

Driving with a Speech Interaction System: Effect of Personality on Performance and Attitude of Driver

Ing-Marie Jonsson, and Nils Dahlbäck

Department of Computer and Information Science, Linköping University, SE-581 83
Linköping, Sweden

ingmarie@ansima.com and nils.dahlback@liu.se

Abstract. Personality has a huge effect on how we communicate and interact with others. This study is one in a series of three that investigates how a speech based in-car system matched with dominant and submissive drivers affects performance and attitude drivers. The study was conducted with 30 participants at Linköping University in Sweden. Data show that using a voice that combines feature from submissive and dominant speech patterns work well for both dominant and submissive drivers. The voice showed the same performance gain as when matching car voice personality with personality of driver, without the negative attitude ratings associated with the submissive car voice found in previous studies. Drivers assessment of the car system show that even though both dominant and submissive drivers find the system helpful, dominant drivers find the system more annoying and more likely to turn the system off. Design implications of in-vehicle systems are discussed.

Keywords: In-car System, Driving Simulator, Driving Performance, Speech system, Attitude, Personality, Dominant and Submissive.

1 Introduction

Humans can easily detect characteristics in a voice and will use that skill when communicating with both humans and speech-based computer systems [1]. The linguistic and para-linguistic properties of a voice can influence people's attention and affect performance, judgment, and risk-taking [2, 3]. Previous studies show that voices used by in-car systems can influence driving performance and driver attitude [4, 5, 6]. Characteristics of the voice affects listeners perception of liking and credibility of what is said, regardless of if the speaker is human or computer-based system [3]. "Speaking is the most social and human thing we do", stated Professor Clifford Nass, professor and director of the Communication between Humans and Interactive Media Lab at Stanford University. "The minute you start speaking or listening to speech, the part of your brain that associates 'humanness' kicks in." [7]

In the context of in-car information systems, Nass et al. [8] show a clear positive effect of matching the emotional characteristics of the in-car voice to the emotional state of the driver. People prefer people to interact with people that are like themselves; it makes it easy to establish common ground and to communicate. Lazarsfeld and Merton [9] showed that most successful human communication will

occur between a source and a receiver who are alike, i.e., homophilous, and have a common frame of reference.

In general terms, theories of similarity-attraction and consistency-attraction [10] would suggest that personality has a huge effect on how we communicate and interact with others. Previous studies show that matching personality when communicating with a computer systems matters [11] and Dahlbäck, Swamy et al. [12] show that even matching accents matters. A system is always rated higher, and the user's perception of the systems performance better in matched cases. For in-car systems and driving performance, Jonsson and Dahlbäck [13], show a clear positive effect on driving performance when matching personality of the in-car voice with personality or driver. There is however a complex interaction between personality, perceived similarity, attitude and performance. Even though performance numbers are better for matched conditions, attitude towards the in-car systems does not necessarily improve with matched conditions.

To further investigate the effects of matching personality of in-car system with personality of driver. The authors designed an in-car system exhibiting properties that can be considered personality neutral, i.e. rating in the neutral zone between dominant and submissive.

The study reported here was designed to investigate if the voice of an in-car system, rated to be neither dominant nor submissive, would be perceived similar enough to trigger positive effects of similarity-attraction on driving performance without exhibiting negative effects on attitude.

2 Study Design and Apparatus

To investigate the effect of a personality-neutral voice on dominant/extrovert and submissive/introvert drivers a study with 30 participants was designed. The study was conducted at Linköping University in Sweden and is a follow-up of a study conducted at Oxford Brookes University in the UK [13].

2.1 Study Design and Participants

The design was a 1 (personality of car voice) x 2 (Personality of driver: dominant, submissive) between subject and gender balanced study.

There were 30 participants in the study (18 assessed as extrovert/dominant and 12 as introvert/submissive) Participants were screened based on the NEO-FFI [14]. It is an abbreviated version of the NEO Personality Inventory Revised (NEO-PI-R) (Costa & McCrae, 1992). It is intended for individuals aged 17 and older and requires a sixth grade reading level. The test items take the form of first person statements which participants are asked to rate on a five point Likert scale ranging from "Strongly Disagree" to "Neutral" to "Strongly Agree." The inventory typically takes 10-15 minutes to complete (Costa & McCrae, 1992).

All participants were students at Linköping University and they were awarded 140SEK for their participation.

2.1 Apparatus

Driving Simulator. The study was done using a driving simulator. This means that results provide an indication rather than a determination of behavior in real cars and real traffic.

There are many factors that motivate the use of driving simulators, the most pertinent being the ability to fully control the experimental setting and driving environment. The average driver will have very few accidents in their lifetime despite the dangers involved in driving. Due to the rarity of incidents, it would be extremely time consuming to set-up an experiment with the characteristics of real driving within the defined parameters of the study, and wait for a significant number of events to occur. Hence, the best way to examine new in-car systems is to challenge people using a driving simulator. Even though the degree of immersion varies with the fidelity of the simulator, the immersive effect is there even for very low fidelity simulators [15].



Fig. 1. STISIM Drive - Driving simulator. Setup Screen and random road scene depicted.

A commercial driving simulator, STISIM Drive model 100 with a 45-degree driver field-of-view, from Systems Technology Inc. was used in the studies. Participants sat in a real car seat and “drove” using a Microsoft Sidewinder steering wheel and two pedals, accelerator and brake. The authors selected a driving experience based on an automatic gearbox, and the simulated driving course was viewed on three large screen monitors in front of participants. The screens were setup as one screen right front, and two screens angled towards the driver on the left and right side respectively.



The view from the driver seat. There are two gauges visualized at the bottom of the screen, a tachometer and a speedometer. Please note the rearview mirror located at top of screen.



Traffic can either be programmed to follow traffic regulations or drive without adherence to traffic regulations. This includes behavior at stop signs, traffic lights and driving speed.

Fig. 2. STISIM Drive – Properties of driving setup and traffic.

Driving scenarios in STISIM Drive consist of a road with objects placed along that road. Note that a driving scenario in STISIM Drive is static. Drivers are driving the exact same road regardless of if they turn left, right or continue straight ahead at any intersection along the way. This ensures a consistent and repeatable driving environment from start to finish for all participants.



Fig. 3. STISIM Drive – Driving scenario with a small village, an intersection and pedestrians.

The driving scenario that was used was the same as in previous studies on personality of voice in cars [13]. It is a varied and realistic road scenario of 52 000

feet (15.85 kilometers), especially designed to take the drivers through rural areas, villages and intersections. In addition to driving the exact same scenario, all properties of the simulator, car, vehicle dynamics, weather conditions and traffic were set to be the identical for all participants.

In-Car System. The authors used the same navigation system as designed for previous studies on personality of voice in cars [13]. It takes the driver to five locations by interacting with drivers at certain locations along the way.

The navigation system consists of 40 utterances. 30 of the utterances are directions or suggestions, and 10 utterances are facts about the immediate surroundings. Directions and suggestions were designed to guide the drivers to the pre-programmed destinations. The facts were added to investigate how much attention drivers were paying to the system and the voice. All 40 utterances were translated to Swedish.

The Swedish voice that was used by the navigation system was selected to be personality neutral, neither rated as dominant, nor rated as submissive. The linguistic features used by the voice were a mix between those used by a dominant and a submissive voice. Choice of words was selected to match the dominant style. Using words such as “will”, “must” and “definitely”, in contrast to submissive style words such as “might”, “could” and “perhaps”. Overall the navigation system used assertive language “You should definitely turn right” in contrast to the submissive language style of “Perhaps you should turn right”. To make the voice and messages less dominant, it was then recorded with lower overall frequency, flat pitch range and slower speed than a typical dominant voice [11]. The male voice used for the systems was reviewed and rated on the same NEO FFI inventory [14] used to screen participants.

3 Procedure and Measures

3.1 Procedure

All participants were informed that the experiment would take one hour and started the experimental session by signing a consent form. This was followed by a five-minute test run of the simulator, where participants could familiarize themselves with the simulator and the controls. This enabled participants to experience feedback from the steering wheel, the effects of the accelerator and brake pedals, a crash, and for us to screen for participants with simulator sickness [16]. None of the signed up participants felt nauseous or discomfort during the training course. All 30 participants proceeded to fill in the first questionnaire consisting of general information such as gender and age and real-life driving experience.

In this study, all participants but one drove the driving simulator from start to finish. One participant retired from the driving session due to simulator sickness. The remaining 29 completed the driving scenario with the exact same navigation system using the same voice, scripted to take the driver to five destinations. During the drive all participants were subjected to the factual information inserted at 10 locations along the road.

After the driving session, participants filled in a set of post driving questionnaires. One of the questionnaires asked participants to assess the voice of the navigation system in terms of how similar it was to them. A second questionnaire asked the driver to assess their driving experience and how the navigation system was perceived to affect their driving performance. The final questionnaire asked participants to recall information volunteered by the navigation system during the drive.

3.2 Measures and Dependent Variables

This study used the same measures for personality, similarity, driving performance and navigation system as used in previous studies on personality of voices in cars, [13]. The authors used these measures in all three personality studies in this suite of studies to ensure consistency and enable comparisons between the different studies.

Personality. Participants were screened based on the NEO FFI inventory [14]. The inventory consists of 60 first person statements which participants were asked to rate on a five point Likert scale ranging from “Strongly Disagree” to “Neutral” to “Strongly Agree.” The NEO inventory measures differences among normal individuals, and will assess individuals on the five factors or dimensions of the five-factor model (FFM) of personality.

Similarity. *similarity-attraction* is an important aspect of how voices influence attitude and perception of spoken messages. Similarity-attraction predicts that people will be more attracted to people matching themselves than to those who mismatch. It is a robust finding in both human-human and human-computer interaction [9, 11]. The theory predicts that users will be more comfortable with computer-based personas that exhibit properties that are similar to their own. Attraction leads to a desire for interaction and increased attention in human-computer interaction [17, 18].

A standard questionnaire on homophily [19] was used to assess similarity. The index for similarity used in the study was constructed as a combination of attitudinal similarity and behavioral similarity. Participants were asked to rate the statements of the inventory based on the question "On the scales below, please indicate your feelings about the person speaking?" Contrasting statements were paired on opposite sides of a 10-point scale such that, 'similar to me' and 'different from me' would appear at different ends.

Driving Performance. This is a collection of measures that consists of accidents and adherence to traffic regulations. The driving simulator automatically collected the data for these measures. *Accidents* is comprised off-road accidents, collisions, and pedestrian incidents. *Adherence to Traffic Regulations* is comprised of speeding, running stop signs, and running red lights.

Because it is much more difficult to drive in a simulator than to drive a real car in real traffic, especially in a low-fi simulator like the one used in this study, the number of incidents are much higher than in real traffic, which makes this a useful measure of driving performance.

Navigation System. This is a collection of measures related to the voice used by the navigation system and how drivers perceive and react to it. The measure *Instructions followed* simply counts how many of the driving instructions drivers followed. There were a total of 30 instructions given by the system to navigate the driver from start to finish. *Time to destination* measures drivers' time to complete the driving scenario to the last destination. *Facts remembered* measures how many of the 10 driving scenario facts that drivers remembered after the driving session ended.

Driver Self-Assessment and Perception of Navigations System. Participants self-assessed their *Normal driving* style based on 8 terms using a 10-point Likert scale. In addition to this, participants also rated the perceived *Influence by navigation system* on their driving performance using a 10-point Likert scale for 9 terms.

Participants were specifically asked assess the driving session and navigation system. The driving session rated in terms of *Fun and Liking*, the navigation system in terms of being *Annoying* and *Helpful*. Finally, participants were asked to disclose their *Willingness to use*, i.e. to install and use in their own cars.

4 Results

The effects of using a “neutral” car voice in a navigation system with personality of drivers were measured by a one (Personality of Navigation System voice) by two (Personality of Driver) between-participants ANOVA.

4.1 Prior Driving Experience

To ensure that there was no bias based on drivers' prior driving experience, data from the two most recent years of driving was collected. The data that included number of miles driven per year, number of accidents, and number of tickets, was averaged for each group of drivers. No significant differences were found across conditions.

4.2 Similarity – Homophily

Data from the similarity assessment show that both groups of drivers felt similar to the car voice. There was no significant difference between the two groups of drivers, dominant drivers felt similar to the person behind the car voice Mean=5.9 SD=1.1, and submissive drivers felt equally similar to the person behind the car voice Mean=5.9, SD=1.0, $F(1, 28) = 0.006$, $p < 1.0$.

4.3 Driving Performance

Bad driving. There was no significant difference between the two groups of drivers on the bad driving indices, accidents and adherence to traffic regulations.

There was no significant difference between *Accidents* for dominant and submissive drivers. Dominant drivers show Mean=3.5, SD=2.0, submissive drivers show Mean=3.1, SD=1.3, $F(1, 28) = 1.8$, $p < 0.7$. Similarly, there was no significant difference between adherence to traffic regulations between dominant drivers (Mean=7.8, SD=5.5) and submissive drivers (Mean=11.1, SD=5.9), $F(1, 28) = 2.1$, $p < 0.2$.

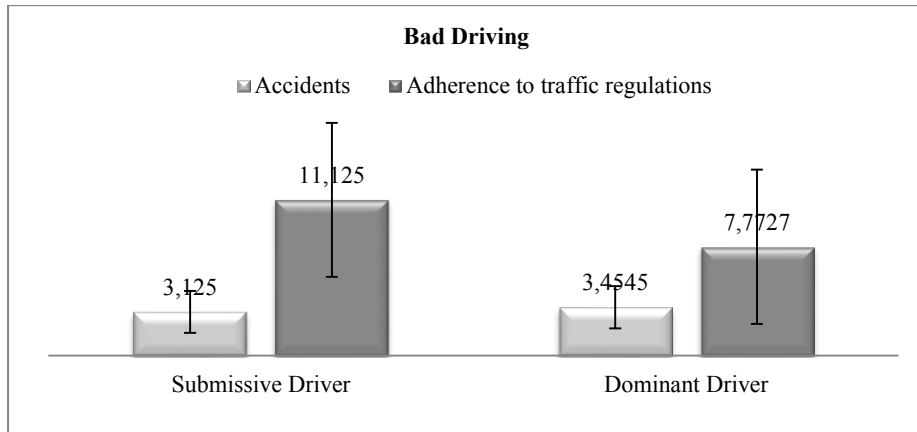


Fig. 4. Bad driving- accidents and adherence to traffic regulations

4.4 Navigation System

Instructions followed. Data show that there was no significant difference between the two groups of drivers when following instructions. Dominant drivers follow the same amount of instructions (Mean=25.1, SD=1.8) as submissive drivers (Mean=24.4 SD=2.0), $F(1, 28) = 0.84$, $p < 0.4$.

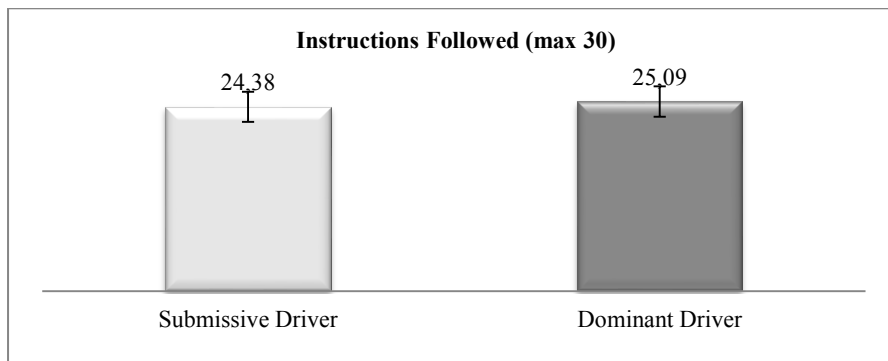


Fig. 5. Instructions followed

Facts remembered. Both submissive drivers and the dominant drivers paid attention to and listened to the voice equally.

Submissive drivers and dominant drivers remembered approximately 80% of the facts uttered during a 25-minute drive. Mean for submissive drivers was 7.9 (SD = 1.2) and mean for dominant drivers was 8.0 (SD = 1.2), $F(1, 28) = 0.04$ $p < 0.9$.

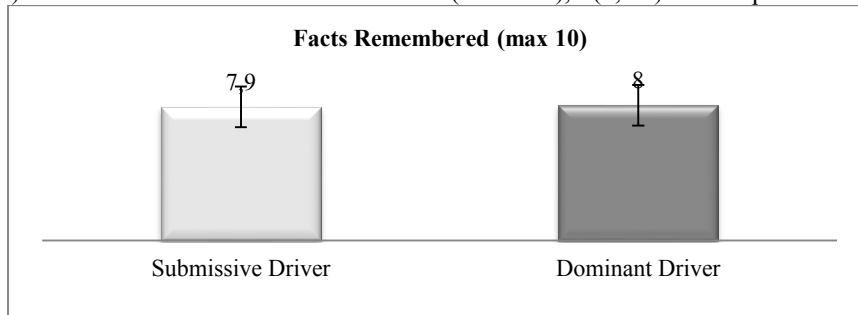


Fig. 6. Facts Remembered

Time to destination. The driving simulator automatically collected completion time, i.e. the time it took for drivers to reach their fifth and final destination. There was no significant difference between how long it took submissive drivers (M=24min. 20sec., SD=3min. 20sec.) and dominant drivers (Mean= 25min. 58sec., SD=3min. 30sec.) to reach the last destination, $F(1, 28) = 1.3$, $p < 0.3$.

4.5 Driver Self-Assessment

Normal driving style. Results from participants' self-assessment of their normal driving style show no significant difference between submissive and dominant drivers. Submissive drivers with Mean=4.8 (SD=0.7) and dominant drivers with Mean=5.3 (SD=1.3), $F(1, 28)=1.1$, $p < 0.3$.

Influence by navigation system. Data from participants rating the influence of the navigation system on their driving performance show no significant difference between the two groups of drivers. Assessing the positive influence, both submissive (Mean= 5.8, SD=1.2) and dominant (Mean=6.5, SD=1.2) drivers perceived that the system made them slightly more safe and careful drivers, $F(1, 28) = 2.2$, $p < 0.15$, than their normal driving style.

Fun and Liking. Data from participants rating their experience with the driving session show that both submissive drivers (Mean=8.2, SD=1.5) and dominant drivers (Mean=7.4, SD=1.5) had fun and liked the driving session.

Navigation System. There was a significant difference in how annoying the two groups of drivers found the navigation system. Dominant drivers found the navigation system to be more annoying (Mean=5.7, SD=2.4), than submissive drivers (Mean 2.8, SD=1.4), $F(1, 28)= 10.6$, $p < 0.005$. Please note that both groups of drivers found the navigation system to be helpful (Mean=7.0, SD=2.3 and Mean=7.4, SD=1.7), $F(1,28)=0.32$, $p < 0.6$.

Willingness to use. When specifically asked if they wanted to install and use the system in their own cars, there were also significant differences between the two groups. Submissive drivers (Mean=8.2, SD=0.9) were more willing than dominant driver (Mean=7.0, SD=1.2) to install in the car, $F(1, 28) = 6.7$, $p < 0.01$.

Dominant drivers (Mean=4.8, SD=2.0) were also more likely to turn the system off than submissive drivers (Mean=2.4, SD=0.8), $F(1, 28) = 10.1, p < 0.005$.

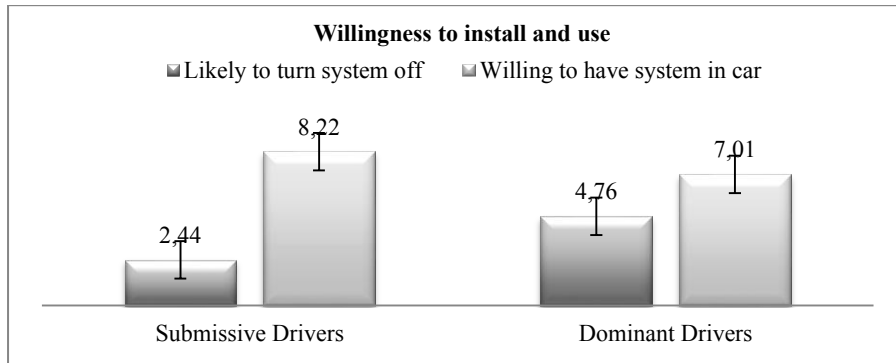


Fig. 7. Willingness to install and use system in own car

5 Conclusions and Discussion

Results from previous studies show that similarity-attraction predicts drivers' performance and attitude, for emotional drivers. The theory however, only partially predicts performance and attitude when matching personality of drivers with personality of car voice [13]. Result from this study show that selecting a car voice neutral on the dimension of dominant/submissive retains the performance benefits seen in matched conditions and lessening the negative attitudinal effects. Both submissive and dominant drivers feel similar to and like the car voice. The differences emerge with the unwillingness and resistance that dominant drivers exhibit in accepting instructions and advice from the system.

Data from a previous study [13] show that both dominant and submissive drivers felt less at ease after driving with the submissive voice, than after driving with the dominant voice. Data from the current study show that submissive drivers had a more positive experience with the personality-neutral car voice than with the submissive car voice in the matched case scenario [13]. Making the car voice personality neutral, reduced the negative impact on perception and willingness to listen, but did not entice dominant drivers to engage and interact. Data furthermore showed that dominant drivers were annoyed by the systems, and that they also were significantly more likely to turn the system off than submissive drivers.

The data from this, and previous studies, show complex interactions between personality, perceived similarity, attitude and performance. It emphasizes that it is important, to find the balance between matching-efforts and efficacy. Having a system that can accurately match drivers' personalities, is a remarkable technological feat, if drivers are not positively influenced by it on all dimensions, it is however a wasteful expense. Previous studies [13] showed that a system could be perceived as annoying and undesirable, regardless of its actual performance. This study shows that it is with a careful voice design possible to create a voice system that in most respects

work for both dominant and submissive drivers. What is especially important is that there are no negative effects on driving performance for either group.

As one study in a suite of personality based studies investigating effects of matching car voice with drivers, this study refines attitudinal results. Even though the data clearly show improvements over the matched cases investigated in a previous study [13], there are still more dimensions to be investigated. Dominant drivers perceive the in-car system tested in this study as a mixed blessing, seen as both helpful and annoying. It is not clear, however, if there is a design solution to this problem.

The bottom line is that even the technologically-best system may not satisfy or help all drivers: While in-vehicle information systems represent exciting technological advances, their deployment should be guided by significant caution.

References

1. Banse, R., Scherer, K. R.: Acoustic profiles in vocal emotion expression. In *Journal of Personality and Social Psychology*, vol. 70, pp. 614 -- 636 (1996)
2. Jonsson, I-M., Nass, C., Endo, J., Reaves, B., Harris, H., Le Ta, J., Chan, N., Knapp, S.: Don't blame me I am only the Driver: Impact of Blame Attribution on Attitudes and Attention to Driving Task. In *SIGCHI*, pp. 1219 – 1222, ACM Press (2004)
3. Nass, C., Brave, S.: *Wired for speech*” how voice activates and advances the human-computer relationship. MIT Press, Cambridge, Mass. (2005)
4. Zajicek, M., Jonsson, I-M.: A Complex Relationship, Older People and In-Car Message System Evaluation. In *Journal of Gerontology*, vol. 6, pp. 66 – 78 (2007)
5. Jonsson, I-M.: Conversational Interfaces and Driving: Impact on Behaviour and Attitude. In *IASTED Human-Computer Interaction*, pp. 224 – 229 (2008)
6. Jonsson, I-M., Dahlbäck, N.: The effects of different voices for speech-based in-vehicle interfaces: Impact of young and old voices on driving performance and attitude. In *Proceedings of the 10th Annual Conference of the International Speech Communication Association, INTERSPEECH 2009*, pp. 2795 – 2798 (2009)
7. La,L., CNET Reviews, iSheep, Fandroids, and why we care so damn much about our smartphones. Online Magazine http://reviews.cnet.com/8301-6452_7-57612654/isheep-fandroids-and-why-we-care-so-damn-much-about-our-smartphones/ November, (2013)
8. Nass, C., Jonsson, I-M., Harris, H., Reaves, B., Endo, J., Brave, S., Takayama, L.: Improving Automotive Safety by Pairing Driver Emotion and Car Voice Emotion. In *SIGCHI*, pp. 1973-1976, ACM Press (2005)
9. Lazarsfeld, P., Merton, R.: Mass Communication, Popular Taste, and Organized Social Action. In *The Communication of Ideas*, pp. 95-188 (1948)
10. Byrne, D.: *The Attraction Paradigm*, Academic Press, New York, NY (1971)
11. Nass, C., Lee, K. M., Does computer-generated speech manifest personality? An experimental test of similarity-attraction. In *SIGCHI*, pp. 329 – 336, ACM Press (2000)
12. Dahlbäck, N., Swamy, S., Nass, C., Arvidsson, F., Skågeby, J.: Spoken Interaction with Computers in a Native or Non-native Language - Same or Different?. In *Proceedings of INTERACT*, pp. 294 – 301 (2001)
13. Jonsson, I-M., Dahlbäck, N.: In-Car Information Systems: Matching and Mismatching Personality of Driver with Personality of Car Voice. In *Human-Computer Interaction. Applications and Services*, pp. 586-595. Springer, Heidelberg (2013)

14. Costa, P. T., Jr., McCrae, R. R. *Revised NEO Personality Inventory (NEO-PI-R) and NEO Five-Factor Inventory (NEO-FFI) professional manual*. Odessa, FL: Psychological Assessment Resources, (1992)
15. de Winter, J., van Leuween, P., Happee, P.; Advantages and Disadvantages of Driving Simulators: A Discussion. In *Proceedings of Measuring Behavior*, pp. 47 – 50 (2012)
16. Brooks, J., Goodenough, R., Crisler, M., Klein, N., Alley, R.: Simulator sickness during driving simulation studies. *Accident Analysis and Prevention*, vol. 42 pp. 788 – 796 (2010)
17. Dahlbäck, N., Wang, Q. Nass, C., Alwin, J.: Similarity is More Important than Expertise: Accent Effects in Speech Interfaces. *SIGCHI*, pp. 1553 – 1556, ACM Press (2007)
18. Nass, C., Lee, K.: Does Computer synthesized speech manifest personality? Experimental tests of recognition, similarity attraction and consistency attraction. In *Journal of Experimental Psychology: Applied* vol. 7 pp. 171 -- 181 (2001).
19. Rubin, R., Palmgreen, P., Sypher, H.: *Communication Research Measures: A Sourcebook*. Guilford Press, New York (1994)