ABSTRACT
There are a number of factors which may influence the validity of experimental studies, including the incentives offered and the instructions provided to participants. These have been little-studied in the driving domain. The aim of this study was to investigate how manipulating these factors influenced participants’ feelings of ‘presence’ (i.e. the extent to which they believed they were actually driving and not in a simulated environment). The findings showed that imposing a penalty system for poor driving performance and providing ‘good driving’ instructions did not affect presence ratings and this can be explained by the inherent need to perform well under test conditions and the small range of performance variability expected in the driving scenario. Participants in the penalty and instructions conditions gave higher ratings for negative effects (related to physically feeling unwell), suggesting that these conditions made them more aware of the physical symptoms of being in a simulator.

CSC Concepts
• Human-centered computing—HCI theory, concepts and models
• Computing methodologies—Virtual reality

Keywords
Driving simulation, Presence, Incentives, Instructions, Validity, Driver behaviour.

1. INTRODUCTION
Driving, as a safety-critical behaviour, is often studied in a simulated environment to minimise risk to participants and researchers and to enable stricter control over experimental variables. Assuming that studies are well-planned and influencing factors tightly controlled, simulator-based research tends to produce high levels of relative validity but is less successful for achieving high levels of absolute validity. Relative validity means that the magnitude of the relationship between two or more dependent variables as measured in a simulator is generally consistent with the magnitude of the relationship between these same variables in a real driving scenario. However, there are various inconsistencies between the simulated environment and the real driving environment, which can reduce the absolute validity of simulator studies. This can occur even when the fidelity of the simulated environment (i.e. the faithfulness with which reality is represented) is considered high, and will affect the overall ecological validity of simulation studies. An important contributing factor is participants’ knowledge that they are in an artificial situation, which reinforces a lack of negative consequences for poor driving performance. With no negative consequences (e.g. safety risks) in the simulator, participants may not be encouraged to drive in a way which is similar to driving on real roads.

There are aspects of the experimental situation which influence the level of ‘buy-in’ to the test situation, potentially increasing participants’ perceptions of the level of ‘reality’ of the test environment. This is also linked to participants’ feelings of presence, which can be described as the measure of the extent to which people believe they are actually driving and not in a virtual environment. Buy-in and presence will be influenced by issues such as the details provided in the study advertisement, the instructions given to participants at the beginning of a test, aesthetic elements of the test environment (e.g. signage, safety measures, separation between simulator and control room, appearance of experimenters), characteristics of the experimenter (e.g. experience, enthusiasm) and methods for rewarding participants for taking part in studies [1-5]. It is difficult to control all of these factors in an experimental setting; however, the purpose of this work was to investigate how some of these can influence the behavioural outcomes of a research study, potentially leading to guidance for control over variables in future studies. In this study, two specific characteristics of experimental design are considered: incentives and instructions.

2. INCENTIVES
Although many research study participants are willing to give their time for free, experimenters sometimes choose to provide an ‘incentive’ to participants for volunteering. This is usually monetary, either as a cash payment, voucher or reimbursement of reasonable travel expenses. It can also be in the form of a gift or in return for course credit, e.g. for student volunteers in university studies. There is a lack of consistency in the reporting and use of incentives for driving studies. To explore this further, the authors undertook a brief survey of articles published between 2014 and 2016 from two of the top journals in the field (Ergonomics and Applied Ergonomics) and from the most recent Automotive User Interface Conference (AutoUI – Nottingham, September 2015). It was found that out of a total of 61 driving study papers, the majority (36) did not state whether or not participants received any incentive. Of the 25 that did provide this information, 17 studies offered monetary incentives, 3 offered university course credits, 2...
Incentives can be effective in driving study participation; however, this is likely to be dependent on the nature of the study but will also be influenced by more practical constraints such as time available, experimenter effort and method of presentation. It is likely that instructions will even vary between individual tests within a single study, due to the involvement of different experimenters, content being forgotten between runs and some inevitable natural variation in delivery of instructions.

Instructions used in psychology experiments have been described as ‘at best ambiguous and at worst internally contradictory’ [11, p.275]. For example, participants are often required to fulfil two or more goals, such as completing tasks and keeping within a time limit, but the study instructions may offer no indication of how these goals should be prioritised. In the driving domain, many studies have tested the performance effects of simultaneously managing a secondary task such as programming a navigation destination and interacting with the primary vehicle controls; however, wording of the instructions can potentially provide implicit cues to participants about which one of these activities should be prioritised and this may not be indicative of ‘natural’ behaviour. Similarly, biased instructions can steer participants to particular behaviours, as has been observed in work on eyewitness identification [12]. Authors have suggested that useful instructions must be designed so that the information can be processed in working memory [13] and so that it can be accepted and translated into goals or intentions by the participant [4]. Interpretation of instructions is also affected by the attitude and expectations of participants; even if thorough instructions are provided to participants there is no guarantee that every person will listen to or read them thoroughly and then act on the information provided [14]. It is very difficult to assess how the design of instructions influences participant performance and there is only a small number of papers on this topic. One of these - [15] - showed that including hints to task success in instructions positively affects subsequent performance on tasks. Another - [16] - presented the results of a collection of studies which have shown that the alteration of just one or two words in the instructions given to participants before a study had a significant impact on subsequent performance: evidence that behaviour under laboratory conditions can be extremely sensitive to small changes in task presentation. This is known as a framing effect [17], whereby the presentation of a task leads to unintended and unpredicted behavioural changes [16].

4. METHOD
The aim of this study was to investigate the influence of two variables on driving performance and experience of ‘presence’ in a driving simulator. The independent variables were financial penalty, with 2 levels: (1) financial penalty based upon the number of assessed driving violations and (2) no financial penalty; and
participant instructions, with 2 levels: (1) detailed instructions and (2) minimal instructions. This design created four conditions:

- Detailed instructions, financial penalty
- Detailed instructions, no financial penalty
- Minimal instructions, financial penalty
- Minimal instructions, no financial penalty

In real driving, drivers are incentivised to drive well by the ‘threat’ of negative consequences of poor driving relating to factors including safety, journey time and fuel costs. In simulated driving none of these factors apply and therefore financial incentives have been used in experimental studies to act as a surrogate to these real life forces, encouraging adherence to more realistic driving behaviours. The financial penalty in this study was designed to increase participants’ perceptions of the negative consequences of driving, to test whether this in turn altered their perceptions of the realism of the driving simulator, i.e. presence. The financial penalty was imposed in the form of a deduction of value from the total reward offered to participants. All participants were offered £10 worth of high street shopping vouchers for completing the study. Half of the participants were allocated to the ‘financial penalty’ conditions: these participants were informed that for every driving violation committed during the drive, £1 would be deducted from the final voucher total. Participants were told that the violations would be assessed by the experimenter according to the UK Driving Test Report [18]. A typical violation in the driving simulator was exceeding the speed limit. At the end of the experiment, regardless of performance level and number of violations, all participants received the full £10 voucher incentive and the experimental manipulation was explained. Participants in the conditions with no financial penalty also received the £10 voucher incentive on completion of the test, but there was no discussion of penalties with these volunteers.

The instructions were based upon the current UK Driving Test Report [18], which is used to assess performance in the UK driving test. Extracts which were relevant to the simulated scenario were adapted from the explanatory notes of the test report to create the detailed instructions. This covered driving precautions, control, use of mirrors, signals, clearance to obstructions, response to signs / signals, use of speed, following distance, progress, positioning, position / normal stops, and awareness / planning. In the detailed instructions conditions, participants were given a printed copy of the instructions to read through in their own time prior to the test. The experimenter reiterated that they should drive as naturally as possible, as they normally would on a journey to and from work, whilst paying attention to the ‘good driving’ instructions they had been given. The instructions were designed to make participants aware of the safety and control aspects of driving in an attempt to stimulate a similar ‘mindset’ to that experienced in real driving. The experimenter then explained the vehicle controls to the participant and talked through the task. In the minimal instructions conditions, participants were not provided with the detailed instructions sheet. In these conditions the experimenter simply asked the participants to drive naturally and introduced the vehicle controls.

Participants
Forty participants were tested in a between-subjects design, with 10 participants per condition. The sample consisted of 15 female and 25 male drivers with a mean age of 30. The sample had a mean driving experience of 10 years and mean annual mileage was estimated at just over 7,000 miles. Participants were randomly allocated to one of the four conditions. Participants were required to hold a full UK driving licence and have at least one year of driving experience on UK roads. A stipulation was that participants had not previously been involved in a driving simulator study as it was thought that this would influence their perceptions of the test environment.

4.1 Procedure
Prior to attending the test, participants were asked to complete an online version of the Driver Behaviour Questionnaire [DBQ; 19]. This was used to assess the frequency of self-reported driving errors and violations committed by participants in the past year in order to provide a baseline performance measure with which performance on the test could be compared. On arrival at the study venue, each participant completed a demographic questionnaire, consent form and simulator sickness assessment questionnaire. This was followed by a practice drive. Participants then drove through a predefined scenario, lasting 20 minutes.

The study was carried out using the Human Factors Research Group driving simulator, at the University of Nottingham. The simulator is a fixed-base system, consisting of a Honda Civic cabin (right-hand drive). The road scene is projected onto a 270° curved projection screen with three projectors providing full peripheral coverage. Separate LCD monitors display feeds for the rear- and side-view mirrors. The simulated environment was generated using STISIM Drive™ version 2 software (System Technology Inc., CA, USA) which also recorded driving performance data. The scenario used in this study consisted of a simple three-lane motorway driving task. The driver started on a slip lane and had to join the main carriageway at the start of the drive. Participants were instructed to maintain position in the inside lane of the motorway at all times during the drive, keeping to a safe speed for the road type. Traffic was included in the scenario in both directions. If the vehicle was driven over the road edge a crash would have been simulated. This feedback is designed to encourage participants to drive in a natural way, demonstrating that there are negative consequences of poor driving performance. In each condition music clips were played in the car at set intervals, interspersed with silent periods of the same length. The participants were told that the aim of the study was to investigate the influence of music on driving; however, the purpose of this was actually to reduce the focus on driving performance. This was intended to make the experience more realistic, as in real driving there would normally be some division of attention between the primary driving task and internal or external distractions, e.g. radio, conversations, personal thoughts. The driver was instructed to pull on to the hard shoulder after 20 minutes via a recorded message played in the vehicle cab: this was the end of the test.

Following the drive, participants were asked to complete two questionnaires to assess presence. The presence questionnaire [20] consists of 32 questions requiring ratings of agreement on a 7-point scale and generates a single total value for overall presence between 0 and 192. The Independent Television Commission - Sense of Presence Inventory [ITC-SOPI; 21] consists of 38 questions requiring agreement ratings on a 5-point scale, the results of which are split into four factor scores: spatial presence, engagement, ecological validity and negative effects. Each score is a mean rating between 1 and 5. Negative effects is a measure of unpleasant feelings including sickness, sleepiness and eye strain and a higher rating indicates more negative effects. Participants also completed a post-trial questionnaire to test for motion sickness symptoms. They were debriefed, given the £10 voucher
incentive for their participation and asked to sign a post-trial consent form. The real purpose of the study was revealed to participants at this stage and they were told that they can withdraw their results if they wish.

4.2 Data Analysis
The effects of the four conditions on driving speed and presence ratings were analysed using the Mann-Whitney and Kruskall-Wallis tests respectively. Correlations (Spearman’s rho) were performed between the two measures of presence and also between the performance measures and the results of the DBQ.

5. RESULTS
5.1 Presence
Presence was measured using the Witmer and Singer and ITC-SOPI questionnaires. Figure 1 shows the mean ratings of presence according to the Witmer and Singer questionnaire. Participants in all conditions rated presence as generally positive as all mean scores were greater than 96 out of a total of 192, however there was little variation between the scores. The ‘minimal instructions, no financial penalty’ condition produced the highest rating of presence, although there was no significant effect of condition on presence according to the Kruskall-Wallis test, $H(3) = 3.94$, $p > .05$.

Figure 1. Presence ratings according to the Witmer and Singer questionnaire.

Mean ratings for the four factors of the ITC-SOPI presence questionnaire are shown in figure 2 (note that for the ‘negative effects’ the higher the rating the more negative effects were experienced). The ‘minimal instructions, financial penalty’ condition produced the lowest mean ratings out of all conditions for spatial presence, engagement and ecological validity / naturalness. It also produced the lowest mean rating of negative effects, which could indicate that lower ‘presence’ results in fewer feelings of discomfort.

Figure 2. Presence ratings according to the ITC-SOPI questionnaire.

Kruskall-Wallis tests showed there was no significant effect of condition on three of the ITC-SOPI measures: spatial presence ($H(3) = 1.79$, $p > .05$), engagement ($H(3) = 3.33$, $p > .05$), and ecological validity / naturalness ($H(3) = 3.24$, $p > .05$). There was, however, a significant effect of condition on ratings of negative effects ($H(3) = 9.02$, $p = .029$). In order to carry out post-hoc analysis, the ‘minimal instructions, no financial penalty’ condition was treated as a baseline as this is probably the scenario similar to most simulator studies. Mann-Whitney tests (with Bonferroni correction to a significance level of $p < .0167$ for multiple comparisons) showed a significant difference between the ‘detailed instructions, no financial penalty’ condition and the baseline, $U = 16.0$, $p = .009$, $r = -.58$, and the ‘minimal instructions, financial penalty’ condition and the baseline, $U = 18.5$, $p = .015$, $r = -.53$, but not between the ‘detailed instructions, financial penalty’ condition and the baseline.

5.2 Mean Speed
There was a significant effect of condition on mean driving speed, $H(3) = 9.27$, $p = .026$, with drivers in the ‘minimal instructions, no financial penalty’ condition having the highest mean speed (72.23 mph) and drivers in the ‘detailed instructions, financial penalty’ condition having the lowest mean speed (69.07 mph): see figure 3. Mann-Whitney tests (with Bonferroni correction to $p < .0167$ for multiple comparisons) showed that mean speed was higher than the baseline (‘instructions, no financial penalty’) for the ‘detailed instructions, financial penalty’ condition ($U = 17.0$, $p = .011$, $r = -.56$) and for the ‘minimal instructions, financial penalty’ condition ($U = 16.5$, $p = .009$, $r = -.57$). There was no significant difference in mean speed between the ‘detailed instructions, no financial penalty’ condition and the baseline. This result suggests that only the financial penalty factor (and not the instructions factor) produced a significant effect on driving performance, with drivers in the penalty conditions driving slower on average (mean = 69.3 mph) than drivers in the no penalty conditions (mean = 71.36 mph).
had no significant effect on all aspects, suggesting that using a financial penalty could make participants adhere more to certain rules and this may be useful in studies where this is necessary, for example to control a possible confounding variable. Based on the performance data, it is difficult to know which experimental condition produced the most ecologically valid behaviour; however, the DBQ data showed that most participants reported that in normal driving they broke the speed limit ‘quite often’ (15 participants) or ‘occasionally’ (15 participants). It may therefore be sensible to assume that in reality people do drive slightly over the speed limit in many circumstances, suggesting that the ‘no penalty’ conditions produced the more ecologically valid behaviour. The correlation analysis of DBQ-rated likelihood to break the speed limit and mean speed for each condition did not support this however, although the sample sizes and variations in mean speeds were considered insufficient for meaningful results in this case. The driving performance effects are also likely to have been dampened by the fact that people tend to perform diligently under test conditions irrespective of task demands [3], or penalties and instructions in this case. In the ‘no penalty’ conditions, it might still be assumed that participants were motivated to drive well by a sense of moral duty [7] or by an expectation that their involvement would contribute to a positive scientific outcome [3]. This would explain the very small differences observed in speed behaviour across the conditions. Furthermore, the activity of driving, particularly in a motorway environment which was used in this study, has a relatively small range of inherent variability, so big differences would not be expected unless the conditions were so drastic as to cause errors, e.g. crashing into another vehicle or veering off the road. These behaviours were not observed in the current study. Deniaud, Honnet [10] encountered similar findings in their study of presence in simulated driving and suggested that simulated driving conditions need to be designed to create distinct levels of attentional demand, involvement and suspension of disbelief to be challenging enough in order to enable detectable variations in presence. Nevertheless, presence is difficult to conceptualise, define and measure [10]. Most studies use self-ratings to assess presence and the two questionnaires used in this study are the most commonly applied in this area, although there have been recent attempts to create alternatives [For example see 22]. Encouragingly, the analysis showed a significant positive correlation between the Witmer and Singer and the ITC-SOPI questionnaires in this study, which indicates that they were measuring the same thing.

5.3 Correlations and Validity

A Spearman’s rho test was applied to investigate whether the different measures of presence correlated. There was a significant positive correlation between the Witmer and Singer ratings of presence and the ITC-SOPI ratings of spatial presence \(r_s = .340, p = .016\), engagement \(r_s = .509, p = .000\) and ecological validity / naturalness \(r_s = .374, p = .009\), measured irrespective of condition. This result strengthens the assumption that the two questionnaires were measuring the same thing. There was also a significant positive correlation between self-reported ratings of likelihood to break the speed limit in motorway driving (according to the DBQ) and mean speed recorded during the simulated drive [Spearman’s rho, \(r_s = .355, p = .022\), N = 40, measured irrespective of condition]. This suggests that for the sample of participants their speed behaviour in the simulator was similar to that which they reported in real driving, evidence of a good level of ecological validity, at least for this particular aspect of driving.

To examine whether condition had any effect on ecological validity as measured by the correlation between DBQ speed score and mean speed, Spearman’s rho correlations were performed for each condition. However, there were no significant correlations between self-reported ratings of likelihood to break the speed limit in motorway driving (according to the DBQ) and mean speed for any condition. It is likely that in this case the samples sizes (10 per condition) were too small and the range of speeds too narrow to show any meaningful results.

6. DISCUSSION

The results show that condition had no significant effect on all ratings of presence with the exception of ITC-SOPI negative effects. This measure relates to feelings of disorientation, tiredness, dizziness, eyestrain, nausea and headache. This result suggests that introducing instructions in this case.

Imposing a financial penalty appeared to make participants drive slower in this study, although no group had an excessively high mean speed (range 69-71 mph). This result is commensurate with previous work on the effect of incentives, suggesting that using a financial penalty could make participants adhere more to certain rules and this may be useful in studies where this is necessary, for example to control a possible confounding variable. Based on the performance data, it is difficult to know which experimental condition produced the most ecologically valid behaviour; however, the DBQ data showed that most participants reported that in normal driving they broke the speed limit ‘quite often’ (15 participants) or ‘occasionally’ (15 participants). It may therefore be sensible to assume that in reality people do drive slightly over the speed limit in many circumstances, suggesting that the ‘no penalty’ conditions produced the more ecologically valid behaviour. The correlation analysis of DBQ-rated likelihood to break the speed limit and mean speed for each condition did not support this however, although the sample sizes and variations in mean speeds were considered insufficient for meaningful results in this case. The driving performance effects are also likely to have been dampened by the fact that people tend to perform diligently under test conditions irrespective of task demands [3], or penalties and instructions in this case. In the ‘no penalty’ conditions, it might still be assumed that participants were motivated to drive well by a sense of moral duty [7] or by an expectation that their involvement would contribute to a positive scientific outcome [3]. This would explain the very small differences observed in speed behaviour across the conditions. Furthermore, the activity of driving, particularly in a motorway environment which was used in this study, has a relatively small range of inherent variability, so big differences would not be expected unless the conditions were so drastic as to cause errors, e.g. crashing into another vehicle or veering off the road. These behaviours were not observed in the current study. Deniaud, Honnet [10] encountered similar findings in their study of presence in simulated driving and suggested that simulated driving conditions need to be designed to create distinct levels of attentional demand, involvement and suspension of disbelief to be challenging enough in order to enable detectable variations in presence. Nevertheless, presence is difficult to conceptualise, define and measure [10]. Most studies use self-ratings to assess presence and the two questionnaires used in this study are the most commonly applied in this area, although there have been recent attempts to create alternatives [For example see 22]. Encouragingly, the analysis showed a significant positive correlation between the Witmer and Singer and the ITC-SOPI questionnaires in this study, which indicates that they were measuring the same thing.

Figure 3. Effect of condition on mean speed.
The main limitation of this study was the difficulty in defining and identifying what represents ‘realistic’ driving behaviour and it was therefore very difficult to draw conclusions on the effect of the independent variables on this. It was also impossible to assess how much attention each participant gave the instructions or the penalty scheme and how this translated to any real performance effect. For example, Orne [3] found that people participating in experimental studies were prepared to carry out a wide range of instructions with a high degree of diligence, even when presented with tasks which were deliberately designed to be frustrating, unrewarding and boring. When people are asked to participate in an experimental situation their motivation to comply with requests seems to change quite dramatically from normal circumstances and it is therefore difficult to infer behavioural effects from experimental studies like simulator trials.

6.1 Implications for the design of simulator studies

It is tempting on first glance to be disappointed in these results for showing minimal differences in presence across conditions; however, we have seen that it is incredibly difficult to assess whether the ecological validity of performance in the driving simulator is affected by factors like incentives and instructions because it is so difficult to know what exactly is indicative of ‘realistic’ driving. Furthermore, the small range over which driving performance can vary in this type of scenario means that large differences will never be induced by changing these incentives and instructions. Perhaps more important than validity, is the issue of consistency in these factors, which will influence reliability of driving studies. When considered from a reliability perspective, the results can be seen as generally positive for the body of literature on driving simulator studies as they suggest that differences in experimental protocol do not actually translate to significant effects on participants’ experiences of realism. This means that results can be generalisable across different studies, despite often resulting from different experimental approaches. Having said this, the authors still support the need for standardised protocols to guide simulator studies. This would follow a similar approach to that outlined in the draft SAE standard, J2944 Driving Performance Definitions [see 23]. This document provides consistent terminology and definitions for driving measures and encourages researchers to describe the method of calculation for any performance statistics reported in a driving study. Adherence to this standard will enable driving studies to be directly compared, as measures applied by different researchers in different institutions and across countries will be consistent. Similarly, it is important that researchers describe the instructions given to participants and any incentives offered or penalties imposed to achieve a level of transparency which will enhance the interpretation and use of driving study data.

7. CONCLUSIONS

The majority of papers on driving studies do not report the incentives and instructions given to the participants in any detail and there has been a distinct lack of previous research into the effects of incentives and instructions on driving studies. This study aimed to address this by investigating the effects of imposing a financial penalty system and providing driving instructions on participants’ behaviours in the simulator and their ratings of presence. The results showed no significant effects of incentives and instructions on most of the presence ratings and this can be explained by the inherent need to perform well under test conditions and the fact that a high level of performance was easily achievable in the driving situation due to the small range of performance variability expected in the scenario. It is possible that a scenario designed to produce more variability in behaviour would produce larger differences between conditions, for example if participants also had to perform a secondary task with high attentional demands at the same time as driving, then this might take priority over adherence to instructions and higher variability in performance may result. This is an area for future research. The results of this study suggest that participants cannot easily be manipulated into more natural behaviours, although the real difficulty is in understanding and measuring what these natural behaviours are. The findings are positive in the sense that differences in incentive and instruction administration between driving studies does not appear to have a large effect on driving behaviours (at least when those behaviours are fairly standard and subject to little natural variation); however the challenge is still how to encourage participants to behave naturally in the simulator and indeed how to test whether they are doing this.

8. REFERENCES

[1] Klintz, B.L., et al., *The experimenter effect*. Psychological Bulletin, 1965. 63(4): p. 223-232.
[2] Rosenthal, R. and D.B. Rubin, *Interpersonal expectancy effects: the first 345 studies*. The Behavioral and Brain Sciences, 1978. 3(3): p. 377-415.
[3] Orne, M.T., *On the social psychology of the psychological experiment: with particular reference to demand characteristics and their implications*. American Psychologist, 1962. 17(11): p. 776-783.
[4] Locke, E.A., *Toward a theory of task motivation and incentives*. Organizational Behavior and Human Performance, 1968. 3(2): p. 157-189.
[5] Groves, R.M., R.B. Cialdini, and M.P. Couper, *Understanding the decision to participate in a survey*. Public Opinion Quarterly, 1992. 56(4): p. 475-495.
[6] Dickert, N. and C. Grady, *What's the price of a research subject? Approaches to payment for research participation*. The New England Journal of Medicine, 1999. 341(3): p. 198-203.
[7] Russell, M.L., D.G. Moralesjo, and E.D. Burgess, *Paying research subjects: participants' perspectives*. Journal of Medical Ethics, 2000. 26(2): p. 126-130.
[8] Wertheimer, A. and F.G. Miller, *Payment for research participation: a coercive offer?* Journal of Medical Ethics, 2007. 33(4): p. 389-392.
[9] Eisenberger, R. and J. Cameron, *Detrimental effects of reward: reality or myth?* American Psychologist, 1996. 51(11): p. 1153-1166.
[10] Deniaud, C., et al., *The concept of "presence" as a measure of ecological validity in driving simulators*. Journal of Technology Science, 2015. 3(1): p. Online.
[11] Edwards, W., *Costs and payoffs are instructions*. Psychological Review, 1961. 68(4): p. 275-284.
[12] Greathouse, S.M. and M.B. Kovera, *Instruction bias and lineup presentation moderate the effects of administrator knowledge on eyewitness identification*. Law and Human Behavior, 2009. 33(1): p. 70-82.
[13] Marcus, N., M. Cooper, and J. Sweller, *Understanding instructions*. Journal of Educational Psychology, 1996. 88(1): p. 49-63.
[14] Oppenheimer, D.M., T. Meyvis, and N. Davidson, *Instructional manipulation checks: detecting satisficing to
increase statistical power. Journal of Experimental Psychology, 2009. 45(4): p. 867-872.

[15] Fernie, G. and R.J. Tunney, Some decks are better than others: the effect of reinforcer type and task instructions on learning in the Iowa Gambling Task. Brain and Cognition, 2006. 60(1): p. 94-102.

[16] Abbink, K. and K. Hennig-Schmidt, Neutral versus loaded instructions in a bribery experiment. Experimental Economics, 2006. 9(2): p. 103-121.

[17] Tversky, A. and D. Kahneman, The framing of decisions and the psychology of choice. Science, 1981. 211(4481): p. 453-458.

[18] Driving and Vehicle Standards Agency, DL25 driving test report form. 2011, HMSO: London.

[19] Reason, J., et al., Errors and violations on the road: a real distinction? Ergonomics, 1990. 33(10/11): p. 1315-1332.

[20] Witmer, B.G. and M.J. Singer, Measuring presence in virtual environments. 1994, United States Army Research Institute for the Behavioural and Social Sciences: Alexandria, VA.

[21] Lessiter, J., et al., A cross-media presence questionnaire: the ITC-sense of presence inventory. Presence: Teleoperators and Virtual Environments, 2001. 10(3): p. 282-297.

[22] Donkor, R., G.E. Burnett, and S. Sharples, Validating driving simulators for distraction research, in International Conference on Driver Distraction and Inattention. 2011: Gothenburg.

[23] Green, P., Standard definitions for driving measures and statistics: overview and status of recommended practice J2944, in Automotive User Interfaces and Interactive Vehicular Applications: AutoUI J. Terken, et al., Editors. 2013, ACM: Eindhoven.