

An MCDM Approach to the Selection of Novel Technologies for Innovative In-Vehicle Information Systems

Isabel C. Lisboa, School of Psychology, University of Minho, Braga, Portugal

Joana Vieira, Centro de Computação Gráfica, Guimarães, Portugal

Sandra Mouta, Centro de Computação Gráfica, Guimarães, Portugal

Sara Machado, Bosch Car Multimedia Portugal, Braga, Portugal

Nuno Ribeiro, Bosch Car Multimedia Portugal, Braga, Portugal

Estêvão Silva, Bosch Car Multimedia Portugal, Braga, Portugal

Rita A. Ribeiro, CTS, Campus FCT-UNL, UNINOVA, Caparica, Portugal

Alfredo F. Pereira, School of Psychology, University of Minho, Braga, Portugal

ABSTRACT

Driving a car is a complex skill that includes interacting with multiple systems inside the vehicle. Today's challenge in the automotive industry is to produce innovative In-Vehicle Information Systems (IVIS) that are pleasant to use and satisfy the costumers' needs while, simultaneously, maintaining the delicate balance of primary task vs. secondary tasks while driving. The authors report a MCDM approach for rank ordering a large heterogeneous set of human-machine interaction technologies; the final set consisted of hundred and one candidates. They measured candidate technologies on eight qualitative criteria that were defined by domain experts, using a group decision-making approach. The main objective was ordering alternatives by their decision score, not the selection of one or a small set of them. The authors' approach assisted decision makers in exploring the characteristics of the most promising technologies and they focused on analyzing the technologies in the top quartile, as measured by their MCDM model. Further, a clustering analysis of the top quartile revealed the presence of important criteria trade-offs.

KEYWORDS

Automotive, Human Factors, Human-Machine Interaction, In-Vehicle Information Systems, Multi-Criteria Decision Making

DOI: 10.4018/IJDSST.2016010103

Copyright © 2016, IGI Global. Copying or distributing in print or electronic forms without written permission of IGI Global is prohibited.

INTRODUCTION

Problem Context

Ever since the first instrument that provided information to the car driver was introduced – the speedometer, invented in the early 1900's but standard in cars only two decades later – the driver's main task of driving the car, the primary task, is shared with multiple secondary tasks which have grown larger in number and complexity. These secondary tasks, whose main objective is to enhance the driving experience while addressing the driver's needs, are mostly related to the interaction with In-Vehicle Information Systems (IVIS) (Damiani, Deregibus, & Andreone, 2009; Harvey & Stanton, 2013).

To the first category of IVIS, *instrumentation*, designed to assist driving by conveying the car's internal state, manufacturers quickly added IVIS for *infotainment* (e.g. radio), *comfort* (e.g. A/C controls) and later, *navigation* (e.g. GPS) – see Damiani et al. (2009) and Harvey & Stanton (2013) for a review. Together, the four types of IVIS should address the driving assistance, the maximization of the safety and pleasure of driving, and should be part of a product differentiation when it comes to marketing any car model.

Societal changes, consumer trends, and multiple technological innovations in the automotive industry – e.g. autonomous driving (Merat & Jamson, 2009; Saffarian, De Winter, & Happee, 2012), large-scale use of smartphones, connected cars (Lu, Cheng, Zhang, Shen, & Mark, 2014; Zhao, 2002) or multimodal interaction (Jæger, Skov, & Thomassen, 2008; Muller, Weinberg, & Vetro, 2011; Spence & Ho, 2012) – pressure towards the introduction of novel devices and technologies inside the car that will be shared by different IVIS and advanced driver-assistance systems (ADAS) (Amditis, A. et al., 2010). Thus, in the always-connected lifestyle of a digital era it is reasonable to expect consumers wishing to extend highly “connected” behaviors into the experience of driving a car, for example, using social media tools by seamless integration of a smartphone when inside an internet-connected car (Harvey & Stanton, 2013; Walker, Stanton, & Young, 2001; Zhao, 2002). Both autonomous driving and multimodal interaction have in common the fact of enabling more complex interactions with IVIS. While manufactures are starting to introduce autonomous self-driving cars in production processes, IVIS' multimodal interactions are already a predominant trend, such as 3D gesture interaction combined with speech recognition (Spence & Ho, 2012). Driving mostly requires visual attention; hence, the combined use of other modalities (audio or tactile), which do not interfere with the modality the driver is already using, might be advantageous. An interesting solution for IVIS is therefore provided by approaches based on multimodal interactions since the disadvantages of a single modality can often be overcome by the cautious combination of other modalities (Muller et al., 2011; Pickering, Burnham, & Richardson, 2007; Spence & Ho, 2012).

Introducing new technologies inside a car, however, comes at considerable risk and a key design challenge of IVIS development remains, as before, the problem of solving the delicate balance of how to introduce novel IVIS technologies to the driver without compromising the primary task of safely driving to a destination. Driving is a complex task that constantly requires “eyes on the road”; it demands attention, different motor responses, and combined perceptual-motor skills (Calhoun et al., 2002; Groeger, 2000). Paradigmatic and sometimes counter-intuitive examples of how difficult it can be to stay on task abound in the literature of interference effects on driving performance: a simple conversation can disturb the driver's attentive scanning and representation of a traffic scene (McCarley et al., 2001); ignoring an incoming phone call (Holland & Rathod, 2013) or even just a phone notification (Stothart, Mitchum, & Yehnert, 2015) can cause significantly more infractions on a road. Finally, even “eyes on the road” can be a misleading idea: a good case study is provided by devices such as Heads Up Displays (HUD), designed specifically to present visual information in the driver's field of view directly while they are looking at road (e.g. by projecting information in the windshield or in a dedicated screen close to it). HUDs can be equally dangerous since, and unlike in the aviation industry, these devices are relatively untested in cars (Tufano, 1997) and are

prone to known effects such as perceptual tunneling (focusing on one aspect of the stimulus ignoring something else in full view) or cognitive capture (focusing on a task or internal thoughts and ignoring the environment). In summary, it is critical that current IVIS technologies and market pressures do not endanger driving performance (Harvey & Stanton, 2013).

Objectives and Motivation

With the goal of identifying novel technologies for next-generation IVIS products we developed a group Multi-Criteria Decision Making (MCDM) approach to measure and sort technologies in terms of their potential for innovative IVIS. We aimed to determine which technologies are innovative but also do not disrupt the delicate balance of primary task vs. secondary tasks inherent in driving – satisfying the two key aspects of our product design problem: (1) consumer needs and market pressures vs. and (2) safety concerns. After rating each decision alternative, using the MCDM model, a clustering analysis of the top quartile was conducted in order to verify for the presence of different criteria trade-offs.

Two classic elements of any multi-criteria decision problem were present in our decision problem: a set of possible alternatives to evaluate (which in our case study corresponded to a large set of technologies) and a set of decision criteria (Ishizaka & Nemery, 2013; Yu, 2013). The criterion set was composed of only qualitative criteria, which we wished to maximize, and they naturally conflicted, demanding careful compromise. We present here the first findings of applying this approach.

Multi-criteria analysis is a valuable tool in this context precisely because it explicitly considers multiple criteria to solve complex decision-making problems (Ishizaka & Nemery, 2013) and allows the exploration of trade-offs between different criteria, resulting in better and more informed choices (Olson, 2008). Previous work in this domain using a MCDM method - AHP pairwise comparison for measuring criteria importance - includes evaluating new advanced driver assistance technologies (Macharis, Verbeke, & De Brucker, 2004) and intelligent transport systems (De Brucker, Verbeke, & Macharis, 2004; Schintler, 2001).

The technologies were subjectively measured on eight qualitative criteria, defined by nine domain experts and using an appropriate group decision-making approach (Bregar, 2014) based on carefully structured meetings: team members were exposed to a presentation of the technology and rated the technology in each criterion individually (without group interaction and using a Likert scale); this was followed by qualitative group discussion of the ratings. To determine criterion importance, we followed what Brucker et al. (2004) labeled as “the verbal statements method” for determining criteria weights: team members were asked, in a separate session, to measure each criterion’s importance using a Likert scale. Criteria rating data for each alternative (i.e. a particular technology) was first aggregated per team member, using the team member specific criteria weights vector and the weighted average operator. After, decision scores were aggregated for each alternative by averaging. Our main goal was not the selection of a technology for immediate product development but the exploration of a heterogeneous set of alternatives (e.g. the differences between a sensor that can be embedded in a system, a prototype, or a complete off-the-shelf product). Moreover, while the decision criteria could be expressed from previous domain knowledge, it was unclear *a priori* how the large set of technologies we collected, from disparate categories, would be organized; our objective was to explore the landscape of current Human-Machine Interface’ technologies with potential for IVIS product development. The characteristics of the most promising technologies were further investigated and explored by concentrating the analysis on the top-quartile technologies (measured by the MCDM model), using a clustering analysis.

METHODOLOGY

We were interested in the overall qualitative pattern or the commonalities vs. differences of the top selected technologies. Thus, two main questions of interest were: (1) are there different decision

criteria tradeoffs when evaluating a set of technologies? And (2) is there an association between particular categories of technologies and specific clusters?

The methodological approach for the decision-making problem of evaluating candidate technologies for innovative IVIS followed the standard decision process of intelligence, design, and choice phases of Turban, Aronson and Liang (2005) – cf. Figure 1 for a schematic representation of the complete process.

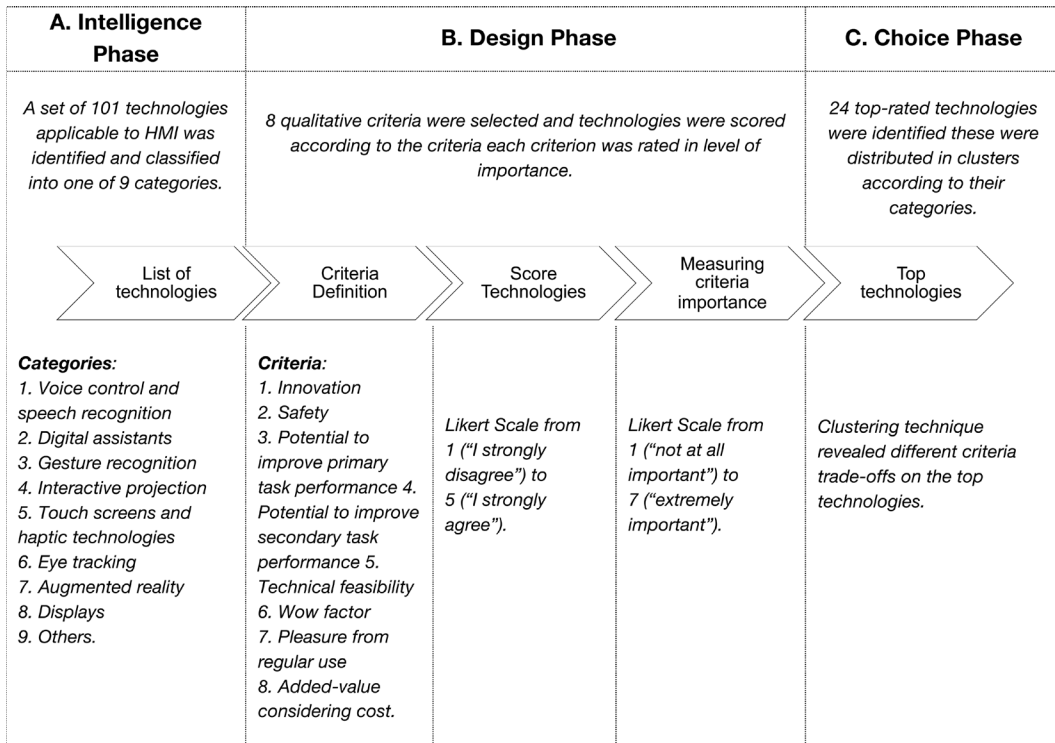
In the intelligence phase, a large survey of technological innovations in the Human Machine Interaction (HMI) field applicable to IVIS was conducted. This first exercise resulted in an initial set of alternatives. In the design phase, the decision criteria were identified and eight qualitative criteria were selected, as specified by the MCDM model. A procedure for measuring criteria importance was established, as well as a procedure for conducting group sessions where participants directly scored the alternatives. In the choice phase, the MCDM model score was computed per alternative, and this was done using the data provided after seven group meetings.

The main analysis consisted in examining in detail the produced rank order, and on focusing this analysis in the top quartile. Each phase will be detailed in the following sections.

INTELLIGENCE PHASE: SELECTION OF TECHNOLOGIES AND MCDM MODEL APPROACH

A set of 101 HMI technologies was identified. Selection of candidate technologies was based on interaction novelty and applicability to IVIS development. Technologies were found by researching market reports and by conducting Internet searches. Inclusion in the set was purposively liberal – candidate technologies were not restricted to the automotive industry, and they ranged in market

Figure 1. Schematic representation of the complete process



availability from prototypes to already available commercial products. In addition, no specific car market segment was considered when making the selection. Finally, technologies could range from complete solutions (e.g. a smartphone integration for the car) to a component (e.g. a new type of sensor for touch screens). Each technology was also classified into nine categories of technology: (1) voice control and speech recognition; (2) digital assistants (hands-free technology, focused on speech-based input and output); (3) gesture recognition; (4) interactive projection; (5) touch screens and haptic technologies; (6) eye tracking; (7) augmented reality; (8) displays (technology based on the visual presentation of information without touch interaction); and (9) others (technology that did not fit into any category).

Because of the large number of alternatives, a group-decision making approach, based on Figueira, Greco, & Ehrgott (2005) was selected, and nine study participants were asked to directly score each technology across all eight criteria using a Likert scale. To avoid the need for a consensus regarding criteria importance, criteria scores were aggregated per alternative and per evaluator, using a weighted average operator (Triantaphyllou, 2000), in which each participant provided their own relative criteria importance. Because we did not have access to a large set of domain experts but were able to bring together a team of nine heterogeneous experts (composed of software and hardware engineers, ergonomists, and experimental psychologists), evaluators were considered as all equally important. To aggregate the multiple final scores on the same alternative, the arithmetic mean operator was selected (Triantaphyllou, 2000).

DESIGN PHASE

Criteria Definition

Decision criteria were derived from meetings with domain experts in the automotive industry and HMI (including a subset of the authors) and the literature on IVIS (e.g. Harvey & Stanton, 2013). Thus, the criterion set included mandatory requirements on innovation potential (the starting point of the decision-making problem), safety, measures of human performance improvements, technological feasibility, and finally, standard qualitative measures from usability related with surprise and pleasure of use. The main goal was identifying potential candidates for medium-term product development of IVIS, irrespective of car market segment and thus the decision criterion set did not include quantitative cost-related criteria. Nevertheless, a proxy was the inclusion of an estimate of technical feasibility. The final eight selected qualitative criteria are described and defined in Table 1.

PROCEDURE FOR DIRECT SCORING OF ALTERNATIVES ON CRITERIA

The evaluating team was heterogeneous and included three ergonomics and usability experts, three engineers, and three experimental psychologists. In total, nine evaluators participated in the study (including a subset of the authors).

Rating was done quantitatively according to level of agreement in a five-point Likert scale from 1 (“I strongly disagree”) to 5 (“I strongly agree”) and the neutral point was 3 (“Neither agree or disagree”). Participants were provided with a set of statements, one per criteria that would define the criteria to be rated – e.g. “This technology can improve driving performance (in the sense of primary task”) or “This technology is innovative in the automotive industry” – and had to express how strongly they agreed with the statement for each particular alternative. The ratings were normalized a-posteriori to the user-friendlier interval [0, 1] in order to facilitate the aggregation of criteria with the simple weighted sum and directly obtaining percentage scores for all alternatives.

Seven evaluation group sessions were conducted; each session took circa 2 hours, and followed a structured approach with the following steps: (1) a short introduction to the technology was presented (typically in the form of a 5-10 minutes video); (2) participants rated individually their level of

Table 1. Definitions of the eight decision criteria

Criteria	Definition
<i>Innovation</i>	Determine how much the technology is perceived as innovative in the automotive industry. This criterion is restricted to the automotive domain.
<i>Safety</i>	The driver's safety may be compromised when performance on secondary tasks interferes, distracting the driver from the primary task. The main objective of this criterion is to rate how much a particular technology may interfere on the primary task, if implemented in a driving context.
<i>Potential to Improve Primary Task Performance</i>	Primary task performance is measured by driving parameters such as lateral and longitudinal control, and the subject's visual behavior. The main objective of this criterion is to detect which of the evaluated technologies has the potential to improve driving performance.
<i>Potential to Improve Secondary Tasks' Performance</i>	Secondary tasks are all tasks that are not directly related to the primary task of driving. Their main objective is to enhance the driving experience while addressing the driver's needs (Harvey & Stanton, 2013). This criterion intends to identify which technology has the potential to improve secondary task's performance, regardless of their impact on the primary task.
<i>Technical Feasibility of Implementation inside a Car</i>	This criterion rates if a technology is feasible to implement inside a vehicle, e.g. in a driving context. This criterion is important if an apparently non-feasible technology inside the car can, in fact, be creatively transformed into a feasible one in a driving context.
<i>User Experience (UX) Surprise Factor ("Wow factor")</i>	The most common definition of "wow factor" states that an object (in this case a technology) has "wow factor" when a set of properties belonging to the object pleasantly surprises the user. This criterion intends to subjectively determine if the evaluated technologies will surprise the user during first or early interactions.
<i>Pleasure from regular use (Delightfulness)</i>	Some IVIS products can become displeasing or dull for the driver after regular use. The main objective of this criterion is to identify technologies that can be pleasurable to use over extended periods of time.
<i>Added-Value Perceived by User Considering Cost</i>	This criterion measures if the user will perceive a specific technology as "worth paying for it". It subjectively measures the price people are willing to pay for a particular technology.

agreement with the criteria statements using five flash cards numbered from one to five – no comments were allowed until all the nine team members scored the technology; (3) a facilitator in the meeting then reread the criteria statements, and asked each participant to reveal their score to the group and recorded these scores; (4) the final step was an open group discussion and this qualitative discussion was also recorded – this reminded participants what the criteria were meant to measure, and were also highly informative in the case of conflicting opinions – data from this last step is not reported here. The format of our group sessions can be seen as approximating the equivalent of individual, independent ratings but our team of evaluators varied considerably in background and also included hierarchical professional relationships. By conducting the rating as we did in step 2 (design phase), evaluators were not pressured to conform, while the group setting benefited the evaluation – team members were constantly reminded of criteria definitions and the discussion allowed gathering data of why some decision alternatives seemed less consensual.

MEASURING CRITERIA IMPORTANCE

Separately from the evaluation sessions, and individually, the same nine evaluators also conducted direct rating of each criteria importance. They were asked to rate, on a seven-point Likert scale, criterion degree of importance, ranging from 1 ("Not at all important") to 7 ("Extremely important"), with 4 ("Neutral") as the middle point. This was done using a short paper questionnaire, where evaluators could again read the criteria definitions (to again remind them of their specific definition) and record

the measurement. The vector of criteria weights, per participant, was obtained by normalizing the rating of criteria importance using the sum of a participant's ratings in the Likert scale.

Note that other weighting procedures could have been used, e.g. (Salo & Hämmäläinen, 2001), but since we are only interested in determining a list of promising technologies and not in selecting the "best" alternative the grouping of experts' opinions seemed reasonable.

CHOICE PHASE

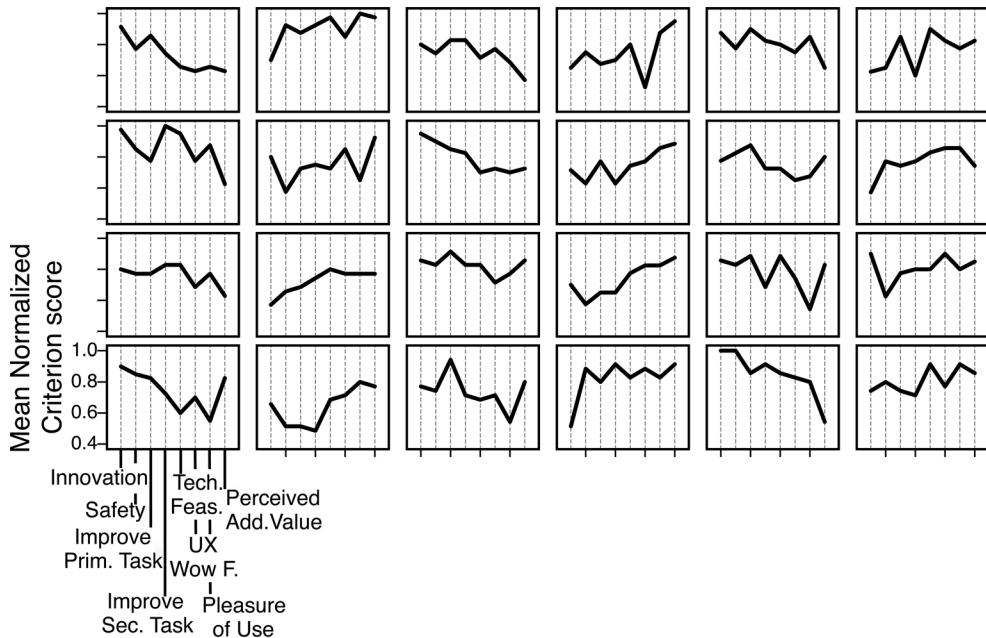
The normalized criteria rating data, per participant and alternative, was aggregated using a weighted average, and the decision scores of multiple participants on the same alternative were aggregated using the arithmetic mean. This generated our final MCDM model decision score and a rank ordering of the alternatives, reflecting the combined opinion of the group of nine experts, going from the technologies with the highest potential to the least appreciated technologies.

In the main analysis, a clustering technique was applied to the top quartile, and criteria scores were tested for the presence of different trade-offs; in addition, the distribution of technology categories was also examined.

RESULTS AND DISCUSSION

From the rank ordered list of all technologies, the top quartile was selected for further analysis (i.e. the technologies on percentile 75 or above), providing twenty-four top-rated technologies, which ranged in their final decision score from 0.725 to 0.903 (see Figure 2) and were the focus of our analysis.

Figure 2. Decision criteria scores for each individual technology in the top quartile of technologies according to the MCDM model; the criteria score shown was obtained by averaging criterion scores across all participants and the technologies are ordered by decreasing value of decision score computed by the MCDM model



Top technologies differed in the way their criterion score was distributed throughout the eight criteria: while all were rated as innovative – likely reflecting the initial selection criteria that was precisely to measure and sort technologies in terms of their potential for innovative products – there was variability in the remaining criterion scores, and the absence of parallel lines in the visualization of Figure 2 suggests the existence of different tradeoffs. Thus, these preliminary results already indicate that these twenty-four technologies were in the top quartile *via* different ways, or different criteria. Due to these differences in the weight distribution of each criterion in the top 24 technologies and also to understand whether there was a predominant category of technology, we conducted a clustering analysis.

A standard k-means clustering technique (with $k = 3$) was applied to the average criterion score data (24 vectors of dimension eight) and three clusters with comparable frequencies were obtained: cluster one accommodated eight technologies, cluster two, seven and cluster three, nine technologies – see Figure 3 and Figure 4 for full results. These three clusters grouped technologies according to

Figure 3. a) Mean criterion score per cluster (data shown are cell means with error bar equal to ± 1 SE); b) Scatter plot of 2nd (x-axis) and 3rd PCA Component regression scores for the top quartile with cluster information overlaid (PCA was applied to individual criterion score data in the complete dataset; three components account for 84% of variance; components shown result from Varimax with Kaiser normalization rotation method); c) Decision score computed by the MCDM model shown per individual cluster member

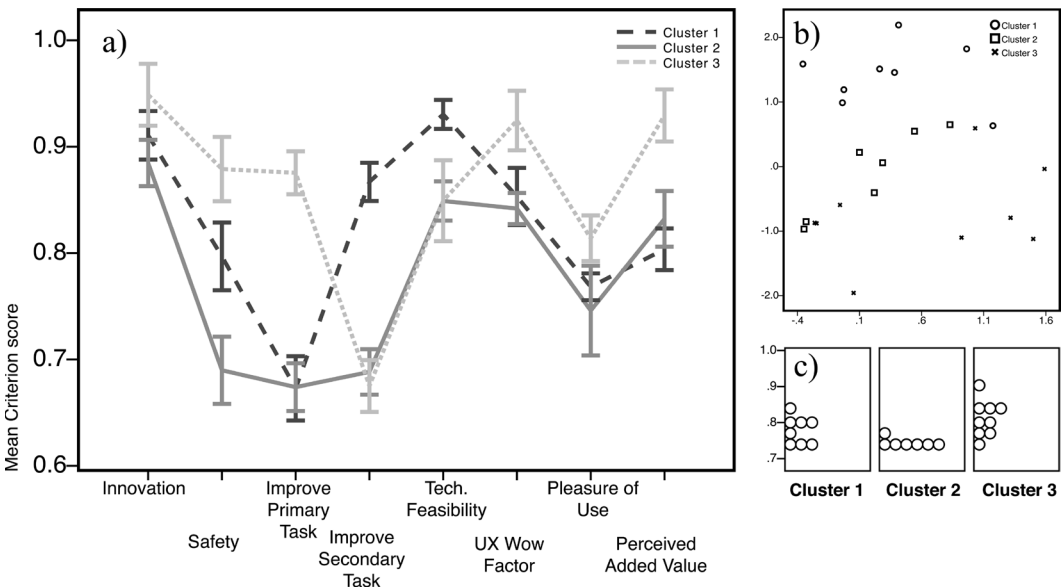
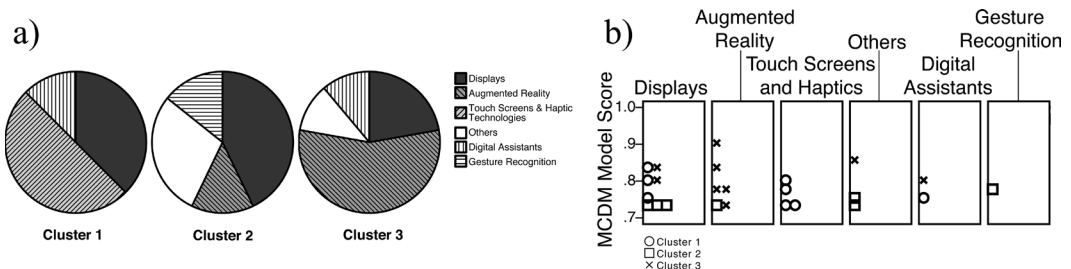


Figure 4. a) Frequency distribution of technology category per cluster; b) Decision score computed by the MCDM model for the top-quartile, per alternative and technology category, with cluster information overlaid



their similarities in criteria scores and allowed the visualization of the commonalities and differences between technology categories. The critical k parameter was selected by applying a hierarchical agglomerative clustering method to the set and inspecting the scree plot.

Overall, only six out of the nine technology categories were represented in the top quartile (see Figure 4). “Interactive Projection”, “Eye Tracking” and “Voice Control and Speech Recognition” categories had lower weighted sum scores and were not present on the top quartile. Some known features of these categories might explain their absence in the top quartile. While “Eye tracking” systems inside the car have the disadvantage of demanding visual interaction (and driving is a “eyes on the road” task) and are also technologically challenging (remote eye tracking under arbitrary lighting conditions with a moving user is non-trivial), “Voice Control and Speech Recognition” can, in principle, reduce the driver’s visual demand which could be an advantage in a driving context. In addition, interaction with a speech-based system seem to be less distracting than conversations, for example, some studies show that drivers have more flexibility to abort an IVIS’ interaction than a conversation with a side passenger (Vetek & Lemmelä, 2011). Nevertheless, “Voice Control and Speech Recognition” systems also produce high cognitive loads in the driver, which can affect the primary task (Shneiderman, 2000). More limitations are also associated with this kind of interaction systems such as the difficulty of some systems in distinguishing acoustically similar words; interference effects with the recognition task from parallel vocalizations (e.g. when other person is talking) or from background noise; and, finally, the drivers’ difficulty to learn and remember speech commands (Bhise, 2011). All of these limitations constitute a tentative explanation for the absence of these categories in the top quartile.

On the contrary, “Displays” category was the only category present in all three clusters. A possible explanation for this result might be that many of the evaluated technologies have what one might call a display.

Another common characteristic between the three clusters, and as expected, is that all clusters scored high on the “Innovation” criteria (cluster 1, $M = 0.861$; cluster 2, $M = 0.835$; cluster 3, $M = 0.899$). Cluster one, consisted mainly of “Touch Screens and Haptics” and had a higher score in the “Improve Secondary Task” ($M = 0.817$) and “Technical Feasibility” ($M = 0.880$) criteria, scoring lower on “Improve Primary Task” ($M = 0.623$). Cluster three, with a high number of “Augmented Reality” technologies, had high scores in all criteria (particularly the “Improve Primary Task” criterion, $M = 0.825$) except “Improve Secondary Task” ($M = 0.625$). “Improve Secondary Task” and “Improve Primary Task” appears to be the main trade-off criteria between cluster one and three. Finally, cluster two scores low on three criteria: “Safety” ($M = 0.640$), “Improve Primary Task” ($M = 0.624$), and “Improve Secondary Task” ($M = 0.638$).

In Cluster three, the category “Augmented Reality” occupies the larger part of the cluster, and in cluster one the main category is “Touch Screens and Haptic Technologies”. Importantly, these differences in technology categories that never share the same cluster also suggest criteria tradeoffs. This was confirmed by examining the characteristics of the cluster centroids and is also evident in Figure 3 panel a). The main result is the different tradeoff of cluster one vs. cluster three: while cluster one contains alternatives with high criterion score on “Improve Secondary Task” but moderate on “Improve Primary Task”, cluster three contains alternatives with high criterion score on “Improve Primary Task” but not in “Improve Secondary Task”. Critically, cluster three technologies were rated as safer and are higher on the three usability criteria (“Wow Factor”, Pleasure of Use” and “Perceived Added Value”), while cluster one technologies were higher on “Technical Feasibility”. Per criterion, cluster two was either similar to cluster one or three, or had the lowest performance.

From this analysis it is clear that no single technology can equally satisfy all criteria, and that technologies arrived to the upper quartile by different means. The need for a compromise is demonstrated in the trade-offs between improving primary and improving secondary tasks (mainly between cluster one and three), which is the main distinction to be made. According to the clustering analysis, it seems that to benefit one criterion is to detriment the other.

CONCLUSION

A perhaps telling indicator of current complexities in car design, above and beyond technological innovation per se, is the current push in the automotive industry for increasing the metrics that guide user experience design, namely adding to the more traditional measures of usability – e.g. usefulness, ease-of-use – new metrics related to the emotional experience of using a car – e.g. joy-of-use (delightfulness), aesthetic pleasure or image-of-use (the social effects from being a user of the product) – and do so at the earliest phases of product design. This is an informative context to understand current automobile industry trends: large pressures for adding novel technologies to the driving experience, ultimately leading to novel forms of human-machine interaction (HMI) that may change the driving experience itself. In fact, HMI is becoming not only a distinctive factor for consumers, that influences car purchases, but also for automobile brands that use HMI as a way to identifying themselves with the consumer.

We were thus interested in searching for technologies that are innovative but secure and feasible to implement and, at the same time, also improve performance. The main objective was to find a methodology that captured innovation, safety, human performance measures, technical implementation, and user experience concerns into a decision model for identifying new technologies for future IVIS products. The combination of a group MCDM model with a clustering technique allowed the construction of a top set of candidate technologies, but most importantly, a deeper understanding of the domain in terms of the technologies' characteristics and trade-offs. One of the main results was the heterogeneity of top technologies in terms of criteria weights, i.e. no unique set of criteria defined the top alternatives. Besides validating the balance between the criteria chosen, the model did not find a dominant technology or even a dominant category of technologies – although it allowed eliminating some categories. Instead, the cluster analysis of the top quartile found one major decision trade-off – the most important is whether the technology contributes to the primary or the secondary task, and what additional criteria increase or decrease depending on the primary vs. secondary task decision. Any future business decision made based on the results reported here will likely be a decision on what side of the trade-off to allocate product development effort, and not between the best technologies. The finding of non-dominated alternatives is thus informative to future product development decisions: this was our key contribution in terms of applicability and business value of the combined clustering and MCDM approach to our problem domain.

However, the selected approach is not without its limitations. The model assumes independence between criteria: future work will include analyzing the trade-offs between criteria and this might yield slightly different results. Group rating of technologies was a long and time consuming process; even using a simple Likert scale to rate each technology, since one hundred technologies were scored. Thus, it is reasonable to assume that a direct scoring method is more adequate for a large set of alternatives, and pairwise comparison approaches are not feasible. The simple weighted average was used to define the ranked list because our objective was to determine the list of more promising IVIS technologies. In future work we intend to perform a comparative study with other outranking models, such as TOPSIS to ensure the robustness of the results when using ordinal classifications. The use of k-means was adequate considering the characteristics our dataset but not necessarily adequate for a general-purpose methodology.

In conclusion, with our MCDM approach to the decision problem of selecting technologies for innovative In-Vehicle Information Systems, it was possible to understand which technologies are good candidates for the car of the future, and to verify where the main difficulties in making a future product development decision will reside. We found one major criteria decision tradeoff in this problem domain by examining the top quartile of alternatives. The primary vs. secondary task tradeoff could simply reflect the fact that some alternatives consisted of solutions or technologies specifically designed to assist in the primary or the secondary task. However, the combination of all criteria in the clustering analysis showed a more complex tradeoff. The cluster that scored high on

improvements to driving was overall superior in safety and usability but at a higher cost in technical feasibility; while the cluster that scored high on improvements to secondary tasks had higher technical feasibility with lower but yet comparable ratings in the usability criteria. Products that have an effect in the primary task will have a higher overall impact in the driver's user experience, but innovations in the secondary task may be more feasible with only a moderate decrease in the other criteria, however, with a cost in safety. Future work will include validation of this hypothesis in our problem domain.

ACKNOWLEDGMENT

This research was supported by: Operational Competitiveness Program – COMPETE, QREN (*Quadro de Referência Estratégico Nacional*), European Regional Development Funds (European Union), R&D project in joint-promotion (HMIEXCEL-2013-2015 36265/2013) HMIEXCEL - *I&D crítica em torno do ciclo de desenvolvimento e produção de soluções multimédia avançadas para automóvel; Strategic program FCT-UID/EEA/00066/2013*. AFP was funded by the Portuguese Foundation for Science and Technology (IF/00217/2013). IL was also partially funded by the Portuguese Foundation for Science and Technology (PD/BD/105966/2014). The authors wish to thank Carlos Silva, João Ferreira, João Lamas, Nuno Silva, Rafael Nascimento, and Rosane Sampaio for their participation in the team of evaluators.

REFERENCES

- Amditis, A., Andreone, L., Pagle, K., Markkula, G., Romera Rue, M., Deregibus, E., . . . Peters, B. (2010). Towards the automotive HMI of the Future; Overview of the AIDE Integrated Project Results. *Intelligent Transportation Systems (Special Issue)*, 11(3), 567 – 578. doi:10.1109/TITS.2010.2048751
- Bhise, V. D. (2011). *Ergonomics in the automotive design process*. CRC Press. doi:10.1201/b11237
- Bregar, A. (2014). Towards a Framework for the Measurement and Reduction of User-Perceivable Complexity of Group Decision-Making Methods. *International Journal of Decision Support System Technology*, 6(2), 21–45. doi:10.4018/ijdsst.2014040102
- Calhoun, V. D., Pekar, J. J., McGinty, V. B., Adali, T., Watson, T. D., & Pearlson, G. D. (2002). Different activation dynamics in multiple neural systems during simulated driving. *Human Brain Mapping*, 16(3), 158–167. doi:10.1002/hbm.10032 PMID:12112769
- Damiani, S., Deregibus, E., & Andreone, L. (2009). Driver-vehicle interfaces and interaction: where are they going? *European transport research review*, 1(2), 87-96.
- De Brucker, K., Verbeke, A., & Macharis, C. (2004). The applicability of multicriteria-analysis to the evaluation of intelligent transport systems (ITS). *Research in Transportation Economics*, 8, 151–179. doi:10.1016/S0739-8859(04)08008-4
- Figueira, J., Greco, S., & Ehrgott, M. (2005). *Multiple criteria decision analysis: state of the art surveys* (Vol. 78). Springer Science & Business Media. doi:10.1007/b100605
- Groeger, J. A. (2000). *Understanding driving: Applying cognitive psychology to a complex everyday task*. Psychology Press.
- Harvey, C., & Stanton, N. A. (2013). *Usability evaluation for in-vehicle systems*. Crc Press. doi:10.1201/b14644
- Holland, C., & Rathod, V. (2013). Influence of personal mobile phone ringing and usual intention to answer on driver error. *Accident; Analysis and Prevention*, 50, 793–800. doi:10.1016/j.aap.2012.07.004 PMID:22871271
- Ishizaka, A., & Nemery, P. (2013). *Multi-criteria decision analysis: methods and software*. John Wiley & Sons. doi:10.1002/9781118644898
- Jæger, M. G., Skov, M. B., & Thomassen, N. G. (2008). You can touch, but you can't look: interacting with in-vehicle systems. *Paper presented at the Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*.
- Lu, N., Cheng, N., Zhang, N., Shen, X., & Mark, J. W. (2014). Connected vehicles: Solutions and challenges. *Internet of Things Journal*, 1(4), 289–299. doi:10.1109/JIOT.2014.2327587
- Macharis, C., Verbeke, A., & De Brucker, K. (2004). The strategic evaluation of new technologies through multicriteria analysis: The ADVISORS case. *Research in Transportation Economics*, 8, 443–462. doi:10.1016/S0739-8859(04)08019-9
- McCarley, J. S., Vais, M., Pringle, H., Kramer, A. F., Irwin, D. E., & Strayer, D. L. (2001). Conversation disrupts visual scanning of traffic scenes. *Paper presented at the 9th Vision in Vehicles Conference*, Brisbane, Australia.
- Merat, N., & Jamson, A. H. (2009). How do drivers behave in a highly automated car. *Paper presented at the 5th International Driving Symposium on Human Factors in Driver Assessment, Training and Vehicle Design*.
- Muller, C., Weinberg, G., & Vetro, A. (2011). Multimodal input in the car, today and tomorrow. *IEEE MultiMedia*, 18(1), 98–103. doi:10.1109/MMUL.2011.14
- Olson, D. L. (2008). *Multi-criteria decision support Handbook on Decision Support Systems 1* (pp. 299–314). Springer. doi:10.1007/978-3-540-48713-5_15
- Pickering, C. A., Burnham, K. J., & Richardson, M. J. (2007). A research study of hand gesture recognition technologies and applications for human vehicle interaction. *Paper presented at the 3rd Conf. on Automotive Electronics*.

- Saffarian, M., De Winter, J., & Happee, R. (2012). Automated driving: human-factors issues and design solutions. *Paper presented at the Proceedings of the Human Factors and Ergonomics Society Annual Meeting*. doi:10.1177/1071181312561483
- Salo, A., & Hämäläinen, R. P. (2001). Preference ratios in multiattribute evaluation (PRIME) - elicitation and decision procedures under incomplete information. *IEEE Trans. on SMC*, 31(6), 533–545.
- Schintler, L. (2001). *Smart flexible integrated real-time enhanced system*. SAFIRES.
- Shneiderman, B. (2000). The limits of speech recognition. *Communications of the ACM*, 43(9), 63–65. doi:10.1145/348941.348990
- Spence, C., & Ho, C. (2012). *The multisensory driver: Implications for ergonomic car interface design*. Ashgate Publishing, Ltd.
- Stothart, C., Mitchum, A., & Yehnert, C. (2015). The attentional cost of receiving a cell phone notification. *Journal of Experimental Psychology. Human Perception and Performance*, 41(4), 893–897. doi:10.1037/xhp0000100 PMID:26121498
- Triantaphyllou, E. (2000). *Multi-criteria decision making methods Multi-criteria Decision Making Methods: A Comparative Study* (pp. 5–21). Springer. doi:10.1007/978-1-4757-3157-6_2
- Tufano, D. R. (1997). Automotive HUDs: The overlooked safety issues. *Human Factors: The Journal of the Human Factors and Ergonomics Society*, 39(2), 303–311. doi:10.1518/001872097778543840 PMID:9302890
- Turban, E., Aronson, J., & Liang, T.-P. (2005). *Decision Support Systems and Intelligent Systems 7 Edition*. Pearson Prentice Hall.
- Vetek, A., & Lemmelä, S. (2011). Could a dialog save your life?: analyzing the effects of speech interaction strategies while driving. *Paper presented at the Proceedings of the 13th international conference on multimodal interfaces*. doi:10.1145/2070481.2070506
- Walker, G. H., Stanton, N. A., & Young, M. S. (2001). Where is computing driving cars? *International Journal of Human-Computer Interaction*, 13(2), 203–229. doi:10.1207/S15327590IJHC1302_7
- Yu, P.-L. (2013). *Multiple-criteria decision making: concepts, techniques, and extensions* (Vol. 30). Springer Science & Business Media.
- Zhao, Y. (2002). Telematics: Safe and fun driving. *IEEE Intelligent Systems*, 17(1), 10–14. doi:10.1109/5254.988442