Learning to drive messaging trial: Simple and low cost behavioural interventions to encourage learner drivers to practise more

Contents

Executive summary

Novice drivers are at an elevated risk of collisions after passing their practical driving test. Their risk of collision has been shown to decrease rapidly with on-road experience. The Department for Transport (DfT) and the Driver & Vehicle Standards Agency (DVSA) conducted research during 2016 and 2017 to test the effectiveness of simple and low cost behavioural interventions. These were designed to encourage learner drivers to gain more on-road experience before taking their practical driving test, and potentially increase their road safety. The research involved online focus groups with learner and novice drivers and also a randomised controlled trial (RCT) which evaluated the impact of the behavioural interventions on the length of the learner driver learning period.

Both the online focus groups and the RCT generated valuable evidence that could be used to inform future road safety initiatives:

- The focus groups highlighted that learner drivers do not seem to explicitly consider road safety as a factor when considering whether or not they are ready for their practical test. The participants had little awareness of the heightened risk they would face on the road post-test. However, once the focus group participants were introduced to facts and figures about road safety, they claimed that road safety would play a more important role in their thinking and test booking behaviour.
- The RCT showed that several behavioural interventions were successful at influencing test booking behaviour at the point of booking, before learners had the opportunity to rebook their test (learners can rebook their test free of charge at any point after booking their test). However, once rebooking behaviour was accounted for the picture changed.
- After accounting for rebooking behaviour, only one intervention was linked with an increased length of learning period. The successful intervention involved sending an email to learner drivers with a 'safety message'. The email also highlighted that it is possible to rebook the test for a later date free of charge.
- The results suggest that, out of the behavioural insights that were tested, messages framed using 'loss aversion' could potentially work in some circumstances to encourage safer driving behaviours. Loss aversion is the insight that people attach greater weight to a potential loss than they attach to a potential equivalent gain. Loss frame messaging seemed to work best when combined with facts and figures about road safety, and when combined with a 'call to action' that suggested how the learner drivers could behave to avoid negative road safety outcomes.
- However, the research also suggested that simple behavioural interventions (like those tested in this research) are not likely to have a large impact on road safety outcomes on their own and should therefore be used in combination with other road safety initiatives.

1. Introduction: Learner driver safety

Background

1.1   The Department for Transport (DfT) has overall responsibility for road safety within the British Government. The United Kingdom has comprehensive driver training and various pre and post-test road safety interventions are in place. Newly qualified drivers, however, still experience an elevated risk of collision after passing their practical test, especially for the first six months after passing their test. The risk is even higher for those under the age of 25.

1.2   Existing estimates indicate that approximately 10 – 20% of new drivers have a crash within the first six months of passing their test[1] and 10% are caught committing a driving offence in their first two years[2]. Furthermore, drivers under 25 are 3 times more likely to be in a reported personal injury accident and 3 times more likely to be killed or seriously injured per mile driven. A fifth of all those killed and seriously injured on our roads are aged between 17 and 24[3]. A large proportion of these passed their driving test less than six months before the collision[4].

1.3   Evidence suggests that young and novice drivers have higher accident rates compared to other drivers for two primary reasons:

-   A lack of experience behind the wheel (leading to skill-based deficits, such as poor hazard perception), and;

-   Common behaviours associated with young and novice drivers (more risk-taking behaviours, such as speeding)[5].

1.4   Limited driving experience before taking the practical test can result in fewer opportunities for the learner driver to encounter risky situations and different road environments and conditions, potentially resulting in an inability to anticipate risk or identify hazards and less time to habituate the actual mechanics of driving. Inexperience is therefore a key contributory factor to novice drivers' heightened risk.

---

[1] DfT Road Safety Research Report No. 81 Cohort II: A Study of Learner and New Drivers Volume 1 – Main Report, 2008; and, Transport Research Laboratory (TRL), Transforming the practical driving test, available from: https://www.gov.uk/government/publications/evaluation-of-the-impact-of-the-new-driving-test
[2] Statistic taken from TRL Research Findings: "Novice drivers: Evidence review and Evaluation Pre-driver training, Graduated Driver Licensing", available from: https://www.gov.uk/government/uploads/system/uploads/attachment_data/file/249282/novice-driver-research-findings.pdf
[3] DfT Road Safety Data: https://data.gov.uk/dataset/cb7ae6f0-4be6-4935-9277-47e5ce24a11f/road-safety-data
[4] See, TRL Transforming the practical driving test: https://www.gov.uk/government/publications/evaluation-of-the-impact-of-the-new-driving-test
[5] Pressley et al., 2016, A review of interventions which seek to increase the safety of young and novice drivers. TRL. Available from: https://www.gov.uk/government/publications/review-of-interventions-to-increase-the-safety-of-young-and-novice-drivers

1.5   If learner drivers could be encouraged to acquire more experience before their test, then the safety of novice drivers could be improved and the number of collisions they are involved in might be reduced.

1.6   A recent evidence review of approaches to learning to drive has shown that *"interventions that target extending the learning period … are successful in reducing injuries."*[6] The same review highlights that greater and more varied pre-test experience lowers the likelihood of a new driver being in an accident post-test. It is worth making clear that it is pre-test practise, not length of learning period, which is the causal factor. However, length of learning period is very highly correlated with pre-test practice[7].

1.7   DfT is currently supporting and trialling a range of initiatives to improve learner and novice driver safety, some of which aim to increase the amount of pre-test practice that learner drivers take[8]. As part of a wider programme of work to improve road safety, DfT developed 6 simple behaviour change interventions with the Driver and Vehicle Standards Agency (DVSA). These interventions all sought to encourage learner drivers to extend their learner period and gain more on-road experience before taking their practical test.

1.8   Each of the 6 interventions was designed using insights from the behavioural sciences literature (for example, insights from behavioural economics, psychology and sociology). Behavioural insights are a policy tool that can help policy-makers create interventions which are sensitive to the ways people actually think and behave. They have been used across government to influence people's behaviour, to help people make better decisions about 'health, wealth and happiness'. It is worth noting that some of the interventions tested in this research also rely on standard tools such as providing information to people about risks, tools which have not previously proven to be that effective[9]. The difference with the interventions tested through this research is that they combine these standard intervention tools, such as providing information about risks, and tools from behavioural insights to try and enhance the effectiveness of the messaging.

1.9   The interventions were delivered through the DVSA online practical test booking system and email communications with learner drivers. The interventions were tested using both exploratory qualitative research (online focus groups) and a randomised controlled trial (RCT) methodology. The upfront qualitative research helped strengthen the interventions by ensuring they were sensitive to the context of learning to drive as experienced by learner drivers. The RCT then provided a robust test of the impact that the interventions had on a number of outcomes. The key outcome was the length of the learning period (the length of time in days between logging into the online system and taking the practical test). A secondary outcome measure was the reported number of hours practise that learner drivers reported taking. One limitation of the trial is that it did not follow up with drivers to see if the interventions had an impact on crash risk. However, current evidence and theory mentioned above suggests increased practice is linked to improved safety.

[6] Ibid.
[7] See for example, DfT Road Safety Research Report No. 81 Cohort II: A Study of Learner and New Drivers Volume 1 – Main Report, 2008.
[8] See Pressley et al., 2016, A review of interventions which seek to increase the safety of young and novice drivers. TRL. Available from: https://www.gov.uk/government/publications/review-of-interventions-to-increase-the-safety-of-young-and-novice-drivers
[9] Ibid.

1.10 Given that the nature of the key outcome measure was the time to an event occurring ("time to booking"), survival analysis was used to analyse the data. The particular model used was a Cox regression frailty model, which allowed us to account for the fact that structural factors, such as test centre waiting times, might influence the outcome of interest, in addition to an individual's own characteristics. A Cox regression frailty model is essentially a random effect model for time variables – it can be thought of as 'multi-level survival analysis'.

1.11 This report brings together the evidence that was gathered through conducting the focus groups and trials, presents the key findings and outlines what the findings might mean for other road safety initiatives.

Objectives and hypothesis

1.12 The objective of the research was to identify simple behaviour change interventions that could encourage learner drivers to increase the amount of practise they do before taking their practical test.

1.13 The research tested the hypothesis that simple behaviour change interventions delivered through the DVSA online practical test booking system and subsequent email communications with learner drivers would:

   i)   Increase the amount of time between booking and taking a practical driving test, and;

   ii) Increase the number of hours practised before taking a practical driving test.

1.14 The research also aims to help inform future communications with learner drivers by generating evidence about what types of messages seem to work and raise the saliency of road safety. More broadly, the trial contributes to a growing evidence base on the contribution that behavioural approaches can have in achieving better policy outcomes.

Structure of the report

1.15 The content of the report is as follows:

- Chapter 2 provides an overview of the behavioural interventions, including the process that was used to design the interventions;

- Chapter 3 describes the methods used for the research, including the exploratory qualitative work, the design of the RCT and the statistical methods employed;

- Chapter 4 details findings from the qualitative focus groups;

- Chapter 5 details findings from the RCT;

- Chapter 6 concludes the report with a discussion of the key findings and implications for applying behavioural insights to the field of road safety.

2. Behaviour change interventions

2.1 The multi-disciplinary field of behavioural science has in recent decades been transforming our understanding of human behaviour. This new understanding has made its way into all areas of policy, including road safety[10]. Much of people's behaviour is automatic, habitual, guided by quick if/then heuristics and mental shortcuts[11]. The behavioural sciences have given us an understanding of these automatic behaviours, and what might be done to change them. One other key thing that the behavioural sciences have shown is that behaviour is very much shaped by the context in which people find themselves, more so than we think and often by just as much as by cognitive processes. Making tweaks to the context, to people's 'choice architectures', has become one popular way of changing behaviour (known as "nudging"[12]).

2.2 The study used this new understanding of behaviour to develop a range of interventions. The interventions were delivered through two 'contact points':

- Contact point 1: The online DVSA booking system for the practical driving test, at the stage when the learner chooses a date for their test.

- Contact point 2: A new email communication which was created for the trial and sent by DVSA 2 weeks before the learner driver was scheduled to take their practical driving test.

2.3 The interventions for contact point 1 included the introduction of a 'default', as well as communications encouraging learner drivers to gain more experience before taking their test. These communications aimed to prompt learners to consider lengthening their learning period if they didn't feel ready to take their test. The communications for contact point 2 prompted / reminded learner drivers that they could re-book their test at no extra cost, delaying if they didn't feel ready.

2.4 The design of the interventions followed a systematic approach. Firstly, the customer journey was mapped (from the point of receiving a learner driving licence to the point of taking the practical test) and the 'contact points' were identified. DfT and DVSA then worked together to develop the interventions. The interventions were informed by the behavioural science literature, in particular the new literature that has developed around 'nudging' and behavioural insights[13]. Some secondary analysis of DfT's Transport Choices segmentation was also undertaken to explore the attitudes of people with a

---

[10] See Christmas, S., Campsall, D., Christie, N., Paving the way: putting behaviour change at the heart of a safe system, available from: http://roadsafetyinsight.com/wp-content/uploads/sites/20/2017/06/rsbs-report-2.1.pdf
[11] See, for example, Gigerenzer, G., Selten, R., eds., 2001. Bounded rationality: The adaptive toolbox. MIT Press.
[12] See Thaler, R., Sunstein, C., 2008, Nudge
[13] See for example, The Behavioural Insights Team's MINDSPACE and EAST frameworks. Available from: https://www.instituteforgovernment.org.uk/our-work/better-policy-making/mindspace-behavioural-economics http://www.behaviouralinsights.co.uk/publications/east-four-simple-ways-to-apply-behavioural-insights/

provisional licence[14]. A number of independent experts[15] also provided feedback on the interventions.

2.5 Following the development of a range of intervention options, we conducted exploratory online focus groups with learner drivers. The focus groups were held in May 2016 and gauged learner driver responses to a long list of options for interventions. They helped tailor the messages to ensure maximum engagement with learner drivers.

Contact point 1: Online booking system interventions

2.6 The key behaviour of booking a date for the practical test takes place at contact point 1. Salience of any intervention at this point is likely to be high as this is the time at which the behaviour takes place. This is the point at which learner drivers will find out what test slots are available and once they've chosen their date and time they must make a debit or credit payment for the test. The time of booking is potentially a window of opportunity to attempt to influence behaviour.

2.7 The main intervention at this point was the introduction of a 'default' into the online booking system. Previous research has shown that people often 'go with the flow' of existing 'choice architecture', using default options rather than seeking other less obvious alternatives. A well-known example is in UK pension policy, where the previous system of employees having to actively opt into a pension scheme has been replaced by one of auto-enrolment for some large companies[16].

2.8 At the time of the research, the online booking system gave a blank box when people select the date for their test. Once the learner driver clicks on the box, the date which learner drivers can select for their test defaults to the date on which they are accessing the online system (i.e., they are defaulted to 'now').

2.9 The trial intervention was to set the default date to a particular time in the future. Applicants were still able to change this date, but the default may influence their decision nonetheless. The default may, for example, act as an 'anchor', making a particular timeframe seem like an appropriate choice. The literature suggests that when we make estimates and judgements, we tend to start with some piece of information known to us (the reference point) and then adjust from there[17].

2.10 A number of different default options were trialled through the online booking system (see Annex 1 for a screenshot of the online system)[18]:

- Intervention 1.1: Default to 8 weeks from today's date

---

[14] Department for Transport., 2011. *Transport choices segmentation: final report*. https://www.gov.uk/government/publications/climate-change-and-transport-choices-segmentation-study-final-report

[15] Professor Michèle Belot from the University of Edinburgh, Professor Catherine Hewitt from the University of York, Dr Michael Sanders from the Behavioural Insights Team and Dr Stephanie Walker (Department for Business, Energy & Industrial Strategy).

[16] See The Behavioural Insights Team, 2010, MINDSPACE. Available from: https://www.instituteforgovernment.org.uk/our-work/better-policy-making/mindspace-behavioural-economics

[17] See, for example: Wilson et al, 1996, A new look at anchoring effects: Basic anchoring and its antecedents. Journal of Experimental Psychology: Vol 125(4).

[18] It is worth noting that there was no default before this research – when learner drivers clicked to select a date, they were defaulted to 'now' (i.e., a default of 0 weeks). Also, these different defaults tested through the RCT were also tested with the focus groups. We tested several different defaults with the focus groups to look at whether there would be a difference between each one and to give some indication of whether a default with a further out date would be more effective than a default with a nearer date.

- Intervention 1.2: Default to 12 weeks from today's date

- Intervention 1.3: Default to 16 weeks from today's date

2.11 The 16 week default was also trialled with three different messages. The messaging was delivered through the online booking system above the box which allowed participants to select a date to search for a test (see Annex 1 for a screenshot of the booking system). Each message contained the same final sentence which was a recommendation that the learner drivers give themselves enough time to practise when choosing a date (see Annex 1 for the messaging). Evidence suggests that when people interpret default options as a recommendation the default is more influential[19].

2.12 Each of the three messages also conveyed different information about the potential consequences of not practising enough before taking the practical test. This aspect of the messaging introduced a 'loss frame'. Framing tailors information to match what we know about people's behavioural and cognitive biases and heuristics. For example, people's responses to incentives are shaped by predictable mental shortcuts, such as strongly avoiding losses.

2.13 People attach far greater weight to a potential loss than they attach to a potential equivalent gain[20]. Research has estimated that people tend to be twice as sensitive to loss compared to equivalent gain. If a policy is to nudge individuals, theory predicts that interventions can be more effective if framed as avoiding losses rather than securing equivalent gains[21]. People have even been shown to make sacrifices to avoid a loss, sacrifices which might actually outweigh the cost of the loss (e.g., learner drivers might spend more time and money on driving lessons to ensure they pass their test).

2.14 The final messages shown alongside a 16 week default were:

- Intervention 2: Just over half of learner drivers fail their practical test. We recommend you give yourself enough time to practise when choosing a date.

- Intervention 3: Each test on average costs over £100 (test fee and car hire). Failing could cost you time and money in the long run. We recommend you give yourself enough time to practise when choosing a date.

- Intervention 4: One in five new drivers has an accident within six months of passing their test. Practising more before your test could potentially be lifesaving. We recommend you give yourself enough time to practise when choosing a date.[22]

2.15 The messaging draws on other behavioural insights, alongside loss aversion. For example, Intervention 4 is framed to activate emotions. Emotions have

---

[19] See, for example: Altmann, Falk and Grunewald, 2015. Incentives and Information as Driving Forces of Default Effects. Available from: https://www.cens.uni-bonn.de/team/board/armin-falk/afg-defaults.pdf

[20] Kahneman and Tversky (eds), 2000, Choices, Values, and Frames

[21] Ibid.

[22] The statistic in this message was based on previous research conducted in 2008. (Wells, P., Tong, S., Sexton, B., Grayson, G. and Jones, E. (2008) *Cohort II: A Study of New and Learner Drivers,* London: Department for Transport: http://webarchive.nationalarchives.gov.uk/20120606213228/; http://www.dft.gov.uk/publications/cohort-ii-a-study-of-learner-and-new-drivers/. Subsequent to this trial launching, new evidence emerged which suggests that the figure is lower than one in five, nearer to one in eleven, this has now been published, see https://www.gov.uk/government/publications/evaluation-of-the-impact-of-the-new-driving-test.

been shown to "powerfully shape our actions"[23]. The message is also made to feel personal by linking to a group that the individual booking the test will identify with – learner drivers. Research has shown that messaging is more effective if respondents perceive it to be personal[24].

2.16 Furthermore, the final sentence of each of the messages acts as a 'call to action'. Rather than just describing an issue (failing the test, losing money, risking a crash) and providing no solution, the messaging suggests a course of action – taking enough time to practise – that might help avoid the negative outcome described. Research has shown the importance of including this type of call to action to indicate to people how they might behave to avoid the issue that has been highlighted[25].

Contact point 2: New email interventions

2.17 DVSA statistics suggest that around 1 in 3 learner drivers cancel their original test to rebook for another date. Around half of these (54%) delay their test. On average, those who rebook their test for a later date do so between 2 and 3 weeks before the date of their original test. When candidates do cancel a booking to move it back, they tend on average to move it back by around 4 weeks[26]. The new email communications ('contact point 2') were designed to encourage this type of behaviour if the learner driver who is receiving the email does not feel ready to take their test.

2.18 The design of the email had a number of elements:

- The main text of the email highlighted a loss frame, and connected this with pre-test practice. Two different loss frames were tested (see Annex 1 for the full interventions):

  - Intervention 5: Email 2 weeks before the booked test – time and money messaging

  - Intervention 6: Email 2 weeks before the booked test – statistics on new driver collision rates

- The emails also highlighted a 'social norm'. Intervening with information describing other people's behaviour can act as a powerful spur to 'follow' that behaviour[27]. The email messages stated that around one in three learners rebook their test, the majority of these delaying their test[28].

- The email also aimed to make the behaviour of rebooking easy for the learner driver[29] by providing the necessary web link and highlighting that rebooking is a simple process.

---

[23] See, for example, BIT's MINDSPACE report: https://www.instituteforgovernment.org.uk/our-work/better-policy-making/mindspace-behavioural-economics

[24] See, for example, BIT's EAST framework: http://www.behaviouralinsights.co.uk/publications/east-four-simple-ways-to-apply-behavioural-insights/

[25] See, for example, BIT's EAST framework: http://www.behaviouralinsights.co.uk/publications/east-four-simple-ways-to-apply-behavioural-insights/

[26] It is likely that the figures presented in this paragraph will vary between test-centres, depending on how long waiting times are. This will be controlled for in the analysis of the trial. See Chapter 4 – Methods, for more information.

[27] See, for example, Cialdini, 2003, Crafting Normative Messages to Protect the Environment. Current Directions in Psychological Science.

[28] The statistic drawn upon in the email interventions ("one in three learners rebook their test") draws attention to the fact that a minority of the population rebook their test. It is therefore possible the intervention could have the effect of discouraging learner drivers from rebooking. Further analysis is needed to investigate whether the intervention encouraged or discouraged people from rebooking.

[29] See BIT's EAST framework; http://www.behaviouralinsights.co.uk/publications/east-four-simple-ways-to-apply-behavioural-insights/

- The email was also timely as it was sent at a 'window of opportunity'. As the date of their practical test approaches, learners may welcome a message which gives them the opportunity to re-book their test free of charge in order to give them more time to prepare.
- The email also addressed the learner driver personally. Personalisation has been shown to increase the salience of the message[30].

2.19 The table below gives an overview of all of the interventions tested in the trial (see Annex 1 for the wording of the interventions):

Intervention name	Behavioural insights used
Default 8 weeks	Default; anchoring
Default 12 weeks	Default; anchoring
Default 16 weeks	Default; anchoring
Default 16 weeks + failure rate message	Default; anchoring; loss aversion; call to action; salience (timely communication)
Default 16 weeks + money message	Default; anchoring; loss aversion; call to action; salience (timely communication)
Default 16 weeks + safety message	Default; anchoring; loss aversion; call to action; salience (timely communication)
Email communication 2 weeks before test date – money message	Loss aversion; social norms; call to action; make it easy; make it attractive; salience (timely communication)
Email communication 2 weeks before test date – safety message	Loss aversion; social norms; call to action; make it easy; make it attractive; salience (timely communication)

Limitations of the behavioural interventions

2.20 From a methodological point of view, there were some limitations to testing the interventions outlined above. Firstly, it was not possible to disentangle if a particular aspect of the messaging was effective. For instance, in the case of intervention 5 and 6, it was not be possible to look at the separate behavioural insights – e.g., loss aversion, social norms – or indeed if the whole was more or less than the sum of its parts. However, it was not deemed practical to test each element of the interventions separately as this would necessitate a substantial increase to the trial sample size.

2.21 Furthermore, the design of the online interventions meant it was not possible to disentangle the online messaging from the 16 week default. It might be possible that the 16 week default 'maxes out' what change in the outcome is possible, and thus the trial would not be able to gauge if the messaging was effective or not. On the other hand, if the messaging had been tested in isolation, it would have given the trial more leeway to examine and isolate the impact of each message.

2.22 Finally, in terms of the email interventions, some individuals did not receive the interventions. The mechanism for sending the email was triggered when the

[30] Ibid.

learner driver was 14 days from their test. This meant, for example, that those who booked their test within 2 weeks of logging into the online booking system did not receive the email (see Annex 4 for more information on this). To account for this, a number of different statistical analyses were conducted – one looking at the total sample, and one looking at only those who received the email intervention (with a matched control group). The method used for this analysis is described further in Chapter 3. The analysis for the email interventions therefore requires more careful interpretation.

3. Methods

3.1 This section outlines the methods used for the research, including the following:

- Online focus groups:
- Randomised controlled trial:
    - o Sampling and randomisation
    - o Trial design
    - o Outcome measures
    - o Caveats and limitations of the methodology
- Statistical methods
- Research ethics

Online focus groups

3.2 DfT commissioned BMG Research to carry out exploratory qualitative research to inform the design of the behavioural interventions. Social research, such as qualitative focus groups, can help identify which particular behavioural insights might work in a given situation, with a given group of people, and also how to tailor those insights to the context of the research. Bringing social research together with behavioural insights can generate a richer picture and more powerful interventions.

3.3 The research comprised six online focus groups which took place from 16th May to 21st May 2016. Online focus groups are an innovative method that is particularly useful for examining people's reactions to online messaging and communications (as the medium of the research is the natural setting for the communications). A report detailing the focus groups has been published alongside this report and contains more information on the sample and methodology of the focus groups, as well as strengths and weaknesses of online focus groups[31].

Randomised-controlled trial

Sample and randomisation

3.4 DVSA's online booking system for the practical test was used as the sampling frame for the RCT. Learner drivers were sampled from this system between 5th and 28th July 2016. As learner drivers entered the system, entering their driver

---

[31] See "Annex: Research findings from focus groups with learner and novice drivers – perceptions of and motivations towards driving" published alongside this report.

number, they were randomly allocated (using a random number generator) to either the control group or one of the trial groups. For technical reasons, the trial sampled only first time takers of the driving test. To check that the randomisation was taking place as expected, DVSA generated automatic updates on the sample size and demographic composition per trial group.

3.5 There was a risk that the interventions might put some learner drivers off booking their test[32]. This meant that if the trial excluded those who didn't book a test from the analysis it would risk causing selection bias (i.e. potentially causing non-representative trial results). The sample therefore included both learners who booked their test and learners who logged into the system but did not book a test there and then. This allowed the research to test whether the interventions put learners off booking their test.

3.6 The sampling ran for a limited period of time, to minimise impact on DVSA's online booking system. The decision rule for stopping the trial was that each trial group should include around 2,800 participants who went on to book their test. This stopping rule was informed by the practical constraints of the trial – how long DVSA could host the interventions, ensuring minimum impact on the system – as well as some preliminary exploratory power calculations to check that this sample size would allow the trial to detect any meaningful changes in the outcomes of interest (see Annex 2 for more information on the power calculations). The sample size per trial group is further detailed in the 'Trial design' section below.

Trial design

3.7 Interventions 2 – 6 were trialled as a complete factorial RCT design. One factor was 'contact point 1', the other factor was 'contact point 2'. In practice this meant that some users received only one intervention, whilst some received one intervention from contact point 1 and one intervention from contact point 2. The design allowed the research to make more efficient use of the sample size that was achievable in the short fieldwork time.

3.8 The three default interventions without messaging (Intervention 1) were trialled as arms separate from the main factorial design. However, they ran in parallel with the factorial design, so could be compared to the same control group.

3.9 Tables 1 and 2 below illustrate the cell structure of the trial, including the final effective sample size achieved per cell. The sample size is described as follows: Total sample (total with no test booked / total with a test booking). For example, for Intervention 1.1 there was a total of 3817 individuals allocated to the trial group, of which 1047 did not book a test and 2770 did book a test.

---

[32] For those who logged in but did not book their test immediately, they would have received the same intervention they were assigned to when they booked their test at a later date.

Intervention	Sample (did not book / booked test)
Intervention 1.1: Default to 8 weeks	3817 (1047 / 2770)
Intervention 1.2: Default to 12 weeks	3802 (1041 / 2761)
Intervention 1.3: Default to 16 weeks	3934 (1073 / 2861)

Table 1: Intervention 1 sample sizes

	No email intervention	Intervention 5: New email 2 weeks before, email message 1	Intervention 6: New email 2 weeks before, email message 2
No online intervention	*Control group:* 3842 (1045 / 2797)	3905 (1073 / 2832)	3891 (1090 / 2801)
Intervention 2: Default to 16 weeks + message 1	3833 (989 / 2844)	3918 (1088 / 2830)	3987 (1062 / 2925)
Intervention 3: Default to 16 weeks + message 2	3955 (1082 / 2873)	3845 (1072 / 2773)	3878 (1066 / 2812)
Intervention 4: Default to 16 weeks + message 3	3869 (1038 / 2831)	3834 (1064 / 2770)	3869 (1040 / 2829)

Table 2: Intervention 2-6 cell structure and sample sizes. Total N=58,179; values are sample size (did not book/booked test)

3.10 Statistical analysis was conducted to test whether the demographics of each group were similar to the control. The control group and all of the intervention groups had a similar age and sex composition. The average demographic makeup across the sample (which statistical testing showed was similar across groups) was as follows: 35% females younger than 25, 35% males younger than 25, 16% females 25 and older, 15% males 25 and older.

Outcome measures

3.11 When learner drivers select a date for their test, they can book up to 20 weeks in advance. Between the booking and the date of their test, they have the opportunity to rebook their test. The trial therefore measured two outcomes:

- Initial booking date ("time to first booking" – not accounting for rebooking behaviour).

- Final booking date ("time to last booking" – accounting for rebooking behaviour, so reflective of when people actually took their test).

3.12 The outcome of interest is the length of time (in days) between logging into the online system and taking the practical test. This measure allowed the trial to investigate if the interventions were effective at encouraging learner drivers to

extend their learning period[33]. The sampling period for the trial ran from 5th July 2016 and finished 28th July 2016. This outcome was chosen because the hypothesis is that the interventions will increase the length of time between logging-in and the test being taken. The data from this initial sampling period allowed for us to capture initial booking behaviour.

3.13 However, to account for the fact that trial participants could rebook their test date, a final measurement of the data took place on 21st December 2016, 20 weeks after the initial sampling period had finished.

3.14 The outcome measures described above capture the length of the pre-test learning period between the date on which the learner booked their test and the date they booked their test for / date they took their test. This outcome measures do not capture the actual amount of pre-test practice hours or any information about their post-test driving.

3.15 It was considered a possibility that the interventions might encourage learner drivers to cram in more practice in the same period of time. For example, without the intervention they might have given themselves 8 weeks to practice, with the intervention they might have given themselves 8 weeks to practice but just practiced more hours within that period. The outcome measures described above would not pick up on this type of behaviour change. Furthermore, it was also deemed a possibility that learner drivers who received the intervention could delay their test, but do not do any additional practice. Again, the outcome measure described above would not pick this up.

3.16 For the reasons described above, the study gathered some descriptive data on how much pre-test practice learner drivers perceive they'd done before taking their test. A survey was sent to trial participants after they had taken their practical test (irrespective of whether they passed or failed their test). The survey asked trial participants how many hours of driving lessons they had with a driving instructor and/or with family or friends between booking and taking their test (see Annex 3 for the survey questions). Although this is a subjective measure, and not as rigorous as other possible methods of recording hours of pre-test practice (i.e. diary recording), it still allowed the study to conduct a sense check on whether the interventions encouraged more in-car practice.

3.17 There were limitations with the survey. Firstly, although the survey was sent to all participants who took their test, only a subsample of these trial participants answered the survey (and the results may not be representative of all participants or reflect what the wider group did). Secondly, due to the timing of the survey, the survey only asked about the learning period so we did not collect any information about post-test driving, including if the participant in the trial had been involved in any collisions. We cannot, therefore assess the effect of the interventions on the likelihood of a young driver to have a collision in the first 6 or 12 months after passing their test.

Caveats and limitations of the methodology

3.18 There are a number of caveats and limitations of the approach described above:

---

[33] It is worth noting that almost all intervention studies that aim to improve novice driver safety have been evaluated against outcome measures that serve as a risk factor for collisions and injuries, such as pre-test practice. This is due to the difficulty of obtaining the relevant collision data and the weaknesses of such data.

- Test centre waiting times vary across the year. If test centre waiting times moderate the impact of the interventions, this could negatively impact external validity. For example, average test centre waiting times were longer in July (the month the trial sampled) than at other points during the preceding year. If a long test centre waiting time lessens the impact of the interventions, then the trial could underestimate the efficacy of the interventions as well as be more likely to falsely accept the null hypothesis that the interventions did not have a statistically significant impact on the outcome measures.

- This sampling method is potentially more sensitive to unknown exogenous factors affecting generalisability (i.e. particular events which change behaviour). For example, a particular news story or a change in technology at a given point in time might alter people's behaviour in relation to booking their test. If this happened, the generalisability of the research might be affected without us knowing. Such events would likely effect all trial groups and would therefore not influence the internal validity of the findings. However, external validity (i.e. generalisability of the findings to situations beyond the trial) could be affected. If the sample had been spread across the year, then the external trial generalisability would be less effected by unknown events.

- The sample only includes learner drivers taking their practical test for the first time. Therefore, the results are statistically generalisable only to this group.

- The trial did not measure the final road safety outcome of collisions.

Statistical methods

3.19 Alongside simple analysis using descriptive statistics, the research used a number of statistical methods. The statistical analysis examined whether learners in the intervention groups gave themselves longer between the day they logged into the DVSA online system and booked their test and the date of their booking. The primary explanatory variable is whether a given learner driver was randomised to receive an intervention or not.

3.20 The main outcomes of interest are time variables (i.e. time between booking and taking the test). This "time to event" data is measured in days. The most suitable form of statistical analysis for this type of data is *survival analysis*. The type of regression used was Cox regression (or proportional hazards regression), which is a method for investigating how multiple independent variables influence the time a specified event takes to happen. Unlike ordinary linear regression, Cox regression for survival analysis can effectively handle the 'censoring of observations'. Observations are called censored when it is unclear how long the time to event is. For example, a learner driver might log into the online DVSA practical test booking system, but not book a test for the duration of the study. With survival analysis, everyone contributes to the analysis irrespective of whether they book a test or not. (Chapter 5 details the percentage of participants who were censored.)

3.21 The data also has a multi-level structure. There are two levels: i) the individual learner driver and ii) the test centre they book through. Factors at both levels are likely to influence the outcomes of interest. For example, the individual outcomes at specific test centres are likely to be correlated with one another. In some areas (e.g. London) learner drivers might not have much choice as to which date they select for their test due to long test centre waiting times. In this case, individual level factors are less likely to influence the date of booking and therefore the effectiveness of the interventions (which target individual decision making) will be limited. To account for this, the analysis used a particular type of survival analysis model – a frailty model. This model is similar to adding a random effect into a linear regression model (which is known as a multi-level or hierarchical linear model).

3.22 Furthermore, it is possible that certain other factors impact on the dependent variable. For example, it is known that gender and age affect attitudes towards learning to drive and road safety. Randomisation should ensure that these factors are balanced across each of the trial groups. The statistical analysis controlled for these variables, assessing if there is significant moderation. The variables that were controlled for were: Gender, age, test centre and test centre wait time[34].

3.23 As mentioned in paragraph 3.25, a survey was conducted on the amount of hours practice that trial participants perceived they had taken between booking and taking their test. This outcome was analysed with a simple multiple regression model. The regression model accounted for: Gender, age, test centre and test centre wait time.

Research ethics

3.24 The research adhered to established ethical guidelines for both the focus groups and the RCT. The Department for Transport has an "ethical sensitivity checklist" that was completed before the research was undertaken to ensure adherence to Government Social Research ethical guidance. Furthermore, the research drew upon ESRC guidance on conducting randomised-controlled trials. A Privacy Impact Assessment was also undertaken by DfT and DVSA.

---

[34] The test centre wait time was defined using the DVSA definition of test centre waiting times.

4. Online focus group findings

4.1 The purpose of the online focus groups was to inform the design of the behavioural interventions. They explored two main areas:

i) Learner drivers' views and perceptions of learning to drive.

ii) Learner drivers' views on a list of options of behaviour change interventions.

This section summarises key findings from the focus groups. For more detailed findings and quotes, please see the additional focus group research report published alongside this report. The focus group report gives additional information on how the focus groups influenced the design and selection of the interventions that were tested in the RCT.

Perceptions of learning to drive, brief summary of findings

4.2 The key findings around learner driver perceptions of learning to drive were as follows:

- Convenience, practicality and socialising with friends and family were the key motivations for individuals to learn to drive. A driving licence was considered valuable for gaining independence, allowing the driver to keep to their own timescales and enabling flexibility when applying for job opportunities.

- Males reported that they were most likely to learn to drive for convenience and due to disliking public transport. Females reported that they were motivated to learn to drive for reasons such as pleasure, socialising or for work commitments.

- The main method of learning to drive was with a professional driving instructor. Some participants also practiced with friends or family members. Both formal and informal methods were considered to work well.

- Participants were asked how many hours of practice they expected to have before they book their driving test. The learners tended to think that between 30 to 40 hours is the ideal number of hours to ensure readiness for the test.

- Many participants mentioned time and money as barriers to having more practise before booking their test. There was a general perception that learning to drive as a whole is too expensive.

- Participants did not spontaneously discuss their road safety as a factor that would influence them booking their test. This suggests that road safety is not necessarily at the forefront of learner drivers' minds. The participants views towards booking their driving test were more focused around their level of confidence, their skills and abilities as a driver (which could relate to safety,

but the explicit link to safety was not discussed by participants) and getting as much practice as possible to be confident they had a chance of passing their test.

- When provided with a list of factors that might influence them to book their driving test, the top three were confidence levels, instructor views and hours of practice.
- Learner drivers were asked what factors would influence whether they reschedule their test or not. Most mentioned they do not think about rescheduling their test. However, the key thing that might make them consider rescheduling is not being confident enough that they are ready to pass. Other factors included weather conditions, time of day of the test and their personal availability.
- Participants were asked what factors they thought would influence learner drivers to ensure they are safe on the road. A range of responses were provided, including the ability to recognise mistakes and improve on them in their lessons, awareness of road signs, sticking to recommended speed limits and being aware of other drivers.

Perceptions of the behaviour change interventions, brief summary of findings

Perceptions of the 16 week default

- There were mixed views amongst participants on the default date being pre-selected to 16 weeks ahead. Some suggest this is too far ahead and they would therefore be more likely to scroll back and find a sooner date.
- Those who felt the 16 week default date was about right suggested they would use the time leading up to their test to practice as much as possible.
- Older participants (aged 25+) and females seemed more accepting of the 16 week default compared with younger participants (aged 17 to 24) and males.

Perceptions of the online messaging delivered alongside the 16 week default

- The focus groups explored three sets of messaging. Each set included three to four option message statements.
- Broadly speaking, the focus groups were positive about the messaging, saying it was informative. They were particularly interested in facts and statistics provided as part of the messaging, e.g. on the number of young learners who fail their test.
- The time and money messages (see intervention 3 in Annex 1 for an example) were perceived to be particularly important in outlining the personal consequences of rushing to take a test. Some suggested these messages would encourage them to think about practising more and make them reflect on how confident they really are to take their test. Both males and females suggested the time and money statement should incorporate the monetary value of the cost of the driving test to have the desired effect of influencing learner drivers to delay booking their driving test. Overall learner drivers

suggested this statement would influence them to think twice about how ready they are to book their test.

- Initial reactions to the safety messaging (see Intervention 4 in Annex 1) suggested that participants, particularly females, perceived some of the wording to be 'scary', 'hard-hitting' and 'fearsome'. However, on reflection many suggested that the safety statements would be the most influential of all of the messages to make individuals think about their level of confidence and readiness to book their practical test. A few participants disliked the use of road deaths statistics and suggested this should be replaced with statistics on accident rates associated with young learners, to be less scary and off-putting. Participants suggested that this statement had the biggest impact as it made them reflect on the importance of road safety, something they might not otherwise have done.

Perceptions of the email communications

- Overall the email communications (see Interventions 5 and 6 in Annex 1) were well received. Most participants suggested they liked the idea of receiving an email before their practical test and said that they found the content beneficial. In particular, they liked that the email reassured them that they could rebook without losing their fee.

- The tone of the email was described as clear, friendly and professional. There were, however, mixed views on the email subject. Although, most suggested it was inviting and felt personal to them.

- Suggested improvements to the email communication included adding facts and statistics, for example on pass or fail rates or information on accidents associated with young learners.

- Participants were asked when they thought the most effective time might be to send the email to learners. Some suggested four weeks may be too far ahead and individuals may not know if they are ready by then. Thus, most suggested receiving the email two weeks before the test was appropriate as this gave them enough time to evaluate their level of readiness and rebook if necessary, without a fee.

- Overall most participants suggested the email communication was influential in making them think about rebooking. They agreed that the email would make them think twice about how ready they were to take their test and they were positive about the ability to reschedule without a fee.

5. Trial results

5.1　This section describes the results from the randomised controlled trial. Full results from the statistical analysis can be found in Annex 4. The randomised controlled trial examined the impact that the interventions had on several outcomes, including time to first booking and time to last booking (see paragraph 3.20 for a definition of these outcomes).

5.2　The section is structured as follows:

- First test booking behaviour, by intervention groups

- Last test booking behaviour, by intervention groups

- Impact of personal characteristics, test centre location and waiting time

- The effect of the treatment on number of hours practised

The effect of the treatment on first test booking behaviour, by intervention groups

5.3　Results are available for 58,179 candidates. Of these, 42,309 candidates booked a test. Across all groups the shortest time to booking was 1 day (i.e. test booked for the following morning). The longest time to booking was 157 days (around 5 months out).

5.4　The bars in Chart 1 display findings from a univariate analysis of the average time to initial booking across each of the trial groups. It indicates that some of the interventions were successful at increasing the time between logging into the online booking system and the date of the first test booking[35]. These results do not account for rebooking behaviour, i.e., learner drivers could potentially rebook their test after receiving the online defaults and/or messaging.

5.5　The univariate analysis shows that those who were randomised to receive the safety message with the 16 week default gave themselves on average an extra 2.5 days to practice for their test, compared to the control group. The chart also shows that the defaults without messaging seem to be effective. The 16 week default seems more effective than the 8 and 12 week defaults. Chart 1 also indicates that the default works better with the messaging[36].

5.6　Statistically significant results from the survival analysis are shown on Chart 1 below with shaded bars[37]. The multivariate analysis indicated that the 12 and

---

[35] Note that as the email interventions were delivered to candidates after their first booking, they are not analysed in this section.

[36] This statement is based on the univariate analysis, not statistical testing.

[37] A caveat is needed for this analysis: The Cox regression model doesn't fit a single parameter as a constant or intercept term, rather a baseline hazard function that is a function of time. Therefore, it is hard to generate averages which have been adjusted for the statistical analysis. Technically speaking, therefore, the statistical significance applies to the hazard functions which can be viewed in Annex 4, not the averages presented in Chart 1. However, for ease of presentation the statistical significance is indicated on the chart.

16 week default and all of the messaging statements had a statistically significant impact on the outcome measure compared to the control group (see Annex 4 for the statistical outputs, including hazard ratios). The clinical significance[38] of the effect size is discussed in the next chapter.

5.7 Statistical testing also indicated that none of the interventions were associated with candidates failing to book their test, i.e. no interventions stopped candidates from booking.

Average time to first test booking, by intervention group
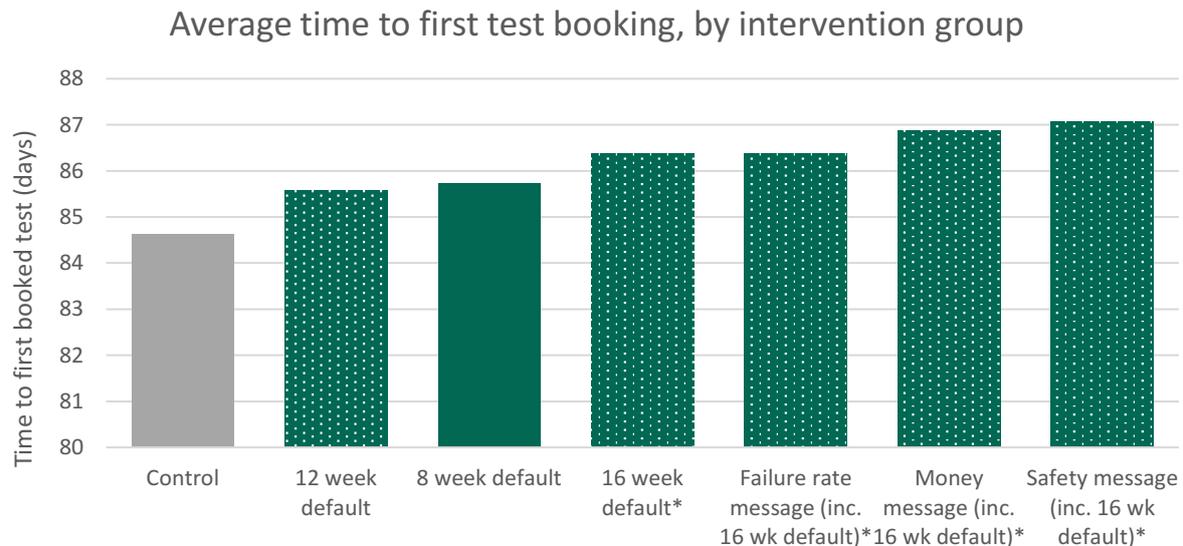
Chart 1: Average time to first test booking, by intervention groups. N=58,179

The effect of the treatment on last test booking behaviour, by intervention groups

5.8 The last test booked data was extracted 20 weeks after the first test booking data, to capture the effect that rebooking behaviour had on the amount of time that trial participants had to practice before their test. The analysis below examines the impact that the email interventions had, as these interventions targeted rebooking behaviour.

5.9 The bars in Chart 2 display findings from the univariate analysis, which shows that when accounting for rebooking behaviour the average time to test for each trial group is slightly longer (compared to first booking behaviour), including the control.

5.10 Statistical significance is reported differently, compared to the previous section. It is not reported on the chart, but rather is reported in Table 3 below (see Annex 4 for full statistical outputs). All participants who were allocated to receive one of the online interventions did receive those interventions. However, not all participants who were allocated to receive the email interventions received them. The emails were sent to participants when they were 14 days before their booked test. This means that some learner drivers allocated to receive the email interventions did not receive them, for example

---

[38] By clinical significance this report means the practical importance of a treatment effect (so, practical rather than statistical significance) - whether the treatment is likely to have a palpable and noticeable effect on road safety, for example.

participants who booked their initial test for within 14 days of first logging into the online booking system. The data for the email interventions is therefore analysed in two ways: i) total sample including all trial participants, those who received the email and those who did not receive the email, ii) a sample of only trial participants who actually received the email. This involved analysing the email intervention groups to identify who actually received the email and then using this information to create a matched control group for purposes of comparability – see Annex 4 for more detail on this. Chart 2 shows averages for the total sample, including those who were allocated to receive the emails but didn't receive them.

5.11 The key findings are as follows:

- ***Online interventions:*** Although the univariate analysis displayed in Chart 2 indicates that learners who receive the online messaging take slightly more time to their test, Table 3 shows that there are no statistically significant effects for any of the online interventions. It seems that when rebooking behaviour is accounted for, the initial nudge isn't as effective. This suggests behaviour change interventions of the type tested in this trial work better if they target one-off decisions that aren't likely to be revisited. If people can revisit their initial behaviour they might correct for the effect that the initial nudge had, i.e. in the case of this trial learner drivers might rebook their test for slightly earlier or not delay their test by quite as much[39]. The interventions therefore could be said to change the expectations of some learner drivers when they book their test, however they do not change the point at which the learner drivers feel ready to take their test and it is likely to be the feeling of readiness that affects the final test date (rather than expectations about when they'll feel ready).

- ***Email interventions:*** The univariate analysis in Chart 2 shows that participants who received the email interventions had longer times to last test booked compared to the control group. The univariate analysis also indicates that the safety email is the more effective of the two emails. This analysis shows that compared to the control group, learners who received the safety email gave themselves on average roughly 2 days longer to practice for their test. The multivariate analysis summarised in Table 3 shows that for the total sample (including both those who received and those who didn't receive the email) neither email intervention is statistically significant at the 5% significance level (although, the safety email is significant at the 10% significance level). However, when the multivariate analysis looks only at those participants who received the email interventions (with a matched control group), the safety email is significant at the 1% significance level. The clinical significance of this effect size is discussed in the next chapter.

---

[39] Note: This is not to say that we suggest it may be better not to allow rebooking – having this choice is important.

Average time to last test booked, by intervention group
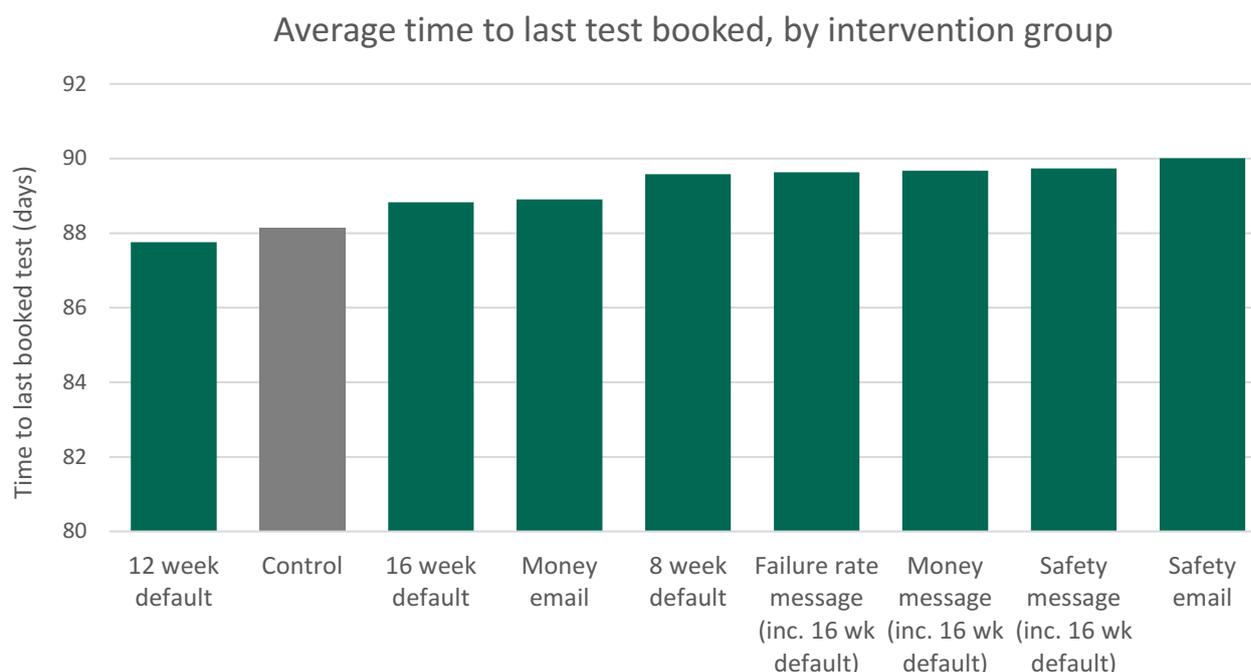
Chart 2: Average time to last test booked, by intervention group. N=58,179

Intervention	Total sample Statistical significance (p-value)	Received email treatment sample (with matched control group) Statistical significance (p-value)
Default 8 weeks	0.32	
Default 12 weeks	0.78	
Default 16 weeks	0.85	
Default 16 weeks + failure rate message	0.28	
Default 16 weeks + money message	0.10	
Default 16 weeks + safety message	0.24	
Money email	0.82	0.32
Safety email	0.09	0.0065**

Table 3: Statistical significance for total sample and received email treatment sample, ** = statistically significant at 1% significance level

Impact of personal characteristics, test centre location and waiting time

5.12 This section first looks at whether gender, age and the waiting time of the test centre that candidates booked through affected time to booking behaviour. It then examines whether personal characteristics of the candidate influence how effective each of the interventions were.

5.13 Chart 4 below shows the average days to last test booking by different gender and age groups. It shows that on average males tend to book earlier than females and that older females tend to book later than younger females. The starkest difference is between males under 25 and females over 25 – there is over 2 weeks difference (16 days) between the two groups. This is an interesting finding given we know that young males are the highest risk group in terms of collision risk[40]. Statistical analysis confirmed that candidate booking behaviour is statistically significantly linked to personal characteristics such as gender, age and the combination of gender and age.

Average days to last booked test, by demographic group
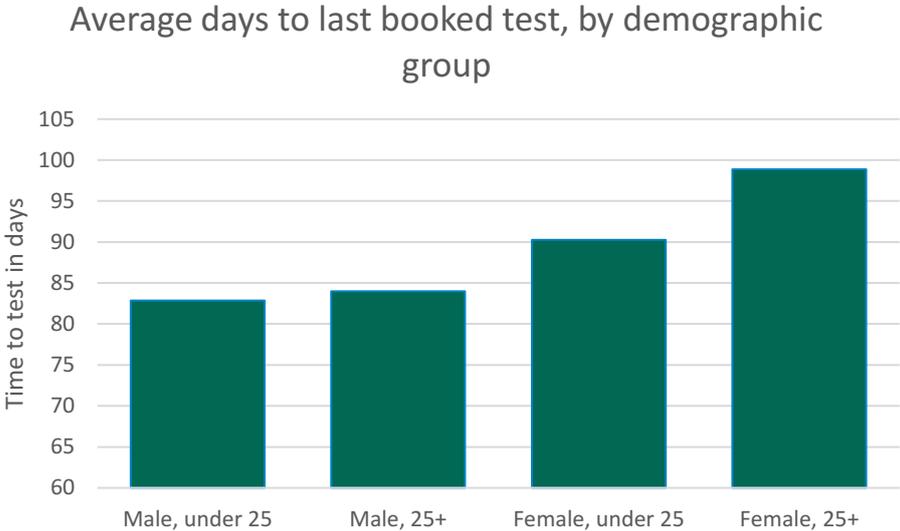
Chart 4: Average days to last booked test, by demographic group. N=58,179

5.14 The test centre and test centre waiting time are also significantly linked to time to booking (see Annex 4 for outputs of the frailty models which show this). Chart 5 below shows that the longer the test centre wait time, after a wait time of about 10 weeks, the longer the time to test.

---

[40] DfT Road Safety Research Report No. 81 Cohort II: A Study of Learner and New Drivers Volume 1 – Main Report, 2008; and, TRL Transforming the practical driving test, available from: https://www.gov.uk/government/publications/evaluation-of-the-impact-of-the-new-driving-test

Average days to last booking by test centre waiting
time (weeks) at first visit
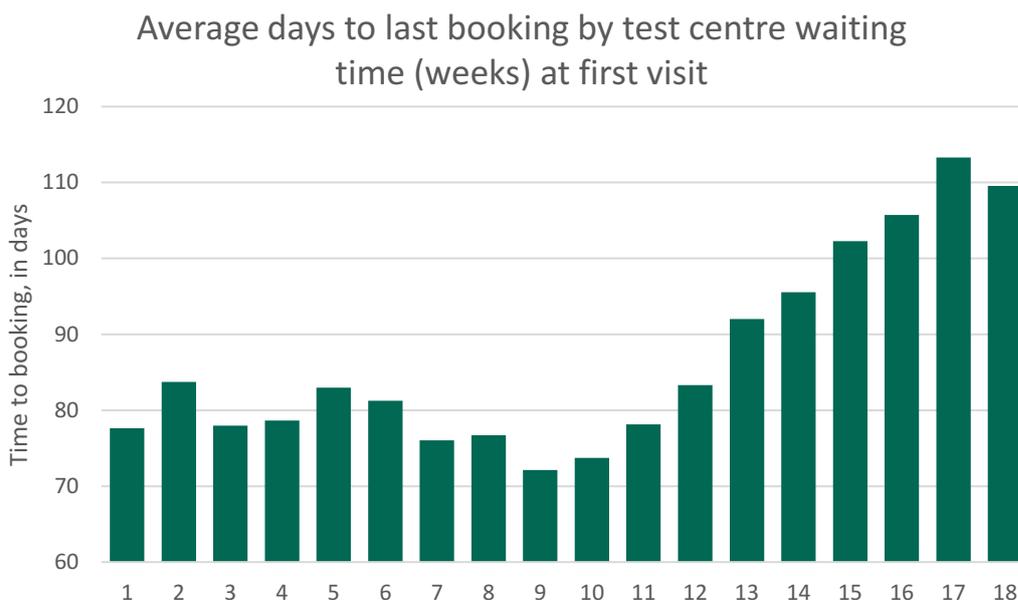
Chart 5: Average days to last booked test, by test centre waiting time (weeks) at first visit. N=58,179

5.15 This analysis is suggestive that there may potentially be unintended consequences of initiatives which try to reduce test centre waiting times. If learner drivers practice more hours the longer their time between booking and taking their test, then it is plausible that if test centre waiting times are reduced, then they will practice less and be less safe on road. However, this is hypothetical and would need confirming through further research.

5.16 ***Impact of interventions, by gender and age:*** Exploratory statistical analysis was conducted on the impact of the interventions, by gender and age. The analysis was conducted by running separate statistical models for each subgroup. The statistical analysis suggests that the interventions had a different impact across demographic groups. Table 4 below summarises these findings, looking at time to last booking[41]. The analysis suggests that the interventions are not effective with the most 'at risk' group, young men, and that they are more effective with lower risk groups such as females. However, as this analysis is exploratory it should be taken only as indicative – further research is required on how simple behavioural interventions influence different groups.

5.17 *Defaults:* Table 4 shows that the 8 week default seemed effective with young and older females, but not males.

5.18 *Online messaging:* Table 4 shows that the online messaging was not effective with males or older females. However, the analysis does suggest that the online messaging is effective with younger females, perhaps because they are less likely to revisit their decision compared to males. Further analysis is needed to examine this in more depth.

5.19 *Email interventions:* The email seem to be effective with females over 25. Further analysis is needed to look at this in more depth.

---

[41] The subgroup analysis for the email interventions used the total sample, rather than just those who received the intervention. Ideally this analysis would also have been conducted for just those who received the email interventions (with a matched control). However this was not possible due to time and resource limitations.

Demographic group	Impact of interventions on last test booking behaviour
Males under 25	None of the interventions had a significant effect
Males 25 and over	None of the interventions had a significant effect
Females under 25	The online messages and the 8 week default were associated with longer times to test booking.
Females 25 and over	The 8 week default was associated with longer times to test booking, as were both email interventions.

Table 4: Impact of interventions on last test booking behaviour, by gender and age

The effect of the treatment on the number of hours practised

5.20 Alongside the key outcome measures analysed above, a survey was also conducted with participants after they had taken their test (irrespective of whether they passed or failed their test). In total, 5,431 candidates responded to the survey. This is a response rate of around 13% of those who booked their test (the response rate was statistically similar across groups)[42].

5.21 The survey suggests that for the control group the average amount of practise between booking and taking a test was 31.4 hours.

5.22 The analysis suggests that none of the interventions were significantly associated with increased hours of practice. However, other factors were linked to the amount of practice taken (see Annex 4 for the full statistical model).

- The wait time at the test centre. The longer the wait time, the more practice undertaken.

- Candidate gender. Females take more practice.

5.23 A caveat is needed for this analysis. The sample size is much smaller than for the primary analysis of time to booking behaviour. Given the limitations of the analysis, further research is required on the link between simple behavioural interventions, such as tested in this trial, and the number of hours practiced.

---

[42] Ideally this analysis would have looked for non-response bias (by comparing those who responded to the overall test taking population, in terms of key demographics) to give enhanced confidence about generalising the findings. However, due to time and resource limitations of the research this was not possible. The analysis of the survey should therefore be taken as indicative.

6. Discussion

6.1 This research set out to examine whether low cost behavioural interventions, designed using insights from the behavioural sciences, could be effective at encouraging learner drivers to give themselves more time before their test to practice.

6.2 The research combined online focus groups, to explore learner drivers' perceptions of driving and the proposed behaviour change interventions, with a randomised controlled trial to examine the impact that the behavioural interventions had across several outcomes. Each aspect of the work has resulted in valuable insights that could be used to inform initiatives to improve road safety amongst learner and novice drivers.

6.3 The research has, however, been limited in time and scope. There are areas which would benefit from further research and analysis. More detailed research, for example, on rebooking behaviour is needed in the future. Furthermore, the trial was also limited in terms of being able to look for relationships between behavioural interventions and number of hours practiced and road safety outcomes – more research is also required in these areas.

6.4 This section brings together all of the key lessons learned, reflects on clinical significance, discusses limitations of the research and describes what research might be conducted in the future.

Key learnings from the focus groups

6.5 Chapter 4 outlined the key findings from the focus groups. For the purpose of this discussion, two findings standout:

- Road safety was not spontaneously considered as an important factor when thinking about readiness to take the practical test or drive on roads. Confidence and driving ability were the main factors that were consciously considered (although these factors might link to road safety, research participants did not draw the link themselves). This suggests that road safety is not necessarily at the forefront of learner drivers' minds.

- However, when prompted to think about road safety, the focus group participants found facts and figures on novice driver safety unexpected and found the statistics on roads deaths shocking and disturbing. They suggested that having an idea of facts and figures around road safety would make them consider more carefully whether they had practiced enough before taking their test and their readiness for independent driving on roads.

Key learnings from the randomised controlled trial

6.6   Chapter 5 outlined the key findings from the RCT. The standout findings are summarised below:

- Most of the online interventions were linked to time to test at the point of first booking (i.e. not accounting for rebooking behaviour). The 16 week default and all of the online messages were associated with longer times to first booking. Trial participants who received the 16 week default with the safety message booked, on average, 2.5 days later compared to the control group. It is not clear, however, whether they booked later to give themselves longer to practise.

- Rebooking behaviour meant that the online messages and defaults did not significantly affect last test booking behaviour. Trial participants who received the online messages and defaults had different rebooking behaviour compared to the control group, which led to a closing of the gap between the trial groups and control group in terms of time to booking. Analysis has yet to explore in-depth how rebooking behaviour was different across groups (see 'Limitations and future research' section below),

- For the last test booked (accounting for rebooking behaviour), when we look at the total sample none of the interventions were statistically linked with increased times to booking. However, not all trial participants allocated to receive the emails actually received the interventions (for reasons described in Chapter 5 and Annex 4). When looking at only those people who received the interventions (with a matched control group) the safety email intervention was linked with increased time to last booking.

- Other factors were found to be linked to increased time to last booking:

    - The test centre chosen and the average waiting time. The greater the wait time, the longer the time to booking.

    - The gender and age of the candidate. Males tended to book earlier, as did younger learner drivers.

- In terms of number of hours practiced, exploratory statistical analysis suggests that none of the interventions were linked to taking more practice before the test. The same statistical analysis suggested that lengthier test centre wait time resulted in more practice, and that females tend to practice more hours than males.

- The fact that test centre waiting times influence both time to booking and the number of hours practiced suggests that a multi-level systems approach is needed to encourage behaviour change. Initiatives that aim to encourage learner drivers to practice more need to consider both social and structural factors alongside psychological factors and individual decision-making.

- The RCT validated findings from the online focus groups which suggested that the safety messaging would be the most impactful of the various types of messaging. The safety messaging comes out as the most impactful across both online and email communications compared to the control group (in terms time to booking). This is potentially contrary to previous evidence around the

use of fear and threat. However, it is worth highlighting a few points: the focus groups suggested that some safety messaging (e.g. around road deaths) was perceived as fear and threat messaging – these messages were not used in the RCT. The messages that were tested in the RCT (around collisions) were perceived to be less fear inducing than statistics on road deaths. Furthermore, the messaging included a 'call to action' to make it practical.

- Previous research has shown that standard (non-behavioural) interventions such as providing information to people about risks have not proved that effective[43]. This research involved combining behavioural interventions with standard interventions, for example by supplying people with information about risks but framing the messaging using theory from the behavioural sciences. The results suggests that a combination of standard intervention tools (providing information) and behavioural intervention tools (e.g., loss aversion framing) may be more effective than relying solely on standard interventions.

- Exploratory analysis suggested that the different demographic groups reacted differently to the interventions – certain interventions could be effective with one group but not another. For example, the email interventions seemed more effective older females, leading to longer times to booking. The interventions seemed to have less effect on males, the higher risk group in terms of road safety outcomes. Due to limitations, this analysis should be taken as indicative – further work is needed to examine the road safety impact that behavioural interventions have across different groups.

Statistical versus clinical significance

6.1 As described above, some of the interventions were found to be related to the outcomes of interest. However, due to limitations of the trial, we are still unsure about clinical significance, i.e. whether the effect size could have a noticeable effect on road safety. To give a definitive answer to this question would require further research which seeks to link the behavioural interventions tested in this trial directly to road safety outcomes, such as collisions or Killed or Seriously Injured statistics.

6.2 In the absence of evidence on safety outcomes, a few points are worth considering:

- Previous studies indicate that learner drivers in the UK typically practice for between 40 – 50 hours before taking their test. Other research has indicated that 100-120 hours is the optimal amount of practice[44]. Given these figures, the effect sizes found in this trial are not likely to be of large clinical significance.

- Having said this, behavioural interventions like those tested in this trial should not operate in a vacuum, as a single solution to the problem of road safety. They should operate together with a range of other initiatives, each contributing to the creation of a safer road network.

---

[43] See, for example: https://www.gov.uk/government/uploads/system/uploads/attachment_data/file/609828/interventions-to-increase-young-and-novice-driver-safety.pdf
[44] Ibid.

- Furthermore, if road safety is a system phenomenon rather than just being about risky people (as the main road safety perspectives argue[45]), then it is possible the effect size is still clinically meaningful. Although behavioural interventions such as tested through this trial might only offer a small benefit to each individual who receives them, arguments from the public health field[46] suggest that shifting the population distribution of a risk factor, even if by a small amount, can be an effective preventative measure. Because individual behaviour takes place within a wider social system, the societal effects of a small distributional shift could potentially be large.

What do these findings mean for initiatives which aim to improve road safety amongst learner and novice drivers?

- It seems that learner drivers give little consideration to safety. However, when they are prompted to think about safety through facts and figures, it seems to play a more salient role in their decision-making. Future initiatives to improve the safety of learner / novice drivers might seek to play on this, to make safety more salient in decision making. It is worth noting that the safety messaging used in this trial was combined with a 'call to action'. Rather than just describing an issue and providing no solution, the messaging suggests a course of action – taking enough time to practise – that might help avoid the negative outcome described. Research has shown that such calls to action are effective at helping changing behaviour.

- One particular behavioural insight drawn upon across the messaging was 'loss aversion'. Loss aversion is the insight that people attach a greater weight to potential loss compared to an equivalent gain. If communications frame outcomes as potential losses from bad behaviour compared to gains from good behaviour, they are more likely to be effective at changing behaviour. The trial indicates that loss framing could potentially be effective for influencing learner drivers and could continue to be used to encourage road safety.

- The trial found that some behavioural interventions (the online interventions) were successful at influencing a one-off decision (first booking behaviour), but when people had the opportunity to revisit that decision (through rebooking) the nudge didn't seem as effective. This has wider implications for the behavioural insights field: if people have the opportunity to revisit their choice, decision, or behaviour easily, then they are likely to need a series of nudges rather than a one-off nudge.

- The trial indicates that low cost behavioural interventions could potentially have some effect, but likely not a large clinically significant effect. Such behavioural interventions would ideally therefore be used in conjunction with other initiatives and programmes, rather than as a stand-alone intervention. They should be applied in an integrated and holistic way, helping reinforce the process of behaviour change.

- Linked to the point above, the findings suggest that a multi-level approach is needed. Individual decision-making is not the whole picture – social and

[45] See Christmas, S., Campsall, D., Christie, N., Paving the way: putting behaviour change at the heart of a safe system, available from: http://roadsafetyinsight.com/wp-content/uploads/sites/20/2017/06/rsbs-report-2.1.pdf

[46] See, for example: Rose, G., 1985. Sick individuals and sick populations. International Journal of Epidemiology;14(1):32-8

structural factors, such as test centre waiting times, also had a large influence on the outcomes explored in this trial. One implication of this is that a multi-disciplinary safer systems approach, in which initiatives attempt to change behaviour through multiple channels and levels, would be the most effective type of approach. More research is needed in this area to determine what would work in terms of influencing behaviours through social and structural factors.

- The trial found that longer test centre waiting times are related to longer time to booking and also the amount of hours that learner drivers practice for. This indicates that any initiatives which seek to reduce test centre waiting times should consider how to mitigate potential unintended consequences of such initiatives.

Limitations and future research

6.1 Whilst this trial has tested a variety of interventions on a large scale, with a large sample size, it is important that a number of limitations of the study are recognised. Linked to these limitations are areas which could be considered for future research:

6.2 **Time of year of the trial:** The time of year during which the trial took place had long waiting times compared to the average for the previous year. It is therefore possible that the interventions may be more effective at other times of year. Future research might seek to explore similar behaviour interventions at other times of year. Future analysis of the data generated for this trial might also investigate how test centre waiting times interacted with the different interventions.

6.3 **Rebooking behaviour:** More detailed investigation of rebooking behaviour is needed, i.e., how rebooking behaviour differed between trial groups and statistical analysis of the demographic makeup of those who rebooked their tests earlier/later. This could help design behavioural interventions which more effectively target rebooking behaviour.

6.4 **Structural versus individual factors:** The trial has demonstrated a link between test centres and test centre waiting times and the risk factors of interest (time to booking, hours practiced). Future analysis could seek to investigate how much variation in the outcomes of interest is explained by structural factors versus individual factors (i.e. through generating statistics such as the 'variance partition coefficient' or the 'intraunit correlation coefficient'). This would help identify areas for policy focus.

Acknowledgements

Annex 1 Behavioural interventions

Intervention 1 introduced a default into the booking system. This intervention does not take people's choice away, as people still get to choose the date they want. It does mean that people might have to scroll backward through test dates rather than forward. Three defaults were trialled:

Within the figure:

GOV.UK   **Book your driving test**

✓ Type of test   ✓ Licence details   **3 Test date**   4 Test centre   5 Test time   6 Your details   7 Card details   8 Confirm & pay   9 Done!

Test date - Car test

Enter a date you would like

All suitable appointments around your chosen date will be shown, including any cancellations

DD/MM/YY

1.1: Default to 8 weeks
1.2: Default to 12 weeks
1.3: Default to 16 weeks

Note: The online messaging (intervention 2-4) appeared just above this box

Instructor's reference number (optional)

▷ What is this?

eg 123456 or 12345

ℹ Your instructor's reference number helps ensure they're not double-booked when accompanying you.

Continue ❯

The 16 week default was also tested with the following messages (as well as in isolation):

Intervention 2: 16 week default + information on test failure rates

Just over half of learner drivers fail their practical test. We recommend you give yourself enough time to practise when choosing a date.

Intervention 3: 16 week default + information on average price of test

Each test on average costs over £100 (test fee and car hire). Failing could cost you time and money in the long run. We recommend you give yourself enough time to practise when choosing a date.

Intervention 4: 16 week default + information on novice driver collision rate

One in five new drivers has an accident within six months of passing their test. Practising more before your test could potentially be lifesaving. We recommend you give yourself enough time to practise when choosing a date.

Two email interventions with different wording were trialled:

Intervention 5: Email 2 weeks beforehand – time and money messaging

Email Subject: Important information about your driving test – are you ready?

Dear [Name],

Your driving test is coming up very soon, on xx/xx/xxxx.

Each test on average costs over £100 (test fee and car hire). Taking your test when you do not feel fully prepared could cost you time and money in the long run. You can rebook your test for a later date before XX/XX/XXXX without losing your fee. One in three learner drivers reschedule their test, the majority of these delaying their test.

To rebook visit https://www.gov.uk/change-date-practical-driving-test and follow the simple steps.

The details of your current test booking are provided at the bottom of this email. I should like to take this opportunity to wish you safe driving for life.

Yours sincerely,

Director of Operations

Intervention 6: Email 2 weeks beforehand – statistics on new driver collision rate

Email Subject: Important information about your driving test – are you ready?

Dear [Name],

Your driving test is coming up very soon, on xx/xx/xxxx.

One in five new drivers has an accident within six months of passing their test.

Practising more before your test could potentially be lifesaving. If you do not feel fully prepared, you can rebook your test for a later date before XX/XX/XXXX without losing your fee. One in three learner drivers reschedule their test, the majority of these delaying their test.

To rebook visit https://www.gov.uk/change-date-practical-driving-test and follow the simple steps.

The details of your current test booking are provided at the bottom of this email. I should like to take this opportunity to wish you safe driving for life.

Yours sincerely,

Director of Operations

Annex 2 Power calculations

Exploratory power calculations were conducted before the RCT launched. The power calculations started with an expected sample size and worked backwards from that to see what effect size would be detectable given a range of different statistical powers (.80 to .95). Power was calculated with the GPower software. A caveat is needed: The power calculations were conducted for a factorial ANOVA (rather than Cox regression survival analysis, which was the final statistical test). The analysis below only presents power calculations for interventions that were included in the factorial design (Interventions 2-6).

Power calculations for factorial design: The power calculations for the factorial design assumed a sample size of 42,000 learner drivers who booked a test. Other assumptions for the factorial ANOVA power calculations are outlined below:
- Contact point 1 – Online communication: 4 groups = 3 degrees of freedom
- Contact point 2 – Email communication: 3 groups = 2 degrees of freedom
- Looking for interactions between contact point 1 and contact point 2: 6 degrees of freedom
- Analysis of data from the DVSA showed that the standard deviation of the outcome measure was 23 days.
- Significance level: This refers to the probability of incorrectly rejecting the null hypothesis. Standard practice is to set this value at .05 (two-tail) – i.e. a one in twenty possibility of incorrectly rejecting the null hypothesis.
- Power: The power of a statistical hypothesis test is the probability that it correctly rejects the null hypothesis when the alternative hypothesis is true.

Hypotheses:
- Null hypothesis 1 (H0.1): There is no difference in the time between logging into the online system and taking the practical test between the groups
- Null hypothesis 2 (H0.2): There are no interactions between the independent variables in the time between logging into the online system and taking the practical test.
- Alternative hypothesis 1 (H1): There is a difference between the control group and the intervention group(s) in the time between logging into the online system and taking the practical test.
- Alternative hypothesis 2 (H2): There is an interaction effect between the online system communications and the email communications in the time between logging into the online system and taking the practical test.

The tables below detail the minimum detectable effect size given the above assumptions.

Power (statistical significance in brackets)	Indicative Main Effect Size of Online Communication	Indicative Main Effect Size of Email	Interaction Effect Size	Sample size
.80 (.05)	~0.016	~0.015	~0.018	42000
.85 (.05)	~0.017	~0.016	~0.019	42000
.90 (.05)	~0.018	~0.017	~0.020	42000
.95 (.05)	~0.020	~0.019	~0.022	42000

The largest minimum detectable size is .022 (with a power of .95). Using the predicted standard deviation of the outcome variable, we can say that this effect size is equal to 1 day difference in the means between the control group and the intervention groups. This is adequate for the purposes of this trial.

Given the lack of evidence of applying behavioural insights within this type of context and to this type of decision, it is difficult to estimate what effect size the interventions will have. However, using effect size conventions, a small effect size (.10) would equal around 5 days, a medium effect size (.25) around 11 days and large effect size (.40) around 20 days.

Annex 3 Survey email and questions

EMAIL TO TRIAL PARTICIPANTS AFTER THEY HAD TAKEN THEIR TEST
Subject: Three quick questions on learning to drive

Dear [name],

You have recently taken your driving test.

We are undertaking some research on learner drivers' experiences. Please could you follow the link below and answer **three simple questions on learning to drive**? Your answers could help improve road safety for everyone.

[Survey monkey link]

Your help is greatly appreciated.

Yours sincerely,
Director of Operations, Driver and Vehicle Standards Agency

SURVEY QUESTIONS
When did you book your driving test?
Don't worry if you don't know the exact date. Please enter your estimated date below in the following format DD/MM/YY.

Since you originally booked your test, how many hours of driving lessons have you had with a driving instructor?
Please enter the number of hours below.

Since you originally booked your test, how many hours of driving lessons have you had with family or friends?
Please enter the number of hours below.

Thank you for taking part in this survey!

Annex 4 Statistical analysis, model outputs

First test booking behaviour: Cox regression survival analysis, model outputs

Variable	Coefficient	se(coef)	se2	Chisq	DF	p	
Test centre wait time	-0.04702	0.002467	0.002449	363.34	1	<0.001	***
Intervention1_8	-0.03735	0.023443	0.023438	2.54	1	0.11	
Intervention1_12	-0.1039	0.023512	0.023506	19.53	1	<0.001	***
Intervention1_16	-0.17469	0.023208	0.023202	56.66	1	<0.001	***
Intervention2	-0.15139	0.015546	0.015543	94.84	1	<0.001	***
Intervention3	-0.1841	0.015626	0.015623	138.8	1	<0.001	***
Intervention4	-0.20365	0.015681	0.015677	168.68	1	<0.001	***
Intervention5_Group	-0.03693	0.013453	0.013451	7.54	1	0.006	**
Intervention6_Group	-0.01531	0.013429	0.013426	1.3	1	0.25	
GenderMale	0.136315	0.009872	0.009869	190.68	1	<0.001	***
Age	-0.00333	0.000634	0.000634	27.55	1	<0.001	***
Frailty (test centre)				47960.24	323	<0.001	***

Last test booking behaviour: Model outputs for Cox regression survival analysis (total sample, including those who received and those who didn't receive the email interventions)

Variable	Coefficient	se(coef)	se2	Chisq	DF	p	
Test centre wait time	-0.01282	0.002534	0.002499	25.57	1	<0.001	***
Intervention1_8	-0.02342	0.023454	0.023447	1	1	0.32	
Intervention1_12	0.00644	0.023491	0.023481	0.08	1	0.78	
Intervention1_16	-0.00442	0.023186	0.02318	0.04	1	0.85	
Intervention2	-0.01665	0.015505	0.015502	1.15	1	0.28	
Intervention3	-0.02592	0.015567	0.015562	2.77	1	0.10	.
Intervention4	-0.01824	0.015564	0.01556	1.37	1	0.24	
Intervention5_Group	-0.00309	0.013424	0.013422	0.05	1	0.82	
Intervention6_Group	-0.02284	0.013378	0.013376	2.91	1	0.09	.
GenderMale	0.272815	0.009897	0.009894	759.81	1	<0.001	***
Age	-0.00709	0.000641	0.000641	122.46	1	<0.001	***
Frailty (test centre)				12498.34	320.1	<0.001	***

Last test booking behaviour: Model outputs for Cox regression survival analysis, including only trial participants who received the email interventions (with a matched control group)

Not all participants allocated to receive the email received the interventions. For example, learner drivers who booked within two weeks of taking their test did not receive the email, not did learner drivers who rebooked their test and were never 14 days from their original test booking. A knowledge of which candidates did not receive the email interventions in the trial groups allowed us to match the control group to the intervention group[47].

Analysis of who in the email intervention trial groups received the emails allowed us to create a matched control group with the following logic:

- If a candidate booked a test 14 or more days after the booking date AND they didn't cancel/rebook that test AND the test date was not after 26/12/2016, they were included in the matched control group.
- If a candidate booked a test 14 or more days after the booking date AND the test was not after 26/12/2016 AND they cancelled/rebooked less than 14 days before the test, they were included in the matched control group.
- If a candidate rebooked a test within 14 days booking date AND NOT within 14 days of the original booking date (before or after), they were included in the matched control group.

We can be confident of a good match between the control and intervention groups. Analysis showed a close match between the control group and email intervention groups in terms of the percentage of people who were knocked out of the analysis due to the process outlined above.

Variable	Coefficient	se(coef)	se2	Chisq	DF	p	
Test centre wait time	-0.01108	0.002849	0.0028	15.11	1	<0.001	***
Intervention1_8	-0.02441	0.025893	0.025883	0.89	1	0.35	
Intervention1_12	-0.01829	0.02587	0.025858	0.5	1	0.48	
Intervention1_16	-0.00732	0.02563	0.02562	0.08	1	0.78	
Intervention2	-0.01808	0.017098	0.017093	1.12	1	0.29	
Intervention3	-0.02659	0.017191	0.017183	2.39	1	0.12	
Intervention4	-0.01595	0.017155	0.017151	0.86	1	0.35	
Intervention5_Group	-0.01476	0.014815	0.014811	0.99	1	0.32	
Intervention6_Group	-0.04017	0.014758	0.014754	7.41	1	0.007	**
GenderMale	0.22283	0.010899	0.010895	417.97	1	<0.001	***
Age	-0.00845	0.000718	0.000717	138.7	1	<0.001	***
Frailty (test centre)				11390.55	316.9	<0.001	***

---

[47] Note: Because the RCT design was factorial, we matched the entire sample.

Survey of number of hours practice taken: Analysis of variance, model outputs

Variable	Df	Sum Sq	Mean Sq	F	value	
Intervention1_8	1	433	433	0.729	0.39	
Intervention1_12	1	13	13	0.023	0.88	
Intervention1_16	1	404	404	0.681	0.41	
Intervention2	1	71	71	0.12	0.73	
Intervention3	1	1817	1817	3.06	0.08	.
Intervention4	1	608	608	1.024	0.31	
Intervention5_Group	1	967	967	1.628	0.20	
Intervention6_Group	1	901	901	1.518	0.22	
GenderMale	1	24762	24762	41.709	<0.001	***
Age	1	232	232	0.39	0.53	
Test centre wait time	1	5499	5499	9.263	0.002	**
as.factor(test centre)	312	195305	626	1.054	0.25	
Residuals	5048	2996853	594			