Feedback in Highly Automated Vehicles: What Do Drivers Rely On?

Joy Richardson[1], Kirsten M.A. Revell[1], Jediah R. Clark[1], Nermin Caber[2], Mike Bradley[2], Theocharis Amanatidis[2], Patrick Langdon[3], Simon Thompson[4], Lee Skrypchuk[4], Neville A. Stanton[1]

[1]University of Southampton, UK, [2]University of Cambridge, UK, [3]Edinburgh Napier University, UK, [4]Jaguar Land Rover, UK

ABSTRACT

SAE (Society of Automotive Engineers) Level 3 vehicles are in development by many manufacturers. In order to deliver increasing amounts and types of information, in-car information systems are becoming more varied and complex. Feedback can now be given to the driver in a wide variety of ways including text and graphics and changing colours across multiple screens, on the windscreen with a Head Up Display, vocal or other audio alerts, ambient lighting and haptics. A high-fidelity simulator study was undertaken in which participants were exposed to all of these feedback modes and then ranked them in terms of reliance. Analysis shows how the feedback modes participants relied on varies widely and how gender can influence the results.

KEYWORDS

Highly Automated, Driving, Reliance

Introduction

What is Reliance?

The concept of reliance is used as a measure in a wide variety of fields such as accountancy (Hampton, C. 2005, & Arel, B. 2010), medicine (Janssens et al, 2004), philosophy (Budnik, 2018), Health Information (Hall et al, 2015), politics (Johnson & Hall 2014), marketing (Mumuni et al 2018), cell Phone technology (Sato, 2013) and epistemology (Fantl et al. 2019). However, the definition is not unanimously defined.

Within the field of philosophy, the difference between reliance and trust has been considered difficult to distinguish, causing much debate (Budnik, C. 2018). Political studies have found that reliance is more complex than the amount of time spent using a particular technology or how frequently it is used, as reliance is also dependent on its usefulness. Reliance therefore is a more robust measure which can also include influence and confidence (Johnson, T. J. & Kaye, B. K. 2014).

Mumuni and colleagues started trying to define reliance by using definitions from the Cambridge Dictionary online and Dictionary.com. The Cambridge Dictionary defines reliance as;
“the state of depending on or trusting in something or someone” or (in Business English)
“the state of needing or depending on something or someone in order to be able to do something”
Dictionary.com defines reliance as “confident of trustful dependence”. They continue to describe reliance graphically through models as the extent to which a person depends on one medium relative to other mediums or technologies, this can include a tendency for a person to value and trust a particular technology (Mumuni, A. G. et al 2019).

Measuring Reliance

Similarly to the disparity in definitions of reliability, there is no standard measure or scale to quantify reliance. Sato and colleagues employed a six-point coding the response to five questions regarding mobile phone usage, 1 responding to feature less frequently used (ranging from 0-1 to 0-25 times per day) up to 6 responding to a feature more frequently used (ranging from >25 to >125 times per day) (Sato et al, 2013).

Likert responses are more frequently used but with varying scales, some using an even figure scale which only allows for positive or negative responses whilst others using an odd figure allowing for a neutral response. Some use the lowest number to mean positive and the highest number meaning negative whilst other researchers employ their scales inversely. Hall and colleagues employ of four-point Likert response scale with one meaning strongly agree and 4 strongly disagree (Hall, A. K. et al, 2015). Johnson & Kaye employ two identical five-point Likert scales with 1 meaning never rely and 5 heavily rely (Johnson, T. J. & Kaye, B. K. 2014). Mumuni and colleagues also employ a Likert response but with a scale of seven points, 1 meaning strongly disagree and 7 strongly agree (Mumuni et al 2019). Hampton also employs a 7 point Likert-type scale with ranges from strongly disagree to strongly agree across five measures of reliance being user perception of agreement, level of confidence in output, user's confidence in their own judgement and finally the reliance on their aid when forming a decision (Hampton 2005).

For this study reliance was taken to mean how beneficial, how credible and how frequently each mode was used. In order to measure this a new scale was developed in which participants could rank how much they relied on the feedback they received from the vehicle. This enabled the participants to compare each feedback mode. They were asked; "Please rank in order (1 to 7) the elements that you relied on for information on the automation system? (1 = least reliance, 7 = most reliance)". Participants inputted their results into an electronic form using radio buttons (Figure.1).

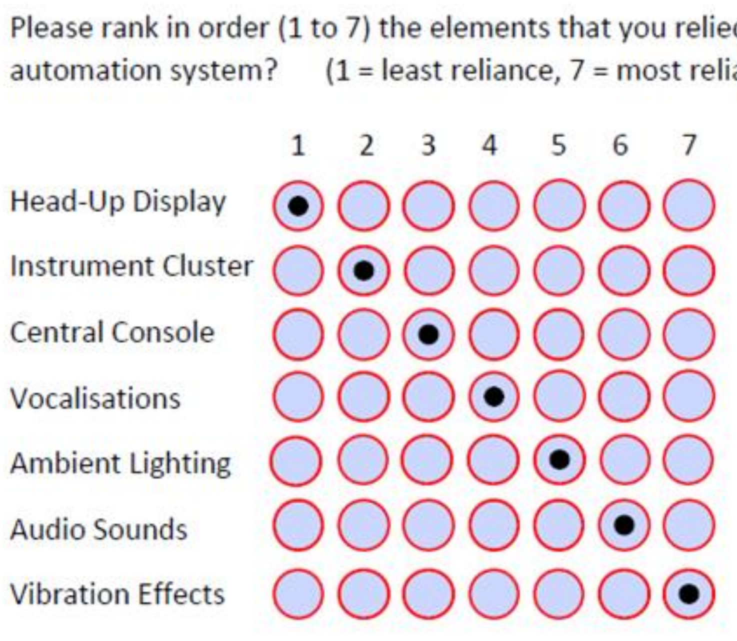

Figure 1: Example of Reliance Ranking Question

Method

Participants

Ethical approval was gained via the University of Southampton ethics panel (ERGO 41761.A3). A total of 66 participants were recruited for this study. In order to be included potential participants needed to have held a valid full UK driving license for more than two years and be generally healthy with no history of motion sickness. The recruited participants were an equal split of female and male and were also divided into the following age groups; Age Group 1 18-34, Age Group 2 35-56 and Age Group 3 57-82.

Design

The study simulated manual and highly automated driving on a UK motorway in a hybrid vehicle, the focus of the study was the transition from the vehicle self-driving to manual control assisted by a customisable HMI (Human Machine Interface). The study consisted of four repeated measures driving trials, each consisting of three handbacks (driver to vehicle), and three handovers (vehicle to driver). For each participant, half of trials were defined as being a 'short' time OOTL (Out-Of-The-Loop) consisting of one minute of automated control (three minutes in total for each short trial). The other half of trials were defined as being a 'long' time OOTL consisting of 10 minutes of automated control (30 minutes in total for each long trial).

Equipment

STISIM (Systems Technology Incorporated Simulator) Drive software was used to simulate a typical UK motorway environment. The simulator represented a hybrid Land Rover Discovery equipped with a single front-view screen replicating the windscreen, with separate digital display wing mirrors and an augmented display for rear-view. The simulator was equipped with a customisable digitalised instrument cluster, central console and head-up display (HUD). The vehicle was also equipped with vocal and audio information streams, ambient lighting to indicate driving mode (orange for manual, blue for automated), and a vibrating seat providing haptic feedback. All of these streams of information were customisable by the participant via the customisation matrix displayed on the central console after each trial. A laptop computer was used by the participants in order to collect questionnaire data.

Procedure

Participants were guided into the driving simulator where they were introduced to the controls and information modes. Drivers were then advised what would happen in the experiment outlining how and when transitions were expected.

During automated control, to simulate a secondary task, the driver played Tetris on a Window's tablet. A visual indicator counting down the time left in automation (from one minute or ten minutes depending on condition) was displayed on all three screens by default. At five, two, and one minute before manual control was expected, an audio tone and a vocalised alert was given to the driver notifying them of time remaining. When the countdown reached zero, the seat vibrated in co-occurrence with an audio and vocal alert. The handover icon animated the requirement to resume driving position. At this stage, the vehicle vocalised questions, and displayed them on each display. Questions were randomly generated from a list of 10 and asked the driver about vehicle status or the driving environment. Each answer was delivered vocally and was categorised as being either correct or incorrect by the researcher. Once the system was satisfied that questions were correctly answered, the vehicle indicated to the driver to take control of the vehicle by vocal and visual communication. After pressing the two green buttons the driver was now in control, audio, vocal, visual alerts were given, the ambient lighting changed to amber and the vibrating seat pulsed one last time to confirm the handover. This process represented one control cycle and was performed three times for each condition.

After each trial, the participant was asked to make changes to the HMI by ticking/unticking boxes in the customisation matrix or using sliding scales. Once satisfied with their decisions, participants saved the matrix and the next trial used these settings. Once three trials had been complete, the driver left the vehicle and took part in a final debrief questionnaire.

Results

During the post-study debrief questionnaire participants were asked the following question; “Please rank in order (1 to 7) the elements that you relied on for information on the automation system? (1 = least reliance, 7 = most reliance)”. The listed elements were aspects of the HMI which could be customised and consisted of; HUD, instrument cluster, central console, vocalisations, ambient lighting, audio tones and vibration. Each element received a score from 1 to 7 allowing for each number only to be used once thereby providing a ranking order of reliance for each participant. 11 people answered the question incorrectly and the score from these participants was removed from the analysis, the remaining participants can be seen in table 1.

Table 1: Revised Participant Demographics

	Age Group 1 (18-34)	Age Group 2 (35-56)	Age Group 3 (57-82)
F	7	10	9
M	9	9	9

The results from the reliance questionnaire have been represented graphically in two ways; Figure 2. Is a bar chart showing the cumulative frequency and distribution of the reliance ranking scores from all participants as previously identified in Table 1. Each of the different elements are listed along the x-axis; HUD, Central Console, (Instrument) Cluster, Vocalisations, Ambient Lighting, Audio Sounds and Haptics. Then each bar represents bow many participants ranked that element which each ranking score from 1-7.

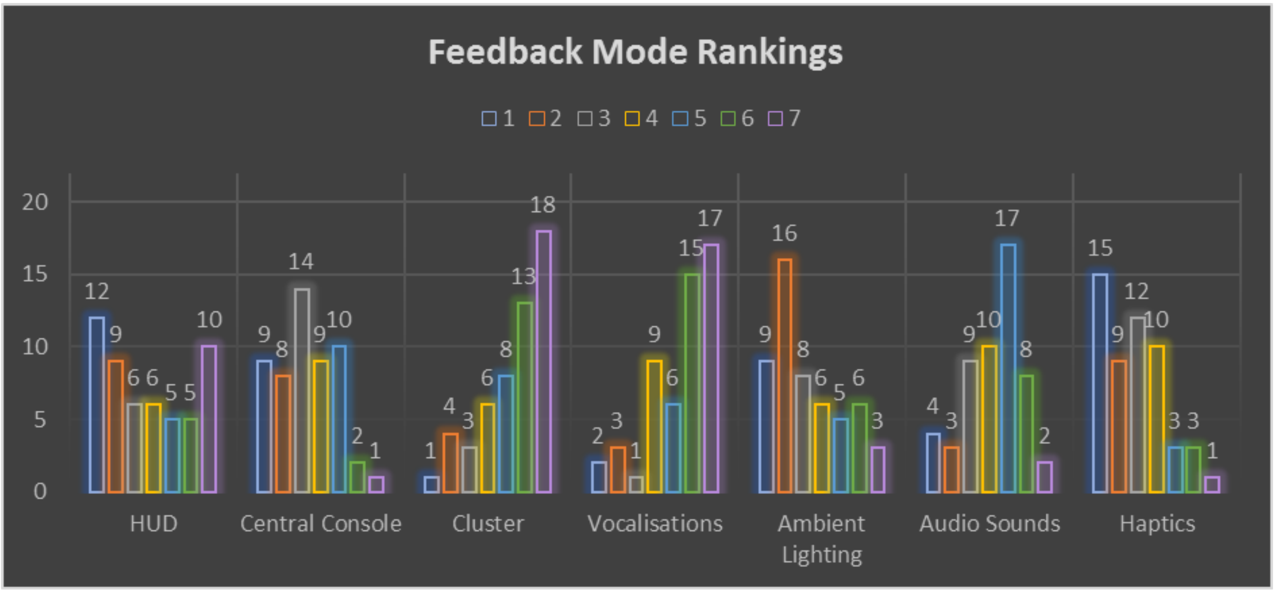

Figure 2: Feedback Mode Ranking Scores

Figure 3. Shows the results from the same ranking question displayed in a different way as a radar graph. Here all of the ranking scores from each participant have been converted into a radar; with

each point of the heptagon being labelled with one of the seven elements listed in the ranking question and it is constructed from 7 concentric shapes, the points of each of which respond to the ranking scores. The points of innermost shape represents a score of 1 and the points of the outermost shape represents a score of 7. All of the individual radars have been uniquely coloured and laid on top of each other to produce this composite graph.

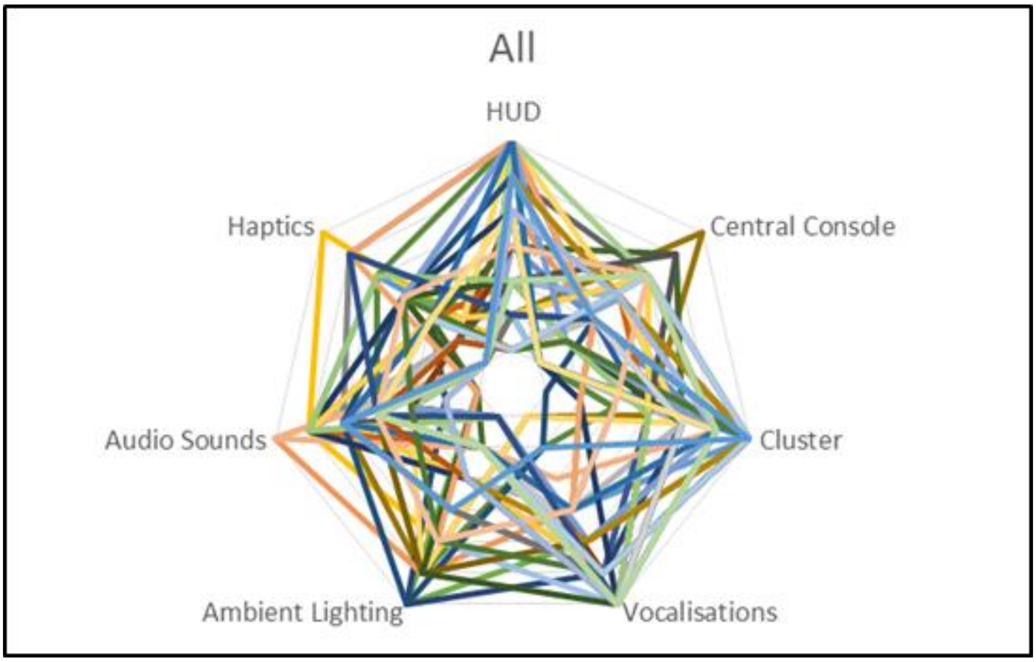

Figure 3: Composite Radar Graph of all Participant Ranking Scores

The individual graphs were divided by gender to create two new graphs as can be seen in Figure 4. The scores from the female participants were generally well distributed across the graph, however there is a noticeable absence of scores 5, 6 or 7, scores relating to the most reliance, for haptics. The scores for the male participants were also generally well distributed however in this case no males scored the central console as either 6 or 7, score relating to the most reliance, for the central console.

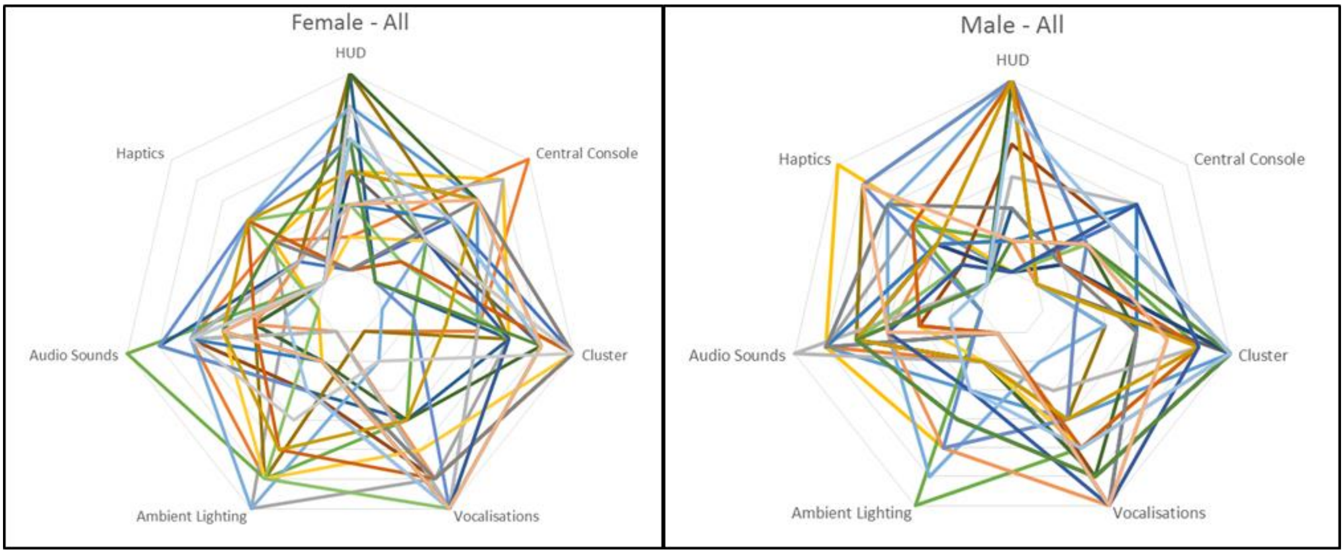

Figure 4: Reliance Scores Divided by Gender

The individual scores were then divided into the three age groups defined in the experiment and new radar graphs created. There were no obvious trends related to age.

Discussion

Figure 2 shows more participants ranked Ambient Lighting (a total of 25 participants) and Haptics (a total of 24 participants) as the elements scoring the lowest two scores of 1 or 2 meaning they were least reliant on these methods for information on the automation system. This reflects the design of the system as these elements are intended to reinforce mode awareness via the ambient lighting or that a handover is either required or has been completed via the haptic feedback seat, they were not designed to be a primary source of information about the automation system.

Figure 2 shows that more participants ranked the Cluster (a total of 31 participants) and Vocalisations (a total of 32 participants) as the elements scoring one of the highest two scores of 5 or 6 meaning they were most reliant on these methods for information on the handover. However, there is wide distribution of ranking scores across all of the feedback modes. Figure 2 also shows there is a wide variety between the ranking scores across the participant group, both graphs also show how every element was scored the highest ranking of 7, therefore relied upon the most, by at least one participant.

Evaluating all the modes the participants had been exposed to, and been able to adapt to suit their preference, throughout all trials highlights those which participants found themselves relying on most. This enabled the analysis of reliance based on gender. This analysis highlighted some differences in the results reported between the male and female participants. Notable is the female participants' seeming trend to not rely on the haptic feedback from the seat. Some previous studies have indicated that females are more sensitive than males in detecting haptic feedback or feel it more intensively (Goff et al 1965, Verillo 1979, Neely & Burnstron 2006, Forta 2009) despite this it has been observed that female participants scored haptic feedback a lower score for satisfaction than their male counterparts when using this as an alert to indicate the need to perform a handover task in an automated driving context (Duthoit et al 2017). This could suggest that whilst the female participants were more sensitive to the haptic alert they did not find this an adequate primary alert for indicating a handover was required. Female participants commented that the haptic seat was the "least effective" mode of communicating the need for a handover or that the positioning of the vibration needed adjustment. By contrast male participants described the haptic system as "effective", "really helped" and for one participant "happy to rely just on the haptics".

Conclusion

The Trust and Acceptance scales more typically employed during the course of a study are effective at demonstrating the opinion of a participant on the overall automation system at different points during the study and can be used to evaluate how the participants' opinions may change as they progress through the multiple trials. However, this newly designed Reliance Ranking Scale can provide a novel insight into the participants' use and opinion of the individual modes of communication used during the study by isolating each individual or grouping them and comparing them against one another.

These results highlight how it is difficult to package feedback modes aimed at particular user groups. The participants all had their own individual profile of feedback mode reliance with no two participants sharing the same profile. The only visible trend was related to the use of the haptic seat. No females relied on this as either a primary, secondary or tertiary mode of feedback. Overall, it seems that haptic feedback could be a useful mode of communicating the need for a handover but as an ancillary means.

Allowing drivers the ability to customise the settings and displays for each drive may allow the driver to select those options which are most pertinent to them completing a safe and effective handover whilst providing a more pleasant and satisfying driving experience.

Acknowledgements

This work was supported by Jaguar Land Rover and the UK-EPSRC grant EP/N011899/1 as part of the jointly funded Towards Autonomy: Smart and Connected Control (TASCC) Programme.

References

Arel, B. (2010) The Influence of Litigation Risk and Internal Audit Source on Reliance Decisions, Advances in Accounting, Incorporating Advances in International Accounting, 26, 170-176

Budnik, C. (2018) Trust Reliance and Democracy, International Journal of Philosophical Studies, 26:2, 221-239

Cambridge Dictionary https://dictionary.cambridge.org/dictionary/english/reliance accessed 19/6/19

Dictionary.com https://www.dictionary.com/browse/reliance accessed 19/6/19

Duthoi, V., Sieffernamm, J-M., Enregle, E., Michon, C. & Blumenthal, D. (2018) Evaluation and Optimization of a Vibrotactile Signal in an Autonomous Driving Context, Journal of Sensory Studies

Fantl, J., McGrath, M. & Sosa, E. Reliance, in Fantl, J., McGrath, M. & Sosa, E. (Eds), (2019) Contemporary Epistemology: An Anthology, John Wiley & Sons

Forta. N. G. (2009) Vibration Intensity Difference Thresholds (PhD Dissertation, University of Southampton)

Goff, G., Rosner, B., Detre, T. & Kennard, D. (1965) Vibration Perception in Normal Man and Medical Patients, Journal of Neurology, Neurosurgery and Psychiatry, 28, 503-509

Hall, A. K., Bernhardt, J. M., & Dodd, V. (2015) Older Adult's Use of Online and Offline Sources of Health Information and Constructs of Reliance and Self-Efficacy for Medical Decision Making, Journal of Health Communication, 20:7, 751-758

Hampton, C. (2005) Determinants of Reliance; An Empirical Test of the Theory of Technology Dominance, International Journal of Accounting Information Systems, 6, 217-240

Janssens, J., Heritier-Praz, A., Carone., M. Burdet, L., Fitting, J., Uldry, C. & Tschopp, J. (2004) Validity and Reliability of a French Version of the MRF-28 Health-Related Quality of Life Questionnaire, Respiration, 71, 567-574

Johnson, T. J. & Kaye, B. K. (2014) Site Effects: How Reliance on Social Media Influences Confidence in the Government and News Media, Social Science Computer Review, 33(2), 127-144

Mumuni, A. G., Lancendorfer, K. M., O'Reilly, K. A. & MacMillan, A. (2019) Antecedents of Consumers' Reliance on Online Product Reviews, Journal of Research in Interactive Marketing

Neely, G. & Burstrom, I. (2006) Gender Differences in Subjective Responses to Hand-Arm Vibration, International Journal of Industrial Ergonomics. 36 135-140

Sato, T., Harman, B. A., Adams, L. T., Evans, J. V. & Coolsen, M. K., (2013) The Cell Phone Reliance Scale: Validity and Reliability, Individual Differences Research, Vol 11(3), 121-132

Verillo, R. T. (1979) Comparison of Vibrotactile Threshold and Supra-Threshold Responses in Men and Women, Perception and Psychophysics. 26 20-24