, the likelihood of a self-reported collision increases. To illustrate what this means in practice, assuming an

18-year-old male who passed his test before the pandemic and therefore had no COVID impact on learning to drive, the model suggests that if mileage is increased from 2,500 miles to 5,000 miles, then the likelihood of a collision is 1.28 times as high (an increase of 28% from 0.176 to 0.226 average collisions).

- Unlike in earlier studies (Forsyth et al., 1995; Wells et al., 2008) but in line with the findings from the more recent DVSA study (Helman et al., 2017), gender was not shown to have a significant impact on the number of self-reported collisions. However, despite the non-significance of this variable, it was decided that the base model should include this variable in order to control for any effects of gender when considering the impact of the interventions in the ITT analysis (see section 4.4.5).
- The age at test pass was not shown to be a significant predictor of collisions. This result is unlike those in previous studies of this type, and is likely due to one or both of the following reasons. Firstly, previous studies included a larger age range (from 17 to over 60), whereas this study focused on a smaller range of ages (from 17 to 24) which may have reduced the variability in number of collisions due to age. Secondly, the COVID-19 pandemic resulted in driving tests being delayed substantially, which may have had an impact on when participants were able to take their tests and the amount of time spent learning, thereby diluting the impact of age at test pass. Similar to gender, it was decided to control for age at test pass for the ITT analysis even though it was not significant in the base model.
- The date of test pass (Pre-COVID-19 or Post-COVID-19) was statistically significant (p<0.01) with a positive coefficient of 0.23 for those who passed their test before the pandemic. This suggests that those who passed their test before the pandemic began were 1.26 times more likely to be involved in a collision (a 26% increase) than those who passed their test after the pandemic began.
- The self-reported impact of the pandemic on the learning to drive process had a significant impact on the number of collisions (p=0.02). The coefficient was positive suggesting that the higher the self-reported impact of COVID-19 pandemic on the learning to drive process for those who passed their test after the pandemic began, the greater is the likelihood of a collision. Assuming an 18-year-old male who drove 5,000 miles and passed his test after the pandemic began, varying the score from -1 (implying the pandemic had a positive impact on a participant's learning to drive process) to 2 (implying the pandemic had a negative impact on a participant's learning to drive process) increases the likelihood of a collision by roughly 1.24 times (a 24% increase from 0.16 to 0.20) average collisions.

4.4.5 Effectiveness of interventions

By adding the intervention groups to the base model, the analysis explores whether the groups (control, telematics, and mentoring agreement) differ in their collision involvement at 12 months post-test, after considering and controlling for the variables in the base model. The purpose of this analysis is to understand the real-world effectiveness of the interventions regardless of the level of engagement with the intervention for each participant. Put simply, it estimates what happens to collision risk at the population level, if such interventions are offered on a voluntary basis. The population in this case is all novice drivers aged 17-24 in Great Britain, although as previously noted, caution is needed when generalising from the sample in the study (which is more female than the population and may differ in other ways).

The variable ‘group’ was added to the base model as a categorical variable and the coefficients are presented in Table 4-10. One of the groups was arbitrarily chosen as the baseline when the project team was blinded for the analysis (to avoid potential bias in model interpretation). After unblinding, it turned out that the control group had been chosen as the baseline.

Table 4-10: Coefficients from the ITT analysis ((*) denotes statistically significant coefficients)

Variable	Coefficient	Standard Error	*p*-value
Intercept	**-4.958**	**0.427**	**<0.01(*)**
Log(Mileage)	**0.362**	**0.036**	**<0.01(*)**
Gender: Male	**-0.088**	**0.067**	**0.18**
Age at test pass	**0.012**	**0.014**	**0.39**
Test pass: Pre-COVID-19 pandemic	**0.236**	**0.085**	**<0.01(*)**
COVID-19 pandemic impact on learning to drive	**0.073**	**0.031**	**0.02(*)**
Group: Telematics	0.005	0.077	0.95
Group: Mentoring agreement	0.068	0.076	0.37

The analysis showed that the groups did not differ significantly on self-reported collisions in the first 12 months of driving post-test. The response variable for the base model and the ITT analysis was also modelled as a binary outcome (no collisions or one or more collisions) using a logistic regression model. However, the results remained unchanged.

The ITT analysis suggests that if the mentoring agreement or telematics app are offered on a voluntary basis, with incentives that match what would be feasible in live roll-out but do not excessively incentivise engagement, there is likely to be no safety benefit in terms of collision reduction at the population level (the population being all novice drivers aged 17-24). This is highly likely to be the case even if the interventions are found to be effective in changing surrogate measures of risk for the small proportion of people who do engage. Put simply, engagement levels are so low under voluntary take-up (just 4% and 16% for the mentoring agreement and telematics interventions respectively, see section 4.5.2) that any impact on safety is likely to be too small to detect at the population level.

4.5 Per protocol (PP) analysis

This section first presents a background to the per-protocol analysis and the rationale for using it. It then presents an analysis of the characteristics of people engaging with the interventions, to check for any bias that might need to be corrected when comparing surrogate measures. Finally, it presents the analysis, checking for the effectiveness of the interventions in changing the surrogate measures.

The key insights from this analysis are as listed below, separately for the mentoring agreement and telematics interventions.

For the mentoring agreement intervention:

- Only 4% of participants assigned to the mentoring agreement group engaged with the intervention.
- There were no significant demographic differences between those who engaged with the mentoring agreement and those who did not.
- In comparison with the control group, those who engaged with the mentoring agreement intervention:
  - Were more likely to report setting voluntary agreements limiting their post-test driving with peer-age passengers in their first six months of post-test driving.
  - Were more likely to report setting voluntary agreements limiting their post-test driving in the dark in their first six months of post-test driving.
  - Reported less of their driving as being in the dark in months 4-6 post-test.

  All of these changes would be expected to have a safety benefit.
- No other differences were found between those who engaged with the mentoring agreement intervention and the control group.

For the telematics intervention:

- Only 16% of participants assigned to the telematics group engaged with the intervention.
- Individuals who engaged with the telematics intervention were more likely to be younger, female, or in full-time education compared with those who were assigned to the intervention but didn't engage. These differences were accounted for in the analysis.
- In comparison with the control group, those that engaged with the telematics intervention:
  - Were more likely to self-report a driving style that was ‘inattentive, careless, irresponsible and risky’.
  - Reported less of their driving as being in the dark in months 7-12 post-test.

  The latter of these changes would be expected to have a safety benefit, while the former would be expected to have a safety disbenefit; it is possible however that the driving style finding reflects better recall of driving (due to feedback from the app) rather than a genuine change in driving style.
- No other differences were found between those that engaged with the telematics intervention and the control group.

4.5.1 *Background*

The per-protocol (PP) analysis looked at those participants in the treatment groups who engaged to some degree with the intervention they were offered. The analysis compared people who actually engaged in some way with their interventions (either the mentoring agreement, or the telematics app) with control participants, on a set of surrogate measures of safety (see section 3.2.3 for details). The impact on number of collisions was not explored in the per-protocol analysis as in the ITT analysis due to the small sample size of participants in this analysis. A small sample like this cannot support a robust analysis if using a measure such as collisions as the outcome variable; the low baseline and higher variability of the collisions measure (compared with the surrogate measures) means that much higher sample sizes are needed to support a robust analysis.

Note that in this section, the sample used is not the same as the trial sample used in the previous ITT analysis. Data on whether participants engaged (or did not engage) with the interventions was available (from delivery partners) regardless of whether they completed any data collection surveys. Therefore, for the matching process described in section 4.5.2, all participants recruited for the novice arm were initially included. For the analysis of effectiveness in section 4.5.3, the sample used depends on the measures being analysed – this is specified in the text where relevant.

4.5.2 *Characteristics of engagers and correcting for engager bias*

An important part of the per-protocol analysis was to ensure as far as possible that, when comparing the treatment engagers with the control participants, this would be comparing 'like for like' in every way possible except the receiving of the intervention. It would not be possible to simply compare engagers in a treatment group with the entire control group; this could well lead to bias as it would be comparing one 'type' of person in a treatment group (the type of person who engages) with at least two 'types' in the control group (people who do, and do not, engage).

Therefore, checks were made for engager bias before running the analysis to assess whether the measured characteristics of people who engaged with the interventions were different to those of people who did not, and to determine whether it was necessary to sample the control group to ensure that the characteristics of the control group participants matched those of the engagers in the respective intervention group. Checking for engager bias was done by comparing engagers and non-engagers for each intervention on a number of different characteristics such as age, gender, education status, social deprivation and whether people do any driving. In addition, a 30-item personality inventory was used in the registration survey to score people on the 'Big Five' personality factors (Openness, Conscientiousness, Extraversion, Agreeableness, Neuroticism). Comparisons performed on these personality variables revealed differences only for the conscientiousness factor for telematics engagers and non-engagers, with a negligible effect size. Therefore, personality variables were not used when sampling the control group to find a matching sample. While a range of measures were used for this checking procedure, it remains possible that the engagers and non-engagers might have differed on other unknown variables.

In all cases the threshold for what constituted being an 'engager' was relatively low. This needs to be considered when interpreting the results; because engagement 'dose' or ‘amount’ of contact with the interventions was low even in those people who did engage, it may be that the effects seen were limited by this.

4.5.2.1 Mentoring agreement

An engager with the mentoring agreement intervention was defined as someone who set at least one agreement using the web-based system. The information used to identify engagers came from RoSPA who provided data with the ‘stage reached’ for each participant who had signed up with them. The stages were (in order): Signed up; Set agreement 1; Set review 1; Set agreement 2; Set review 2; Set agreement 3; Completed.

An engager was therefore defined as any participant with ‘stage reached’ recorded as ‘Set agreement 1’ or later. Any participant who had only reached the ‘signed up’ stage was classed as a non-engager. There were also a number of participants who never signed up with RoSPA and were therefore not included in their data, these participants were also classed as non-engagers. Table 4-11 shows the breakdown of participants in the mentoring agreement group by level of engagement. Of the 4,183 participants in the mentoring agreement group, only 4% (148 participants) were identified as engagers. That is to say – of the people given the opportunity to use the mentoring materials, only 10% registered with the provider, and only 4% used the materials, with less than 1% using them throughout their entire learning to drive journey, as had been intended (50 participants set two agreements, 42 participants set three agreements and only 31 completed the final review).

Table 4-11: Summary of engagement for mentoring agreement group

	Number of participants	% of participants in mentoring agreement group
Total in mentoring agreement group	4,183	100%
Total signed up with RoSPA	410	10%
Total who set at least one agreement (engagers)	**148**	**4%**

The characteristics of the 148 engagers were compared with those of the 4,035 non-engagers to determine if the ‘type’ of person who engaged with the mentoring agreement differed from the ‘type’ who did not. The analysis found no differences between engagers and non-engagers for any of the variables compared. This meant that the mentoring agreement engagers could be compared with the whole control group in further analysis and there was no need to use a matching process to enable a comparison; there was confidence that the comparisons would be ‘like for like’ (apart from receiving the intervention) based on the measures used.

Charts showing the full results of all the comparisons carried out are presented in Appendix E.

Table 4-12 shows the number of mentoring agreement intervention engagers and participants from the control group sample who completed each survey, and the total number of mentoring agreement engagers and control group sample participants included in the per-protocol analysis.

Table 4-12:Summary of participants included in per-protocol analysis for the mentoring agreement intervention

Survey	Number of mentoring agreement engagers	Number of control group sample participants
Test pass	148	3,373
3-month	145	2,908
6-month	138	2,669
12-month	134	2,473
Total filling out all surveys	131	2,289
Total included in PP analysis - answered at least one survey	**148**	**3,460**

4.5.2.2 Telematics

The telematics provider supplied engagement data to indicate the level of engagement for the participants who had entered their programme. Participants were assigned to one (or more) of five stages to indicate their level of engagement with the telematics programme:

- Stage 1 – Entered the programme (registered with the provider) but did not download the app
- Stage 2 – Entered the programme and downloaded the app but no evidence of any engagement with the app (No mileage or journeys recorded)
- Stage 3 – Entered the programme, downloaded and paired the app, and evidence of journeys recorded (the degrees to which participants in this group engaged are broken down further into Stages 3a and 3b – all participants in Stage 3 were also in either Stage 3a or Stage 3b.)
    - Stage 3a – This is a subset of Stage 3. The app was paired but less than 10 miles was recorded throughout the time spent in the programme
    - Stage 3b – This is a subset of Stage 3. This group were more engaged with higher levels of mileage (more than 10 miles) and higher numbers of treats awarded to them.

An engager with the telematics intervention was defined as anyone who was assigned to Stage 3 – that is they downloaded and paired the app, and had some journeys recorded. Any participants assigned to Stage 1 or Stage 2 were classed as non-engagers. Also, any participant who did not enter the programme at all was classed as a non-engager. Table 4-13 shows the breakdown of participants in the telematics group by level of engagement. Of the 4,180 participants in the telematics group, 689 (16%) were identified as engagers.

Table 4-13: Summary of engagement for telematics group

	Number of participants	% of participants in telematics group
Total in telematics group	4,180	100%
Total entered telematics programme	3,545	85%
Stage 1 total	515	12%
Stage 2 total	2,341	56%
Stage 3 total	689	16%
Stage 3a total	379	9%
Stage 3b total	310	7%
Total engagers with telematics programme (Stage 3 total)	**689**	**16%**

The characteristics of the 689 engagers were compared with those of the 3,491 non-engagers. There were three variables where differences were identified: age, gender, and education status (shown in Figure 4-18, Figure 4-19 and Figure 4-20 respectively). Charts showing the results for the other variables examined are presented in Appendix D.

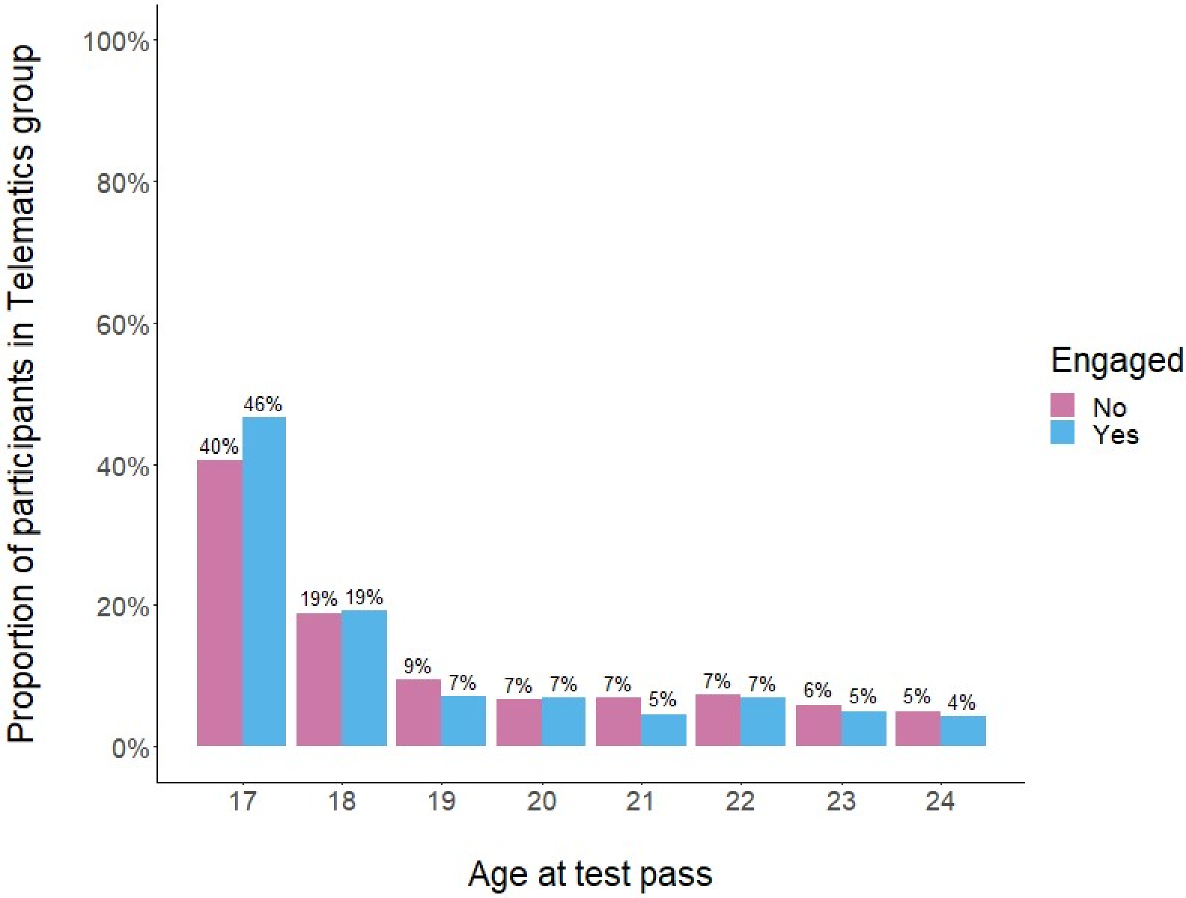


Figure 4-18: Proportion of engagers and non-engagers in the telematics group by age (689 engagers and 3,491 non-engagers)

Figure 4-18 shows that the proportion of participants in the 17-year-old category was higher for engagers than non-engagers (46% and 40% respectively), suggesting that younger people were more likely to engage with the telematics intervention. When comparing the mean age for the two groups, this effect was less pronounced; the mean age for non-engagers was 18.9 and for engagers it was slightly lower at 18.6.

Statistical tests showed that there was a significant difference in the distribution of engagers and non-engagers by age at test pass (*p*=0.04), albeit with a small effect size of 0.05.

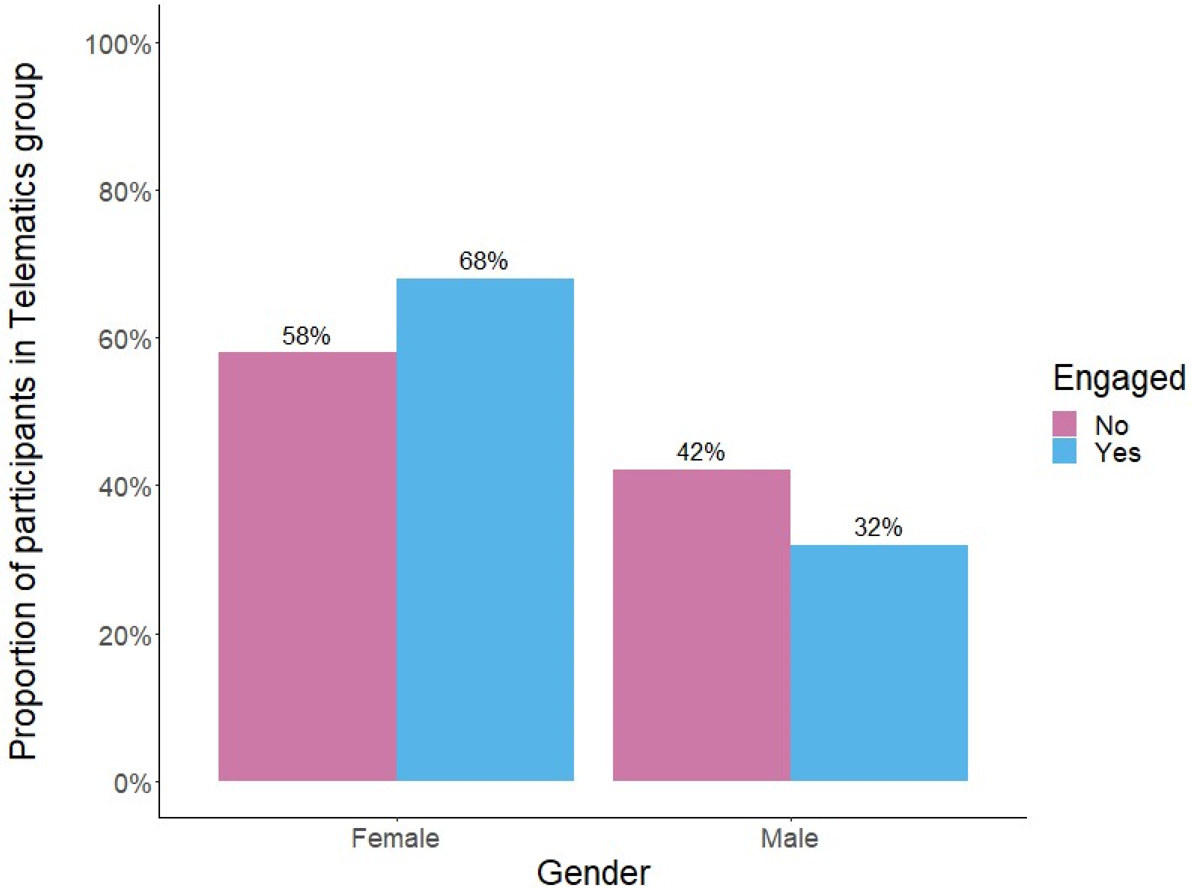


Figure 4-19: Proportion of engagers and non-engagers in the telematics group by gender (689 engagers and 3,491 non-engagers)

The results in Figure 4-19 show that engagers with the telematics intervention were more likely to be female than male; 68% of engagers were female, compared with 58% of non-engagers.

Statistical tests showed a significant difference in the distribution of engagers and non-engagers by gender (*p*<0.01) with a small effect size of 0.07.

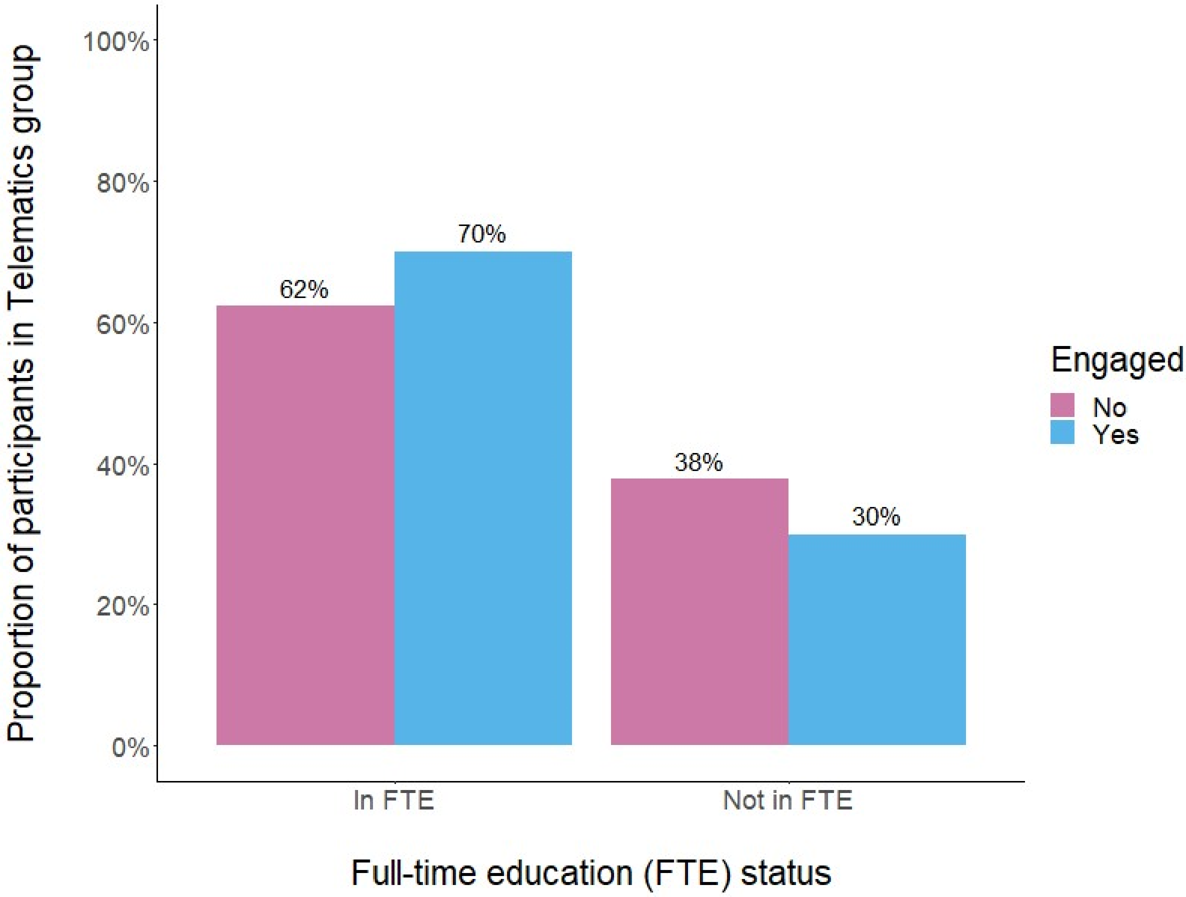


Figure 4-20: Proportion of engagers and non-engagers in the telematics group by education status (689 engagers and 3,491 non-engagers)

Figure 4-20 shows that the proportion of participants who were in full time education when they began the trial was higher for engagers (70%) than non-engagers (62%).

Statistical tests showed that the difference in distribution of responses between engagers/non-engagers and those in full-time education was significant ($p<0.01$) with a small effect size of 0.05.

The differences identified between engager and non-engager characteristics, despite the small effect sizes, meant that it was necessary to correct for engager bias in the control group, in order for comparisons between engagers and control participants to be robust. Therefore, a nested case control method was used to create a sample of control group participants with characteristics that matched those of the participants who had engaged with the telematics intervention. For details on this process, see Appendix F.2.

Table 4-14 shows the number of telematics engagers and participants from the control group sample who completed each survey, and the total number of telematics engagers and control group sample participants included in the per-protocol analysis.

Table 4-14: Summary of participants included in per-protocol analysis for the telematics intervention

Survey	Number of telematics engagers	Number of control group sample participants
Test pass	666	2,392
3-month	583	2,081
6-month	542	1,910
12-month	514	1,752
Total filling out all surveys	489	1,636
Total included in PP analysis - answered at least one survey	**669**	**2,447**

4.5.3 Effectiveness of interventions

This section presents the findings on the effectiveness of the two interventions in changing the relevant surrogate outcomes defined. In other words, for people who engage, what does the intervention change?

The engagers with each intervention were compared with the relevant control group participants over each of the reporting periods to assess what difference the intervention made to a variety of surrogate measures. As noted in section 2.2.1, some measures assessed were intervention-specific (e.g., number of limits set on driving in certain conditions or driving style) and some were general surrogate safety measures (e.g., number of near misses).

The number of engagers and control group participants in each comparison varies for each reporting period because some participants did not answer all surveys. However, the results for the analysis of variance (ANOVA) statistical tests (which are performed to look at differences by time, group and the interaction between time and group) in this section only use the data from participants filling out all surveys, as they require no missing entries. The number of these participants is given in Table 4-11 and Table 4-14 for each of the intervention groups.

4.5.3.1 Mentoring agreement

This section presents the results of analysis which seeks to answer the question: for people who engage, what does the mentoring agreement intervention change? Mentoring agreement engagers and control group participants were compared using the following measures to try to answer this question:

- Self-reported near misses – testing whether there were fewer for those who engaged with the mentoring agreement intervention
- Self-reported Driving Events scale score – tested whether this was lower (indicating lower frequency of unsafe driving events) for those who engaged

- Self-reported frequency of breaking the speed limit when driving – tested whether this was less frequent for those who engage
- Self-reported limits set on driving in certain conditions – tested whether more limits were set by mentoring agreement intervention engagers
- Self-reported proportion of mileage driven in certain conditions (with at least one passenger of similar age, in the dark) – tested whether this was lower for mentoring agreement intervention engagers
- Self-reported attitudes towards new driver restrictions and more general enforcement – tested whether greater agreement was seen from mentoring agreement engagers, especially on new driver restrictions.

All comparisons in this section have been made between 148 participants who engaged with the mentoring agreement intervention and all 3,460 control group participants that filled in at least one survey; however, numbers of participants vary by reporting period. See section 4.5.2.1 for the definition of a mentoring agreement intervention engager, and a summary of the number of participants in each group included in the analysis of each reporting period.

Near misses

At each time point, participants were asked how many 'near misses' they had experienced during that reporting period. A near miss was described as the impression of only just avoiding an accident.

Figure 4-21 shows the distribution of number of near misses (corrected for mileage) reported by participants who engaged with the mentoring agreement intervention and control group participants, during each of the reporting periods (months 1-3, months 4-6 and months 7-12). This is presented as a boxplot which displays the inter-quartile range (25^{th} and 75^{th} percentiles) as the top and bottom of the box, and the median (50^{th} percentile or 'average') as the horizontal line in the box. Values outside of the 95% confidence interval are shown as individual data points. The mean number of near misses per mile reported for both groups across the three periods is shown in Table 4-15.

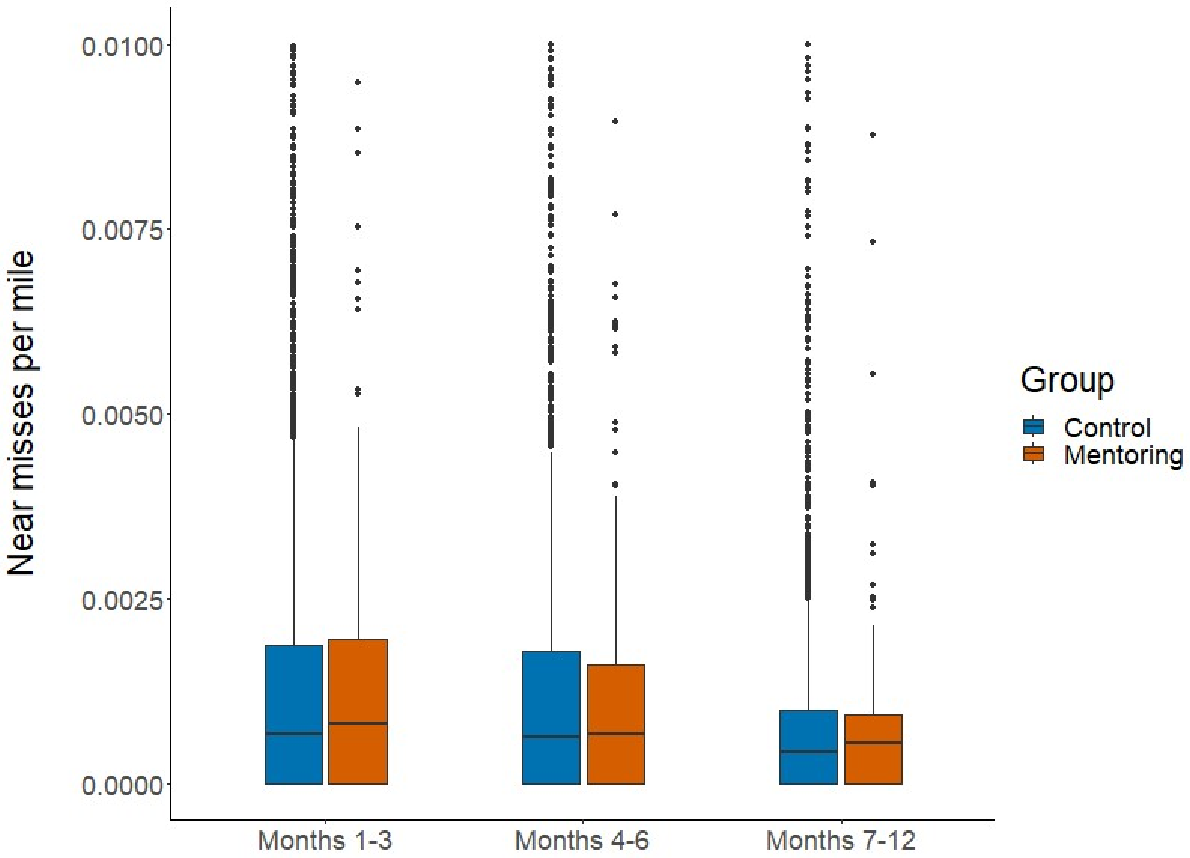


Figure 4-21: Distribution of number of reported near misses per mile in each post-test reporting period for mentoring agreement engagers and control group participants (Months 1 to 3: 145 engagers, 2,908 control group. Months 4 to 6: 138 engagers, 2,669 control group. Months 7 to 12: 134 engagers, 2,473 control group).

Table 4-15: Number of reported near misses per mile in each post-test reporting period for mentoring agreement engagers and control group participants (Months 1 to 3: 145 engagers, 2,908 control group. Months 4 to 6: 138 engagers, 2,669 control group. Months 7 to 12: 134 engagers, 2,473 control group)

	Mean number of near misses per mile driven		
Group	**Months 1-3**	**Months 4-6**	**Months 7-12**
Control group participants	0.0009	0.0008	0.0005
Mentoring agreement engagers	0.0011	0.0009	0.0005

Engagers with the mentoring agreement intervention and control group participants all reported low numbers of near misses per mile in all reporting periods. The rates per mile are shown in Table 4-15. For context, in months 1-3 and 4-6 people in both groups combined were reporting around one near miss for every 1,000 miles of driving, with this dropping to around one every 2,000 miles of driving in months 7-12. The number of near misses per mile varied slightly between the intervention and control groups in months 1-3 and months 4-6 but was the same in months 7-12.

Both groups reporting decreasing numbers of near misses per mile over time. The number of near misses per mile reported by mentoring agreement engagers was slightly higher than that reported by control participants in the first two reporting periods but in months 7-12 both groups reported the same number of near misses per mile.

Statistical tests showed no significant differences between the two groups (p=0.80). The differences over time were significant (p=0.03) for both groups with a very small effect size of 0.0009. The interaction effect was not significant (p=0.35), meaning that the effect of time was the same for both groups.

Driving events score

In each post-test survey participants were asked about how often each of six driving events had occurred whilst they were driving during that reporting period:

- Braking sharply to avoid a collision with the vehicle ahead because it had slowed
- Pulling out to overtake or turn right not noticing another vehicle in their 'blind spot'
- Failing to notice someone waiting at a pedestrian crossing
- When cornering, finding they were traveling too fast to negotiate the bend safely and having to brake
- Failing to give way when entering a roundabout to a vehicle already on the roundabout
- Braking or swerving suddenly to avoid an accident.

Participants indicated how often each of the above events had occurred using a six-point scale ranging from ‘never’ (a score of 1) to ‘nearly all the time’ (a score of 6). The sum of these scores across all six events was then calculated to give an overall ‘driving events score’ A lower driving events score indicates fewer occurrences of the driving events and therefore theoretically safer driving.

Figure 4-22 shows the distribution of driving events scores for each reporting period for participants who engaged with the mentoring agreement intervention and participants in the control group. The corresponding means are shown in Table 4-16.

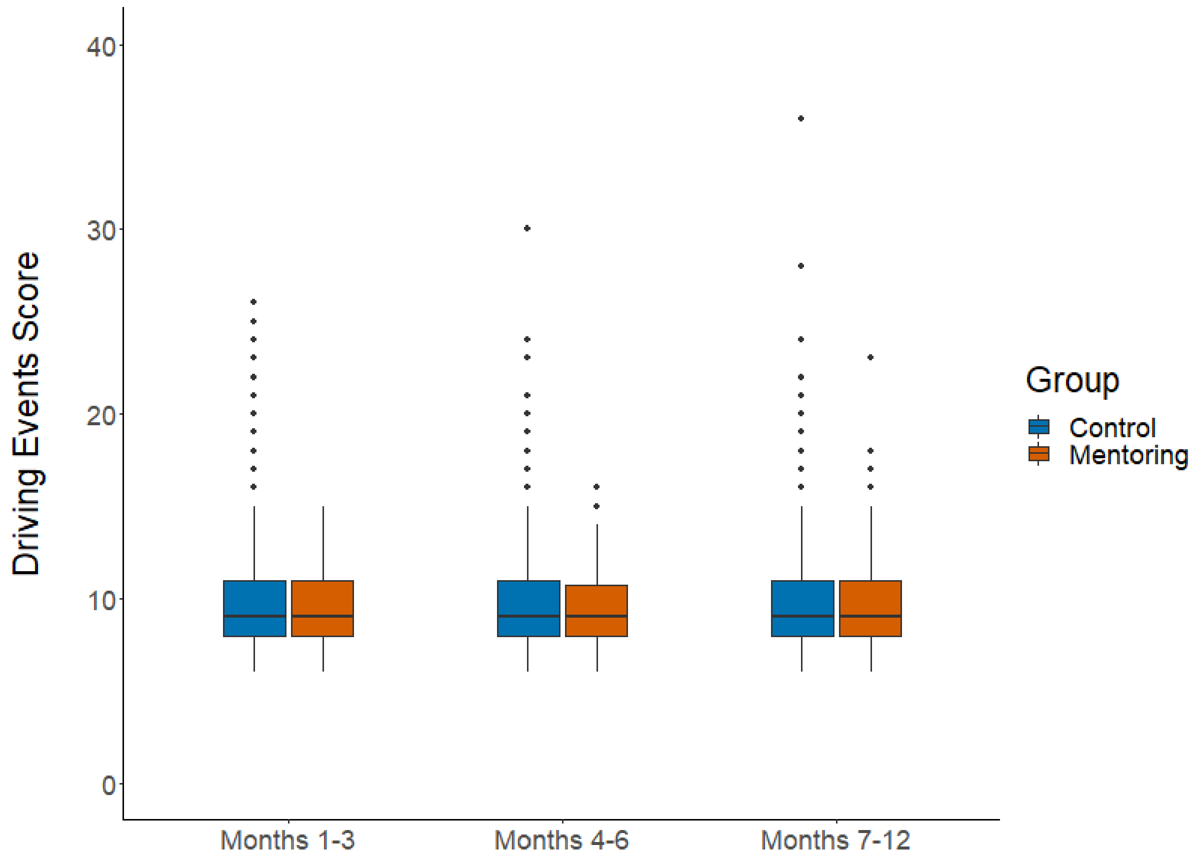


Figure 4-22: Distribution of driving events scores for each post-test reporting period for mentoring agreement engagers and control group participants (Months 1 to 3: 145 engagers, 2,908 control group. Months 4 to 6: 138 engagers, 2,669 control group. Months 7 to 12: 134 engagers, 2,473 control group).

Table 4-16: Mean driving event score in each post-test reporting period for mentoring agreement engagers and control group participants (Months 1 to 3: 145 engagers, 2,908 control group. Months 4 to 6: 138 engagers, 2,669 control group. Months 7 to 12: 134 engagers, 2,473 control group).

	Mean driving event score		
Group	Months 1-3	Months 4-6	Months 7-12
Control group participants	9.24	9.49	9.71
Mentoring agreement engagers	9.40	9.32	9.65

For all reporting periods, the distributions of driving events scores were very similar for mentoring agreement engagers and control group participants, with the mean scores between 9 and 10 across all reporting periods for both groups.

There were no clear patterns seen in the mean driving events score across reporting periods and all differences between groups and periods were very small. In months 1-3 mentoring agreement engagers reported higher driving events scores than control group participants (i.e., the driving events listed above happened more frequently for the mentoring agreement group), but in months 4-6 this pattern was reversed. In months 7-12, the mean score is lower for the mentoring agreement engagers than the control participants (indicating lower frequency of driving events for mentoring agreement engagers). Both groups reported higher mean scores in months 7-12 than in the earlier reporting periods. Statistical tests showed no significant differences between groups (p=0.92). Taking intervention and control groups together, the difference seen in the mean driving events score over time was significant (p<0.01) with a very small effect size of 0.001. The interaction between group and time was not significant (p=0.42) meaning that the effect of time was the same for both groups.

Limits set on driving in certain conditions

At the end of each reporting period participants were asked a question about setting limits on their driving. The question asked whether a parent, guardian or other mentor had been involved in setting limits on their driving in a number of situations. Two of these situations related directly to the mentoring agreement intervention: driving with passengers of a similar age and driving at night or in the dark.

Figure 4-23 shows the proportion of mentoring agreement intervention engagers and the proportion of control group participants who reported setting limits on their driving in the situations mentioned above.

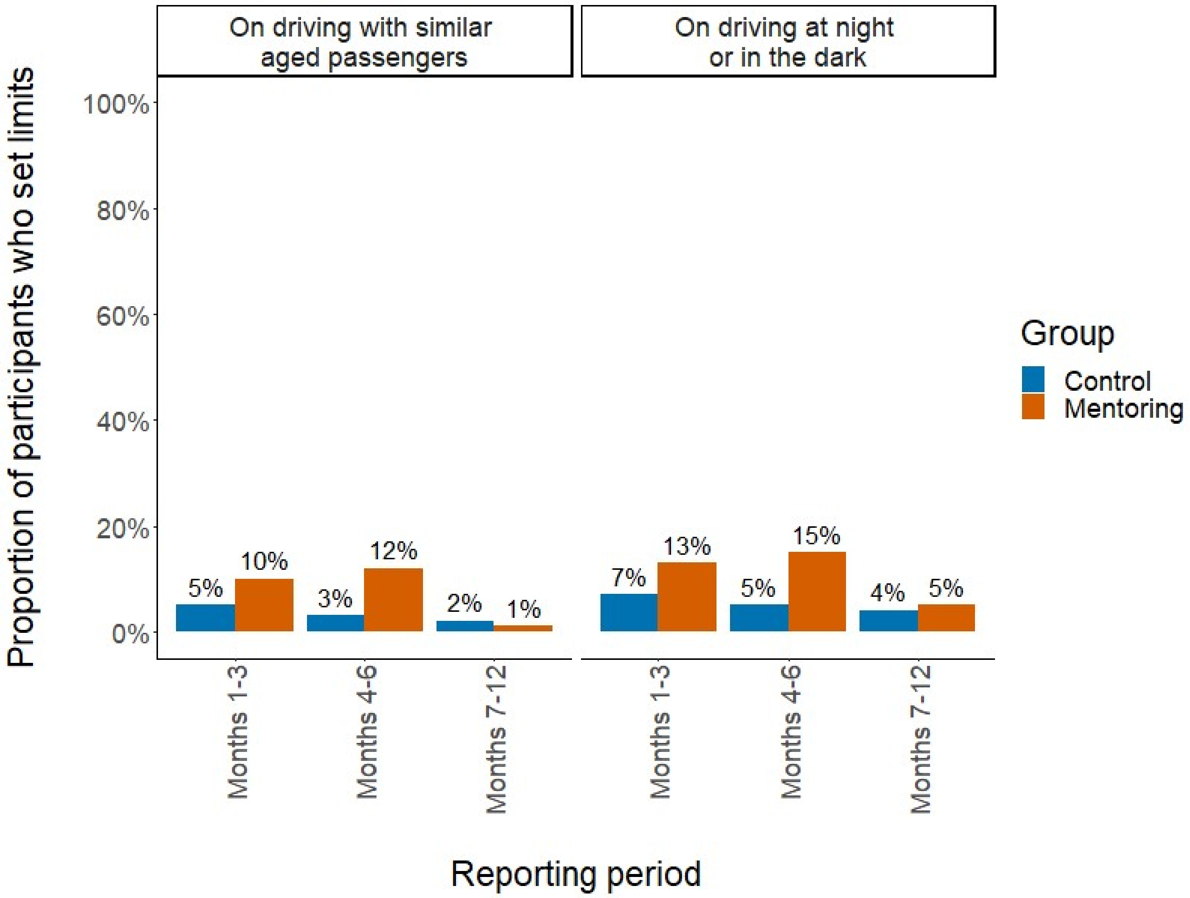


Figure 4-23: Proportion of mentoring agreement engagers and control group participants who reported setting limits on driving with passengers and on driving at night or in the dark in each post-test reporting period (Months 1 to 3: 145 engagers, 2,908 control group. Months 4 to 6: 138 engagers, 2,669 control group. Months 7 to 12: 134 engagers, 2,473 control group).

Figure 4-23 shows that the proportion of mentoring agreement engagers who set each type of limit included in the mentoring agreement was generally higher than the proportion of control group participants who set the same limits, at least in the first six months post-test.

Whilst the proportion of control group participants setting each of the limits decreased over time, the same is not true for the mentoring agreement engagers. For both limits on passengers and limits on driving at night or in the dark, the highest proportion of engagers setting these limits is seen in months 4-6 (15% and 12% respectively), with the proportion in months 1-3 being slightly lower (10% and 13%) and in months 7-12 being considerably lower (1% and 5%).

It is interesting to note that the proportion of mentoring agreement engagers who reported setting each type of limit is generally low (below 20% in all reporting periods). This is despite these engagers being defined as people who set at least one agreement using the RoSPA system (see Table 4-13). In comparing the mentoring agreement group engagers and the control group however, the system data cannot be relied on as corresponding data do not exist for control participants. Therefore, it is necessary to rely on the self-reported setting of limits, which is likely to be an underestimate for both groups (given that the mentoring agreement group were all known to have set limits of some kind within the system). There is a chance that some of the differences in self-reported limits reflect a social desirability bias on the part of the mentoring agreement participants, although as noted the number reporting limits is actually lower than the number known to have done so. The limitations associated with self-report data, noted in section 2.2.1, should be considered when interpreting the results.

Despite the relatively low overall proportions of mentoring agreement engagers who reported setting limits, the mentoring agreement engagers still differ noticeably from the control group participants in the first two reporting periods. A much higher proportion of mentoring agreement engagers set limits on driving with passengers compared with control group participants in months 1-3 (10% compared with 5%) and months 4-6 (12% compared with 3%). The same was true for limits on driving at night or in the dark (13% compared with 7% in months 1-3 and 15% compared with 5% in months 4-6). The pattern was different in months 7-12, with smaller differences seen between the two groups and a higher proportion of control group participants than mentoring agreement engagers setting passenger limits.

A number of statistical tests were conducted to test for differences in the responses between the control and mentoring agreement group at the three time points. In the case of driving with similar aged passengers, there was a significant difference between groups with a higher proportion of the mentoring agreement engagers reporting limits at months 1-3 (p<0.01, effect size=0.04), and in months 4-6 (p<0.01, effect size=0.10). However, this difference did not appear in the last reporting period (p=0.90).

Looking at limits around driving at night or in the dark, statistical tests showed a significant difference between groups with a higher proportion of the mentoring agreement engagers reporting limits in months 1-3 (p<0.01, effect size=0.05) and in months 4-6 (p<0.01, effect size=0.09). There was no significant difference in the last reporting period (p=0.43).

Mileage driven with at least one passenger of similar age, and in the dark

Participants were asked what proportion of their mileage in each reporting period had been driven on various road types and in various other conditions. There were two driving conditions which were particularly relevant to the mentoring agreement: driving with at least one passenger of a similar age, and driving in the dark. The proportion of mileage driven in each of these conditions in each reporting period is compared between mentoring agreement engagers and control group participants in Figure 4-24.

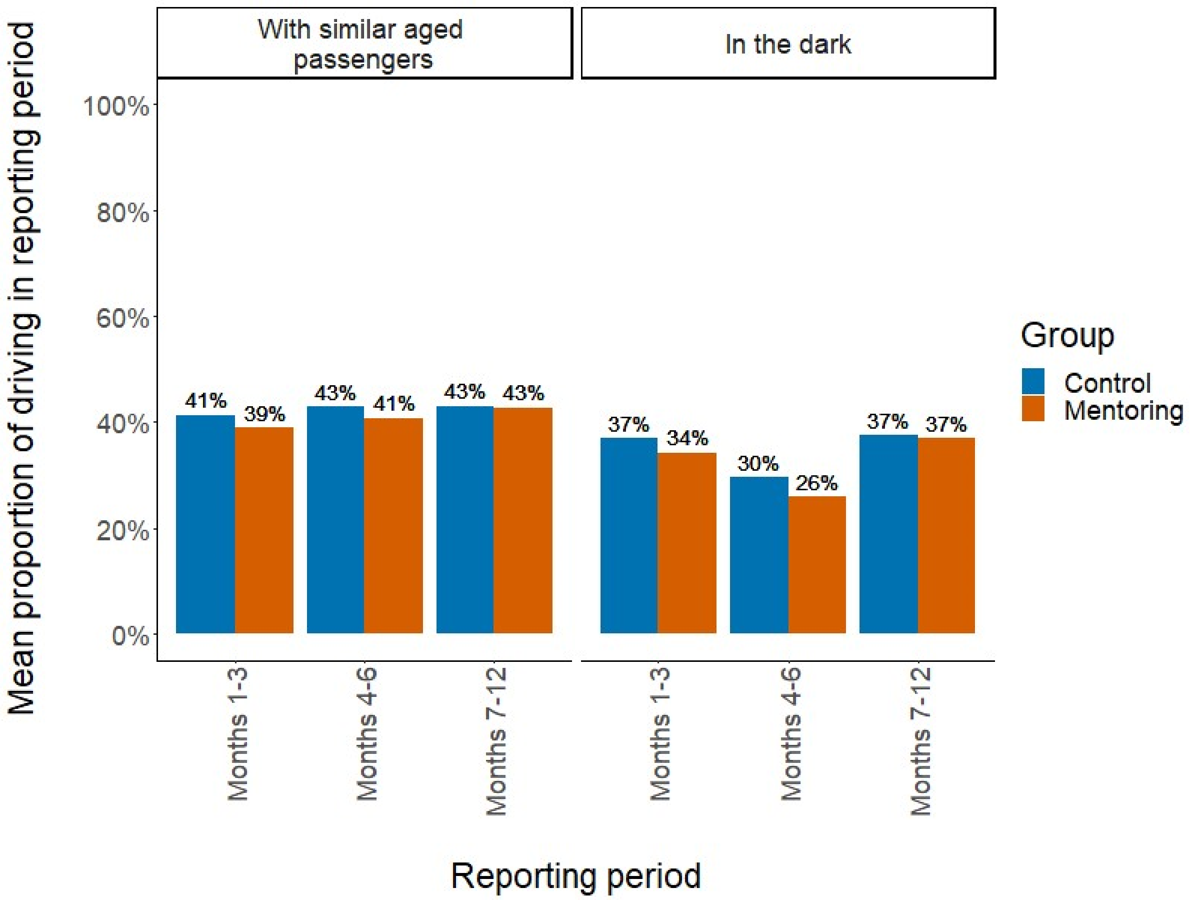


Figure 4-24: Reported proportion of miles driven by mentoring agreement engagers and control group participants with passengers, and in the dark in each post-test reporting period (Months 1 to 3: 145 engagers, 2,908 control group. Months 4 to 6: 138 engagers, 2,669 control group. Months 7 to 12: 134 engagers, 2,473 control group).

The proportion of driving reported as being with passengers showed a general upward trend over time, with some differences in months 1-3 and 4-6; however, none of the group differences were statistically significant.

Mentoring agreement engagers reported a lower proportion of mileage driven in the dark than control group participants in all reporting periods, but the difference was largest in months 4-6. In months 4-6 the proportion of mentoring agreement engager mileage driven in the dark was 26%, compared with 30% of mileage driven by control group participants; this difference was significant (p=0.02) while others were not. For both groups, a higher proportion of driving in the dark was done in months 1-3 and months 7-12 than in the middle reporting period; this pattern was also seen in across all participants in all groups in the exploratory analysis.

Exceeding speed limit

In each post-test survey, participants were asked how often they thought they had exceeded the speed limit whilst driving during the relevant reporting period. Participants indicated their frequency of exceeding speed limits using a number from 0 (never exceeded the speed limit) to 100 (exceeded the speed limit all the time). The distributions of these frequency scores for mentoring agreement engagers and control group participants in each reporting period is shown in Figure 4-25 and the mean frequency scores are shown in Table 4-17.

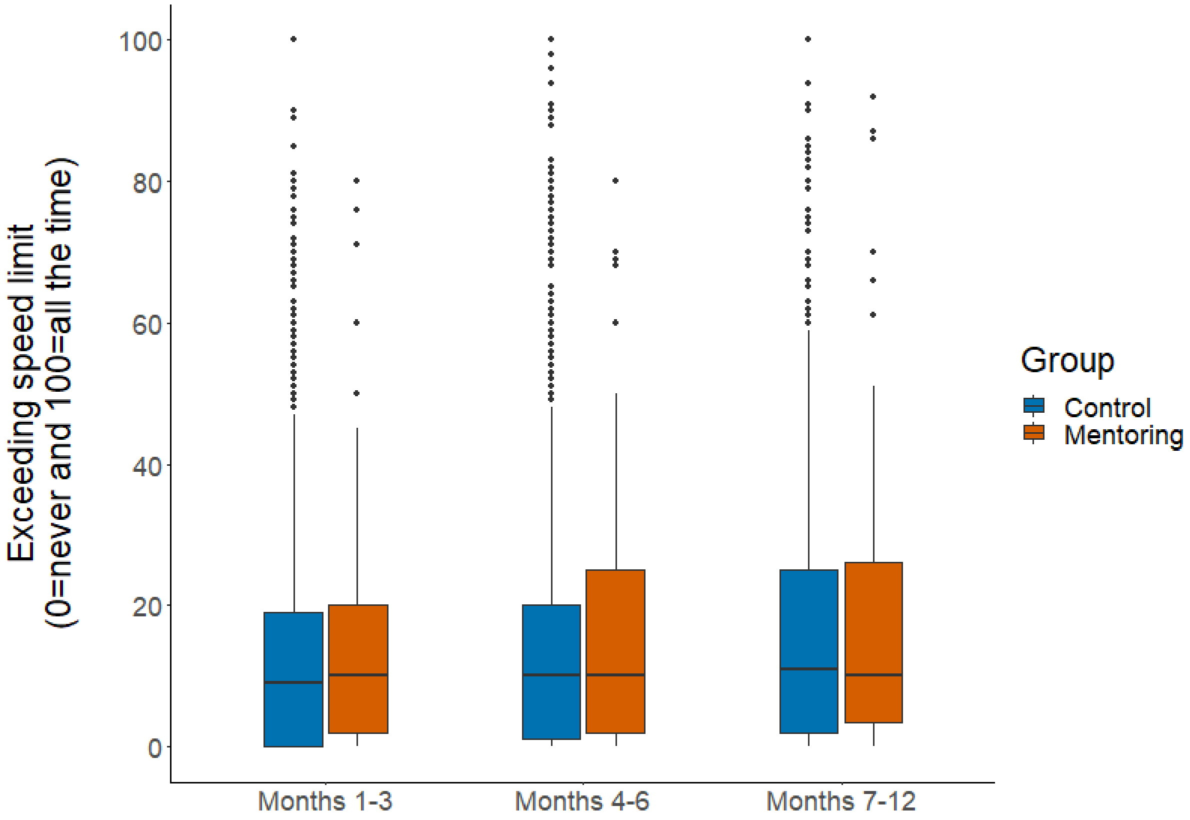


Figure 4-25: Distribution of how often participants thought they exceeded the speed limit for mentoring agreement engagers and the corresponding control group sample (Months 1 to 3: 145 engagers, 2,908 control group. Months 4 to 6: 138 engagers, 2,669 control group. Months 7 to 12: 134 engagers, 2,473 control group).

Table 4-17: Mean exceeding speed limit frequency score in each post-test reporting period for mentoring agreement engagers and control group participants (Months 1 to 3: 145 engagers, 2,908 control group. Months 4 to 6: 138 engagers, 2,669 control group. Months 7 to 12: 134 engagers, 2,473 control group).

Group	Mean exceeding speed limit frequency score		
	Months 1-3	Months 4-6	Months 7-12
Control group participants	12.9	15.4	17.5
Mentoring agreement engagers	13.6	16.0	17.9

The distributions of exceeding the speed limit frequency scores are generally similar between groups, although mentoring agreement engagers reported slightly higher mean frequency scores than control group participants in all three reporting periods.

For both groups, the frequency of exceeding speed limits increased over time. Participants in the control group reported a mean frequency score of 12.9 (on a scale of 0 to 100) in months 1-3, increasing to 17.5 in months 7-12. The mean scores for mentoring agreement engagers were slightly higher but showed the same pattern, increasing from 13.6 to 17.9 over time.

Statistical tests showed there was no significant difference between groups (p=0.61). The difference over time was significant for both groups combined (p<0.01) with a small effect size of 0.003; however, the interaction between group and time was not significant (p=0.95), meaning the effect of time was the same for both groups.

Attitudes towards new driver restrictions and more general enforcement

After passing their driving tests and in each of the reporting periods (months 1-3, months 4-6 and months 7-12) participants were asked to select their level of agreement towards speed enforcement and new driver restrictions. These included 7-point Likert-scale statements around stronger enforcement of speed limits, drink-driving and alcohol, minimum practice hours and more supervision for six months after passing their test.

The data were suitable for the statistical approach of factor analysis. This technique is used to simplify data when groups of questions (in this case attitude statements about enforcement) are thought to be measuring a smaller number of underlying constructs or 'factors'. The logic of factor analysis is that rather than running group comparisons on many items that measure a smaller number of underlying factors, one can establish first what the underlying factors are, and then compare groups on these directly (more information on factor analysis is provided in Appendix D). In this case the analysis identified a three-factor solution, shown in Table 4-18.

As these questions were repeated in each of the surveys, the analysis presents a comparison of how novice driver attitudes towards speed enforcement and new driver restrictions changed over time. In order to ensure that the results were comparable over time, the coefficients for the factor scores were created using the test pass survey and applied to the responses from each of the reporting periods. This ensured that the results were presented using the same factor coefficients and any differences could be attributed to a change over time.

For all factors, there were no statistically significant differences between the control and mentoring agreement groups, and no statistically significant changes over time.

Table 4-18: Factors for the attitudes to driver restrictions and enforcement scale

Factor 1 – attitudes to learner driver requirements	Factor 2 – attitudes to general enforcement	Factor 3 – attitudes to post-test restrictions
Learner drivers should have a minimum learning period of at least 12 months before they can take their test Learner drivers should need to have at least 100 hours of on-road practice before they can take their test	Speed limits should be more strongly enforced Drink/drug driving laws should be more strongly enforced Laws to stop the use of mobile phones and other distracting devices when driving should be more strongly enforced For 6 months after passing their test, new drivers should not be allowed to drink any alcohol if driving	For 6 months after passing their test, new drivers should not be allowed to drive in the dark without a supervising driver For 6 months after passing their test, new drivers should not be allowed to carry passengers under 25 without a supervising driver

4.5.3.2 *Telematics*

This section presents the results of analysis which seeks to answer the question: for people who engage, what does the telematics intervention change? Telematics engagers and control group participants have been compared using the measures below to try to answer this question:

- Self-reported number of near-misses – tested whether there were fewer for those who engaged with the telematics intervention
- Self-reported Driving Events score – tested whether this was lower (indicating lower frequency of unsafe driving events) for those who engage
- Self-reported frequency of exceeding the speed limit – tested whether this was less frequent for those who engaged with the telematics intervention
- Self-reported proportion of mileage driven in the dark – tested whether this was lower for telematics intervention engagers
- Self-reported driving style – tested whether driving styles were different in those who engaged with the telematics intervention
- Self-reported attitudes towards new driver restrictions and more general enforcement – tested whether there was greater agreement from telematics engagers.

All comparisons in this section have been made between 669 participants who engaged with the telematics intervention and 2,447 participants in the matched control group sample; however, numbers of participants vary by reporting period. See section 4.5.2.2 for the definition of a telematics intervention engager, a description of the control group sampling method, and a summary of the number of participants in each group included in the analysis of each reporting period.

Near misses

Figure 4-26 and Table 4-19 show the near misses per mile reported by participants who engaged with the telematics intervention and participants in the corresponding control group sample during each of the reporting periods (months 1-3, months 4-6 and months 7-12). A near-miss was described as the impression of only just avoiding an accident.

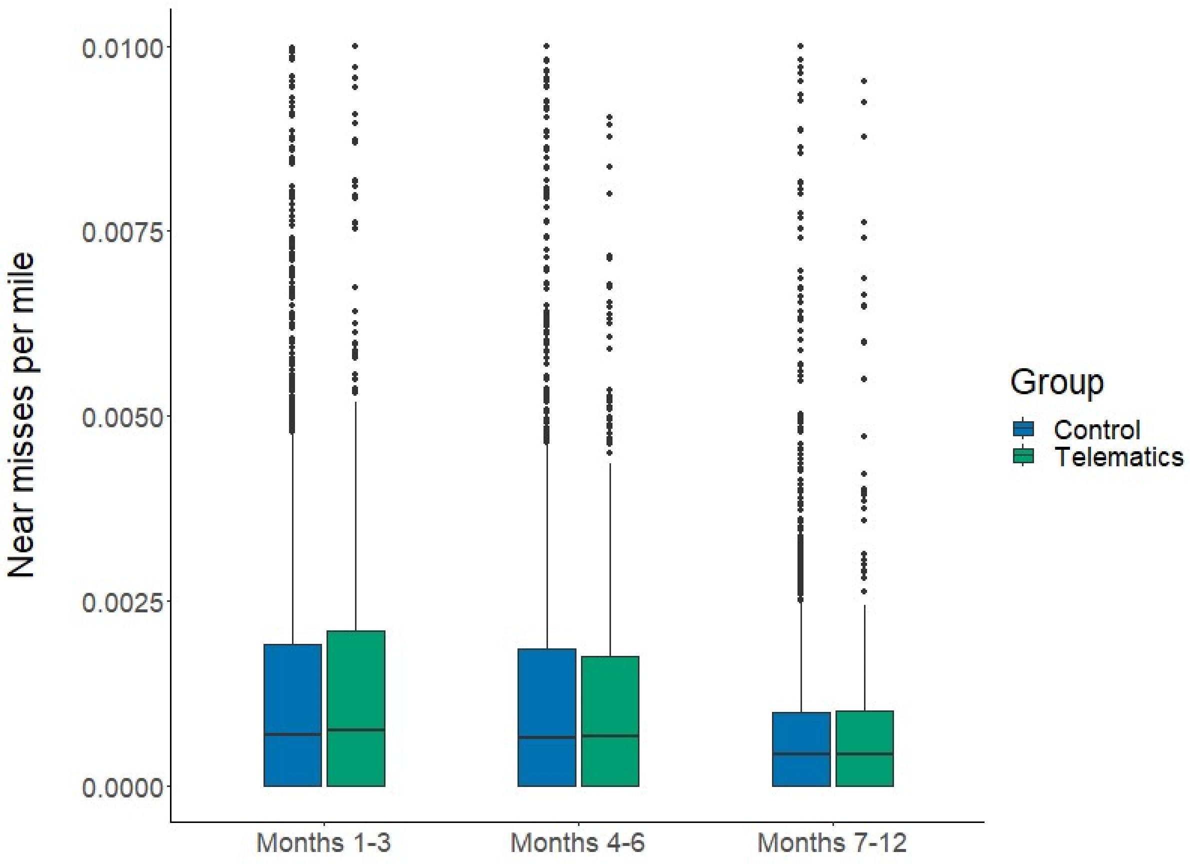


Figure 4-26: Distribution of reported number of near misses per mile in each post-test reporting period for telematics engagers and the corresponding control group sample (Months 1 to 3: 583 engagers, 2,081 control group. Months 4 to 6: 542 engagers, 1,910 control group. Months 7 to 12: 514 engagers, 1,752 control group).

Table 4-19: Mean number of reported near misses per mile in each post-test reporting period for telematics engagers and the corresponding control group sample (Months 1 to 3: 583 engagers, 2,081 control group. Months 4 to 6: 542 engagers, 1,910 control group. Months 7 to 12: 514 engagers, 1,752 control group).

	Mean number of near misses per mile driven		
Group	Months 1-3	Months 4-6	Months 7-12
Control group participants	0.0009	0.0009	0.0005
Telematics engagers	0.0011	0.0009	0.0005

Both control group participants and telematics group engagers reported low numbers of near misses per mile travelled in all reporting periods. The rates per mile are shown in Table 4-19. For context, in months 1-3 and 4-6 people in both groups combined were reporting around one near miss for every 1,000 miles of driving, with this dropping to around one every 2,000 miles of driving in months 7-12. For both groups, the number of near misses per mile decreased over time.

The number of near misses per mile travelled reported by the control group participants decreased from 0.0009 in months 1-3 to 0.0005 in months 7-12, with an average of 0.0009 near misses reported during the period in between (months 4-6). Participants who engaged with the telematics intervention reported very slightly higher numbers of near misses per mile in months 1-3 (0.0011) but the same number as control group participants in the other two periods. Statistical tests showed no significant differences between groups (p=0.94). The difference over time was significant for both groups combined (p<0.01), albeit with a small effect size of 0.01, and the interaction between group and time was not significant (p=0.84), meaning the effect of time was the same for both groups.

Driving events score

In each post-test survey participants were asked about how often each of six driving events deemed unsafe had occurred whilst they were driving during that reporting period. Their responses were converted into the 'driving events' score described earlier (section 4.5.3.1), with a lower score indicating a lower frequency of unsafe driving events.

Figure 4-27 shows the distribution of driving events scores for each reporting period for participants who engaged with the telematics intervention and participants from the corresponding control group sample. The corresponding means are shown in Table 4-20.

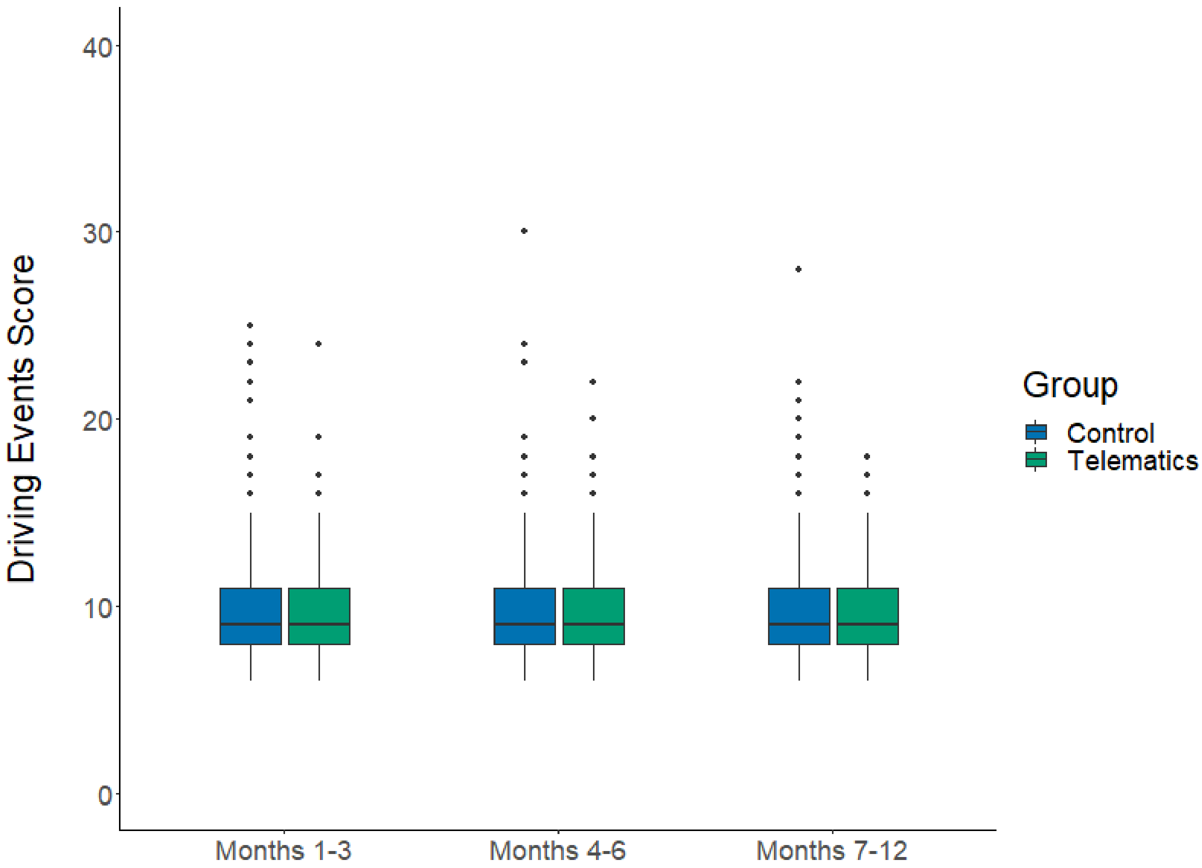


Figure 4-27: Distribution of driving events scores for each post-test reporting period for telematics engagers and the corresponding control group sample (Months 1 to 3: 583 engagers, 2,081 control group. Months 4 to 6: 542 engagers, 1,910 control group. Months 7 to 12: 514 engagers, 1,752 control group).

Table 4-20: Mean driving event score in each post-test reporting period for telematics course engagers and control group participants (Months 1 to 3: 583 engagers, 2,081 control group. Months 4 to 6: 542 engagers, 1,910 control group. Months 7 to 12: 514 engagers, 1,752 control group).

	Mean driving event score		
Group	Months 1-3	Months 4-6	Months 7-12
Control group participants	9.23	9.48	9.71
Telematics engagers	9.25	9.52	9.77

For all reporting periods, the distributions of driving events scores were very similar for participants who engaged with the telematics intervention and participants in the control group sample. This suggests that the telematics intervention did not have an effect on how frequently the events included in this measure occurred. For both groups (as with the mentoring agreement group) the mean driving events score increased over time, with lower mean scores in months 1-3 (9.23 for the control group and 9.25 for the telematics group) and higher mean scores in months 7-12 (9.71 for the control group and 9.77 for the telematics group). This means that, on average, the frequency of the various driving events included in this measure increased over time.

Statistical tests showed no significant difference between groups (p=0.63). The difference over time was significant for both groups combined (p<0.01), albeit with small effect sizes (0.01), and the interaction effect was not significant (p=0.9), indicating that the effect of time was the same for both groups.

Exceeding the speed limit

Participants were asked to indicate how often they thought they exceeded the speed limit during each reporting period. Their responses ranged from 0 (never exceeded the speed limit) to 100 (exceeded the speed limit all the time). Figure 4-28 shows the distribution of responses to this question for participants who engaged with the telematics intervention and participants from the corresponding control group sample. The mean frequency scores are shown in Table 4-21.

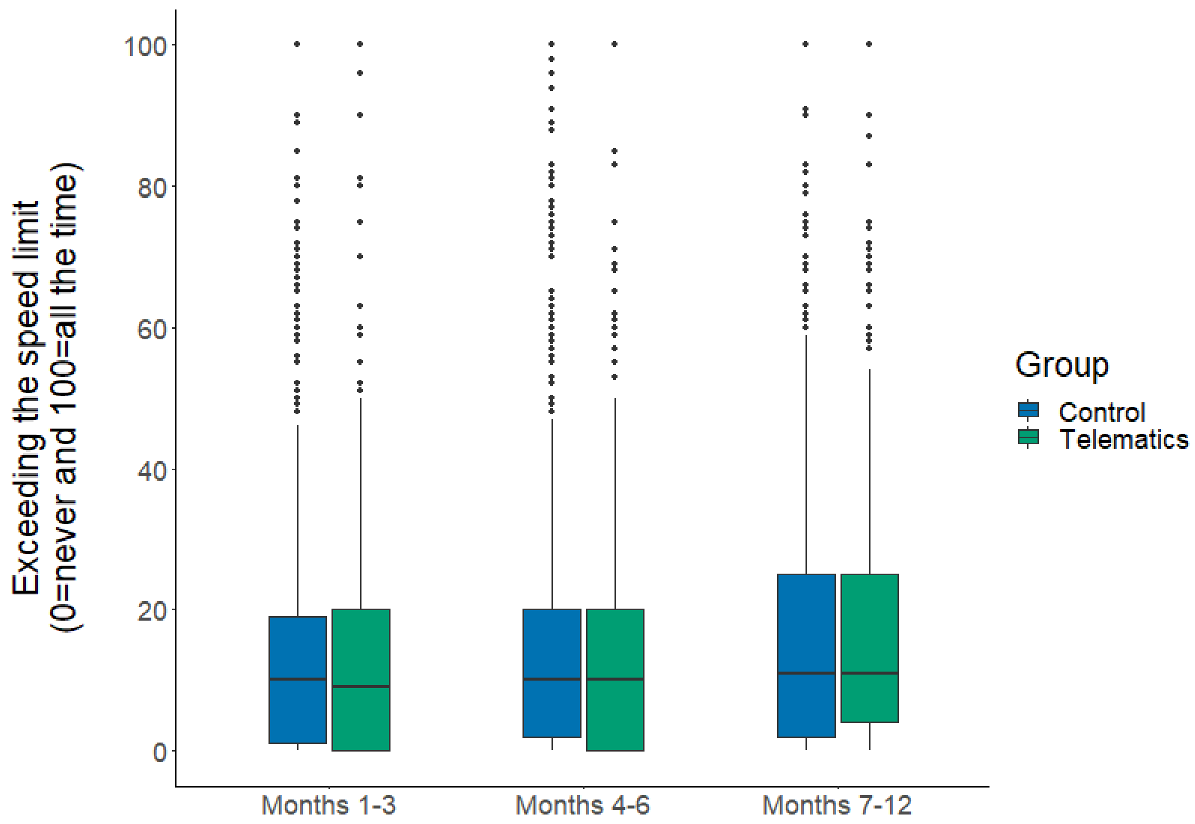


Figure 4-28: Distribution of how often participants thought they exceeded the speed limit for telematics engagers and the corresponding control group sample (Months 1 to 3: 583 engagers, 2,081 control group. Months 4 to 6: 542 engagers, 1,910 control group. Months 7 to 12: 514 engagers, 1,752 control group).

Table 4-21: Mean exceeding speed limit frequency score in each post-test reporting period for telematics engagers and the corresponding control group sample (Months 1 to 3: 583 engagers, 2,081 control group. Months 4 to 6: 542 engagers, 1,910 control group. Months 7 to 12: 514 engagers, 1,752 control group).

	Mean frequency score		
Group	Months 1-3	Months 4-6	Months 7-12
Control group participants	13.1	15.5	17.1
Telematics engagers	12.9	14.4	17.9

In general, the distributions of exceeding the speed limit frequency scores were similar between the telematics engagers and the corresponding control group sample, and the mean scores were low in all reporting periods (between 12 and 18 on a scale of 0 to 100). For both groups, the frequency of exceeding speed limits increased slightly over time.

The difference in responses between groups was not significant (p=0.96). There was a significant difference in responses over time for both groups combined (p<0.01) with a small effect size of 0.01. However, the interaction effect was not significant (p=0.20), indicating that the effect of time was the same for both groups.

Driving in the dark

Figure 4-29 shows the proportion of post-test mileage that was driven in the dark by telematics engagers and participants in the corresponding control group sample. The proportions are shown for miles driven during each reporting period.

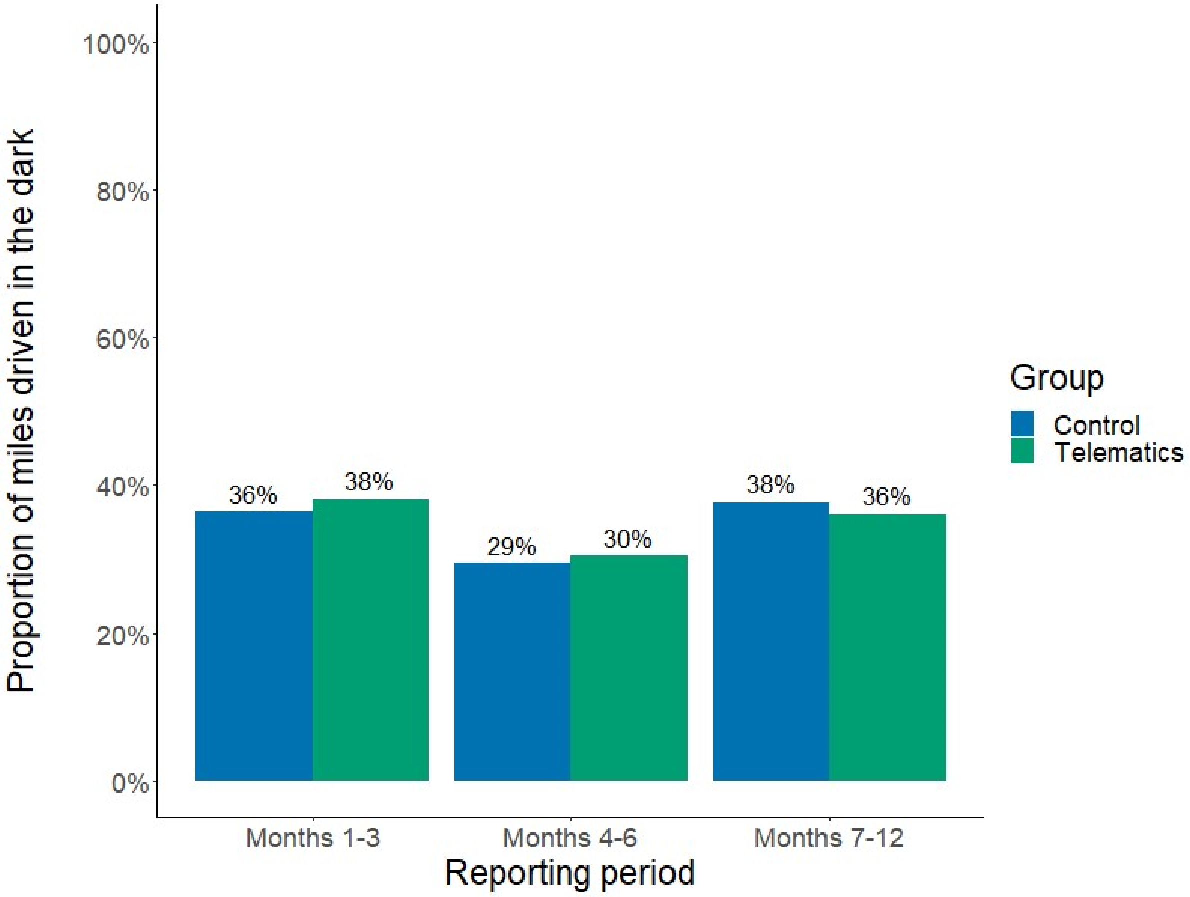


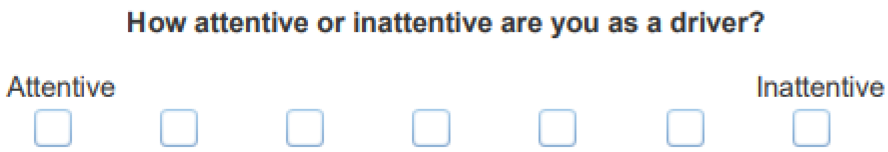
Figure 4-29: Reported proportion of miles driven in the dark during each post-test reporting period by telematics engagers and the corresponding control group sample (Months 1 to 3: 583 engagers, 2,081 control group. Months 4 to 6: 542 engagers, 1,910 control group. Months 7 to 12: 514 engagers, 1,752 control group).

The proportion of mileage driven in the dark was very similar between groups. Both groups reported lower proportions of driving in the dark in months 4-6 compared with the other two reporting periods (this pattern agrees with what was seen across all participants in all groups in the exploratory analysis in Figure 4-10).

Statistical tests showed that there was a small statistically significant (p=0.04) difference in the proportion of driving in the dark by group for months 7-12; however, this had a negligible effect size of 0.06.

Driving style

Participants were asked a range of questions about their driving styles shortly after passing their test and in each of the reporting periods (months 1-3, months 4-6, and months 7-12). Driving style was measured using the scales developed by Guppy et al. (1990), which comprise of seven points on 12 bipolar scales such as ‘Attentive – Inattentive’ or ‘Safe-Risky’ on which participants were asked to rate themselves. For instance, participants were presented with the driving style depicted below and asked to rate themselves on the scale between the anchors ‘attentive’ and ‘inattentive’:


How attentive or inattentive are you as a driver?

Attentive Inattentive

These scales have been used in numerous studies, and the 12 scales typically reduce to three factors characterising particular driving styles. As noted in section 4.5.3.1 (with more detail in Appendix D), factor analysis is useful to reduce multiple items to a more manageable set of underlying factors, resulting in fewer group comparisons. The data were suitable for factor analysis and a three-factor solution was identified (see Table 4-22).

Table 4-22: Factor structure for the driving style scales

Factor 1	Factor 2	Factor 3
attentive/inattentive	placid/irritable	decisive/indecisive
careful/careless	patient/impatient	experience/inexperienced
responsible/irresponsible	considerate/selfish	confident/nervous
safe/risky	tolerant/intolerant	fast/slow

As these questions were repeated in each of the surveys, the analysis presents a comparison of how novice driver attitudes change over time. In order to ensure that the results were comparable over time, the coefficients for the factor scores were created using the test pass survey and applied to the responses from each of the reporting periods. This ensured that the results are presented using the same factor coefficients and any differences can be attributed to a change over time. Due to differences in sample sizes over time, the average factor scores presented are only done so for each reporting period.

Figure 4-30 presents the results for Factor 1. A higher score on the scales for Factor 1 indicates a self-reported driving style that is more ‘inattentive, careless, irresponsible and risky’.

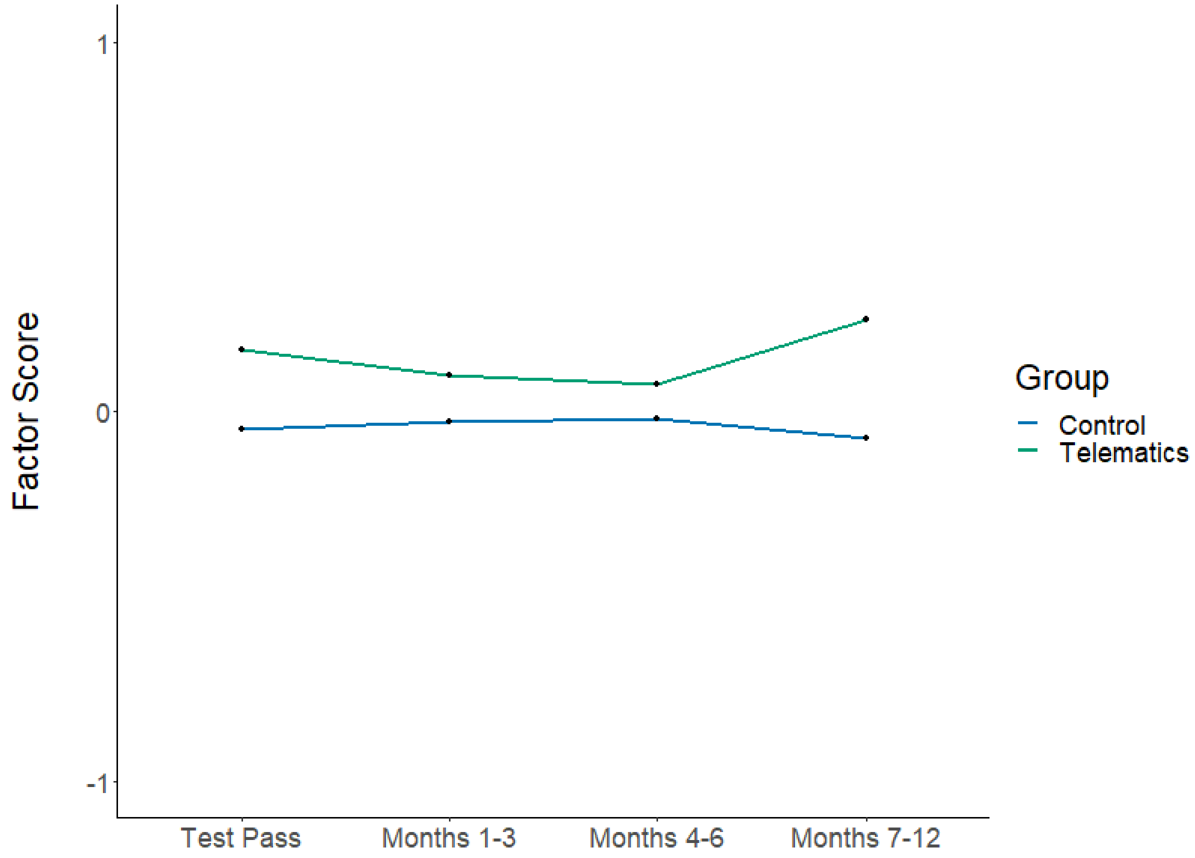


Figure 4-30: Scores for a driving style that is 'inattentive, careless, irresponsible and risky' over time, and by group (Test Pass: 666 engagers, 2,392 control group. Months 1 to 3: 583 engagers, 2,081 control group. Months 4 to 6: 542 engagers, 1,910 control group. Months 7 to 12: 514 engagers, 1,752 control group).

There were significant differences between the two groups, albeit with very small effect sizes. The telematics group reported a driving style that was slightly more inattentive, careless, irresponsible, and risky. Statistical tests showed a significant difference between the groups (p=0.03) albeit with a an effect size close to zero. There were no differences over time for either group (p=0.51) or due to the interaction between group and time (p=0.63). The difference in Factor 1 scores between the groups is potentially concerning although is also very small and may not reflect an actual change in risk; instead, it may reflect people in the telematics group being more aware of their true driving style.

Statistical tests showed no significant difference in self-reported driving style for Factor 2 between the control and telematics groups (p=0.06), or over time (p=0.80). The interaction between group and time was also non-significant (p=0.90). Finally, there was no significant difference in self-reported driving style for Factor 3 between groups (p=0.93). The difference over time was significant for both groups combined (p=0.04 with a very small effect size close to 0) but the interaction effect was not significant (p=0.23), meaning and time effect was the same for each group.

Attitudes towards new driver restrictions and more general enforcement

After passing their driving tests and in each of the reporting periods (months 1-3, months 4-6 and months 7-12) participants were asked to select their level of agreement towards speed enforcement and new driver restrictions. These included 7-point Likert-scale statements around stronger enforcement of speed limits, drink-driving and alcohol, minimum practice hours and more supervision for six months after passing their test.

The data were suitable for factor analysis and a three-factor solution was identified – the same as in Table 4-18. Similar to the process mentioned in section 4.5.3.1 under attitudes to new driver restrictions, the factor scores from the test pass survey were applied to each reporting period to enable comparisons over time (more information is provided in Appendix D).

For all three factors, statistical tests showed no significant differences between groups, time periods or due to the interaction effect.

5 Discussion and conclusions

The novice driver arm of the Driver2020 project aimed to test the real-world effectiveness of two interventions, on both collision involvement, and a number of surrogate safety measures, in the first 12 months of post-test driving.

The main findings from the study are discussed below, first for the effect of the interventions on collisions, and then on surrogate measures. It should be noted that the COVID-19 pandemic affected either learning to drive, post-test driving, or both during the study for the majority of participants. Thus, some caution should be exercised in how generalisable the findings are to future circumstances.

5.1 Effect of interventions on collisions

Engagement with the two interventions was low. For the mentoring agreement intervention, around 4% of people offered the intervention actually engaged with the materials in some way, with less than 1% fully engaging throughout the intervention (1% of participants set two agreements, 1% set three agreements, and less than 1% completed the final review). For the telematics app, 16% of participants offered the app engaged with its features to some degree (9% of participants recorded less than 10 miles throughout the intervention; 7% recorded more than 10 miles).

The intention to treat analysis, which compared collision involvement rates between the intervention and control groups for all participants (regardless of engagement) showed that the groups did not differ in collision risk after mileage, age, gender and COVID-19 effects were controlled. This is not surprising given the low engagement rates; even if the interventions were extremely effective for those who engaged, they would only be having an effect on a small proportion of the intervention group. The analogy is that if a medicine exists that cures a disease with 100% efficacy, and only 1% of people take the medicine, then the real-world effectiveness of the medicine will be 1%. Unfortunately, it is not possible to examine differences in collisions solely in the people who engaged with the interventions, as the sample sizes were insufficient to support such an analysis; this is why surrogate measures were used for the per-protocol analysis (see below).

Based on the trial sample, this finding suggests that if the mentoring agreement and telematics interventions trialled in the novice driver arm of the Driver2020 study were offered on a similar voluntary basis to novice drivers aged 17-24 in Great Britain they would be unlikely to lead to any measurable reduction in collision risk in this population.

5.2 Effect of interventions on surrogate measures

To examine the potential impacts on those participants who did engage, groups were compared on a number of surrogate measures of risk. This ‘per protocol’ analysis included only those who engaged with the two interventions, and corresponding control participants.

The amount of engagement with interventions in those who did engage was also relatively low, with the ‘threshold’ for being described as an engager typically ‘any use’ of the intervention in question. This needs to be considered when interpreting the results; because engagement 'dose' or 'amount' of intervention was low even in those people who did engage, it may be that the effects seen were limited by this.

5.2.1 Mentoring agreements

Two positive findings were seen for the mentoring agreement.

Firstly, this intervention, relative to the control group, led to increases in the proportion of people self-reporting that they had set limits for driving with peer-age passengers (10% compared with 5%, and 12% compared with 3%, for months 1-3 post-test and months 4-6 post-test respectively). Similar increases were seen for limits on driving in the dark (13% compared with 7%, and 15% compared with 5%, for months 1-3 post-test and months 4-6 post-test respectively).

Secondly, at 4-6 months post-test, engagers in the mentoring agreement group reported a slightly lower proportion of their mileage being in the dark than the control group (26% versus 30%).

These differences are in line with the findings seen in similar interventions (notably the 'Checkpoints' programme in the US) which have been reviewed in Pressley et al. (2016). Such studies have tended to show modest increases in the ways in which novice drivers manage their driving risks on being offered appropriate materials, although few major differences are seen in more objective measures – the notable exception being Simons-Morton, Hartos, Leaf and Preusser (2006) who showed a reduction in traffic violations in the first 12 months of driving.

One general measure not affected by the setting of limits was the number of risky driving events recorded. These increased over time for all groups. Similar findings have been seen for violations in other studies (for example Wells et al., 2008; McKenna, 2018); violations recorded increase in novices over their early driving period, even at a time when they are becoming safer in terms of collision risk. It is possible that, as drivers adapt to the driving environment, their increased confidence leads them to adopt a slightly riskier style (McKenna, 2018).

It is promising that an easily accessible set of web-based materials can lead to greater engagement with voluntary limits for these high-risk situations, and some small reduction in driving in the dark, in those participants who chose to engage.

Unfortunately, the engagement rate with the mentoring agreement was very low. This presents the question of how it could be improved. The findings in Hitchings et al. (2024) will be useful in helping to answer this question. Another area that may prove fruitful is in considering which behaviour change techniques might be used alongside the materials to an even greater degree than is already the case. Sullman (2017) found that for adolescents, the techniques of 'prompting self-monitoring of behaviour', 'providing feedback on performance' and 'prompting specific goal setting' were among the most successful techniques in non-driving domains. The mentoring agreement intervention itself includes elements of these and other behavioural change techniques; it might be that a specific intervention that sits alongside the mentoring materials, with the express aim of embedding those behaviour change elements, could increase engagement.

5.2.2 Telematics intervention

Engagement with the telematics intervention was associated with two statistically significant changes in surrogate measures relative to the control group. Firstly, there was a small reduction in the proportion of mileage reported as being in the dark, for the period 7-12 months post-test. This is promising, although is smaller in magnitude to the change seen in the mentoring agreement group in months 4-6.

Secondly, there was a very small increase in one risky driving style ('inattentive, careless, irresponsible and risky'). One possible reason for this is that the intervention in some way really did promote a driving style that is self-reported as very slightly more 'inattentive, careless, irresponsible and risky'. Another possibility is that simply having a device which provides feedback on risky driving might provide an awareness in drivers which makes their ratings of their driving style closer to what is really the case. This is similar to the mechanism suggested for the increased number of self-report collisions seen in Helman et al. (2017) in those with telematics-based insurance policies, compared with those without; it might simply be easier to recall the real number of collisions one has experienced when one's insurance policy comes with 'black box' technologies that detect such events.

While telematics-based approaches to 'pay how you drive' car insurance for young novice drivers remain popular with providers, the Driver2020 study was unable to demonstrate any major obvious benefits of such an approach outside of an insurance context in a controlled study, even for those who chose to engage. Greater understanding of the use of behavioural feedback in such technology-driven approaches is needed if their promise is to be realised outside of settings that are constrained by commercial interests.

5.3 Other findings of note

5.3.1 Impact of COVID-19

COVID-19 had a major impact on the Driver2020 study, and also seems to have had an impact on the risks faced by newly-qualified drivers when they begin driving. In the novice arm of the study, those who passed their driving test before the pandemic began in March 2020 were 1.26 times more likely to report a collision in their first year of post-test driving than those who passed after this point.

There are likely several mechanisms at play in explaining this effect. One is that people who passed their test after the pandemic were more likely to experience a delay in licensure, and thus be both older and potentially have had more on-road practice when they finally began driving; as noted in Section 1 and Appendix A.2, both of these changes would be expected to lead to lower collision risk (Wells et al., 2008; Forsyth et al., 1995; Maycock et al., 1991). Another potential contributing factor is that the driving environment itself changed fundamentally after the pandemic, making it more likely that newly-qualified drivers would experience collisions. For example, in many countries (including Great Britain) average traffic speeds increased during the initial lockdowns (Wegman & Katrakazas, 2021); this would be expected to increase risk for all drivers, and newly-qualified drivers might be especially prone to the increased task demand that higher speeds bring (Fuller, 2011). Further analyses of the data would be required to understand the detailed impact of the

COVID-19 pandemic, and the mechanisms by which it occurred, on newly-qualified drivers in Great Britain during this period.

A further complication of the COVID-19 effects in the novice arm specifically is that for those who passed their test after the pandemic, the amount they reported their learning being negatively affected by it was associated with an increase in collision risk. Further analyses would need to look at the novice arm specifically to understand the mechanisms involved.

5.3.2 The effects of post-test experience

In the novice arm, the analysis showed that as mileage increased, although the number of self-reported collisions went up, the *rate* of increase per mile went down. This confirms the importance of post-test experience in reducing risk, building on the evidence base summarised in Section 1 and Appendix A.2.

5.4 Conclusion – novice driver arm

Neither of the two interventions offered to novice drivers reduced collisions in the first 12 months of post-test driving when offered under a voluntary approach.

An important finding that helps in the interpretation of this main conclusion is that when such interventions are offered under the conditions of voluntary engagement used in this study, levels of engagement are extremely low.

Analysis of surrogate measures with the people who engaged to some degree suggests that small to modest changes in some variables are possible, particularly from the mentoring agreement intervention.

Neither of the interventions, however, show any sign of improving safety for young and novice drivers to the extent shown by stronger, legislative approaches such as the introduction of hazard perception testing in Great Britain in 2002 (Wells et al., 2008), or stronger approaches to licensing seen in other countries (Russell et al., 2011; Kinnear et al., 2013).

Further work with the Driver2020 dataset is expected to help further elucidate the effects of the COVID-19 pandemic on learning to drive and on early post-test driving in young and novice drivers. It also provides a recent dataset that can be used to examine this group, and the long-understood road safety challenge they present.

References

Begg D and Stephenson S (2003). Graduated driver licensing: the New Zealand experience. *Journal of Safety Research*, 34(1), 99-105.

Brown ID, Groeger JA and Biehl B (1987). Is driver training contributing enough towards road safety? In J. A. Rothengatter & R. A. de Bruin (Eds.), *Road Users and Traffic Safety*. Assen: Van Gorcum.

Christie R (2001). *The effectiveness of driver training as a road safety measure. A review of the literature.* Victoria, Australia: Royal Automobile Club of Victoria.

Clarke DD, Ward P and Truman W (2002). *In-depth accident causation study of young drivers* (TRL542). Crowthorne: Transport Research Laboratory.

Clinton K and Lonero L (2006). *Guidelines for evaluating driver education programs.* Washington, DC: AAA Foundation for Traffic Safety.

Department for Transport (2015). *Working together to build a safer road system: British road safety statement.* Stationary Office. https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/487949/british_road_safety_statement_web.pdf

Department for Transport (2021). https://www.gov.uk/government/statistical-data-sets/driving-test-statistics-drt, accessed 12/01/22

Forsyth E, Maycock G and Sexton B (1995). *Cohort study of learner and novice drivers. Part 3: Accidents, offences and driving experience in the first three years of driving.* Crowthorne: Transport Research Laboratory.

Fylan F (2017). *Using Behaviour Change Techniques: Guidance for the road safety community.* RAC Foundation.

Glad A (1988). *Phase 2 driver education, effect on accident risk.* Oslo, Norway: Transport Institute.

Goldstein LG (1972). Youthful drivers as a special safety problem. *Accident Analysis and Prevention*, *4*(3), 153-189.

Gov.uk website deprivation data (2019). [online]. [Accessed March 2022]. Available from World Wide Web: https://www.gov.uk/government/uploads/system/uploads/attachment_data/file/467774/File_7_ID_2019_All_ranks__deciles_and_scores_for_the_Indices_of_Deprivation__and_population_denominators.csv/preview

Gov.scot website deprivation data (2020). [online]. [Accessed March 2022]. Available from World Wide Web: https://www.gov.scot/publications/scottish-index-of-multiple-deprivation-2020v2-data-zone-look-up/

Gov.wales website deprivation data (2019). [online]. [Accessed March 2022]. Available from World Wide Web: https://statswales.gov.wales/Catalogue/Community-Safety-and-Social-Inclusion/Welsh-Index-of-Multiple-Deprivation

Guppy A, Wilson P and Perry J (1990). Driving attitudes and driving experience. In Benjamin, T. (ed.) *Driving Behaviour in a Social Context.* Caen, France: Paradigme

Helman S, Grayson GB and Parkes AM (2010). *How can we produce safer new drivers? A review of the effects of experience, training, and limiting exposure on the collision risk of new drivers* (INS005). Crowthorne: Transport Research Laboratory.

Helman S, Wallbank C, Chowdhury S, Hammond J, Kinnear N, Buttress S, Jenkins R and Grayson G (2017). *Transforming the practical driving test: final report* (PPR828). Crowthorne: Transport Research Laboratory.

Helman S and Weekley J (2024). *Driver2020 – an evaluation of interventions designed to improve safety in the first year of driving.* Report D4: Summary of findings (PPR2012). Commissioned and funded by the Department for Transport. Crowthorne: Transport Research Laboratory.

Hitchings J, Holcombe A, Christie N, Weekley J and Helman S (2024). *Driver2020 - an evaluation of interventions designed to improve safety in the first year of driving. Report D3: Delivery of interventions and engagement by novice and learner drivers* (PPR2011). Commissioned and funded by the Department for Transport. Crowthorne: Transport Research Laboratory.

House of Commons (2021). Road safety: young and novice drivers. Fourth Report of House of Commons Transport Committee Session 2019-2021, HC169. London: House of Commons Transport Committee.

Isler RB and Starkey NJ (2008). The 'frontal lobe' project: A double-blind, randomized controlled study of the effectiveness of higher level driving skills training to improve frontal lobe (executive) function related driving performance in young drivers. Final report by the Traffic and Road Safety Research Group, Psychology Department, University of Waikato. Hamilton, New Zealand: University of Waikato.

Jones B (1993). *The effectiveness of skid-car training for teenage novice drivers in Oregon.* Salem, OR: Drive and Motor Vehicle Services.

Katila A, Keskinen E, Hatakka M and Laapotti S (1996). Skid training for novice drivers – benefit for adults, pitfall for youngsters. *Paper presented at the International Conference of Traffic and Transport Psychology, Valencia, Spain.*

Kinnear N, Hammond J, Smith L, Ramnath R, Simms G, Hitchings J and Helman S (2020). *The potential impact of Graduated Driver Licensing in Northern Ireland* (PPR934). Crowthorne: Transport Research Laboratory

Kinnear N, Lloyd L, Helman S, Husband P, Scoons J, Jones S, Stradling S, McKenna F and Broughton J (2013). *Novice Drivers: evidence review and evaluation: pre-driver education and training, graduated driver licensing, and the New Drivers Act* (PPR673). Commissioned and funded by the Department for Transport. Crowthorne: Transport Research Laboratory

Kinnear N, Kelly SW, Stradling S and Thomson JA (2013). Understanding how Drivers Learn to Anticipate Risk on the Road: A Laboratory Experiment of Affective Anticipation of Road Hazards. *Accident Analysis and Prevention*, 50, 1025-1033.

Kinnear N and Wallbank C (2020). *Graduated driver licensing in the Isle of Man* (PPR952). Crowthorne: Transport Research Laboratory.

Maycock G, Lockwood CR and Lester J (1991). *The accident liability of car drivers* (RR315). Crowthorne: Transport and Road Research Laboratory.

Mayhew DR, Simpson HM and Pak A (2003). Changes in collision rates among novice drivers during the first months of driving. *Accident Analysis and Prevention*, 35(5), 683-691.

Mayhew DR, Simpson HM and Robinson A (2002). The safety value of driver education and training. *Injury Prevention*, 8(Supp. II), ii3–ii8.

Mayhew DR, Simpson HM, Williams AF and Ferguson SA (1998). Effectiveness and role of driver education and training in a graduated licensing system. *Journal of Public Health Policy*, 19, 51–67.

McCartt AT, Shabanova VI and Leaf WA (2003). Driving experience, crashes and traffic citations of teenage beginning drivers. *Accident Analysis and Prevention*, 35, 311–320.

McCartt AT, Mayhew DR, Braitman KA, Ferguson SA and Simpson HM (2009). Effects of age and experience on young driver crashes: review of recent literature. *Traffic Injury Prevention*, 10(3), 209-219.

McCartt AT, Teoh ER, Fields M, Braitman KA and Hellinga LA (2010). Graduated licensing laws and fatal crashes of teenage drivers: a national study. *Traffic Injury Prevention*, 11, 240-248.

McKenna, FP (2018). Do low levels of confidence suppress the tendency to violate? *Transportation research part F: traffic psychology and behaviour*, 55, 136-140.

Montori VM and Guyatt GH (2001). Intention-to-treat principle. *Cmaj*, 165(10), 1339-1341.

O'Cathain A, Croot L, Duncan E, et al. (2019). Guidance on how to develop complex interventions to improve health and healthcare. *BMJ Open*; 9:e029954.

Ouimet MC, Pradhan AK, Brooks-Russell A, Ehsani JP, Berbiche D and Simons-Morton BG (2015). Young drivers and their passengers: a systematic review of epidemiological studies on crash risk. *Journal of Adolescent Health*, 57(1), S24-S35.

Quimby AR, Maycock G, Palmer C and Buttress S (1999). *The Factors that Influence a Driver's Choice of Speed: A Questionnaire Study* (TRL325). Crowthorne: Transport Research Laboratory.

Pressley A, Fernández-Medina K, Helman S, McKenna FP, Stradling S and Husband P (2016). *A review of interventions which seek to increase the safety of young and novice drivers* (PPR781). Commissioned and funded by the Department for Transport. Crowthorne: Transport Research Laboratory.

Roberts IG and Kwan I (2001). School-based driver education for the prevention of traffic crashes. *Cochrane Database of Systematic Reviews*, Issue 3. Art. No.: CD003201. https://doi.org/10.1002/14651858.CD003201

Russell KF, Vandermeer B and Hartling L (2011). Graduated driver licensing for reducing motor vehicle crashes among young drivers. *Cochrane Database of Systematic Reviews*, 10, CD003300.

Sagberg F (1998). Month-by-month changes in accident risk among novice drivers. *Paper presented at the 24th International Congress of Applied Psychology*, San Francisco, August 9–14.

Senserrick TM and Williams AF (2015). Summary of literature of the effective components of graduated driver licensing systems (No. AP-R476/15).

Sexton B and Grayson G (2010). *The accident history and behaviours of new drivers who pass their first practical driving test* (PPR427). Crowthorne: Transport Research Laboratory.

Simons-Morton BG, Hartos JL, Leaf WA and Preusser DF (2006). The effect on teen driving outcomes of the Checkpoints Program in a state-wide trial. *Accident Analysis and Prevention*, 38, 907-912.

Smith JD, Li DH and Rafferty MR (2020). The Implementation Research Logic Model: a method for planning, executing, reporting, and synthesizing implementation projects. *Implementation Science*, 15, 84. https://doi.org/10.1186/s13012-020-01041-8

Soto CJ and John OP (2017). Short and extra-short forms of the Big Five Inventory–2: The BFI-2-S and BFI-2-XS. *Journal of Research in Personality*, 68, 69-81.

Sullman M (2017). *Young Driver Safety: A review of behaviour change techniques for future interventions*. London: RAC Foundation

Tong S, Lloyd L, Durrell L, McCrae-Mckee K, Husband P, Delmonte E and Buttress S (2015). *Provision of telematics research* (PPR755). Crowthorne: Transport Research Laboratory.

Vaa T, Høye A and Almqvist R (2015). Graduated driver licensing: searching for the best composition of components. *Latin American Journal of Management for Sustainable Development*, 2(2), 160-176.

Vernick JS, Li G, Ogaitis S, MacKenzie EJ, Baker SP and Gielen AC (1999). Effects of high school driver education on motor vehicle crashes, violations and licensure. *American Journal of Preventative Medicine*, 16(1), 40–46.

VicRoads (2017). *Examination of the impact of the graduated licensing system on young novice driver safety: summary report.* Melbourne: VicRoads. ISBN 978-0-7311-9176-5

Weekley J and Helman S (2024). *Driver2020 – an evaluation of interventions designed to improve safety in the first year of driving. Supplementary appendix – data collection surveys and intervention logic models* (PPR2034). Commissioned and funded by the Department for Transport. Crowthorne: Transport Research Laboratory.

Weekley J, Helman S, Makosa H, Harpham N and Hutton J (2024b). *Driver2020 – an evaluation of interventions designed to improve safety in the first year of driving. Report D2: Effectiveness of interventions delivered to learner drivers* (PPR2010). Commissioned and funded by the Department for Transport. Crowthorne: Transport Research Laboratory.

Wells P, Tong S, Sexton B, Grayson G and Jones E (2008). Cohort II: a study of learner and new drivers. Volume 1: main report. Road Safety Research Report No. 81. London: Department for Transport.

West SG, Duan N, Pequegnat W, Gaist P, Des Jarlais DC, Holtgrave D, Szapocznik J, Fishbein M, Rapkin B, Clatts M and Mullen PD (2008). Alternatives to the randomized controlled trial. *American Journal of Public Health*, 98(8), 1359–1366. https://doi.org/10.2105/AJPH.2007.124446

Williams AF and Ferguson SA (2004). Driver education renaissance? *Injury Prevention*, 10 4–7.

Williams AF (2017). Graduated driver licensing (GDL) in the United States in 2016: A literature review and commentary. Journal of safety research, 63, 29-41.

Appendix A What is known about young and novice drivers – the starting point for the Driver2020 project

This section expands on the summary provided in Section 1.

A.1 Young and novice drivers represent a long-standing road safety challenge

Half a century ago, Goldstein (1972) noted “That youthful drivers…are over-represented in accidents…considerably beyond their proportion in the driving population, has been well known for several decades…”. Within the same context Goldstein also drew attention to the widely observed fact that those with little experience of a new task tend to make more errors and show less dependable skill and judgement in its execution than those with more experience. Thus, it would not be a stretch to claim that the safety challenge presented by young and novice drivers is something that has been known for three-quarters of a century, and the group is still over-represented in fatal and serious crashes (e.g. House of Commons, 2021). Evidence from research into this group has confirmed that age and experience both play a role in an increased risk of being involved in road collisions.

A.2 Increased age and increased experience are associated with reduced risk

Studies in multiple countries have shown that the collision risk of new drivers is greater than that of more experienced drivers (Wells, Tong, Sexton, Grayson & Jones, 2008; Mayhew, Simpson & Pak, 2003; McCartt, Shabanova & Leaf, 2003; Sagberg, 1998; Forsyth, Maycock & Sexton, 1995; Maycock, Lockwood & Lester, 1991). These studies also show that the younger drivers are at a greater risk of collision than older drivers (although it should be noted that risk rises again in old age). Several reasons are offered as to what it is about younger age that leads to greater risk, including those associated with lifestyle (for example driving while under the influence of alcohol, and with friends in the car who distract the driver) and neuroscience (for example the underdevelopment of the frontal lobes – see Isler & Starkey, 2008).

Figure A-1 shows data reproduced from Maycock et al. (1991), from Great Britain. These data are modelled from self-reported collisions, with exposure kept constant. The dotted lines in the figure show (separately for males and females) the first-year accident liability for drivers passing their test at a given age. The solid lines show accident liability for people who pass their test (and therefore begin driving) at age 17, as they get older and accumulate on-road driving experience. Around 90% of the collisions in this dataset and others like it in UK studies are so-called ‘damage only’ collisions. The remaining 10% is dominated by collisions in which slight injuries occurred. Serious injury collisions tend not to be frequent enough to study statistically in this way in sampled datasets.

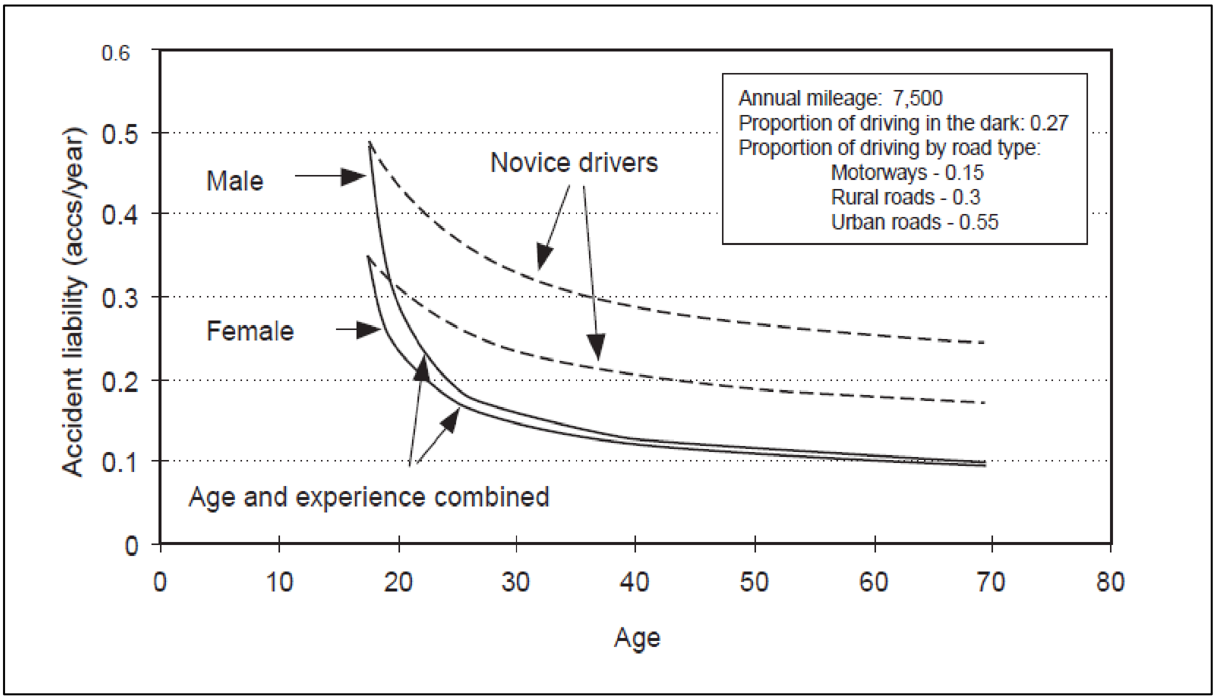


Figure A-1: The effects of age and experience on collisions risk (figure reproduced from Maycock et al., 1991)

While from data collected 30 years ago, this figure demonstrates the broad pattern of findings with respect to age and experience seen in modern datasets. All other things being equal, the younger a driver is when they begin unsupervised driving, the greater their risk of being involved in a collision, and new drivers of all ages become safer as they accumulate on-road experience. There are exceptions; for example, a subset of first-time passers in Sexton and Grayson (2010) were shown to have a lower crash risk despite being younger and reporting a driving style that has in general been associated with greater risk. However, the general protective effects of maturation and on-road experience are seen in multiple countries. Estimates suggest that most of the improvements in safety arising from experience come in the first 1,000-3,000 miles of independent driving (McCartt et al., 2003; see also Kinnear, Kelly, Stradling & Thomson, 2013).

It should be noted that collisions in the learner period, when drivers are supervised, are at a very low level compared with the levels reached when unsupervised driving begins (for example see Vicroads, 2017).

A.3 Approaches that target age and inexperience in higher risk situations have worked well to reduce risk in young and novice drivers

The most successful approaches to lowering the risk of collisions in this group have been those licensing approaches known collectively as 'Graduated Driver Licensing' (GDL). Such approaches focus explicitly on increasing the age at which drivers become licensed, and on increasing levels of on-road experience in safer conditions both before licensure (supervised driving), and afterwards (supervised driving, or solo driving in lower risk conditions). The evidence is summarised here; for readers interested in more detail, other reviews are available (Kinnear et al., 2020; Kinnear & Wallbank, 2020; Kinnear, Lloyd et al., 2013; Russell, Vandermeer & Hartling, 2011).

Age has been targeted in several ways in GDL systems. The first way is to simply define the age at which someone can begin to drive unaccompanied, and the evidence suggests that when this is set to an older age, this increases safety (McCartt, Mayhew, Braitman, Ferguson & Simpson 2009). A second way in which age is targeted is to increase the minimum time for which someone needs to remain in the 'learner' phase of the licensing system. Evaluations of this have consistently found that longer periods lead to greater safety benefits (Senserrick & Williams, 2015).

A longer learning period also gives more opportunity for on-road practice. This has also been targeted directly in some licensing systems through the setting of minimum requirements; evidence from Australia suggests that if at least 120 hours of on-road practice can be achieved, this probably has safety benefits (Senserrick & Williams, 2015), although the evidence on this issue is less well established than on increasing licensure age and the length of learner periods.

The approach to managing post-licence experience in GDL is built on an understanding of specific higher risk situations for young and novice drivers. The reasoning is that if experience can be allowed to build up initially either in low-risk conditions, or when supervised in higher risk conditions, the probability of collisions can be reduced. The two higher risk situations that are most well understood are driving when carrying peer-age passengers and driving at night.

For drivers aged under 25, carrying at least one passenger has been shown by a systematic review of the literature (Ouimet et al.,2015) to increase fatal crashes by 1.24 to 1.89 times relative to solo driving. The risk estimate for carrying two or more passengers was 1.70 to 2.92. It has been shown that passenger restrictions in GDL systems lead to overall reductions in risk (Senserrick & Williams, 2015; Williams, 2017; Vaa et al., 2015; Begg & Stephenson, 2003).

Younger drivers are also known to be over-represented in collisions at night in the UK (Clarke, Ward & Truman, 2002) and it is known that the more night-time driving is restricted in GDL systems in the US, the greater the reductions in collisions (McCartt et al., 2010).

There are other components that are often found in GDL systems. These include alcohol restrictions, speed limits, vehicle power restrictions and the use of vehicle identifiers to signal licensing status. The evidence on these has proven difficult to isolate from the effects of other GDL components (for a recent summary see Kinnear et al., 2020).

Overall, the evidence shows that GDL components that focus on age and on-road experience, both before and after licensing, are effective at reducing collision risk in novice drivers.

A.4 Approaches based on education, training, technology, and other mechanisms have not fared as well in reducing risk

The literature on interventions described using labels such as 'driver education' and 'driver training' is very large. Consequently, there have been numerous systematic and narrative reviews of it (Kinnear et al., 2013; Clinton & Lonero, 2006; Mayhew, Simpson & Robinson, 2002; Roberts & Kwan, 2001; Christie, 2001; Vernick, Li, Ogaitis, Mackenzie, Baker & Gielen, 1999; Mayhew, Simpson, Williams, & Ferguson, 1998; Brown, Groeger, & Biehl, 1987). These reviews have all come to the same conclusion, an example of which is the following quote from Helman, Grayson and Parkes (2010):

> "The only direct benefits imparted by broad driver education and training would appear to be the basic vehicle-control skills and knowledge of road rules necessary for entering the driving population. According to the evidence, it has no measurable direct effect on collision risk, and its continued use should therefore be set against much lower expectations in terms of what it can contribute directly to the safety of new drivers." (Helman et al., 2010, p8).

Such approaches can even cause harm; examples can be found in the literature regarding safety interventions such as skid-training courses that appear to promote over-confidence and risk-taking (Katila et al., 1996; Jones, 1993; Glad, 1988), and of driver education courses that can increase risk through allowing earlier licensure (Williams & Ferguson, 2004).

The conclusion from Helman et al. (2010) is necessarily in need of regular review since the specific theoretical approaches used in training and education (for example behaviour change techniques – see Fylan, 2017) and the technologies available to support them develop over time.

One promising approach, which illustrates the fast pace of technology development and the need for sound evaluation, is the use of 'telematics'. Typically, such technologies are used as part of insurance policies for young drivers. Tong et al. (2015) reviewed the literature on such approaches and concluded that no convincing evidence existed that they reduce risk. According to Tong et al., one of the main issues with the literature is that it is not feasible to utilise insurance data alone in evaluating such technologies, because of the self-selection and insurer-selection biases inherent in such datasets. Without properly matched control groups of drivers who do not have such policies, it is not possible to properly evaluate any change in risk associated with having such a policy independently of other effects, such as the types of drivers who typically have such policies.

One study that does have data on telematics policies, with a comparison group – albeit not one free from the biases mentioned above – is that of Helman et al. (2017). In a dataset from over 4,000 novice drivers in the UK, the study examined those factors associated with collision risk at six months post-test, when age and exposure were controlled. Those drivers in the sample with a telematics-based insurance policy reported *more* collisions at six months post-test than those without such a policy (the increase in risk after correcting for exposure and age was 50%). This concerning finding may have arisen due to biases in the groups and illustrates the importance of undertaking research from which causality can be inferred; innovation without controlled evaluation is not enough.

Appendix B Email used for recruitment (general format)

From: Research at DVSA
To: Email address
Subject: Take part in newly-qualified driver research

Hi [Name],

Congratulations on passing your driving test on [Test date] at [Driving Test Centre] driving test centre. Now that you've passed, you can take part in some exciting research called 'Driver2020'.

It's looking at ways to help people like you to be safer, more confident and more skilful in their first year of driving.

What's involved
You'll have to fill in some **online** surveys during your first year of driving.
You may also be asked to either:

- take part in some free training
- download an app

Why you should take part
You'll get paid for filling in the surveys.

You'll also be entered into a prize draw to win free car insurance for a year, a tablet and vouchers.
And it might help you to become a more skilful driver.

Coronavirus (COVID-19) (Note this section was only used during the pandemic)
The Government has provided guidance regarding staying at home and avoiding non-essential travel. Please continue to follow this guidance. For further details please visit:
https://www.gov.uk/coronavirus

In light of the current situation we understand that you will likely be driving less than expected. Don't worry about this – you can still take part in this project.

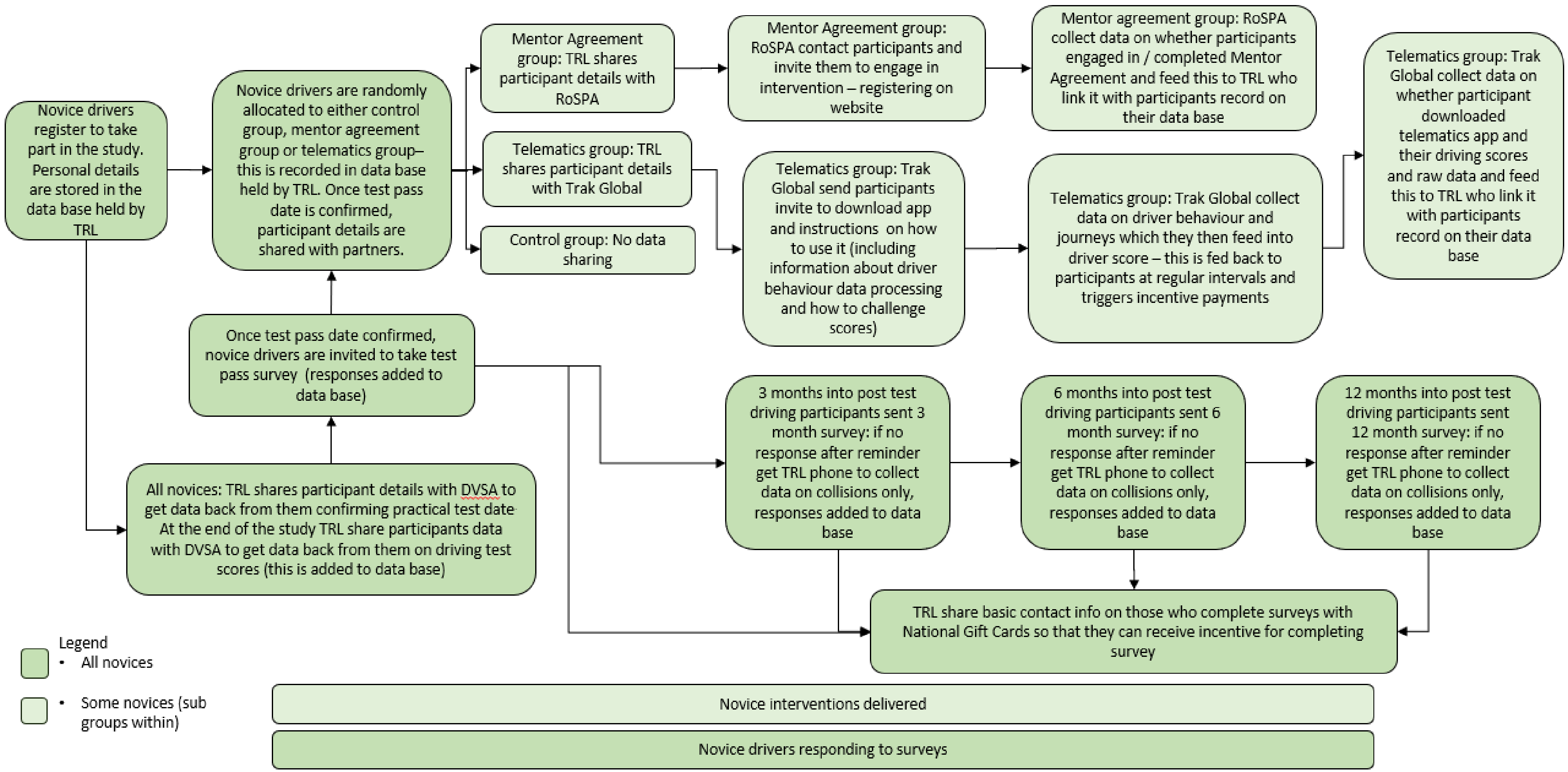
Find out more
Find out more or register to take part at:
https://www.gov.uk/guidance/take-part-in-newly-qualified-driver-research-driver2020

You can leave the research at any time you want if you decide to register. All of your information will be kept safe and secure.

We really hope that you can get involved.

Thank you, Driver and Vehicle Standards Agency (DVSA)

Appendix C Novice interventions – data flows



Appendix D Statistical testing

Statistical tests were used throughout this study to test for differences between groups (for instance, control, logbook, hazard perception training or education course groups, or between pre-pandemic and post-pandemic groups).

In all cases where statistical tests were used to compare data, the convention from the behavioural sciences of reporting *p*-values was adopted. The probability value, *p*-value, is used to determine statistical significance in a hypothesis test. If the probability of obtaining the value of the test statistic is less than 0.05, then the null hypothesis (or the assumption that the groups are equal on the same measure) is rejected and the result is termed 'statistically significant' (less than 5% likely to have occurred by chance alone). Tests where $p<0.05$ are therefore considered significant, however exact *p*-values have been reported throughout except where $p<0.01$.

Statistical significance alone can be misleading in trials with very large samples because it is influenced by the sample size. Increasing the sample size makes it more likely to find a statistically significant result, even for very small effects. As the sample size increases, the size of the effect that can be detected on a given measure decreases (i.e., one can identify very small or minor changes in behaviour). Therefore, this study presents the effect sizes where the results were statistically significant. Effect sizes help understand the magnitude of differences found. Both are essential to understand the full impact of the interventions.

A number of different tests were used in this study, depending on the type of data collected. Depending on the type of data, the tests could assume independence (responses are obtained from different individuals) or repeated measures analysis (the same individual responded multiple times). For instance, tests looking at the differences over the three reporting periods were repeated measures tests, whereas tests comparing the overall differences between the control and intervention groups assumed independence. Both parametric and non-parametric approaches were used based on the satisfaction (or not) of the parametric test assumptions; parametric approaches are more powerful but require the data to meet a number of criteria before they can be used.

The meaning of an effect size varies on the type of statistical test that is being used; the same effect size value has different meanings for different statistical tests. To give an indication of real-world impact, values of a given effect size statistic are described as 'negligible', 'very small', 'small', 'medium' or 'large'.

Techniques used include:

- Chi-square tests for independence used to compare the relationship between categorical variables. This is a non-parametric test. The effect size was determined using Cramer's V. Chi-square tests were used extensively during the analysis, for example to compare demographic distributions between groups in sections 4.3.1 and 4.5.2.
- Independent sample t-tests to compare mean values of a continuous variable for two groups or the non-parametric alternative, Mann-Whitney U tests. Instead of comparing means like the t-test, the Mann-Whitney test converts the scores into ranks and compares the median of the two groups. Effect sizes for the Mann-Whitney U tests were determined using Cliff's Delta. These tests were used in

section 4.5.3 as part of the per protocol analysis. Mann Whitney U tests were also used to compare the distribution of the personality variables in section 4.5.2.

- One-way analysis of variance (ANOVA) to compare the change in mean values of a continuous variable across an independent variable with more than two 'levels' (for example, intervention group). Generalised eta square was used to determine the effect size. These tests were used in section 4.5.3 as part of the per protocol analysis to assess differences between groups at each time point. The non-parametric alternative, Kruskal Wallis, was used if the assumptions of the parametric test were not satisfied; the effect size in this case was estimated as the eta squared based on the H-statistic. Kruskal Wallis tests were used in section 4.3.4 to assess differences between groups in learning and driving conditions.
- Repeated measures analysis of variance (ANOVA) to compare the change in mean values of a continuous variable over time (and across independent variables). Generalised eta square was used to determine the effect size. These tests were used in section 4.5.3 as part of the per protocol analysis to assess differences over time within the same group. The non-parametric alternative, Friedman test, was used if the assumptions of the parametric test were not satisfied, with effect sizes estimated using Kendall's W value. Freidman tests were also used in section 4.5.3.

D.1 Factor analysis

In addition to the tests described above, factor analysis was used to reduce some of the behaviour and attitudinal questions from a large number of related variables into a smaller number of linearly independent variables, reflecting the same underlying information. Based on common practice for these scales, factor analysis was used to combine multiple high correlated items into a single variable (factor). The output from the factor analysis for each question was the number of factors and the items relating to each factor – for example for the 'attitudes to enforcement and restrictions' (eight items combined into three factors) and 'driving style' (12 items combined into three factors).

The number of factors was determined using the 'scree test'. This involves looking at a plot of eigenvalues for factors and determining where the shape of the curve changes. For both sets of items, a three-factor solution was identified.

To determine the items relating to each factor, the factor loadings were used. The loadings (scores between -1 and 1) indicate the strength and direction of the association between each factor and item. No items related strongly to more than one factor. Scores for each factor were then calculated using the coefficients from the factor analysis.

The factor analysis used data from the test pass survey only to determine the number of factors, the survey items and the coefficients relating to each factor. These were then used to calculate the factor scores for participants across all surveys. This ensured that the results are presented using the same factor coefficients and any statistical differences (analysed using the tests described above) can be attributed to a change over time.

Appendix E Characteristics of engagers and non-engagers

As discussed in section 4.5, as part of the per-protocol analysis that was carried out, the characteristics of the engagers for each intervention group were compared with the characteristics of the non-engagers to determine if the ‘type’ of person who engaged differed from the ‘type’ who did not. Where differences were identified for a characteristic variable, the results were shown in the main body of this report. This appendix shows the full results of all the comparisons carried out where no differences were identified.

E.1 Mentoring agreement intervention

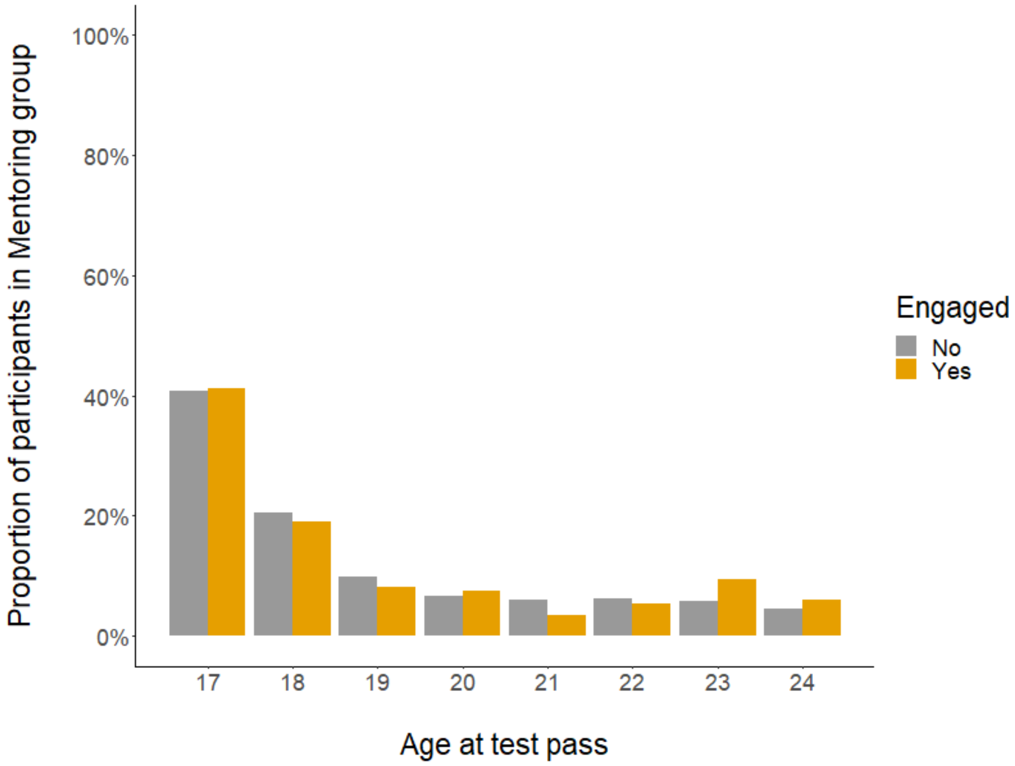


Figure E-1: Proportion of engagers and non-engagers in the mentoring agreement group by age (N=4,183)

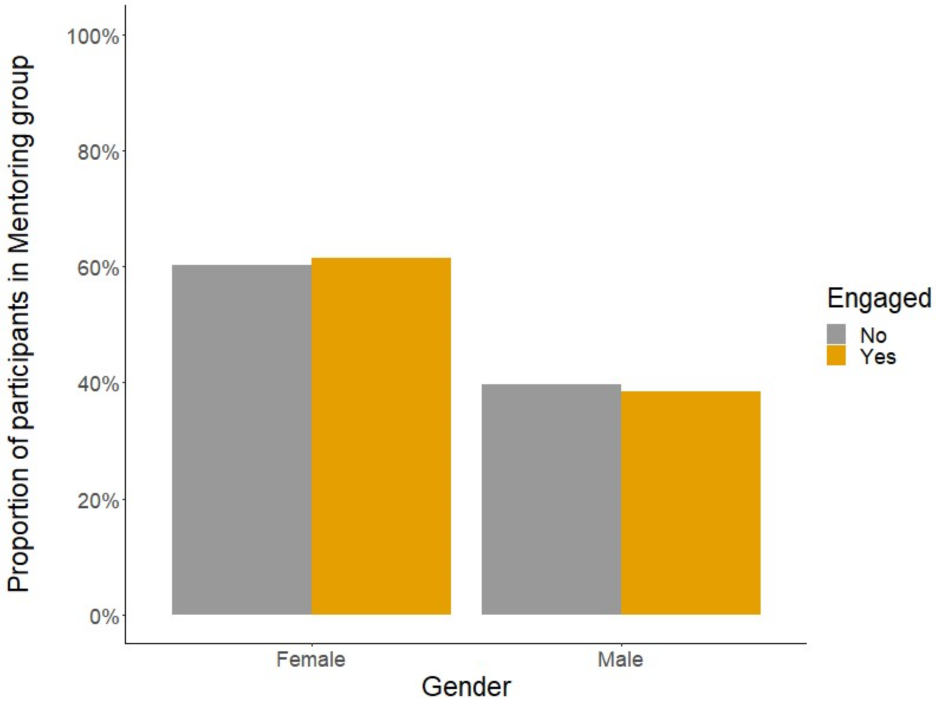


Figure E-2: Proportion of engagers and non-engagers in the mentoring agreement group by gender (N=4,183)

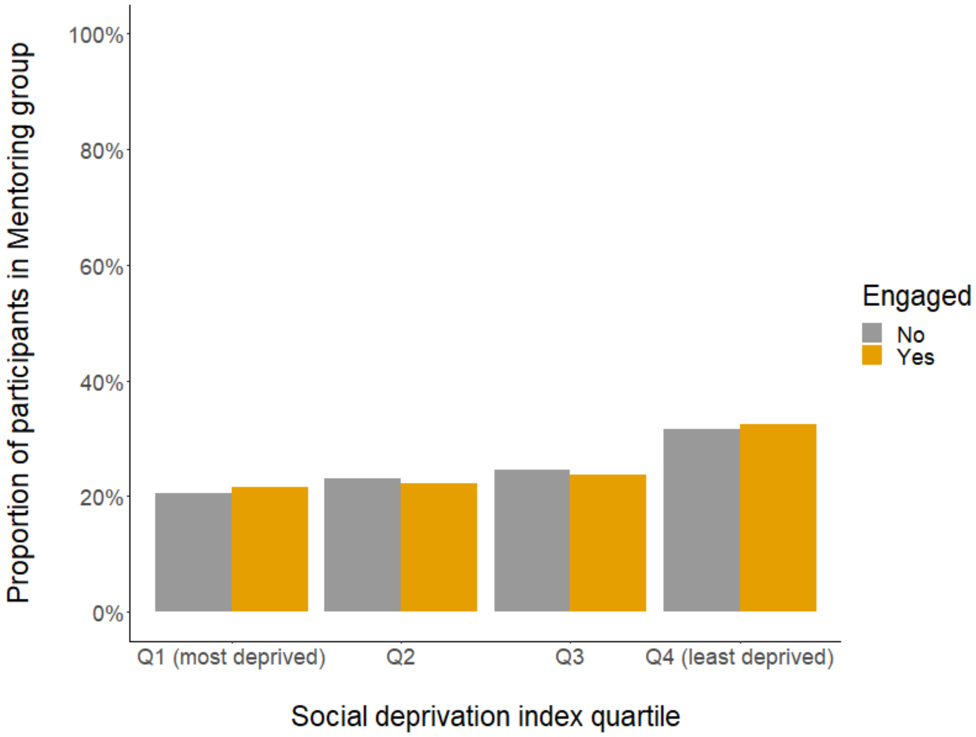


Figure E-3: Proportion of engagers and non-engagers in the mentoring agreement group by social deprivation index quartile (N=4,183)

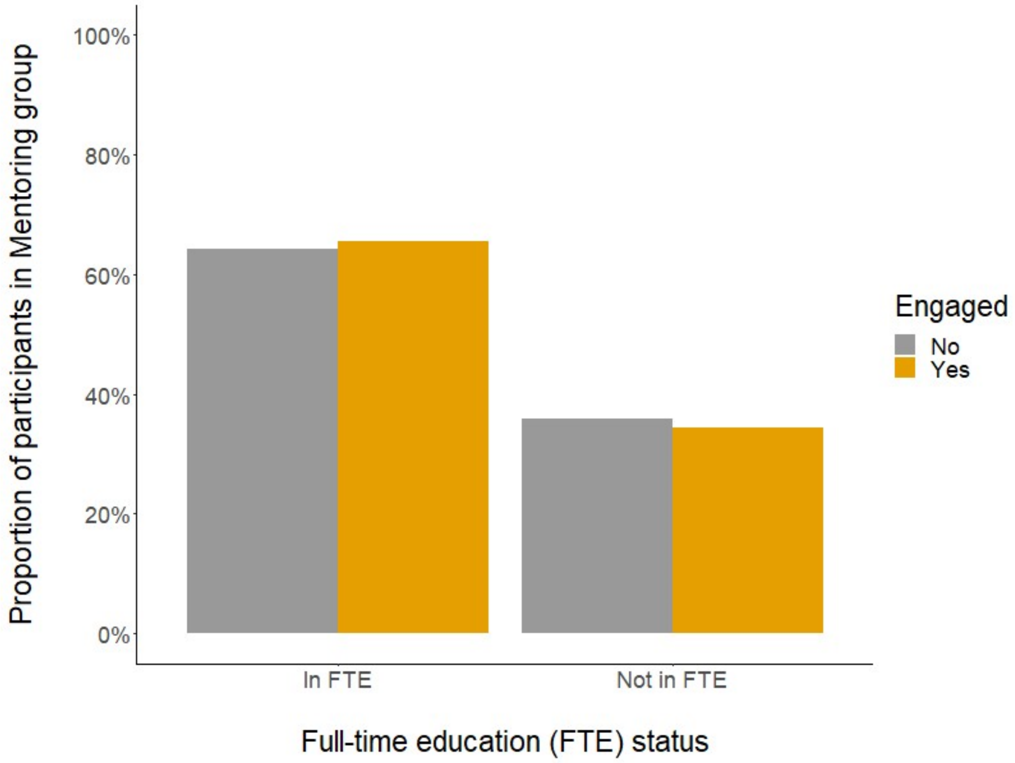


Figure E-4: Proportion of engagers and non-engagers in the mentoring agreement group by education status (N=4,183)

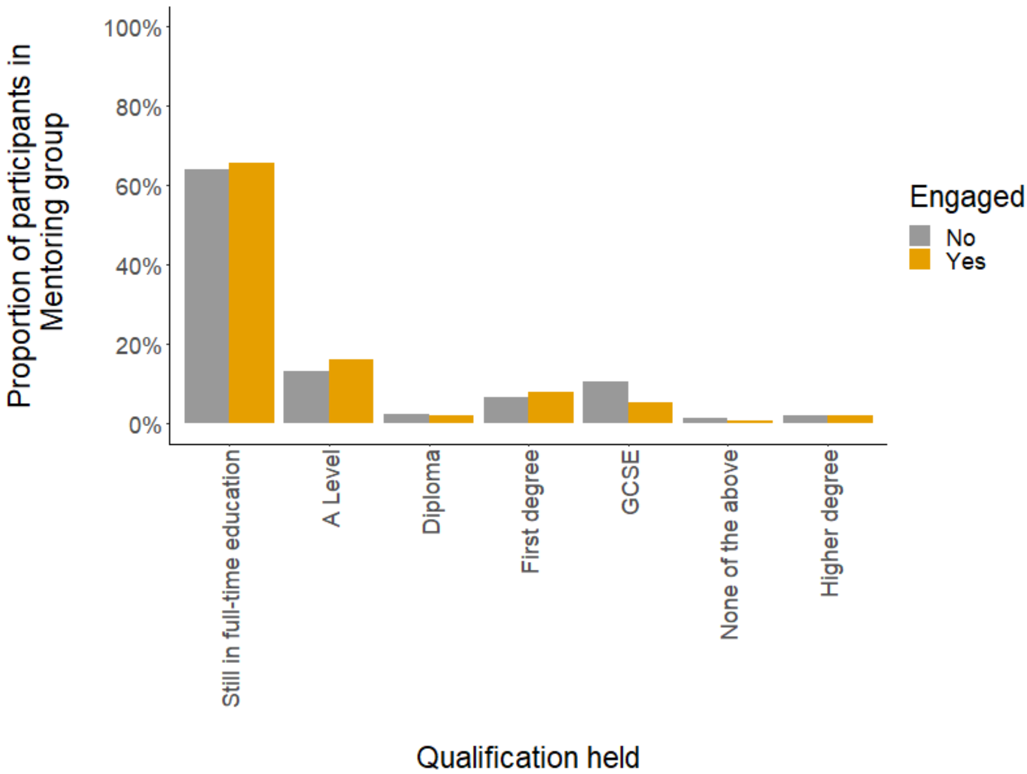


Figure E-5: Proportion of engagers and non-engagers in the mentoring agreement group by highest education qualification held (N=4,183)

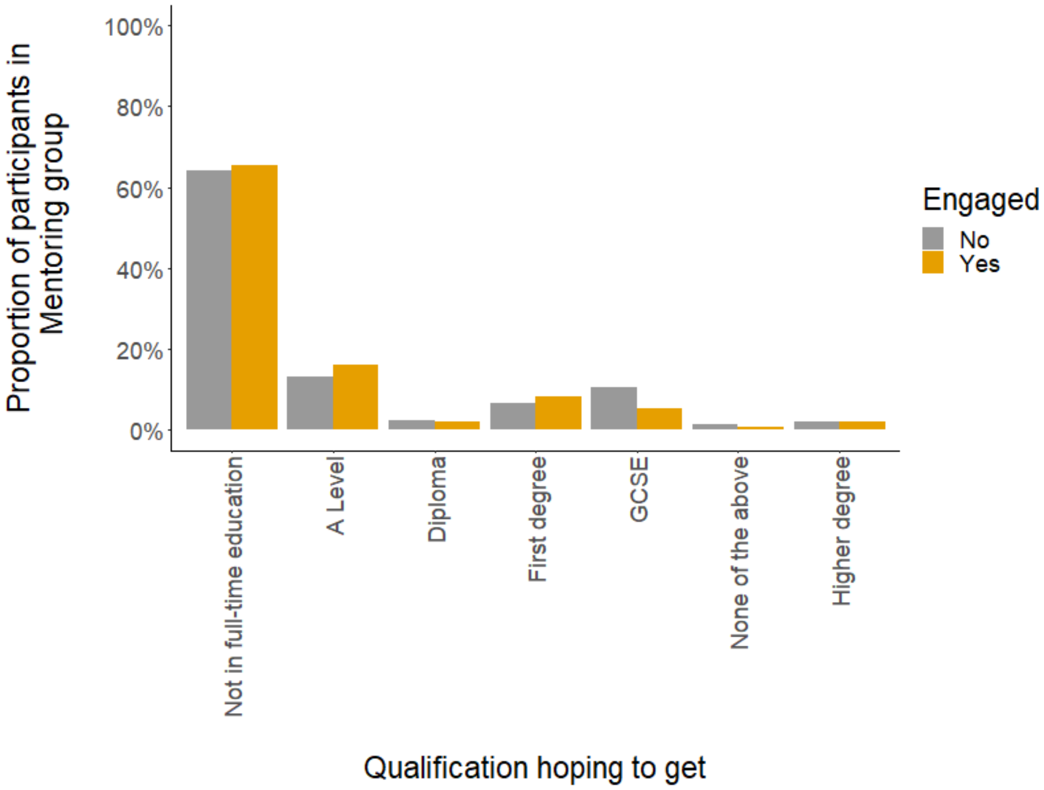


Figure E-6: Proportion of engagers and non-engagers in the mentoring agreement group by education qualification hoping to get (N=4,183)

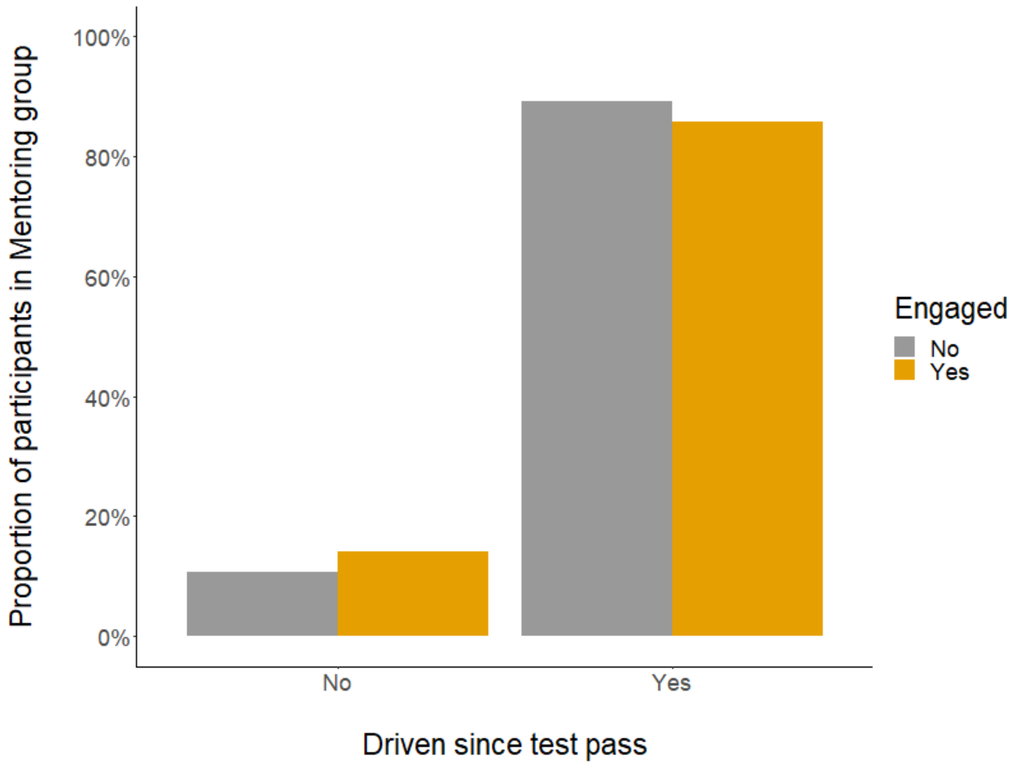


Figure E-7: Proportion of engagers and non-engagers in the mentoring agreement group who did and did not drive after passing their test (N=4,183)

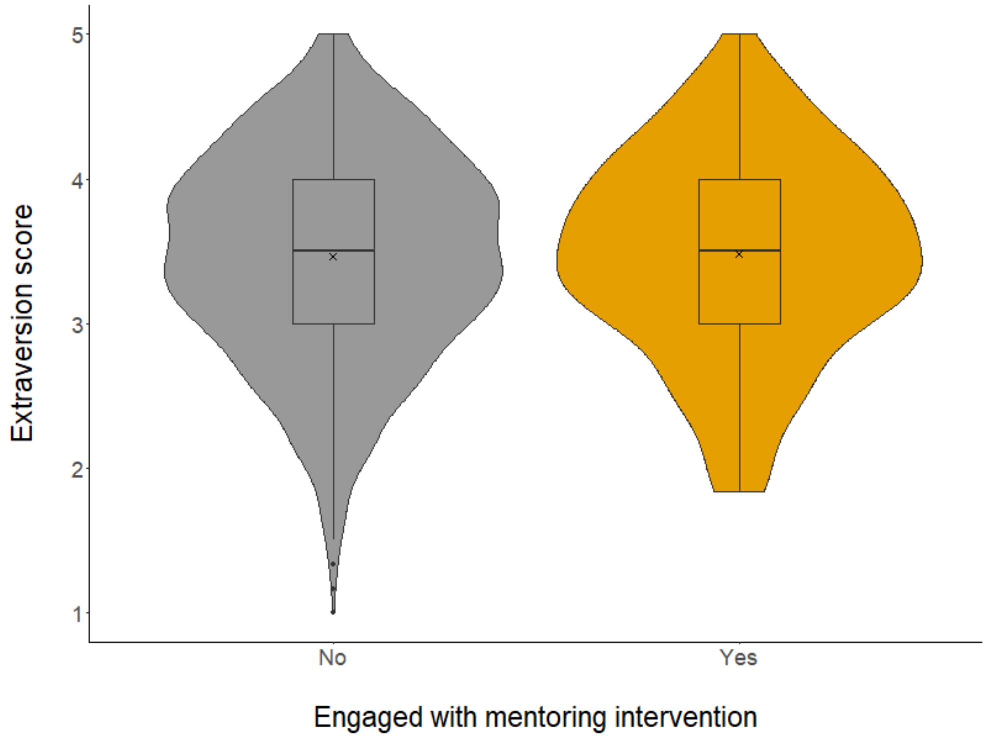


Figure E-8: Distribution of extraversion score by engagers and non-engagers in the mentoring agreement group (1=least, 5=most) (N=4,183)

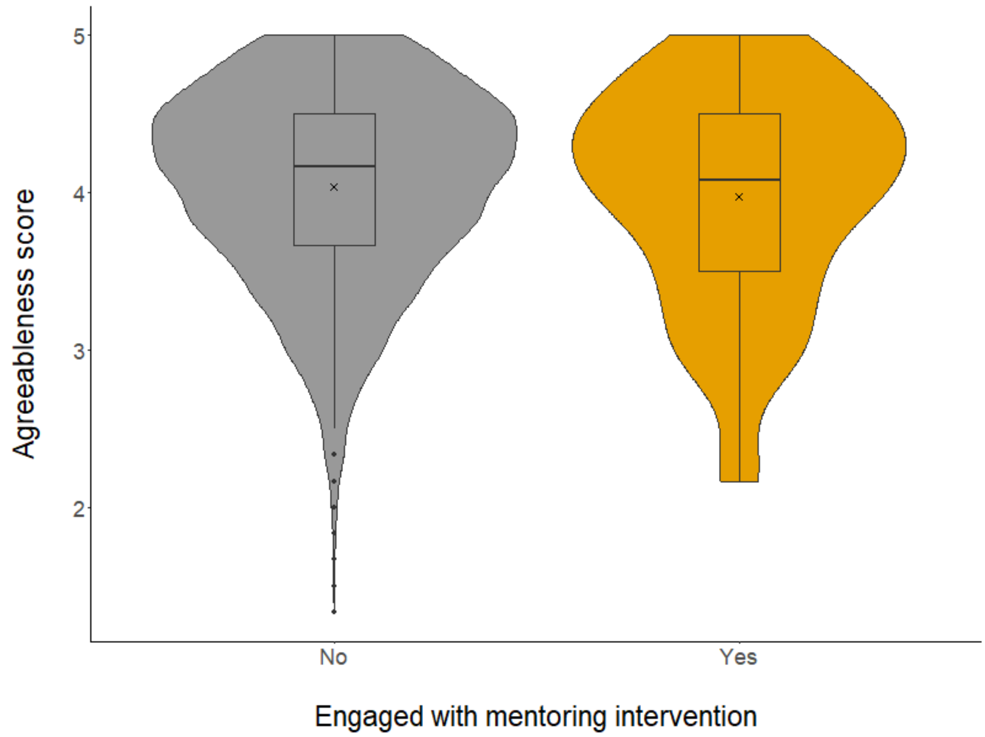


Figure E-9: Distribution of agreeableness score by engagers and non-engagers in the mentoring agreement group (1=least, 5=most) (N=4,183)

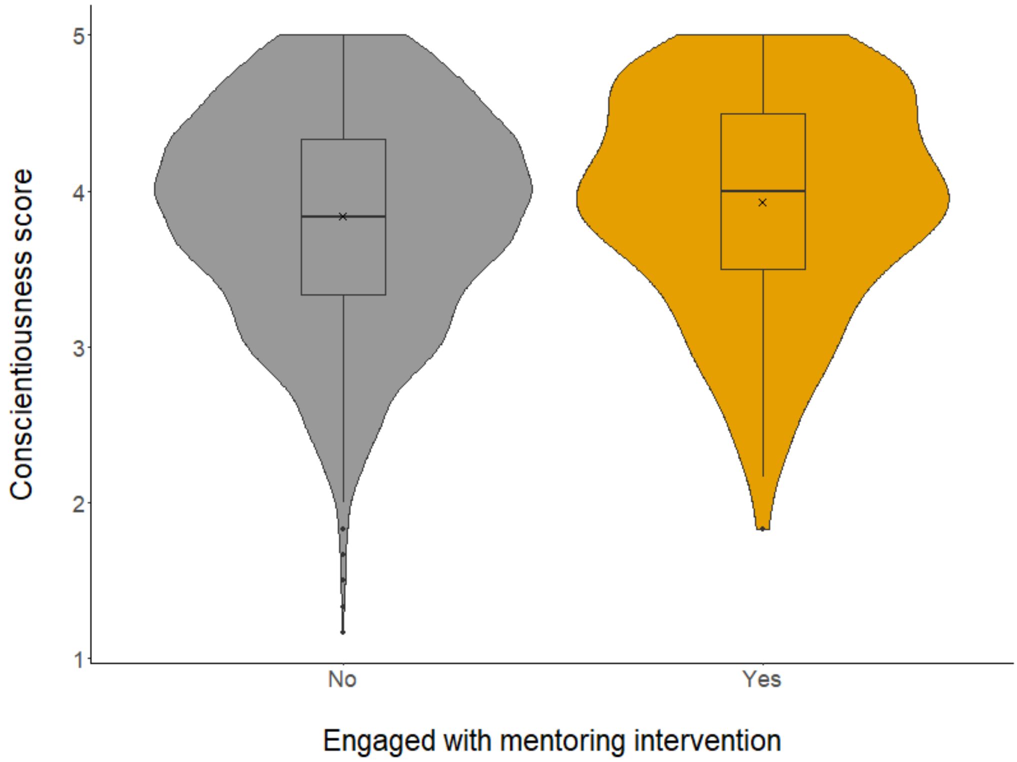


Figure E-10: Distribution of conscientiousness score by engagers and non-engagers in the mentoring agreement group (1=least, 5=most) (N=4,183)

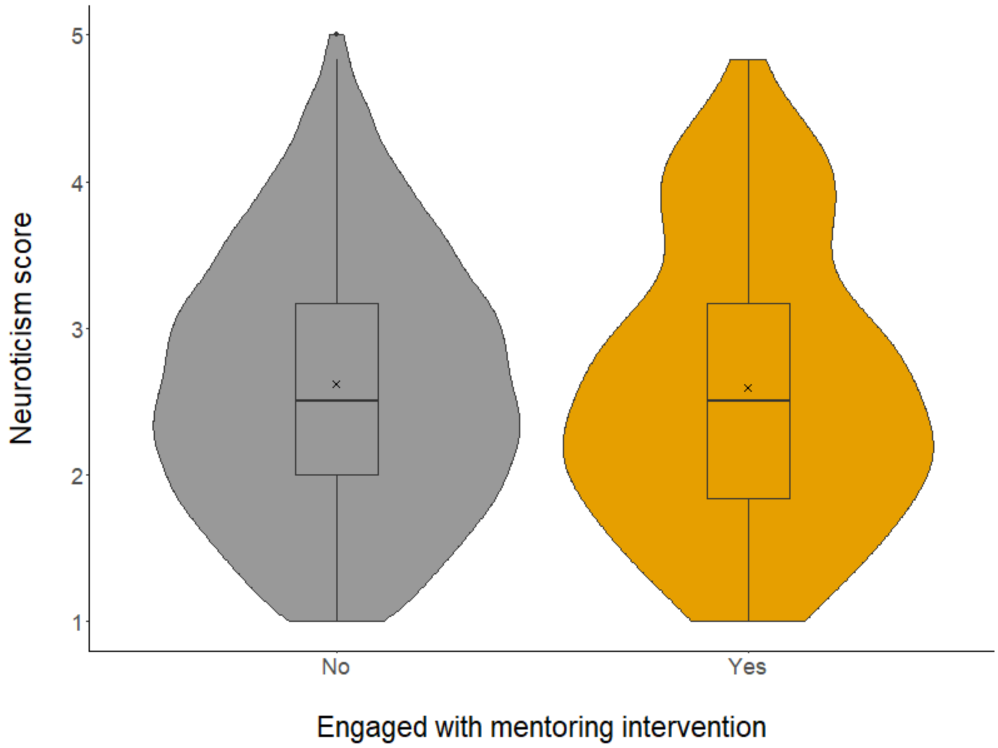


Figure E-11: Distribution of neuroticism score by engagers and non-engagers in the mentoring agreement group (1=least, 5=most) (N=4,183)

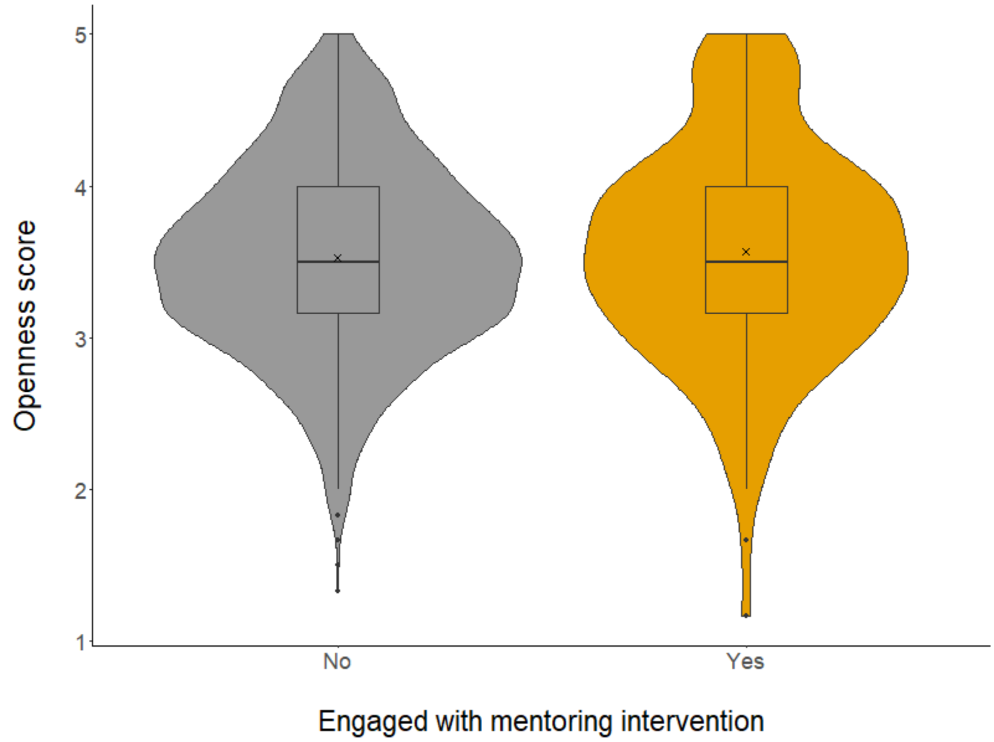


Figure E-12: Distribution of openness score by engagers and non-engagers in the mentoring agreement group (1=least, 5=most) (N=4,183)

E.2 Telematics intervention

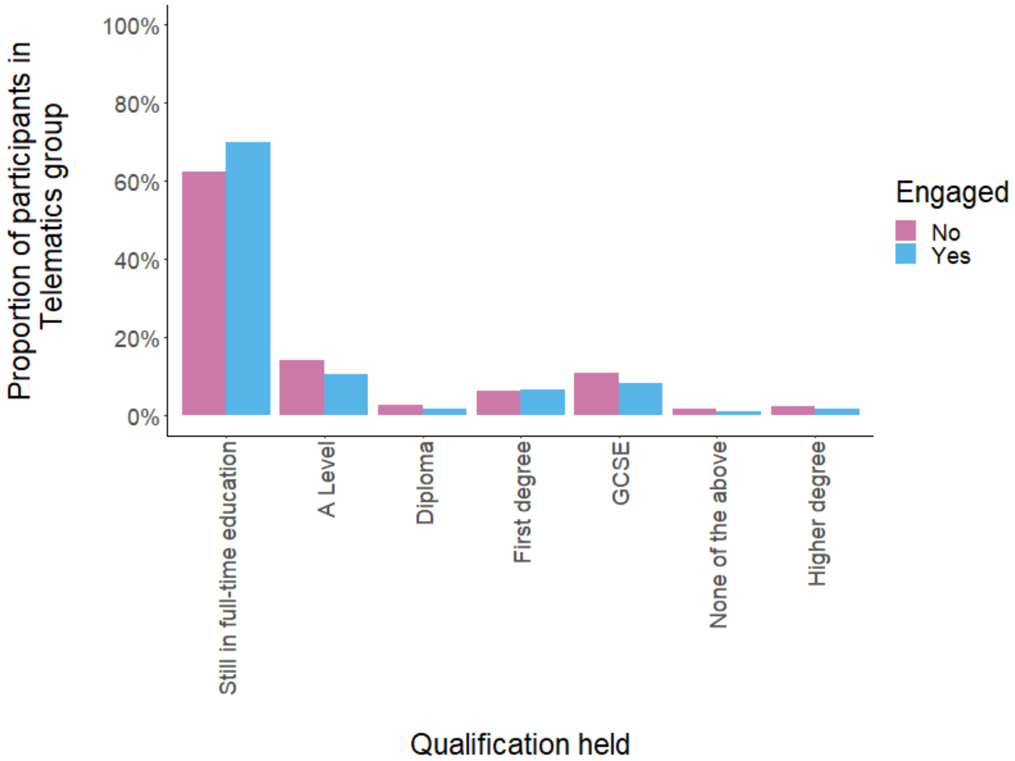


Figure E-13: Proportion of engagers and non-engagers in the telematics group by highest education qualification held (N=4,180)

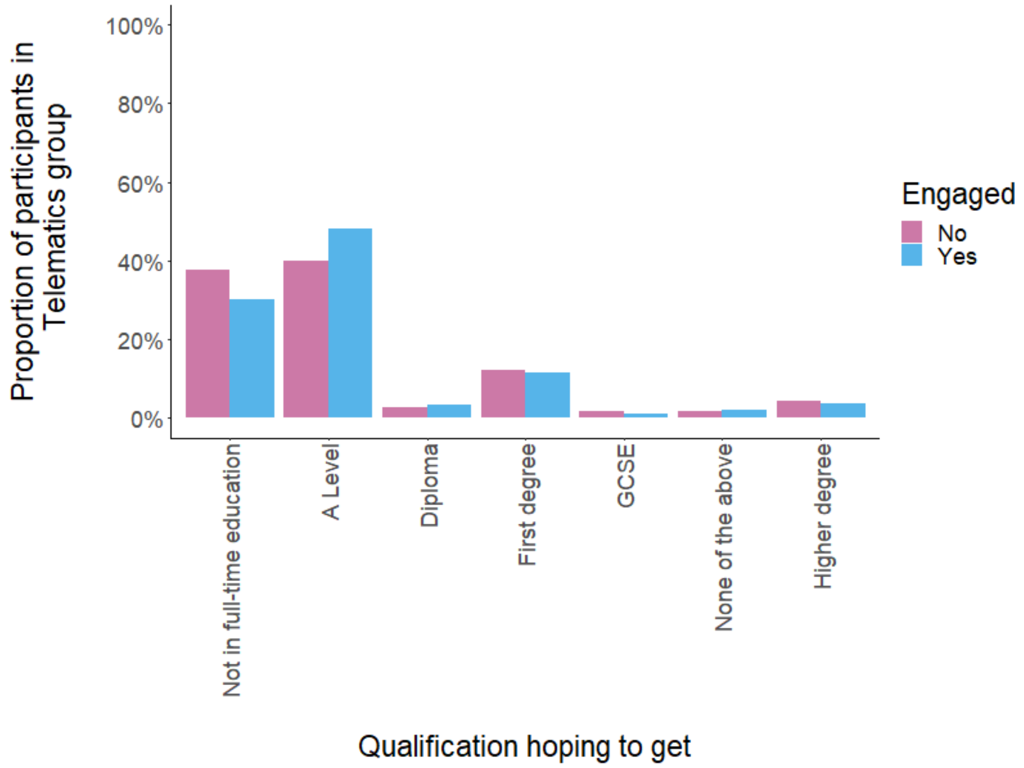


Figure E-14: Proportion of engagers and non-engagers in the telematics group by education qualification hoping to get (N=4,180)

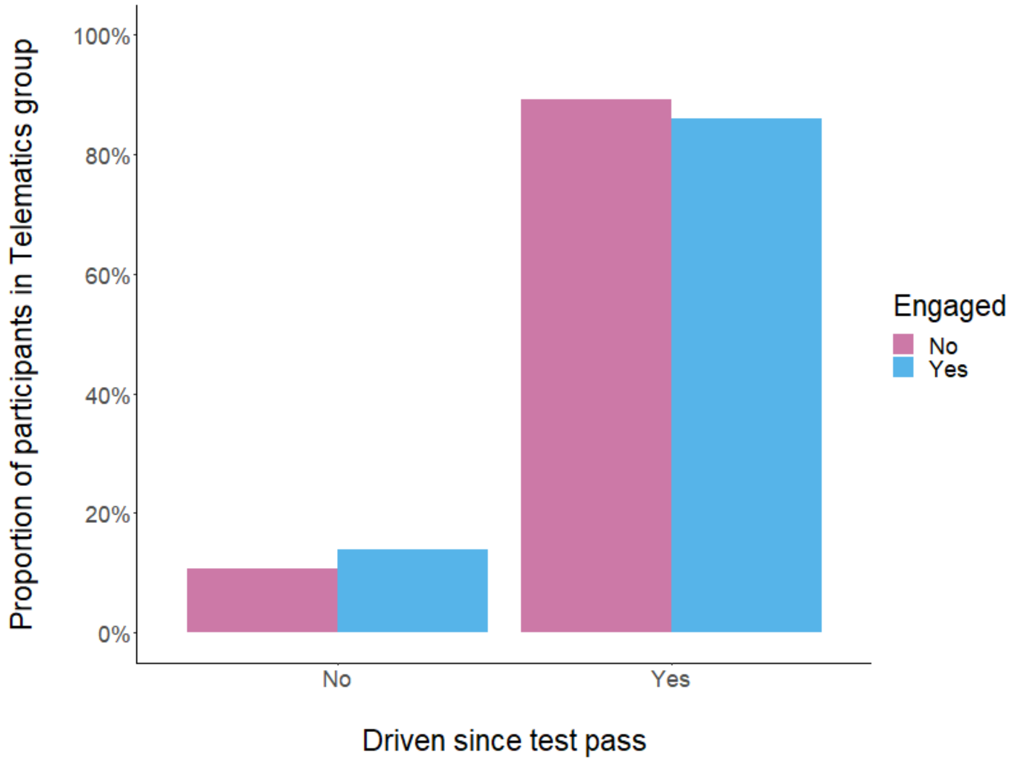


Figure E-15: Proportion of engagers and non-engagers in the telematics group who did and did not drive after passing their test (N=4,180)

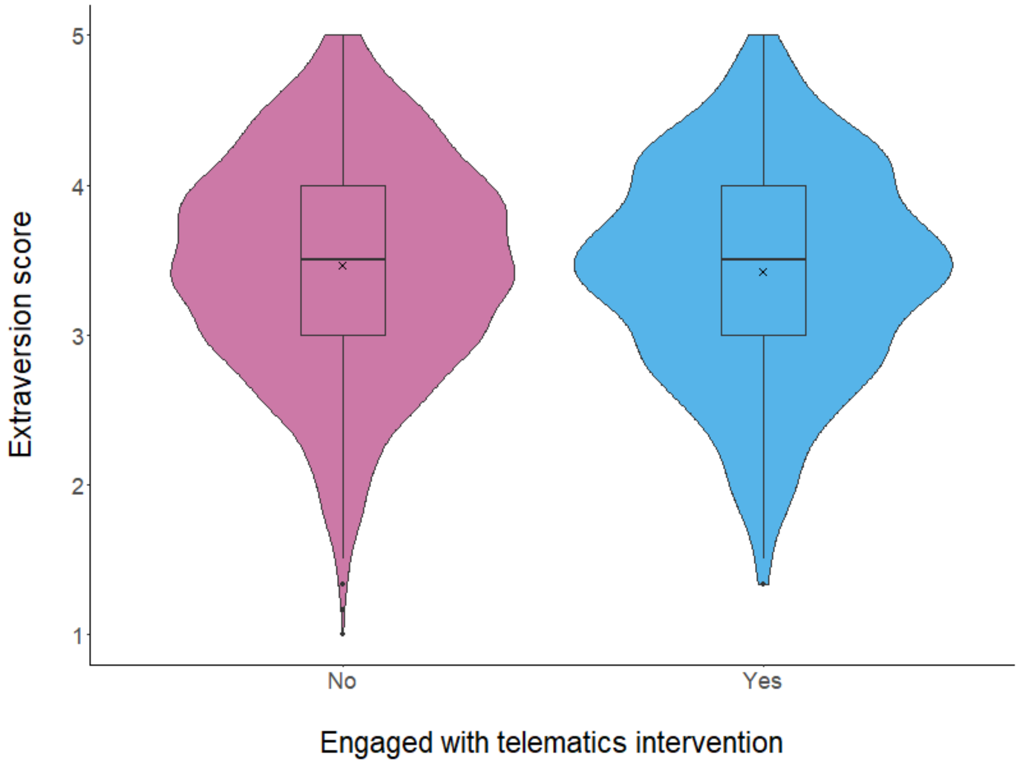


Figure E-16: Distribution of extraversion score by engagers and non-engagers in the telematics group (1=least, 5=most) (N=4,180)

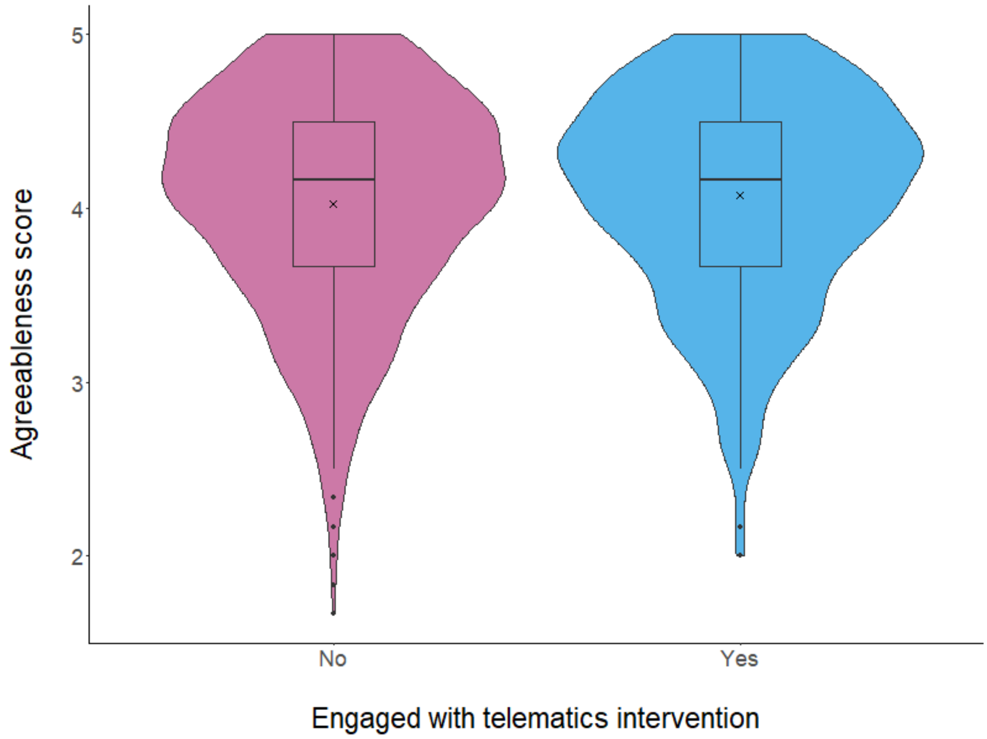


Figure E-17: Distribution of agreeableness score by engagers and non-engagers in the telematics group (1=least, 5=most) (N=4,180)

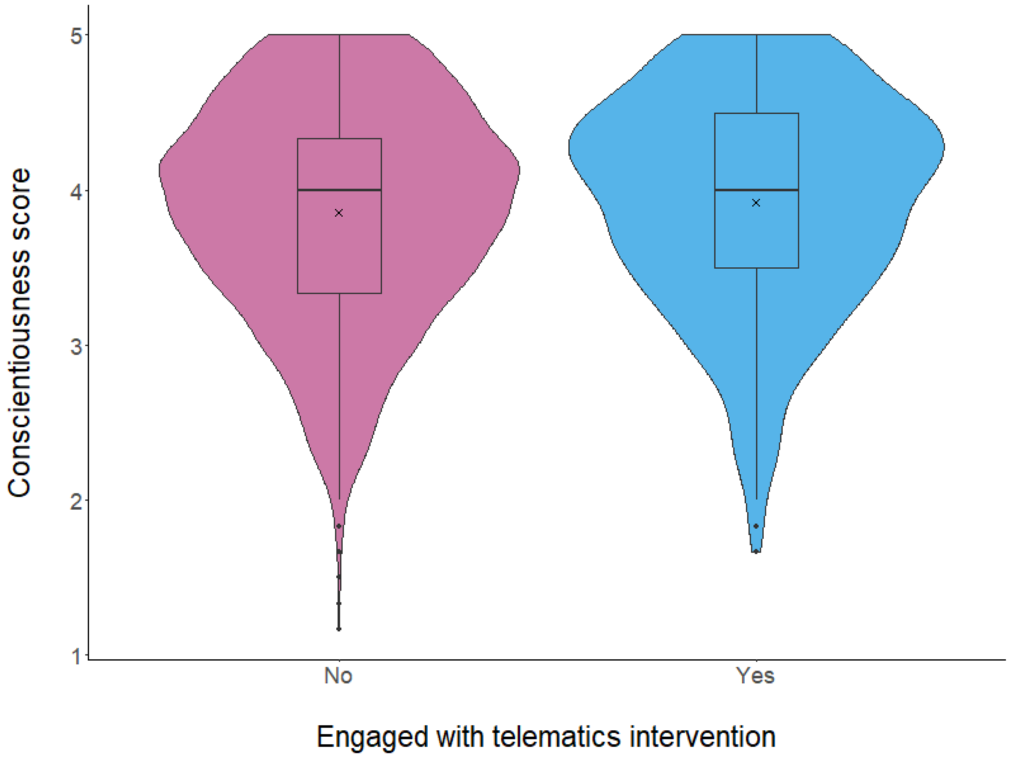


Figure E-18: Distribution of conscientiousness score by engagers and non-engagers in the telematics group (1=least, 5=most) (N=4,180)

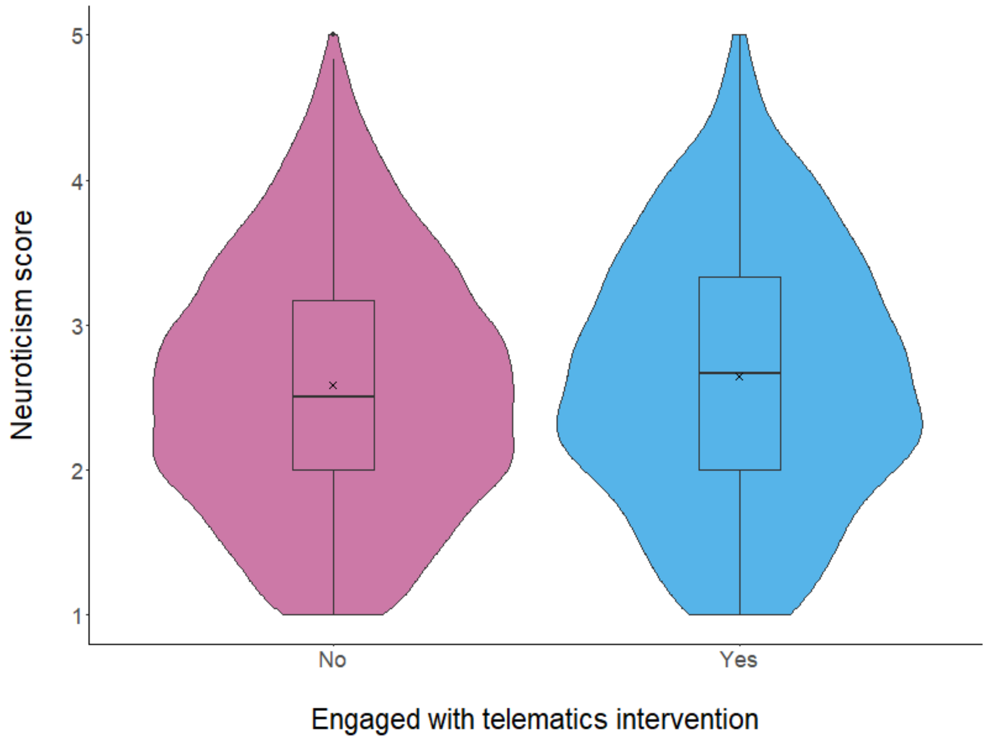


Figure E-19: Distribution of neuroticism score by engagers and non-engagers in the telematics group (1=least, 5=most) (N=4,180)

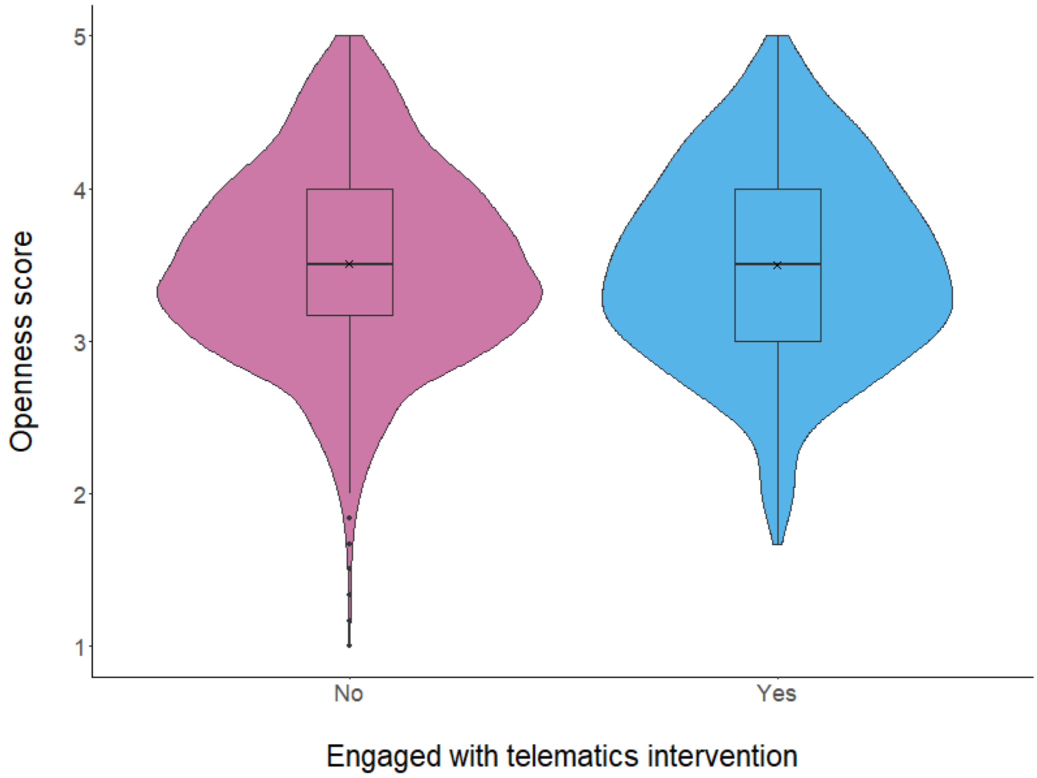


Figure E-20: Distribution of openness score by engagers and non-engagers in the telematics group (1=least, 5=most) (N=4,180)

Appendix F Per-protocol analysis – matching process

F.1 Mentoring agreement intervention

There was no requirement for a matching process for this intervention as there were no significant differences between the control group and the mentoring agreement engagers.

F.2 Telematics intervention

Participants were matched on the three variables that had been found to be different for engagers and non-engagers: age, gender, and education status. These three variables were used to sort the engagers into 32 'cells', defined as the combinations of the different options for each of the variables (for example, male 17-year-olds in full-time education). The control group participants were also sorted into the same cells.

To create the control group sample for use in the per-protocol analysis, four control group participants were randomly chosen from each cell to correspond with each engager in that cell. So, for example, for every engager who was a male 17-year-old in full-time education, four male 17 year olds in full-time education were randomly selected from the control group and were included in the control group sample.

This process resulted in a control group sample of 2,738 participants with age, gender, and education status characteristics which matched those of the participants who engaged with the telematics intervention. This is because there were two cells where there was an insufficient number of control group participants in that cell to achieve the desired 1:4 ratio of engagers to control participants. Note that the total used in the per-protocol analysis was 2,447 as there were 291 participants in the control group sample who did not answer any surveys (apart from the registration survey).

Driver2020 - an evaluation of interventions designed to improve safety in the first year of driving

The Driver2020 study evaluated five interventions designed to reduce collisions and improve safety in young novice drivers in their first year of driving. The study was split into two arms; one covered interventions delivered to learner drivers, and the other covered interventions delivered to novice drivers (after they had passed their test). This report covers the novice arm. Two interventions were delivered to novice drivers who were aged 17-24 when they began driving. The interventions were a web-based mentoring agreement designed to help drivers (ideally with the help of a mentor) set limits in their early driving on higher risk driving situations such as driving in the dark or carrying peer-age passengers, and a 'telematics' smartphone app that provided feedback on driving style and time of day. Novice drivers who registered for the study were assigned randomly to one of the two intervention groups, or to a control group. The intervention participants were then invited to engage with the intervention if they wished. All participants completed self-report surveys after they passed their test, and at 3, 6 and 12 months post-test. These surveys asked people to report the driving they undertook, any collisions post-test, and various surrogate measures of safety targeted by the interventions. Neither of the interventions were associated with a reduction in collisions, and engagement levels with the interventions were very low. In participants who engaged with the mentoring agreement there was an increase in the proportion of people reporting setting limits on driving in the dark and with peer-age passengers, and a reduction in reported proportion of driving in the dark in the first six months of driving – both promising from a safety perspective. Engagement with the telematics app was associated with less reported driving in the dark in the second six months of driving, but also with a slightly riskier self-reported driving style.

Other titles from this subject area

PPR2010 Weekley J, Helman S, Makosa H, Harpham N and Hutton J (2024b). Driver2020 – an evaluation of interventions designed to improve safety in the first year of driving. Report D2: Effectiveness of interventions delivered to learner drivers. Commissioned and funded by the Department for Transport. Crowthorne: Transport Research Laboratory.

PPR2011 Hitchings J, Holcombe A, Christie N, Weekley J and Helman S (2024). Driver2020 – an evaluation of interventions designed to improve safety in the first year of driving. Report D3: Delivery of interventions and engagement with them by novice and learner drivers. Commissioned and funded by the Department for Transport. Crowthorne: Transport Research Laboratory.

PPR2012 Helman S and Weekley J (2024). Driver2020 – an evaluation of interventions designed to improve safety in the first year of driving. Report D4: Summary of findings. Commissioned and funded by the Department for Transport. Crowthorne: Transport Research Laboratory.

PPR781 Pressley A, Fernández-Medina K, Helman S, McKenna FP, Stradling S and Husband P (2016). A review of interventions which seek to increase the safety of young and novice drivers. Commissioned and funded by the Department for Transport. Crowthorne: Transport Research Laboratory.

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

ISSN: 2514-9652

doi: 10.58446/qddo4799

PPR2009