
CONTENT-BASED INTERACTION DESIGN – AN EXAMPLE FROM SOLVING INTERACTION PROBLEMS IN CARS

P. Saariluoma and T. Kujala

Keywords: HMI, content-based analysis, capacity-based analysis, interaction design, In-Vehicle Information Systems

1. Introduction

Human machine interaction (HMI) design is a problematic area for design research. The lack of real unity in this area is a well-known fact, although a number of different paradigms and styles have been developed. One may, for example, speak of human device interaction, human-computer interaction, usability, ergonomics, and human factors [ISO 9241-11 1998, Kantowitz & Sorkin 1983, Nielsen 1993, Preece et al. 1994]. There is thus a whole family of approaches, which open in some sense different but also overlapping ways of meeting design problems.

Another division in the field is formed by machine- and human-driven approaches. Usability engineering is a typical machine-driven approach. In usability engineering, designers refine the use and usability of existing or emerging technologies [Rosson and Carroll 2002]. This type of design is highly dependent on the development of technological innovations. The problem is that users are often made to adapt to technological environments rather than that the developing understanding of users would direct the technical design.

The opposite pole to machine-driven design can be called user centered design [Norman and Draper 1986]. The key idea in this view is that the study of users, their needs and capacities, is the prime moving force of design. Even here one can find a number of possible ways of viewing things. A popular method is to begin with activities and the social interaction of users. Another way of looking at things can be based on psychological analysis of human mentality, i.e., motives, emotions, cognitive limitations and socio-cultural interaction [Saariluoma 2004]. This approach has been termed user psychology [Moran 1981, Oulasvirta and Saariluoma 2004, 2006, Parkkola, Saariluoma and Berki, in press, Saariluoma 2004].

In developing user psychological analyses, one must meet a number of important challenges. One of these is to find explanatory concepts, which can best explicate the underlying mental processes [Saariluoma 2004]. Different aspects of human behavior can be explained with different types of concepts. If a person has lost her ability to speak, we look for neural damages. If a person cannot speak a foreign language fluently, we normally improve training. The explanations for these two types of deficits are different, and subsequently, so are our conclusions.

In the investigation of modern human machine interaction design, a very powerful explanatory basis has been the limited information processing capacity of the human. This mostly refers to either attention or working memory [Baddeley 1986, Cowan 2000]. The basic idea in capacity-based thinking is that the situational demands exceed the limited capacity of the human operator and for this reason people make errors. The underlying assumption is that people have specific limited subsystems such as attention or working memory which cause problems in overly demanding situations [Broadbent 1958, Cowan 2000, Miller 1956, Navon & Gopher 1979]. Capacity-based thinking has influenced human factors research and, consequently, human machine interaction design very strongly. It has been typical to use dual task experiments, which are based on the premise that the secondary tasks detract capacity or resources from the primary task and thus uncover the possible consequences of the capacity limitation, such as deteriorated object and event detection [e.g. Schlegel, 1993, Wierwille et al. 1996, Girard et al. 2006, Groeger 2000, Jahn et al. 2005, Atchley & Dressel 2004]. The basic theoretical background for these experiments can be found in modular working memory literature [Baddeley 1986, 1996, Cowan 2000]. In accordance with the basic theory, human factors research has systematically observed that secondary tasks either make people err or slow down their performance in the primary task.

The capacity-based research tradition has been very influential and successful. However, it is not evident that all kinds of interaction problems can be solved by using this explanatory framework. We may have to adopt alternative explanatory grounds for solving different types of interaction problems. One alternative is content-based thinking, which assumes that mental contents can explain some aspects of human behavior [cf. Saariluoma and Nevala 2006]. It is known in expertise research that all the errors of experts cannot be explained in terms of limited capacity. Instead a part of them seem to be connected to mental contents [Saariluoma 1992, 1995]. This is why it would be useful to investigate the possible roles of mental contents in the context of interaction design oriented tasks. Here, we have selected interaction with systems while driving a car as our topic area and present an example of how the concepts of mental contents can be utilized in interpreting drivers' dual task performance.

2. Method

2.1 Subjects

The 16 volunteer subjects were recruited via public university e-mail lists. They included 9 women and 7 men between the ages of 20 and 33 (avg 24.31, sd 3.93). They all had a valid driving license and driving experience from 2 to 100 thousand kilometers (avg 45.31, sd 39.83). 8 of them were classified as experienced drivers ($\geq 30\ 000$ km, 4 men, 4 women) and 8 as novice drivers ($\leq 20\ 000$ km, 3 men, 5 women). All subjects had vision that was normal or corrected to normal.

2.2 Design and procedure

The experimental design consisted of a driving task with a driving simulator and a series of visual secondary tasks. Secondary tasks consisted of spaced or compressed text, which made the discriminability of the text vary between groups (Figure 1). The design imitated situations where the driver is reading an e-mail message with an in-car internet system while driving. The message could for example contain information about the locations of meetings to which the driver is heading. The experiment included trials without the secondary task and with the secondary task. The design was thus within subject design over secondary task condition and between subjects design over text types.



Figure 1. Different ways to represent the information on an in-car display.

The instructions for the driving task were to keep the blue bonnet of the vehicle between the white lane markers and to keep the velocity of the vehicle between 40-60 km/h. In the secondary task, subjects were asked to answer questions displayed in the upper part of the display based on the secondary task texts. The text changed after each correct answer. There was no time pressure in completing the secondary tasks but the trial lasted as long as the subject had completed the total number of five secondary tasks.

In the beginning subjects were given a query about general background information. After this, a helmet-mounted eye tracker was calibrated for the subject and overall instructions for the experiment were given. The experiment started with a practice trial. The practice driving task was performed without any secondary tasks. After this, the subject completed trials with and without secondary tasks. The order of the trials was counterbalanced. The subject was asked to fill in a reduced NASA Task Load Index (NASA-TLX) questionnaire (no weighting) [Hart & Staveland 1988] after both trials. Before the trial with the secondary task, the subject was given one secondary task without driving for practice. Prior to trials, the subjects were informed that the ten most accurate subjects in the driving tasks would be rewarded with movie tickets, and that accuracy is defined by the time spent outside the instructed areas (lane/speed zone). After the secondary task trial, the subjects were interviewed for discovering arguments for their NASA-TLX-answers, their strategies for time-sharing during the trial and whether they were able to keep to their lane with ambient vision [Summala et al. 1996].

2.3 Apparatus and measurements

The tools used in the experiment included the driving simulator comprised of a high-definition data projector, simulator computer, speakers and steering wheel with force feedback and pedals (Figure 2). The driving simulation software was open-source-based car simulation software called Racer (www.racer.nl). The car used in the experiment was Ford Focus RS with automatic gears and it was adjusted for a realistic driving experience. The driving view included a speedometer just above the steering wheel. The road used in the practice was a track-like circuit, while in the actual trials a more road-like environment simulating Polish countryside was used. Other equipment included a helmet-mounted iView X HED-eye-tracking system with a 50Hz sampling rate, a computer for controlling the secondary task 17" display, a microphone, a screen capturer, a mixer, a laptop for capturing mixed video, NASA-TLX questionnaires and query sheets.



Figure 2. The simulator environment.

Independent variables for analysis included the secondary task condition (within groups design) and the text type for the secondary task (between groups design). The dependent variables included driving errors (number and duration), which were defined to occur when the car was outside the instructed areas in the driving task; means, variance and maximum lengths of glance durations; the number of glances while driving in curves; and the NASA-TLX ratings.

Variance in glance durations has been used previously by Wikman et al. [1998, 2004, 2005] as a measure for time-sharing efficiency. The number of glances while driving in curves was selected for the analysis to measure subjects' situation dependent time-sharing efficiency. With this variable, we measured subjects' abilities to assess the difficulty of the driving situation and the extent to which they were able to adapt their task switching according to this information. Difficulty factors affecting task bandwidth in the driving situation have been previously defined by Wickens et al. [2003] and Horrey et al. [2006] to include speed (static in our experiment), width of the road (static) [Wikman et al. 1998], wind or other similar factor affecting the position of the car (static), curvature of the road and visibility to the road ahead. Changes in these variables should have an influence on the average 1.6 s glance time on secondary displays [Wierwille 1993] if the time-sharing models are effective. There is to some extent a general agreement that an average driver is capable of dividing visual attention efficiently between the driving task and a visual secondary task in a way that the duration of a single glance time at a secondary user interface typically stays under 1.6 seconds, and that the drivers are capable of proportioning the durations of glances according to the demands of the driving situation [Wierwille 1993].

The controlled variables included driving experience and gender (balanced between groups), user interface properties other than the one varied (same font sizes and text locations in different designs, also the point of the information searched for varied within text between tasks) and the order of the trials, which was counterbalanced within groups. Mixed video from the eye-tracking system and screen capturer were scored frame-by-frame with advanced video scoring software for behavioral research. Other data included the interview notes of the experimenter. Questionnaires were analyzed for means and variances within and between groups. T-tests with paired and independent samples were used in order to find statistical significance in the results.

3. Results

The text type did not have a significant effect on the number or duration of driving errors. However, the secondary task condition had a significant effect on the number of driving errors (Table 1).

Table 1. Number of driving errors, means (standard deviations). Spaced = Spaced text, Compressed = Compressed text. T-test, two-tailed, equal variances not assumed.

Group	Without secondary task	With secondary task	p-value
Spaced	4,00 (2,93)	14,50 (8,54)	0,007**
Compressed	6,13 (4,88)	15,25 (13,74)	0,038*
p-value	0,313	0,898	

The analysis of glance durations uncovers the efficiency of time-sharing strategies and respective models. The Compressed group had significantly greater means, maximums and standard deviations of glance durations at text compared to the Spaced group (Table 2). The total duration of glances at the secondary display during tasks was not significantly different between the text groups.

Table 2. Glance durations, means in seconds (standard deviations). Spaced = Spaced text, Compressed = Compressed text, TGT = Total Glance Time. T-test, two-tailed, equal variances not assumed.

Group	TGT	Mean	Standard deviation	Max
Spaced	209,92 (79,88)	0,97 (0,34)	0,59 (0,14)	3,47 (1,25)
Compressed	249,82 (101,81)	1,46 (0,50)	1,33 (0,60)	8,89 (5,63)
p-value	0,399	0,040*	0,010**	0,029*

However, these measures do not tell anything of whether the subjects were able to adapt their time-sharing and task switching appropriately to the demands of the driving task. The analysis of the eyes-off road in curves -variable gives information about the appropriateness of time-sharing strategies during the dynamic driving task. The results in Table 3. can be seen as evidence for difference between the efficiencies of the time-sharing models in the two groups.

Table 3. Number of glances, means (standard deviations). Spaced = Spaced text, Compressed = Compressed text. T-test, two-tailed, equal variances not assumed.

Group	Total	Glance duration > 2s	Glance duration > 2s in curve
Spaced	230,50 (85,07)	13,38 (10,97)	2,00 (1,93)
Compressed	174,50 (64,26)	38,75 (21,89)	8,88 (6,27)
p-value	0,161	0,015*	0,017*

The glances were started usually on a straight road, but especially in the Compressed group the subjects did not restore their eyes onto the road until they drove into a curve. There was a significant difference (T-test, two-tailed, equal variances not assumed, $p = 0.000$) between the number of driving errors committed in curves (avg 4.63, sd 3.61) and the number of driving errors committed on a straight road (avg 0.44, sd 0.89) in the trial without secondary tasks, which means that driving in a curve was more demanding than driving on a straight road. Glance

durations of more than 2 seconds while driving in a curve can be considered as unsafe (for similar measures, see Wikman et al. 1998).

Nasa-TLX measured the experienced demands of the tasks subjectively. There were no differences between groups in the results of the NASA-TLX questionnaire. Instead, except for physical demand, there were differences between trials within groups in all reported scales, including mental demand ($p = 0.000$), temporal demand ($p = 0.000$), effort ($p = 0.000$), performance ($p = 0.000$) and frustration ($p = 0.000$). Every participant in both groups rated the trial with the secondary tasks more demanding on these scales than the trial without the secondary task.

In the post experiment interviews, it was found that the subjects tried to follow time-sharing strategies by, for instance, attempting to allocate their visual attention to the display only while they were driving on a straight road or when the speed was easy to keep constant (no up or down hills), or trying to maintain the lane position with their ambient vision while reading the text. More than half of the subjects reported that they found it impossible to concentrate on reading the text and on maintaining the lane position and speed at the same time, especially in curves.

Finally, age, gender or driving experience did not have significant effects on any of the variables.

4. Discussion

The secondary task had a significant effect on driving behavior, which illustrates that capacity-based thinking can explain errors in our experiment. This was naturally expected, since extensive literature on secondary tasks with driving has demonstrated this [e.g. Zwahlen et al. 1988, Schlegel 1993, Salvucci 2001, Jamson & Merat 2005, Klauer et al. 2006].

However, the text type did not have a significant effect on the number or duration of driving errors. Instead, text type had an effect on the mean duration of glances, maximum durations and variance of the durations. Finally, text type affected the number of overlong glances in curves. The first finding indicates that gaze behavior is not in direct relation with capacity based driving errors. As a matter of fact, one can find in literature numerous articles which suggest that the measures for driver's performance, such as deviations in lane position, do not necessarily tell much about the differences of use risks between different secondary system designs per se, because the risks do not necessarily manifest themselves as driving errors in experiments or in real traffic [e.g. Jahn et al. 2005, Tijerina 2001].

The other findings indicate that glance durations are essentially affected by the differences in texts. Glance durations are indicative with respect to the time-sharing models. There were significant differences even between experienced drivers in their skills in time-sharing. Previously these differences have been found between novice and experienced drivers [Wikman et al. 1998], between young and aged drivers [Wikman & Summala 2005] and between healthy drivers and drivers with cerebral lesions [Wikman et al. 2004]. In addition, against expectations, the qualities of in-car user interfaces, in this case a quality difference as subtle as text spacing, was seen to have a clear effect on the time-sharing efficiency of the drivers. This means that glance duration models are essential in analyzing risk factors in interaction between the driver and in-car information system.

The overall meaning of our experiment is that capacity does not give particularly helpful information for designers. It is not sufficient to know that behaviors are different with respect to some gross measure such as number of errors because this does not yet tell what precisely should be changed. Instead, it is necessary to have accurate qualitative information about how behaviors have changed. In our case, the essential changes can be found in glance duration patterns and gaze direction strategies in varying driving conditions.

The basic finding has consequences, when considering interaction analysis for design. The traditional capacity-based interaction analysis is insufficient with respect to the quality of interaction. Take for example the following European Commission's [2006] design principles:

- 4.3.4.2.: The system should not require long and uninterruptible sequences of manual-visual interactions.
- 4.3.4.3: The driver should be able to resume an interrupted sequence of interactions with the system at the point of interruption or at another logical point, and
- 4.3.4.4.: The driver should be able to control the pace of interaction with the system.

These kinds of principles do not yet specify what kind of interaction information should be collected to be able to follow them. However, considering our results, it is hard to satisfy how the interaction patterns should be investigated on the grounds of quantitative error numbers. They do not specify what the interaction patterns should be like. Instead, by considering the specific nature of the patterns qualitatively, we can get information which is relevant for making decisions concerning the suitability of planned interaction models.

The number of In-vehicle Information Systems (IVIS) is increasing and, in addition, the use of mobile devices for different purposes while driving is getting more popular. This underlines the importance of finding ways to assess these systems' impact on driving safety in an ecologically valid way. Another important question is how to design safe user interfaces or driving modes for these systems. Safe user interfaces guide the driver in constructing a safe and efficient way to combine the driving task and the secondary tasks. This is achieved by dividing attention between the driving task and the secondary tasks appropriately by having the possibility to take into account the varying demands of the driving situation.

In order to be able to design increasingly complex environments, it is essential to find ways of analyzing the interaction patterns accurately. Very often, the gross numerical information does not have sufficient power of expression. For this reason, it is essential to develop effective means for analyzing mental contents. This is the only way to get an exact picture of the risks and difficulties people have with new technical environments.

References

- Atchley, P. & Dressel, J., "Conversation Limits the Functional Field of View". *Human Factors*, Vol.46, No.4., 2004, pp 664-673.
- Baddeley, A.D., "*Working Memory*", Oxford University Press Oxford, 1986.
- Baddeley, A.D., "Exploring the Central Executive", *The Quarterly Journal of Experimental Psychology*, Vol. 49, No.1., 1996, pp 5-28.
- Broadbent, D.E., "*Perception and Communication*". Pergamon London UK, 1958.
- Cowan, N., The Magical Number 4 in Short-Term Memory: A Reconsideration of Mental Storage Capacity. *Behavioral and Brain Sciences*, Vol.24, No.1., 2000, pp 87-185.
- Commission of the European Communities, "*Recommendation on Safe and Efficient In-vehicle Information and Communication Systems: Update of the European Statement of Principles on human machine interface*", C(2006) 7125 final, Commission of the European Communities Brussels, 2005.
- Girard, J-M, Tricot, N. Younsi, K. and Popieul, J-C, "When Does the Driver Workload Reach Its Limits?" *Proceedings of the IEEE ITSC*, Toronto, 2006, pp 578-583.
- Groeger, J.A., "*Understanding Driving: Applying Cognitive Psychology to a Complex Everyday Task*", Psychology Press Hove, 2000.
- Hart, S. G. and Staveland, L. E., "Development of NASA-TLX: Results of Empirical and Theoretical Research", *Human Mental Workload*, Elsevier Amsterdam, 1988, pp 139-183.
- Horrey, W.J., Wickens, C.D. and Consalus, K.P. "Modeling Drivers' Visual Attention Allocation While Interacting With In-Vehicle Technologies", *Journal of Experimental Psychology: Applied*, Vol.12, No.2, 2006, pp 67-78.
- ISO 9241-11, "*Guidance on Usability*", International Organization for Standardization, 1998.
- Jahn, G., Oehme, A., Krems, J.F. and Gelau, C., "Peripheral Detection As a Workload Measure In Driving: Effects of Traffic Complexity and Route Guidance System Use In a Driving Study", *Transportation Research Part F*, Vol.8, 2005, pp 255-275.
- Jamson, A.H. and Merat, N. "Surrogate In-Vehicle Information Systems and Driver Behaviour: Effects of Visual and Cognitive Load In Simulated Rural Driving", *Transportation Research Part F*, Vol. 8, 2005, pp 79-96.
- Kantowitz, B.H. & Sorkin, R.D. "*Human Factors: Understanding People-System Relationships*", John Wiley & Sons NY, 1983.

Klauer, S. G., Dingus, T. A., Neale, V. L., Sudweeks, J.D. and Ramsey, D.J. "The Impact of Driver Inattention on Near-Crash/Crash Risk: An Analysis Using the 100-Car Naturalistic Driving Study Data", DOT HS 810 594, Virginia Tech Transportation Institute, 2006.

Miller, G.A., "The Magical Number Seven, Plus or Minus Two: Some Limits on Our Capacity for Processing Information", *Psychological Review*, Vol.63, 1956, pp 81-97.

Moran, T.P., "Guest Editor's Introduction: An Applied Psychology of the User", *ACM Computing Surveys*, Vol.13, No.1., 1981, pp 1-11.

Navon, D. and Gopher, D., "On the Economy of the Human-Processing System", *Psychological Review*, Vol.86, No.3., 1979, pp 214-255.

Nielsen, J. "Usability Engineering", Academic Press Boston, 1993.

Norman, D.A. and Draper, S.W., "User Centered System Design: New Perspectives on Human-Computer Interaction", Lawrence Erlbaum Associates Hillsdale NJ, 1986.

Oulasvirta A. and Saariluoma P., "Long-Term Working Memory and Interrupting Messages In Human-Computer Interaction", *Behaviour and Information Technology*, Vol.23, No.1., 2004, pp 53-64.

Oulasvirta, A., and Saariluoma, P., "Surviving Task Interruptions: Investigating the Implications of Long-Term Working Memory Theory", *International Journal of Human-Computer Studies*, Vol.64, No.10, 2006, pp 941-961.

Parkkola, H., Saariluoma, P. and Berki, E. "Action-Oriented Classification of Families' Information and Communication Actions: Exploring Mothers' Viewpoints", *Journal of Behaviour & Information Technology*, (in press).

Preece, J., Rogers, Y., Sharp, H., Benyon, D., Holland, S. and Carey, T., "Human-Computer Interaction", Addison-Wesley Wokingham, 1994.

Rosson, M.B. and Carroll, J.M., "Usability Engineering: Scenario-Based Development of Human-Computer Interaction", Morgan Kaufmann Publishers Inc. San Francisco, 2002.

Saariluoma P., "Error in Chess: The Apperception-Restructuring View", *Psychological Research*, Vol.54, No.1., 1992, pp 17-26.

Saariluoma P. "Chess Players' Thinking: A Cognitive Psychological Approach", Routledge London UK, 1995.

Saariluoma P., "Explanatory Frameworks For Interaction Design", *Future Interaction Design*, Springer London, 2004, pp 67-83.

Saariluoma, P. and Nevala, K., "The Focus of Content-Based Approach to Design Engineering - a Reply to Eder", *AEDS Workshop proceedings*, University of Western Bohemia Pilsen, 2006, pp 105-110.

Salvucci, D.D., "Predicting the Effects of In-Car Interface Use On Driver Performance: An Integrated Model Approach", *International Journal of Human-Computer Studies*, Vol.55, No.1., 2001, pp 85-107.

Schlegel R.E., "Driver Mental Workload", *Automotive Ergonomics*, Taylor & Francis London UK, 1993, pp 359-382.

Summala H., Nieminen T. and Punto M., "Maintaining Lane Position With Peripheral Vision During In-Vehicle Tasks", *Human Factors*, Vol.38, No.3., 1996, pp 442-451.

Tijerina, L., "Issues in the Evaluation of Driver Distraction Associated with In-Vehicle Information and Telecommunications Systems", S 01 - 0926, Transportation Research Center East Liberty, 2001.

Wickens, C. D., Goh, J., Helleburg, J., Horrey, W. J., and Talleur, D. A., "Attentional Models of Multi-Task Pilot Performance Using Advanced Display Technology", *Human Factors*, Vol.45, No.3., 2003, pp 360-380.

Wierwille, W.W., "An Initial Model of Visual Sampling of In-Car Displays and Controls", *Vision in Vehicles IV*, Elsevier Amsterdam, 1993, pp 271-279.

Wierwille L., Tijerina S., Kiger T., Rockwell T., Lauber E. and Bittner A.Jr., "Heavy Vehicle Driver Workload Assessment Task 4: Review of Workload and Related Research", DOT HS 808 467, U.S. Department of Transportation, 1996.

Wikman, A.S., Haikonen, S., Summala, H., Kalska, H., Hietanen, M. and Vilkki, J., "Time-Sharing Strategies In Driving After Various Cerebral Lesions", *Brain Injury*, Vol.18, No.5., 2004, pp 419-432.

Wikman, A.S., Nieminen, T. and Summala, H., "Driving Experience and Time-Sharing During In-Car Tasks on Roads of Different Width", *Ergonomics*, Vol.41, No.3., 1998, pp 358-372.

Wikman, A.S. and Summala, H., "Aging and Time-Sharing In Highway Driving", *Optometry and Vision Science*, Vol.82, No.8, 2005, pp 716-723.

Zwahlen, H.T., Adams, C.C. and DeBald, D.P., "Safety Aspects of CRT Touch Panel Controls In Automobiles", *Vision in Vehicles II*, Elsevier Science Publishers B.V. Amsterdam, 1988, pp.335-344.

Pertti Saariluoma, Professor, Cognitive Science
 University of Jyväskylä, Faculty of Information Technology
 P.O. Box 35, 40014 University of Jyväskylä, Finland
 Telephone +358 14 260 3095, Telefax +358 14 260 4400, e-mail ps@jyu.fi