Available online at www.sciencedirect.com

ScienceDirect

Procedia Manufacturing 3 (2015) 2642 – 2649

Procedia MANUFACTURING

6th International Conference on Applied Human Factors and Ergonomics (AHFE 2015) and the Affiliated Conferences, AHFE 2015

Evaluation of a radio tuning task on Android while driving

Michael Krause*, Cajetan Angerer, Klaus Bengler

Institute of Ergonomics, Technische Universität München, 85747 Garching, Germany

---

Abstract

We implemented the AAM reference radio tuning procedure with some modifications for Android and tested it against a hardware radio on a tablet and on a smartphone. The used measurement methods involved an AAM track with driving metrics and eye tracking, a Lane Change Test (LCT) and occlusion testing. The metrics are reported and compared. The glance metrics seem to be potentially influenced by a carry-over effect from a training task toward longer glance durations. Nevertheless, the results are promising for further testing and validation experiments. Due to the easy setup, more uniform hardware (Android tablet) and a procedure without experimenter announcements, the open source application could be a valuable tool in experiments.

Keywords: Radio tuning task; AAM; Touch screen; Reference task; Driver distraction; LCT; Occlusion; Glance duration

---

1. Introduction

The AAM guideline [1] specifies a radio tuning task as a socially accepted reference for assessing the suitability of tasks while driving, regarding driving metrics. When the procedure is applied in experiments, one often has to adapt the procedure to the available equipment. These adaptions, together with the differences in hardware, can affect most of the driver distraction metrics (see [2]) and, therefore, counteract the basic idea of a reference. Therefore, we implemented the procedure with a modification for Android and tested it against a hardware radio.

---

* Corresponding author. Tel.: +49-89-28915404; fax: +49-89-28915389.
E-mail address: krause@tum.de

doi:10.1016/j.promfg.2015.07.334

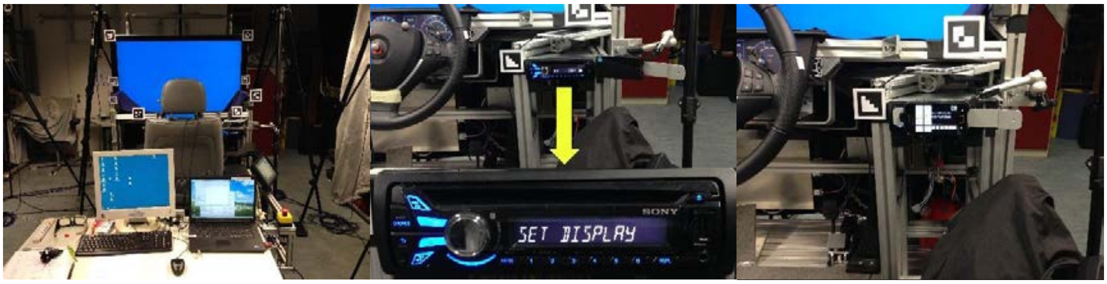

Fig. 1. Experimental setup (a) complete (b) driver space and radio (c) smartphone radio task.

2. Methods

2.1. Procedure

First, the subjects filled in a demographic questionnaire. Then, the subjects were introduced and trained on each of the three devices (hardware radio, smartphone, tablet; see section 2.3) for the radio tuning task until they performed five, error-free frequency tunings on each device. The experiment involved three measurement methods (AAM driving+ eye tracking, LCT and Occlusion; see section 2.4). The order of the three measurement methods was permutated. Due to an error, the order of the three devices within one measurement method was not fully permutated (first device for Occlusion is always the radio, for LCT the tablet and for AAM the smartphone). For LCT and AAM, baseline driving was permutated first or last condition. In Occlusion the devices started alternating with unoccluded or occluded trial. The experimental session needed about one hour per person.

2.2. Experimental setup

For the experiment we used a seating mock-up with a 55”-LCD screen (see Figure 1). The hardware radio was about 30° below the normal line of sight. If a 180cm male adjusted the steering wheel, the location was 45cm left to the centre of the steering wheel, 10 cm below and 20cm behind. The tablet and the smartphone were mounted directly in front of the hardware radio to be at the same location.

2.3. Radio tuning task

2.3.1. Android application (smartphone / tablet)

For the smartphone (Samsung Galaxy Ace S5830, 3.5”, 480x320, Android 2.3.6) and the tablet (Ampe A65, 6.5”, 800x480; Android 4.3) the same app was used, but rendered to a different output (Figure 2). The implemented app was inferred from the specification in AAM pp.46. Modifications and rationales are explained afterwards.

The app includes 10 music and 10 speech signals which are randomly assigned as radio stations. The frequency tuning itself (not counting presses of the radio button and band switch) needs 40 to 44 button presses. It seems the specification has a literal error in the AM radio band: “AM 530 to 930 kHz (steps of 5 kHz, approx. 200 steps)”.

Fig. 2. Radio tuning app (a) smartphone (b) tablet (applies some vertical distortion).

This would mathematically lead to a frequency band of (930kHz – 530kHz) / 5kHz only 80 steps. We changed it to a reasonable range of 540kHz to 1610kHz for American radios. The specification also states: “A radio with several push-buttons and a display can be used or simulated.” This statement implies that touch screens are not in the originally-intended scope. Different statements of the specification state that the task data (band and frequency) is announced by the examiner. In the implementation, this is achieved visually “Tune XY”. This changes the task slightly. It removes the hard-to-standardize auditory announcement, diminishes the cognitive workload of remembering data, takes the (variable) examiner out of the experimental loop and prevents requests to the examiner (“Sorry, which frequency?”). The font size is specified with at least 5mm in AAM. The size of the button labelling and the “Tune”-instruction is rendered 2.8mm on both devices; this is in between the minimal and acceptable specification of font size in a distance of 65cm (ISO 15008:2009).The current band and frequency has a height of 4.2mm, this being above the recommended font size of ISO15008:2009 for 65cm distance. The buttons have a width and height of 1.6cm on the tablet and a width of 1.4cm and a height of 1.2cm; on the smartphone. The current (modified) version and source code can be found at the website of the Institute of Ergonomics [3].

2.3.2. Hardware radio

The specification states [1]: “If a real radio is used, it should provide reasonable approximation of these features.” Therefore, the authors of AAM seemed to be aware that their specification would sometimes need modifications. We even run into a problem, when we try to buy a new radio that is close to the specification; as a radio that could be switched to manual tuning with separate up/down buttons could be hard to find. For the experiment we used a SONY CDX GT570UI.

The radio display has a font size of 7mm and the frequencies to tune, are communicated with an iPhone right next to the radio with a font size of 2.25m (minimal size according to 15008:2009 at 65cm distance). The subjects were instructed and trained to get into the manual tuning mode (1s long press on the up or down button) and to use single button presses to tune. The radio uses a time-out to get out of manual tuning mode if no manual tuning is performed for a moment. For the experiments, it is needed that the task can be performed over and over again without interruptions. For this reason, the AAM procedure, to switch between CD and radio, was discarded. The radio has five (FM1, FM2, FM3, LW, MW) bands, not four (AM, FM1, FM2, Weather) according to AAM; this somehow counterbalances the discarded switch from CD to radio. The target band and frequencies were all on FMx and 40 to 44 steps apart (the radio uses 0.05MHz steps; AAM assumes 0.1MHz). At the target frequency, music was played via a FM transmitter. After hitting the target frequency and 2-second delay, the examiner showed a new target band and frequency on the iPhone.

2.4. Measurement methods

2.4.1. AAM driving and eye tracking

The driving task is to follow a leading vehicle (constant speed 80km/h) at a safe distance of 50m.With slight modifications for the metrics, widths and appearance (e.g. white lines, no rumble strips, etc.) for German Autobahn specifications, this task resembles the main aspects of the AAM driving scenario. The driving simulation is SILAB 4 (WIVW GmbH, Würzburg) the used eye tracker Dikablis with D-LAB 2.1 (Ergoneers GmbH, Manching). Participants were instructed for safe driving and prioritization of the driving task, as well as fast and accurate radio tuning. For the measurement condition ‘AAM’, the subjects drove at least 3 minutes to get familiar with the simulation. After that, they entered at least one phone number into the smartphone while driving, to get accommodated to a dual task setting. For the measurement trials, the simulation was started every time again from stand still. Then, the subjects needed some time to settle (into the following distance). After 30 seconds, the examiner gave the verbal command to start the three sequential radio tunings; these first 30 seconds are unconsidered. When the third tuning was finished, the recording was stopped. For the baseline run, 2 minutes of driving where recorded and the first 30 seconds were also discarded for calculations. The AAM condition is used for the metrics: Standard Deviation of Lane Position (SDLP), Standard Deviation of Time Headway (in this case SD of tip-tail Headway or SD Time Gap), Total Time on Task while driving, Total Glance Time to the device, Mean

Single Glance Duration to the device, Number of Glances to the device and Glance Frequency (Hz) to the device. Lane exceedances are normally very rare events and were therefore not assessed.

2.4.2. Lane change task (LCT)

Ten LCT tracks were programmed into SILAB 4 according to original LCT implementation and ISO/DIS 26022:2007.The driving data was converted with a script and analysed with LCT- Analysis V3.03 (Daimler AG, 2011) regarding MDEV and adaptive MDEV. To get familiar with the LCT, subjects drove one track (3 minutes) and entered phone numbers. In the measurement trials the subjects started the tuning of three radio frequencies at the LCT start sign; the recording was stopped after finishing the third tuning. For the LCT baseline run, a whole track was driven. The subjects were instructed for fast and accurate lane changes, but the artificial, fast behaviour mentioned in the ISO standard was not enforced (adaptive MDEV should account for this).

2.4.3. Occlusion

For occlusion, PLATO spectacles (Translucent Technologies) are used with system paced 1500ms open and 1500ms closed intervals according to ISO 16673:2007. Participants were instructed that operation of the device is possible during occlusion, as is fast and accurate radio tuning. The experimental setting was the seating mock-up (see Figure 1) in driver position. Participants get familiar with occlusion by entering at least one phone number. Times were recorded with a stop watch for three sequential radio tunings. Total Shutter Open Times (TSOT) were derived from stopped times by modulo calculation (see also [4]).

2.5. Subjects

In this experiment, 24 subjects participated voluntarily without compensation. All had a valid driver’s licence. The group involved 17 males and 7 females, with an average age of 26 years (SD: 3.6 years), all with an academic background (students/staff). Two persons are left-handed. A red-green deficit was self-reported by two persons. The frequent use of a touch smartphone was reported by all, except for one person. The usage of touch tablets, touch satnav and touch car infotainment were each bimodal with peaks at the upper and lower end of the usage frequency scale.

3. Results

In the following figures, we visually compare the relationship between interquartile ranges, medians and min/max values; and use a high p-value level (preferably > 0.25; two-tailed t-tests) as an indication that differences are unlikely. Pearson Correlations (r) are also reported.

3.1. Occlusion metrics

As can be seen in Figure 3, the total task time unoccluded on the devices is partly different. For the hardware radio, this is clearly visible. The most conservative reference would be the tablet with a median of 41.72s, therefore, resulting in about 14s per frequency tuning.

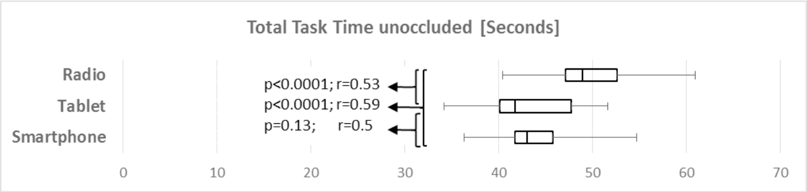

Fig. 3. Total Shutter Open Time unoccluded for three sequential tuning tasks (minimum, Q1, median, Q3 and maximum).

The Total Shutter Open Time (Figure 4) is also the most conservative (shortest, if used as a reference) for the tablet. The median is 27s. Therefore, one tuning would need 9s TSOT. Dividing the 85th percentile value for the tablet (30.28s) by 3 would yield about 10s.

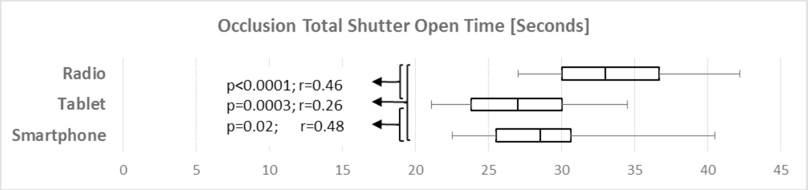

Fig. 4. Total Shutter Open Time (TSOT) for three sequential tuning tasks (minimum, Q1, median, Q3 and maximum).

The R-ratios (TSOT divided by TTT unoccluded) in Figure 5 are an indicator of the interruptability, resumability as well as blind-operability. The graphic shows no great differences. Nevertheless, the median of the tablet is the lowest, most conservative reference.

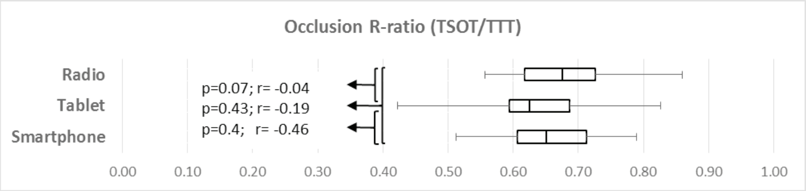

Fig. 5. R-Ratio (minimum, Q1, median, Q3 and maximum).

3.2. Eye tracking metrics while driving (AAM)

Figure 6 shows the total task on time (TTT) while driving. Again, the tablet would be the most conservative (fastest reference) with a median value of 54.56s, therefore, resulting in about 18s for one tuning.

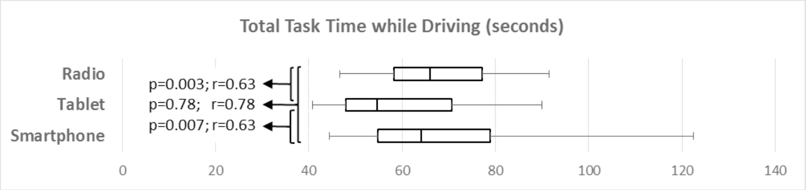

Fig. 6. Total Task Time for three sequential tuning tasks (minimum, Q1, median, Q3 and maximum).

The Total Glance Time (TGT) at the device in Figure 7 is also the smallest (most conservative) for the tablet. With a median of 30.42s, a single tuning should need around 10s glance time. Dividing the 85th percentile value of the tablet (40.85s) by 3 would yield about 14s.

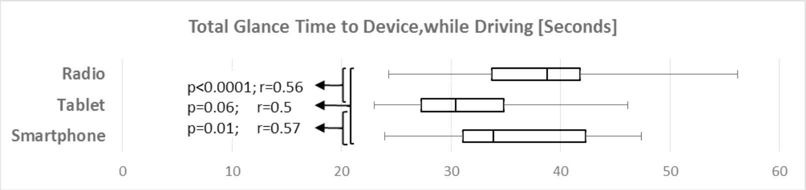

Fig. 7. Total Glance Time for three sequential tuning tasks (minimum, Q1, median, Q3 and maximum).

The number of glances at the device is reported in Figure 8. The median of the tablet is the lowest with 19.5 glances, therefore, resulting in about 6-7 glances per tuning.

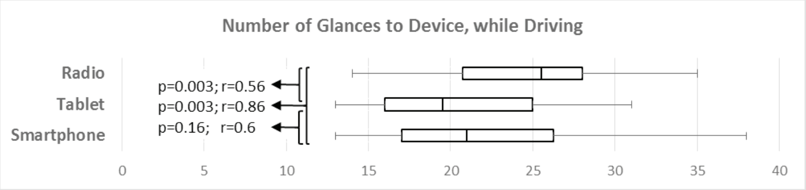

Fig. 8. Number of Glances for three sequential tuning tasks (minimum, Q1, median, Q3 and maximum).

The mean single glance duration for the three devices is about the same (1.68s median for tablet). The more important 85th percentile values calculated according to AAM would be: Radio 2.07s, Tablet 1.91s, Smartphone 2.06s; surprising high.

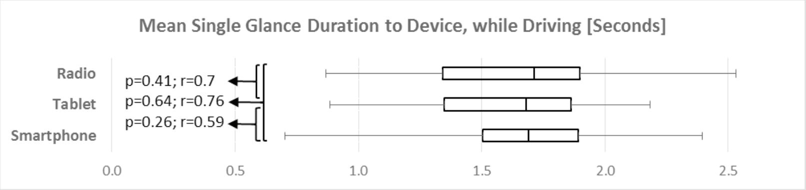

Fig. 9. Single Glance Duration (minimum, Q1, median, Q3 and maximum).

For the glance frequency (Hz), the smartphone has a slightly lower median value (0.32 Hz), indicating the gaze is directed at the devices about every third second.

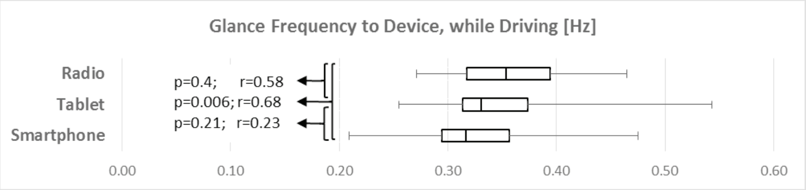

Fig. 10. Glance Frequency to the Device (minimum, Q1, median, Q3 and maximum).

3.3. Driving metrics (AAM)

For the SDLP (Figure 11) and deviation in the time headway (Figure 12), the results are roughly comparable. The LCT MDEV values (Figure 13 and 14) can also be seen as SDLP values to some extent.

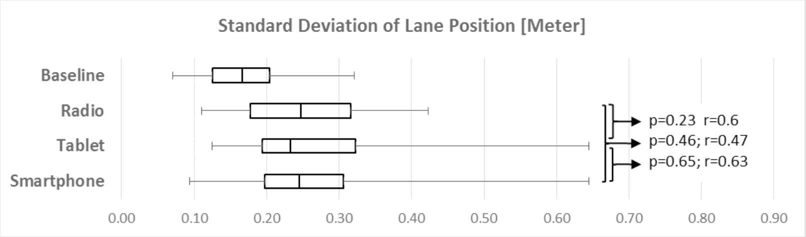

Fig. 11. Standard Deviation of Lane Position (minimum, Q1, median, Q3 and maximum).

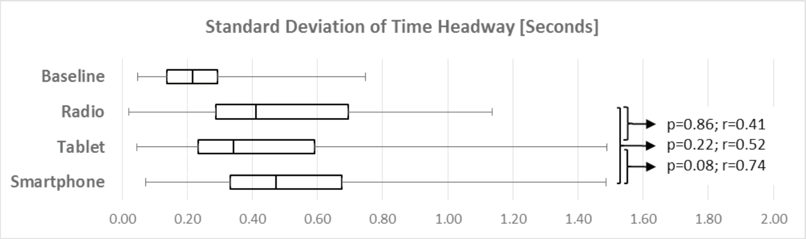

Fig. 12. Standard Deviation of tip to tail Time Headway (minimum, Q1, median, Q3 and maximum).

3.4. Driving metrics (LCT)

One person was excluded from analysis (N=23); the error-integral (adaptive MDEV 1.34m for tablet) rendered this data set an extensive outlier. The person missed a lane change, which would have been over two lanes, whereas other people only missed single lane changes. Due to the reduced LCT distance (three radio tunings), this has a high impact. The four subjects with baseline MDEV > 1.2m are not excluded. We attribute the high value to the non-enforced artificial steering behaviour and more realistic driving dynamic, and emphasize in this case the adaptive MDEV. Nevertheless, the devices yield comparable results in Figure 14 for the classic MDEV as well as for the adaptive MDEV (Figure 14).

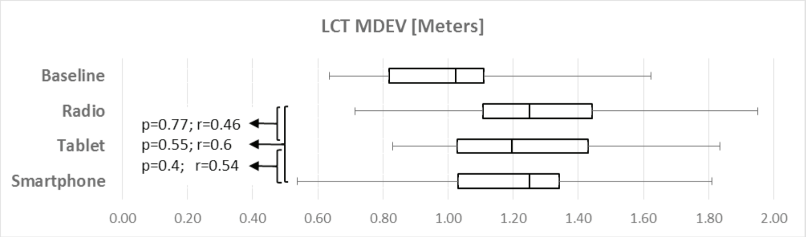

Fig. 13. Lane Change Test Mean Deviation (minimum, Q1, median, Q3 and maximum).

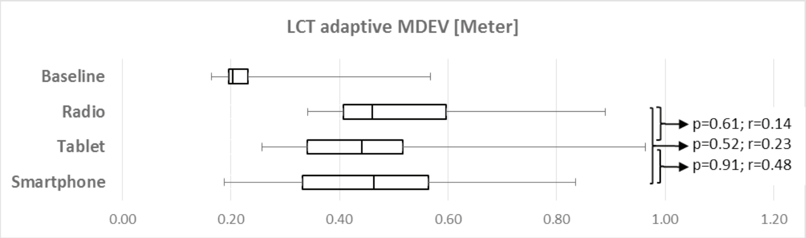

Fig. 14. Lane Change Test adaptive Mean Deviation (minimum, Q1, median, Q3 and maximum).

4. Discussion

For all major metrics, the tablet would yield comparable or even more conservative values than the radio. Three metrics (TSOT, TTT while driving and TGT) are significantly different between tablet and smartphone. The main difference between both should be the button size. Therefore, the size of the device seems to be of concern. We recommend a tablet with at least 6.5”, 800x480 and 160dpi (dots per inch) or better, mounted at around 30° below normal line of sight or higher. The radio now limits itself to 800x432 device independent pixels (dip) and the app internal specifications are now in dip as well (buttons 95dip x 95dip). ‘Dip’ is an Android concept to specify sizes for a default device with 160 dpi. If the app is used on better/bigger devices, the app is programmed to use the centre region. On smaller devices, or in portrait mode, some sizes might be downscaled. The app should be inspected visually and with a scale before used; e.g., some device manufacturers can apply zooms or slight changes.

One plea to this study could be the younger age of the test group, of about 26 years, compared to the 45-65 years of the recommended AAM test sample. First, this study is a with-in subject design and therefore best suited to detect differences between the tested devices (radio, tablet, smartphone). Additionally, in [5] we found no significant differences between these age groups when operating a touch screen pointing task while driving.

In different graphics from [2], a task time of typically about 20s could be found for button-operation of the AAM tuning while driving; this is comparable to the 18s for a single tuning in Figure 6. In [2], the TGT is typically around 8-10s for a tuning; this is comparable to the 10s in Figure 7. The mean SGD in [2] is typically about 1.2s; this seems different from the higher 1.7s of this study (Figure 9). We did a fast check with two people in real traffic on the open road and got SGDs similar to [2] with the radio app. We also used the app in another simulator experiment (currently under analysis) and found mean SGDs of 1.36s (N=19; AAM P85th SGD 1.55s); glance frequency 0.33Hz. Therefore, one hypothesis is that the phone number task, which was used to accommodate dual task settings, potentially introduced a carry-over effect toward a gaze strategy with longer glances. This procedure for accommodation was inspired by: *"The participants should get an opportunity to get experience with the dual task situations. To this end a typical secondary task which is not included in the set of secondary tasks under test can be used (e.g., "repeatedly enter a four digit number into a cellphone").*" (ISO ISO/DIS 26022:2007)

The glance frequencies seem different too. In this study, it is around 0.35Hz (Figure. 10); [2] reports about 30/min (0.5Hz) for driving on a real track and around 45/min (0.75Hz) for simulated driving. This simulated driving involved a peripheral detection task in the driving scene [2] and is, therefore in our opinion, not suited for comparisons. We assume that the driving environment (real/simulated) accounts partly for the differences in glance frequency.

5. Conclusion

The application with the modified procedure would have the benefits of taking the examiner (announcements) out of the loop, making the software the same and diminishing the influence of different hardware. As a convenient side effect, the open source availability enables modifications for experiments. An example is the control via an external Bluetooth keyboard used in [4]. Highly standardized setups exist, e.g., the Surrogate Reference Task (SuRT) and Critical Tracking Task (CTT) in ISO/TS 14198:2012, when data are compared between laboratories. If the radio task is further assessed and developed, it could be a supplement of this intentionally artificial tasks. It is fast to apply (e.g., with a car cradle) and thus would also facilitate other things, e.g., teaching driver distraction courses about eye tracking, LCT, AAM.

Further experiments with real and simulated driving would help to clarify. The use in different laboratories and setups would help especially to get a data basis, judge if the task is applicable as a reference and how it compares to AAM hardware setups.

References

[1] AAM 2006. Driver Focus-Telematics Working Group 2006, Statement of Principles, Criteria and Verification Procedures on Driver Interactions with Advanced In-Vehicle Information and Communication Systems

[2] M. Perez, J. Owens, D. Viita, A. Angell, T.A. Ranney, G.H.S. Baldwin, E. Parmer, J. Martin, W.R. Garrott, E.N. Mazzae, Radio tuning effects on visual and driving performance – Simulator and test track studies. NHTSA Report No. DOT HS 811 781. May 2013

[3] M. Krause, Manual Radio Tuning Task on Android for Driver Distraction Studies. http://go.tum.de/882480 (accessed March 23, 2015)

[4] M. Krause, N. Donant, K. Bengler, Comparing Occlusion Method by Display Blanking to Occlusion Goggles. AHFE 2015 (In press)

[5] A.S. Conti, F. Kremser, M. Krause, D.J. An, K. Bengler, The effect of varying target sizes and distances between target and non-target elements on goal-directed hand movement times while driving. AHFE 2015 (In press)