

# Change the way to manage an in-vehicle menu selection and thereby lower cognitive workload?

Nikolai Pärsch<sup>1\*</sup>, Martin Baumann<sup>2</sup>, Arnd Engeln<sup>3</sup>, Lutz W.H. Krauß<sup>1</sup>

<sup>1</sup> UI/UX Display & Interaction, Dr. Ing. h.c. F. Porsche AG, Germany

<sup>2</sup> Institute of Psychology and Education Dept. Human Factors, Ulm University, Germany

<sup>3</sup> Interaction Design and User Experience (IDUX), Stuttgart Media University, Germany

\*[nikolai.paersch@porsche.de](mailto:nikolai.paersch@porsche.de)

**Abstract:** The research reported here aims to investigate in more detail cognitive workload of in-vehicle information systems (IVIS). Various operating concepts for one specific task are tested. In detail, a function selection is implemented as a hierarchical menu selection or as a search function with text input. The text input modes are varied between speech, touch keyboard and touch gesture (handwriting) and tested with a driving simulator study. Main findings are, that cognitive workload of search via speech input is lower than the other alternatives being tested. Compared to workload of n-back levels, speech input is lower and manual interactions have a cognitive workload that is comparable to a level between 1- and 2-back tasks. Training effects are mainly observed at menu selection as well as text input by handwriting. Impact of operating errors on cognitive workload seems to be high and should be researched in further studies.

## 1. Introduction

There are several alternatives to manage a function selection within the car for designers of IVIS. Those alternatives may be divided into two groups: a selection via hierarchical menu structure or a selection via key word search and text input. For users each version seems to have its pros and cons. Lee [1] explains the main advantages of a hierarchical menu: the states of the program are displayed explicitly, so the action for the user is more recognition than recall. Search based functions on the other hand, don't need users to adapt to a certain logic but also require them to have specific keywords in mind.

The suitability of these systems for the driving context has been examined very often by measuring their *visual* workload (see e.g. Heinrich [2]). Today, there are approaches that also consider *cognitive* workload in the vehicle. Strayer et al. [3] show, that cognitive workload varies depending on task type (e.g. calling) or the mode of interaction (center stack, auditory vocal, center console).

Concerning menu and search design alternatives, there is a lot of research, which should be summarized in the following lines.

### 1.1. Menu-Driven Systems

Hierarchical menus have a long history in computer systems. Lee [1] defines menus as user-selectable data. These can be found in most technical products with graphical user interface, also in passenger vehicles. Many guidelines exist about relevant factors for designing a good hierarchical menu system. Some of the factors are stated in Norman [4], for example: "depth versus breadth", "organization of lists", "clustering" and "item meaningfulness and distinctiveness". These design factors for menus could have implications to the visual workload needed (e.g. Burnett et al. [5], Hornof et al. [6]) but also on the cognitive workload (Matsuo [7]).

Depth versus breadth addresses the question, how many items should be displayed at one page and how many pages should follow on the next layers. Burnett et al. [5] evaluate different combinations and show, that at structured menus (arranged alphabetically), breadth is favored over depth. For unstructured menus (arranged randomly), that finding applies conversely.

Organization of lists addresses the order of list entries. The adequate ordering method may differ depending on the specific use case. In short, there is alphabetic, numeric, chronological, cognitive, semantic and an ordering by frequency of use.

Clustering means the organization of list items. To find an optimal list there are two ways to cluster content: top-down or bottom-up. Regarding top-down, the designer starts with first-order categories and divides the entries step by step until he arrives at the last level. Bottom-up the designer looks at all items and clusters them by similarity, then groups them step by step into larger groups until all the groups are combined.

Item meaningfulness and distinctiveness concerns the verbalization of items. Items should transfer information but also should be distinct to each other. As an addition the use of graphics e.g. icons could be suitable to solve these issues at some points.

### 1.2. Search-Driven systems

Search driven concepts are also well-known in technical systems, e.g. the probably most common example: the google search. Users provide keywords to retrieve their desired information or get to the desired stage in the interactive system. The search function always consists of a text input, where keywords could be typed in. The way of carrying out the text input, often differs between the context of use and operating device. In the vehicle the most common types of text input are done by rotating wheel, by touchscreen keyboard, by handwriting gesture on a touchpad as well as a text input via speech.

Graf et al. [8] show the suitability for this kind of search function in the context of IVIS. They compare two kinds of search functions: a quick search, where users can freely type in search terms and a categorical search, where users narrow down their search results by choosing a corresponding category. They analyze, that the search approach seems to be equally suitable or even superior to menu driven interaction. But how about different ways to carry out a text input? This comparison is made for instance in Kujala et al. [9]. They compare touch keyboard, handwriting and text input by voice recognition by their workload. Voice recognition shows the lowest values, followed by keyboard and handwriting input. Haslbeck et al. [10] also compare touch keyboard and handwriting amongst other modalities and find in addition several factors, that influence workload during driving, for example the interruptibility and the size of touch areas or handwriting input.

### 1.3. Research questions

To put it in a nutshell, a lot of research has been carried out on both domains: menu-driven and search-based systems. Guidelines exist about important factors, that influence the quality of each approach. Comparing these two approaches, less research could be found. Especially when focusing on cognitive workload, no study results are available that compare these different approaches to manage a menu selection in the vehicle. This research gap will be addressed in the following. The research reported here aims to investigate in more detail cognitive workload of a hierarchical menu selection and a search function with text input. The text input modes are varied between speech, touch keyboard and touch gesture (handwriting). The driving simulator study (conducted in December 2017) proves, if there are differences of these operation variations concerning cognitive workload.

## 2. Method

### 2.1. Subjects

Participants were recruited by newsletter for all employees of Porsche AG at Weissach, Germany in December 2017. In sum 36 persons participated in the study, all persons had no connection to IVIS development. Ten cases were excluded because of simulator sickness or data logging issues

The final sample consisted of 26 persons, 17 males and 9 females. One person was below 25 years old, seven participants were between 25-39, 15 between 40 and 55 years and 3 persons were beyond 55 years.

Concerning experience with the interaction methods analyzed, participants were well experienced with touchscreen and touchscreen keyboard interaction. Speech interaction was used more rarely and text input by handwriting was mostly unknown to our participants.

### 2.2. Apparatus

The experiment was conducted in the driving simulator of Porsche AG with motion dynamics. The mock-up was equipped with two stacked touchscreens in the center

console. The lower screen was used for text input by touch gesture, other interactions were executed on the higher screen. The IVIS software prototype was especially programmed for this experiment. Touch gesture input was processed by automatic text recognition, speech input recognition was realized as a Wizard-of-Oz approach directed by a research assistant.

Gaze data (Dikablis Professional binocular eye-tracker), driving data and IVIS events were collected with a 60 Hz sampling rate and were logged within the D-Lab 3.45 software suite (time synchronized).

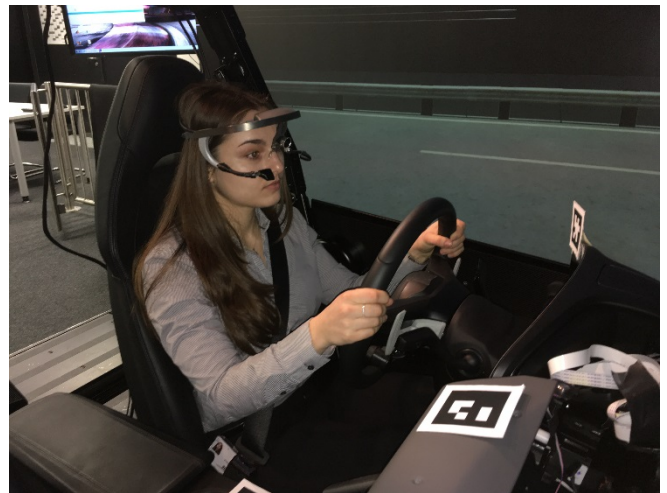


Fig. 1. Impression of setup and eye-tracking measurement

### 2.3. Tasks & procedure

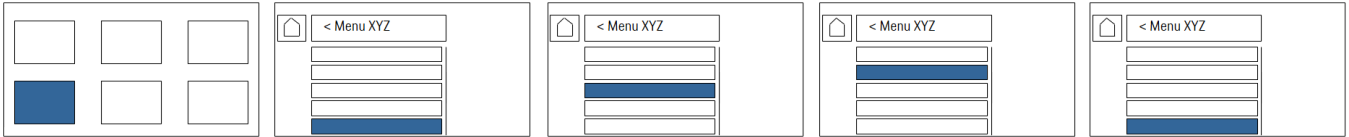
As primary task, the participants were driving on a three-lane German highway, following a lead-vehicle. The lead-vehicle travelled with a speed varying between 65 and 75 mph. Participants were instructed to keep a constant distance between the leading car (similar driving task in Large et al. [11]).

The secondary tasks were arranged in two blocks: n-back tasks and IVIS tasks. In this form of n-back tasks, digits were presented auditory and the delayed response of the participant was carried out verbally. The higher the delay of the recall task, the higher was the cognitive workload. For further information of the n-back tasks please see Mehler et al. [12]. The n-back tasks were used to generate benchmark data to compare with the IVIS tasks. Three different levels were used: 1-back, 2-back and 3-back. For this experiment the translated version and audio files by the Chair of Ergonomics, Technical University of Munich, were used.

The IVIS tasks consisted of four different approaches to manage a menu selection: search via a menu hierarchy, search via auditory vocal text input, search via text input over keyboard on touchscreen and search via handwriting text input gesture on touchscreen. An example for a task is "Please change the interior lighting color to blue". For exemplary procedure please see Figure 2.

The menus were developed under consideration of the presented guidelines in the instruction. Menus were created with a cognitive perspective (according areas in the car) and items were sorted with respect to expected frequency of use. Five menu items were presented per page, allowing a

## Search via menu hierarchy



## Search via text input: speech, touch keyboard, touch handwrite gesture



**Fig. 2.** Exemplary Procedure

reasonable touch area size. Items per menu level varied between two and sixteen entries. However, use case items were always shown on the first or second page. Regarding menu depth, the final item selection was always on the fifth and last level of the menu. Overall, menu selection use cases took 5 operating steps.

The quick search by text input use cases were constructed as follows: users selected the main menu by touch, then started entering characters (by keyboard or handwriting). After two characters were entered, the result was presented in the result list on the right side of the screen. When selected, the second to last menu page was shown and the last two items had to be selected. Overall, operating steps were comparable to those in the menu hierarchy (five steps). When entering text by speech, users had to tap the microphone button above the keyboard and then proceeded with speech input. When providing a right keyword, the research assistant forwarded the screen to the desired menu (interaction of research assistant hidden from participants). All in all, this interaction consisted also of five operating steps.

The procedure of the experiment started with a training phase to get used to the n-back and IVIS tasks. After a 5-minute test drive without secondary tasks, four blocks of secondary tasks followed: Block A with n-back tasks, Block B1 with IVIS tasks, Block B2 with a repetition of the IVIS tasks and Block B3 with a second repetition of one of the IVIS tasks. Between the subjects, Block A and B and the tasks within the Blocks were in randomized order. Between the tasks there were recovery phases without secondary tasks. The experiment had a duration of approximately 75 minutes.

### 2.4. Data analysis

Cognitive workload was measured by three different types of measurement: physiological data, performance metrics and subjective ratings. (O'Donnell and Eggemeier [13])

Regarding physiological data, blink-related measures (Marquardt et al. [14]) were recorded. However, due to several data-logging issues, this data is not part of the analysis. In order to measure performance within the primary task, driving data was observed. The standard deviation of distance to the lead vehicle and the standard deviation of lane position was measured (Rauch & Gradenegger [15]). Concerning secondary task performance, error-rate, number and duration

of IVIS interaction events were measured. Error-rates were calculated as follows: the optimum count of operating steps was subtracted from the overall count of operating steps at this task. This balance was divided by the optimum count of operating steps to form the final error-rate. Regarding the subjective ratings, the mental dimension of the NASA TLX (Hart & Staveland [16]) was used.

To analyze differences between n-back tasks, IVIS modes and IVIS repetitions, non-parametric tests (Wilcoxon) were executed and can be found in the appendices.

In order to explore differences between interaction modes but without the effect of operating errors, a subset of error-free interactions was created. Therefore, only those tasks were considered, that had the lowest possible value of operating steps (in number 5).

## 3. Results

Results can be split up according to three different research questions: 1) comparing the cognitive workload of the different interaction methods to select a function; 2) examining differences in training effects of tested alternatives; 3) analysing the effect of fault tolerance and regarding faultless executions of use cases.

### 3.1. Cognitive workload of interaction methods

Table 1 presents the results of the n-back tasks as well as the first cycle of interaction use cases. As cognitive workload measurements, the NASA TLX mental dimension, the variability of lane position and distance keeping and error rates are reported.

NASA TLX values and error-rates seem to be quite robust indicators for the increase in cognitive workload regarding the three n-back levels, as can be seen in Table 1 and Table 3. There are significant differences between level 1 and 2 and between level 2 and 3. The variability of lane position and distance keeping on the other hand are not showing a linear increase over these three levels. 1-back and 3-back have comparable variabilities whereas 2-back shows a lower variability of these two metrics. This result should be considered when interpreting data from these measurements. Concerning the first cycle of interaction use cases, cognitive workload when searching via speech text input is

Task	N	NASA TLX [mental]		Distance	Lane position	Error-Rate	
		Mean	SD	SD	SD	Mean	SD
1-back	23	6.4	2.6	45.8	0.18	0.05	0.09
2-back	25	12.6	4.0	35.9	0.26	0.14	0.14
3-back	25	15.8	4.0	47.1	0.19	0.31	0.27
Menu 1	26	9.5	5.1	33.7	0.26	0.45	0.61
Menu 2	26	7.4	4.5	22.3	0.19	0.35	0.69
Menu 3	7	10.7	6.6	30.6	0.28	0.30	0.41
Keyboard 1	26	8.5	4.1	27.4	0.29	0.55	0.66
Keyboard 2	26	7.0	4.0	19.3	0.24	0.52	1.01
Keyboard 3	6	5.6	4.0	17.2	0.21	0.10	0.17
Gesture 1	26	8.3	4.7	38.8	0.31	1.57	1.70
Gesture 2	26	6.5	3.5	29.3	0.25	0.33	0.55
Gesture 3	7	5.0	2.7	23.4	0.25	0.00	0.00
Speech 1	26	5.6	4.2	23.2	0.24	0.01	0.20
Speech 2	26	4.5	2.6	21.6	0.20	0.00	0.13
Speech 3	6	5.7	2.6	15.2	0.14	0.06	0.13

**Table 1.** Cognitive workload of interaction methods

significantly lower than the workload while performing the other IVIS interactions. This difference is shown by the subjective measurement and the error-rates, as well as partly by the variability of the distance (differences between SDS and HWR, Menu). Regarding the remaining variants, especially input by handwriting seems to be more cognitive demanding due to its higher error rate, that also results in a higher variability of distance keeping. During the second cycle of interaction use cases, results from first cycle remain mainly constant: speech interaction is significantly less demanding regarding the subjective measurements and the error-rates. Concerning the variability of the distance, handwriting shows more variability than input by touch keyboard.

### 3.2. Comparing cognitive workload of n-back and IVIS

N-back tasks are useful to interpret the measurement values of cognitive workload. It is known, that 1-back represents a low to moderate cognitive workload whereas 2-back usually represents a higher workload. Compared to the n-back tasks there are following results (see Table 4).

Interaction via speech is cognitively less demanding than both N-Back levels. This is reported by subjective measurements, distance keeping variability and partly error rate. Lane keeping variability however is significant higher than at the 1-back task.

The cognitive workload of interaction with the menu hierarchy seems to be between the workload of the 1-back and 2-back levels. Concerning first round of interaction, subjective workload is higher than 1-Back and lower than 2-back. The variability of lane position is lower but the error-rate higher than at both n-back tasks. The second round of interaction has a lower workload: subjective workload is comparable to 1-back and lower to 2-back, variability of distance keeping is lower, variability is lower than 1-back and comparable to 2-back and error-rate is higher than 1-back and comparable to 2-back.

The cognitive workload of keyboard text input is more hardly to interpret, because results between the measurement methods are not homogeneous. First round of interaction is subjective more demanding than 1-back and less demanding than 2-Back. Variability of distance is lower than both n-back

levels, variability of lane position and error rates are higher than both levels. The workload of the repetition is comparable to 1-back and less demanding than 2-back. Variability of distance keeping is lower than both levels and variability of lane position and error rates are comparable to 2-back.

Regarding workload of handwrite recognition, the first round of interaction, subjectively workload is comparable to 1-back, variability of distance keeping is comparable to both levels and variability of lane position and error rates are higher than both levels. The second round of interaction with handwrite interaction shows similar results. Only variability of lane position is now comparable to 2-back and error rates are comparable to both levels.

Cognitive workload of all IVIS interactions seems to be below the workload of 3-back tasks. This is shown by subjective measurements and variability of distance.

### 3.3. Training effects concerning cognitive workload

Interactions with the IVIS were repeated for two times, the cognitive workload measurements for these trials are shown in Table 1 and Table 5. Between first and second trial, cognitive workload decreases especially at the menu interaction and the handwriting task. Regarding the menu task, there are significant decreases at the subjective measurement, as well as the variability of distance and lane position. Concerning handwriting, the differences of the subjective measurement and the error rates are significant lower. Regarding the touch keyboard task, there are only significant decreases at the lane position variability. Speech interaction on the other hand shows no significant differences and remains mostly on the same level.

Regarding trials two and three, cognitive workload seems to rise at some tasks, but these differences are not significant (could be due to small sample size in third trial). There is only one significant drop in subjective cognitive workload within the keyboard task repetitions.

### 3.4. Comparing error-free trials

Text-input by speech was mostly error-free due to its wizard-of-Oz approach. In order to focus on the differences on the way of interaction and not on the error-rate, results presented here, are only focusing on error-free trials (Table

2). Due to the unfamiliarity with the system, numerous errors occurred especially in the first trial. Therefore, results are presented for the second trial with a higher sample size and statistical significant differences (Table 6): speech is less demanding than menu and the keyboard task.

Task	N	NASA TLX [mental]	Distance	Lane position
		Mean	SD	SD
1-back	16	6.1	49.4	0.20
2-back	6	12.1	35.1	0.11
3-back	0	-	-	-
M2	11	6.8	15.8	0.15
K2	8	6.4	15.3	0.20
G2	7	5.0	22.9	0.16
S2	19	4.7	20.9	0.20

**Table 2.** Cognitive workload of error-free second trials (M=Menu, K=Keyboard, G=Gesture, S=Speech)

Concerning variability of distance keeping there are no significant differences, concerning variability of lane position there are significant differences between second trial of menu and speech tasks.

#### 4. Discussion and Conclusions

The study examines the cognitive workload of several methods to manage a function selection with an IVIS: selection via menu hierarchy or selection via search and text input. Text input methods are varied between speech, touch keyboard and touch gesture handwriting. Results show, that cognitive workload of the search function with text input mode via speech is lower than of the remaining variants. Strayer et al. [3] also found a difference between voice interaction and interaction via center console, which supports this finding.

There seem to be no differences in cognitive workload between the haptic interactions presented in this experiment, although conceptual differences seem to be quite large. Only handwriting input is partly more demanding, especially due to its significant higher error rate.

Comparing workload of n-back tasks and IVIS tasks, speech interaction is less demanding than both n-back levels. Interaction via menu hierarchy has a workload between 1-back and 2-back. A comparison of handwriting input and keyboard input with n-back levels is difficult, because measurement methods are varying strongly. When looking at subjective measurements, keyboard input as well as handwriting input is comparable to the workload of 1-back.

Regarding training effects of IVIS interactions, cognitive workload of menu interactions and handwriting interactions are decreasing significantly. Maybe hierarchical menus need some training to know more about the logic of the structure and touch gesture inputs need some training how characters can be recognized by the system. The implemented speech task was quite fault-tolerant because of its Wizard-of-Oz approach. Finally, text input by touch keyboard is often used by smartphone users, which could explain, that there are not so many training effects concerning these variants.

An additional analysis of error-free task trials shows, that the impact of operating errors on the cognitive workload should be considered. Workload between the alternatives is more aligned in comparison to trials that includes errors. When discussing this topic, it should be kept in mind, that fault tolerance also is often a characteristic of an interaction alternative. The freedom of design when creating menus is more dynamic and leaves a higher risk of decreasing fault tolerance than solely technical implementations of text input methods.

Nevertheless, the impact of operating errors on the cognitive workload should be addressed in further studies. An example would be a fault-tolerance-factor compared to the well-known age-factor in key-stroke modellings.

#### 5. References

- [1] Lee, E. S., & Raymond, D. R. (1993). Menu-driven systems. *Encyclopedia of Microcomputers*, 11, 101-127.
- [2] Heinrich, C. (2013). Fighting driver distraction: worldwide approaches. In *Proceedings of the 23rd International Technical Conference on the Enhanced Safety of Vehicles (ESV)*.
- [3] Strayer, D., Cooper, J., Goethe, R., McCarty, M., Getty, D., Biondi, F. (2017). *Visual and Cognitive Demands of Using In-Vehicle Infotainment Systems*. Salt Lake City, UT: University of Utah.
- [4] Norman, K. L. (2008). Better design of menu selection systems through cognitive psychology and human factors. *Human factors*, 50(3), 556-559.
- [5] Burnett, G. E., Lawson, G., Donkor, R., & Kuriyagawa, Y. (2013). Menu hierarchies for in-vehicle user-interfaces: Modelling the depth vs. breadth trade-off. *Displays*, 34(4), 241-249.
- [6] Hornof, A. J., & Halverson, T. (2003, April). Cognitive strategies and eye movements for searching hierarchical computer displays. In *Proceedings of the SIGCHI conference on Human factors in computing systems* (pp. 249-256). ACM.
- [7] Takashi, M. (2008). Investigating the cognitive loads involved in searching hierarchical menus using eyeblinking as an index. *The Japanese Journal of Cognitive Psychology*, 6(1) 1-10.
- [8] Graf, S., Spiessl, W., Schmidt, A., Winter, A., & Rigoll, G. (2008, April). In-car interaction using search-based user interfaces. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems* (pp. 1685-1688). ACM.
- [9] Kujala, T., & Grahn, H. (2017, September). Visual Distraction Effects of In-Car Text Entry Methods: Comparing Keyboard, Handwriting and Voice Recognition. In *Proceedings of the 9th International Conference on Automotive User Interfaces and Interactive Vehicular Applications* (pp. 1-10). ACM.

[10] Haslbeck, A., Popova, S., Krause, M., Pecot, K., Mayer, J., & Bengler, K. (2011, July). Experimental evaluations of touch interaction considering automotive requirements. In International Conference on Human-Computer Interaction (pp. 23-32). Springer, Berlin, Heidelberg.

[11] Large, D. R., Burnett, G., Anyasodo, B., & Skrypchuk, L. (2016, October). Assessing Cognitive Demand during Natural Language Interactions with a Digital Driving Assistant. In Proceedings of the 8th International Conference on Automotive User Interfaces and Interactive Vehicular Applications (pp. 67-74). ACM.)

[12] Mehler, B., Reimer, B., & Dusek, J. A. (2011). MIT AgeLab delayed digit recall task (n-back). Cambridge, MA: Massachusetts Institute of Technology.

[13] O'Donnell, R. D., & Eggemeier, F. T. (1986). Workload assessment methodology.

[14] Marquart, G., Cabrall, C., & de Winter, J. (2015). Review of eye-related measures of drivers' mental workload. Procedia Manufacturing, 3, 2854-2861.

[15] Rauch, N., & Gradenegger, B. (2007). Das Konzept des Situationsbewusstseins und seine Implikationen für die Fahrsicherheit. FAT-Schriftenreihe, (210).

[16] Hart, S. G., & Staveland, L. E. (1988). Development of NASA-TLX (Task Load Index): Results of empirical and theoretical research. Advances in psychology, 52, 139-183.

## 6. Appendices

**Table 3.** Results of wilcoxon tests comparing cognitive workload of interaction methods (M=Menu, K=Keyboard, G=Gesture, S=Speech)

Task	NASA TLX [mental]	SD Distance	SD Lane position	Error-Rate
<b>n-back</b>				
1 - 2	.000	.784	.248	.004
2 - 3	.000	.391	.214	.005
<b>IVIS 1</b>				
M – K	.354	.469	.517	.537
M – G	.646	.166	.353	.028
M – S	.002	.038	.292	.003
K – G	.852	.058	.648	.044
K – S	.010	.367	.080	.001
G – S	.003	.001	.269	.000
<b>IVIS 2</b>				
M – K	.722	.585	.166	.474
M – G	.569	.115	.657	.948
M – S	.000	.829	.778	.023
K – G	.852	.030	.957	.447
K – S	.000	.620	.191	.001
G – S	.000	.264	.326	.004

**Table 4.** Results of wilcoxon tests comparing cognitive workload of interactions methods and N-Back tasks (M=Menu, K=Keyboard, G=Gesture, S=Speech)

Task	NASA TLX [mental]	SD Distance	SD Lane position	Error-Rate
<b>1-back vs. IVIS 1</b>				
M	.009	.130	.000	.002
K	.025	.018	.000	.001
S	.124	.001	.003	.396
G	.137	.304	.001	.002
<b>1-back vs. IVIS 2</b>				
M	.533	.002	.006	.050
K	.737	.000	.001	.007
S	.015	.003	.002	.108
G	.879	.114	.003	.089
<b>2-back vs. IVIS 1</b>				
M	.011	.276	.032	.011
K	.004	.006	.023	.005
S	.000	.001	.069	.028
G	.002	.657	.020	.002
<b>2-back vs. IVIS 2</b>				
M	.000	.007	.192	.520
K	.000	.000	.074	.126
S	.000	.003	.174	.349
G	.000	.241	.162	.005
<b>3-back vs. IVIS 1</b>				
M	.000	.150	.000	.338
K	.000	.002	.000	.306
S	.000	.002	.002	.000
G	.000	.600	.000	.007
<b>3-back vs. IVIS 2</b>				
M	.000	.002	.004	.436
K	.000	.000	.001	.475
S	.000	.002	.002	.531
G	.000	.022	.003	.000

**Table 5.** Results of wilcoxon tests comparing cognitive workload of IVIS repetitions (M=Menu, K=Keyboard, G=Gesture, S=Speech)

Task	NASA TLX [mental]	SD Distance	SD Lane position	Error-Rate
<b>IVIS 1 vs. 2</b>				
M	.012	.013	.009	.145
K	.062	.073	.034	.361
S	.107	.551	.292	.763
G	.038	.242	.074	.006
<b>IVIS 2 vs. 3</b>				
M	.715	.735	.397	.109
K	.027	.345	.600	.285
S	1.000	.753	.463	.317
G	.068	.686	.686	.655

**Table 6.** Results of wilcoxon tests comparing cognitive workload of IVIS second trial, error-free (M=Menu, K=Keyboard, G=Gesture, S=Speech)

Task	NASA TLX [mental]	SD Distance	SD Lane position
<b>IVIS 2</b>			
M – K	.180	.686	.500
M – G	.180	.655	.180
M – S	.042	.953	.028
K – G	.655	.655	.655
K – S	.016	.779	.401
G – S	.197	.753	.515