

Assessment of cognitive workload of in-vehicle systems using a visual peripheral and tactile detection task setting

Klaus Bengler^{a*}, Martin Kohlmann^a, Christian Lange^b

^a*Institute of Ergonomics, Technische Universität München, Boltzmannstr. 15, D-85747 Garching, Germany*

^b*Ergoneers, GmbH, Mitterstraße 12, D-85077 Manching, Germany*

Abstract. The increase of driver information and infotainment systems includes also interaction technologies like speech interaction that minimize visual-manual demand and put the focus to cognitive demand. The question is whether this could lead to distraction effects and decreased traffic safety. This study presents an evaluation method for cognitive demand based on different detection paradigms in a dual task setting. A listening and a backward counting task are realized on three difficulty levels as simulations of cognitively loading secondary tasks and investigated using a visual versus a tactile detection paradigm. The results show that both detection paradigms are able to discriminate the task levels and that subjects successfully apply compensation strategies in the dual task setting especially during the listening task.

Keywords: Detection Response Task, Driver Distraction, Cognitive workload, Mental workload, Peripheral Detection, Tactile Detection

1. Introduction

The increase of driver information and infotainment systems, that in many cases have benefits for traffic safety like e.g. navigation systems, often motivates the question whether they could lead to distraction effects and a decrease of traffic safety if used while driving. While numerous projects like CAMP, HASTE, AIDE and ADAM ([1],[2],[3],[4]) already addressed the assessment of visual-manual demand, this study focuses on the evaluation of cognitive demand. A valid assessment of cognitive demand would be helpful to support the developer in the layout of these interaction concepts, as speech interaction and the processing of vocal information increases [5]. Based on the multiple resource theory of attention one could assume that negotiation of a cognitive task should have an effect on a simultaneous primary task in a dual task setting [6] [7]. Still the question is whether drivers are able to compensate for resource limitations by task switching strategies.

Already [8] and [9] showed the potential of detection paradigms to assess cognitive secondary in-vehicle tasks following the multiple resource model.

The method is based on the idea of limited resources and the hypothesis that increased workload leads to a reduction of the driver's field of view, as well as decelerated and more selective and attentive processing. [7]

In the sense of the European Statement of Principles [10] this methodological development is necessary to enable the assessment of interaction concepts during the development process in a reliable, valid and efficient way.

2. Material and Methods

For a subject-based study the experimental setup of [8] was partially replicated. It consisted of a tactile detection task and a visual peripheral detection task.

24 subjects had to react to a visual stimulus (PDT) realized by five red LEDs (0°, symmetrically 11° and

* Corresponding author: bengler@ife.mw.tvm.de

23°) one meter in front of the test person 3° below their line of sight or to a small electrical vibrator at the wrist of the subdominant hand of the test person (TDT). Responses were given in both conditions by pressing a switch held in the dominant hand of the test person.

21 subjects were male, three were female. The average age was 30.8 (StdDev 10.3). There were 22 right-handed subjects.

The experimenter was present in the laboratory room during the whole experimental session and gave pre-defined instructions to each participant.

After being seated in front of the PDT setup the seat was adjusted so that the stimuli were 3° below and a yellow dot on the subject's line of sight. The instructions for the PDT were as follows: "Please fixate the yellow dot in the middle and press the switch instantly after you notice a signal."

The instructions for the TDT were as follows: "Please press the switch instantly after you feel a vibration."

One run (PDT or TDT) took about two minutes. After each run the subjects filled out a NASA-TLX questionnaire. Each subject performed a baseline measurement as single task condition for PDT, TDT, counting and listening each.

Each task was introduced by a short explanation and training.



Fig 1. Experimental setup for the PDT condition with the central yellow fixation stimulus being visible. The OCR patterns are fixed for additional eyetracking recordings but are not necessary for the PDT



Fig 2. Experimental setup for the TDT condition with the vibration unit fixed to the subjects forearm



Fig 3. Handheld response button

The interstimulus interval was 3-5 seconds in both conditions. The stimulus duration was varied between 1-2 seconds for the PDT and was 1 second for the TDT. While being engaged in one of the detection tasks the subjects had to perform a counting task on three task levels (C1-C3): counting forward by 1, forward by 7 and backward by 6 and 7 alternating. The experimenter recorded any error in the counting task in the protocol.

The second task under investigation was a listening task which simulated cognitively loading secondary in-vehicle tasks. Subjects had to listen to a person who had been recorded while counting using the same task levels (L1-L3) as for the counting task (up by 1, up by 7 and down by 6 and 7 alternating). The recording included errors that should be noted by the subject saying "error".

The conditions PDT vs. TDT and counting vs. listening were randomized with the baseline conditions.

Following this experimental plan all subjects conducted counting and listening on three difficulty levels under PDT and TDT conditions under dual task conditions plus a baseline for any of the tasks under single task conditions.

The dependent variables were response times and misses for the detection tasks and counting performance and error detection rate in the secondary tasks.

3. Results

The analysis of this dual task experiment is divided into the analysis of the primary task performance (PDT vs. TDT) and the secondary task performance (counting vs. listening).

The frequency of counting errors shows clearly the difference between the different task levels and gives a good indication that subjects focused onto the operation of the counting task as there is no significant difference to the baseline. This is different for the listening task where the baseline condition differs significantly in the L2 situation from the dual task conditions.

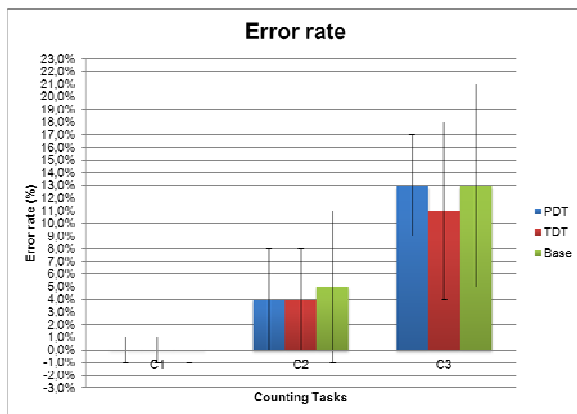


Figure 4. Error rates during counting task (level C1 to C3) in PDT, TDT and Baseline condition (+/- 1StdDev)

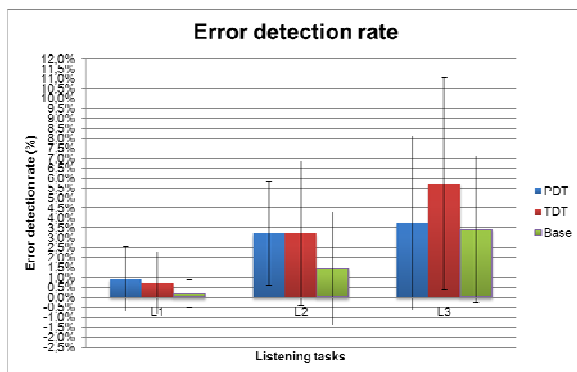


Figure 5. Error rates during listening task (level L1 to L3) in PDT, TDT and Baseline condition (+/- 1StdDev)

In all conditions the hit rate in the detection tasks was above 70% which was set as a quality criterion for the overall task performance. The analysis of the PDT and TDT condition shows a significant difference in reaction time between the three difficulty levels of the counting task ($p < 0.05$). This significant main effect is comparable for PDT (PDT-C) as well as for the TDT (TDT-C). This significant difference is also detectable for the comparison to the baseline.

There is no significant difference in reaction time for the listening task under TDT conditions (TDT-L) except for the comparison to the baseline reaction time. For the PDT condition there is a significant difference between L1 and L3 and between all levels to the baseline ($p < 0.05$).

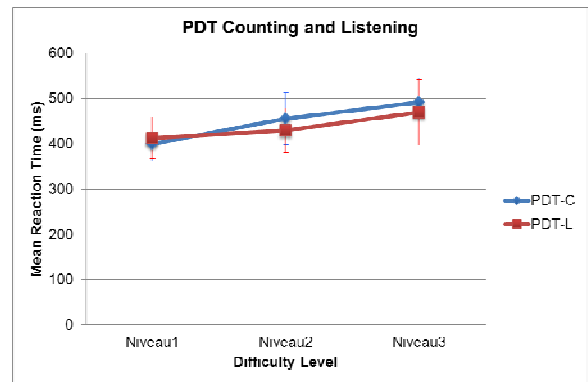


Figure 6. Mean reaction times for all three levels of counting (PDT-C) and listening task (PDT-L) under PDT condition (+/- 1StdDev)

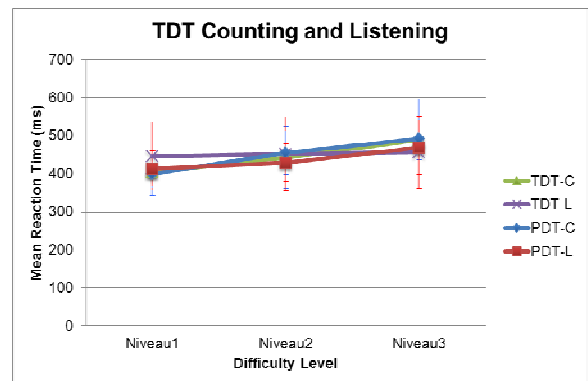


Figure 7. Mean reaction times for all three levels of counting (TDT-C) and listening task (TDT-L) under TDT condition (+/- 1StdDev)

In figures 8 and 9 it can be seen that the effects of the dual task paradigm on reaction times are comparable for both detection tasks.

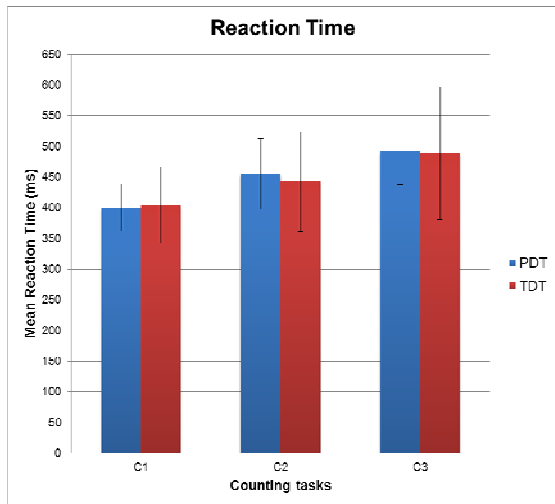


Figure 8. Mean reaction times for levels C1-C3 under PDT and TDT condition (+/- 1StdDev)

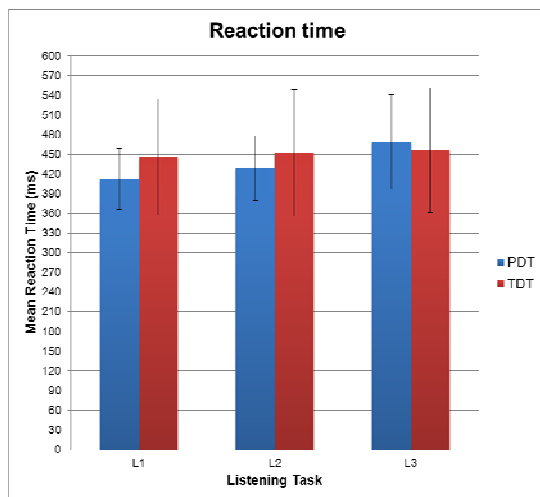


Figure 9. Mean reaction times for levels L1-L3 under PDT and TDT condition (+/- 1StdDev)

4. Discussion

This study shows that experimental detection paradigms are able to discriminate in a dual task setting secondary tasks that differ in their cognitive demand. For the tactile detection task this is a very good replication of Merat et al. Moreover it shows that a visual peripheral detection paradigm is a valid setup to test cognitive workload and might be even more sensitive to “passive” listening tasks. Reaction time proved to be a sensitive parameter for the evaluation of differ-

ent task difficulties. The results are valued as important input to current standardization activities.

The analysis of reaction times during the listening task shows that this metric is only of limited value. This might be due to the fact that in comparison to the “active” counting task (user paced) that “passive” listening task (system paced) requires different resources and indicate that subjects successfully applied compensations strategies in the listening condition, i.e. slowing down working on the listening task while being active in the detection task.

Further experiments should be conducted to evaluate secondary tasks that are more natural like listening and speaking in natural dialogs or interacting with devices that include cognitively loading features like speech interaction to adapt this methodology for the evaluation of interaction concepts and devices that are intended for use while driving. Moreover the study will now be replicated in a dual task driving or driving simulator study to correlate the detection rate to relevant performance parameters of the driving task and a triple task setting to check for the compatibility of detection tasks with driving-like settings.

References

- [1] L. Angell, et al., “Driver workload metrics project task 2, final report,” National Highway Traffic Safety Administration (NHTSA), Tech. Rep. DOT HS 810 635, 2006.
- [2] Carsten, O., Merat, N., Janssen, W., Johansson, E., Fowkes, M., and Brookhuis, K., HASTE Final Report, Contract No. GRD1/2000/25361 S12.319626, Human Machine Interface and the Safety of Traffic in Europe (HASTE) Project, 2005.
- [3] Martin BAUMANN, Diana RÖSLER, Georg JAHN, and Josef F. KREMS Assessing Driver Distraction using Occlusion Method and Peripheral Detection Task
- [4] Breuer, J., Bengler, K., Heinrich, C., & Reichelt, W. (2003). Development of advanced driver attention metrics (ADAM). In H. Strasser, K. Kluth, H. Rausch, & H. Bubb (Eds.). Quality of Work and Products in Enterprises of the Future. Proceedings of the 50th – Anniversary Conference of the GfA and the XVII Annual ISOES Conference in Munich, 2003 (S. 37-40). Stuttgart: Ergonomia.
- [5] Bengler, K. (2000) Automotive Speech-Recognition - Success Conditions Beyond Recognition Rate. Proc. LREC 2000, Athens, (pp. 1357-1360)
- [6] Wickens, C. D. (1984). Processing resources and attention. In R. Parasuraman & R. Davies (Eds.), Varieties of Attention . New York: Academic Press.
- [7] Wickens, C. D., & Hollands, J. G. (1999). Engineering psychology and human performance. (Third ed.). Upper Saddle River, NJ: Prentice Hall.
- [8] Merat, N. and Jamson, A.H. 2008. The Effect of Stimulus Modality on Signal Detection: Implications for Assessing the Safety of In-Vehicle Technology. Human Factors, Vol. 50, No. 1, pp. 145–158.
- [9] Engström, J.; Aberg, N.; Johansson, E. & Hammerbäck, J. (2005). Comparison between visual and tactile signal detection

tasks applied to the safety assessment of In-vehicle information systems (IVIS). Proceedings of the Third International Driving Symposium on Human Factors in Driver Assessment, Training and Vehicle Design. Rockport, Maine, 2005.

[10] European Statement of Principles on the Design of Human Machine Interaction (ESoP, 2005): Draft, European Commission Information Society and Media Directorate-General – G4 ICT for Transport, 2005.