Brigham Young University
BYU ScholarsArchive

International Congress on Environmental Modelling and Software

4th International Congress on Environmental Modelling and Software - Barcelona, Catalonia, Spain - July 2008

Jul 1st, 12:00 AM

An Analysis of Interface Quality for Decision Support

Timothy Maciag

Daryl H. Hepting

Follow this and additional works at: https://scholarsarchive.byu.edu/iemssconference

Maciag, Timothy and Hepting, Daryl H., "An Analysis of Interface Quality for Decision Support" (2008). *International Congress on Environmental Modelling and Software*. 157.
https://scholarsarchive.byu.edu/iemssconference/2008/all/157

This Event is brought to you for free and open access by the Civil and Environmental Engineering at BYU ScholarsArchive. It has been accepted for inclusion in International Congress on Environmental Modelling and Software by an authorized administrator of BYU ScholarsArchive. For more information, please contact scholarsarchive@byu.edu, ellen_amatangelo@byu.edu.

iEMSs 2008:International Congress on Environmental Modelling and Software
Integrating Sciences and Information Technology for Environmental Assessment and Decision Making
4th Biennial Meeting of iEMSs, http://www.iemss.org/iemss2008/index.php?n=Main.Proceedings
M. Sànchez-Marrè, J. Béjar, J. Comas, A. Rizzoli and G. Guariso (Eds.)
International Environmental Modelling and Software Society (iEMSs), 2008

An Analysis of Interface Quality for Decision Support

Timothy Maciag [a] **Daryl H. Hepting** [a]

[a]*Department of Computer Science, University of Regina, 3737 Wascana Parkway, Regina, CANADA (maciagt@cs.uregina.ca, dhh@cs.uregina.ca)*

Abstract: This paper examines the quality of user interfaces in an application to choose products. A high-quality interface should enable users to perform prescribed tasks quickly and without errors. The interface should also enable users to find products which match their personal values. A high-quality interface can especially empower those diverse users who might otherwise be overwhelmed by technology but who need information to ground their actions. By examining user performance in such regards, insight into the quality of interface may be acheived. For this purpose, we examined user performance while utilizing different interfaces for product selections. An analysis and discussion of the examination is presented, along with opportunities for future work.

Keywords: decision support; usability; user evaluation

1 INTRODUCTION

Over the last few decades, many authors have discussed usability and utility in terms of interface design [Grinstein et al., 2003]. If a tool is not usable but has utility, users may not utilize the system to its full potential. If it is usable without utility, people will not come back. In this regard, Rosson and Carroll [Rosson and Carroll, 2002] describe three levels of designing user interfaces with various degrees of usability and utility, these being *effective, comprehensible, and satisfying* with *satisfying* being the level that interface designers strive to achieve in their designs. A satisfying user interface could be one that encapsulates the notions of effectiveness, efficiency, and comprehensibility.

When considering design solutions for the development of user interfaces which provide satisfying user interactions, the diversity of end-user may play an essential role. As such, an interface should be designed to incorporate aids for both beginners and experts alike, as outlined by Raskin's [2000] *myth of the beginner-expert dichotomy*. In such regards, Fischer [Fischer, 2001] describes *high-functionality applications* (HFA); those which enable beginners to learn the system over time while allowing experts to harness the full potential of available resources.

Specifically we are interested in designing satisfying HFAs in the form of decision support tools. Decision support tools, which include software tools that assist end-users' in tasks involving decision formulations, seek to aid end-users by highlighting key information relevant to the task, providing a medium for comparison, aiding query formation, and by providing an interface between the user and system data [Phillips-Wren and Forgionne, 2006; Pu and Chen, 2005]. To paraphrase Richard Hamming [1962], "the purpose of decision support is to easily find the product that matches one's values, not to easily find some product which does not interest me." For this purpose, we consider designing *augmented* [1991] HFAs which provide a balance between the capabilities of the end-user and core system functionality. This paper presents work towards these goals.

The rest of the paper proceeds as follows. Section 2 describes the decision support tools used. Section 3 describes the methods used and the metrics collected during two evaluations for product decision support systems. Section 4 describes the analysis and interpretation of the study data. Section 5 provides conclusions and areas for future work.

2 TOOLS FOR EVALUATION

For our analysis, we conducted an in-depth usability analysis of two support tools, one developed by the United States Environmental Protection Agency (US-EPA) and the other developed by us called *cogito*. Both tools provided varying interface designs for searching a database of 29 environmentally preferable cleaning products distinguished by eight environmental attributes: *skin irritation*, *food chain exposure*, *air pollution potential*, *product contains fragrances*, *product contains dye*, *product uses recyclable packaging*, *product is a concentrate*, and *product minimizes exposure to concentrate*. We examined the quality of the user interface in terms of measuring user effectiveness and efficiency, hypothesizing that higher levels of each may indicate greater satisfaction of use.

2.1 US-EPA Tool

The US-EPA tool consisted of three searching tools: a single attribute ranking tool (SART), a multiple attribute ranking tool (MART), and a weighted attributed ranking tool (WART). SART enables users to filter products using one attribute only. Users choose a single attribute and are able to sort the attribute's values in terms of those that are most environmentally preferable to least (ascending order). MART, illustrated in Figure 1, enables users to filter products using two methods: by assigning priority of up to four of the attributes and/or by setting thresholds for specific attributes. WART enables users to filter cleaning products by assigning weights to specific attributes. Users assign attributes weights by choosing numerical values between 0 and 1000[1]. Higher weights indicate greater importance. Although the functionality of the search/filter procedure differed for each US-EPA EPP search tool, the results displayed by each tool used a tabular format.

2.2 Cogito

Cogito [Hepting, 2003], illustrated in Figure 2, is an implementation of an augmented HFA. End-users are provided the functionality to scan and filter items within an information space. The interface is comprised of rectangular cells displaying system objects based on end-user queries. *Cogito's* built-in querying functionality equips end-users with the basis to form queries by selecting system attributes and applying thresholds accordingly.

To evaluate the *cogito* system with respect to the US-EPA tools we created two *cogito* interface representations, a textual interface which displays products using textual data in the form of HTML, and a graphical interface which displays products using a nightingale rose illustration, see Figure 1. Each interface representation utilizes the functionality of the *cogito* system similarly; only the visual representation of the products differ.

3 EXPERIMENT DESIGN

For our analysis, we conducted two experiments. The goal of the first experiment was to evaluate the three decision support tools provided by the US-EPA in order to formulate which of the tools provided users with the most adequate support. The goal of the second experiment was to evaluate the *cogito* interfaces with respect to the best of the US-EPA tools.

[1]The range *0 to 1000* was the one originally provided by the US-EPA and was not changed for the purposes of the evaluation

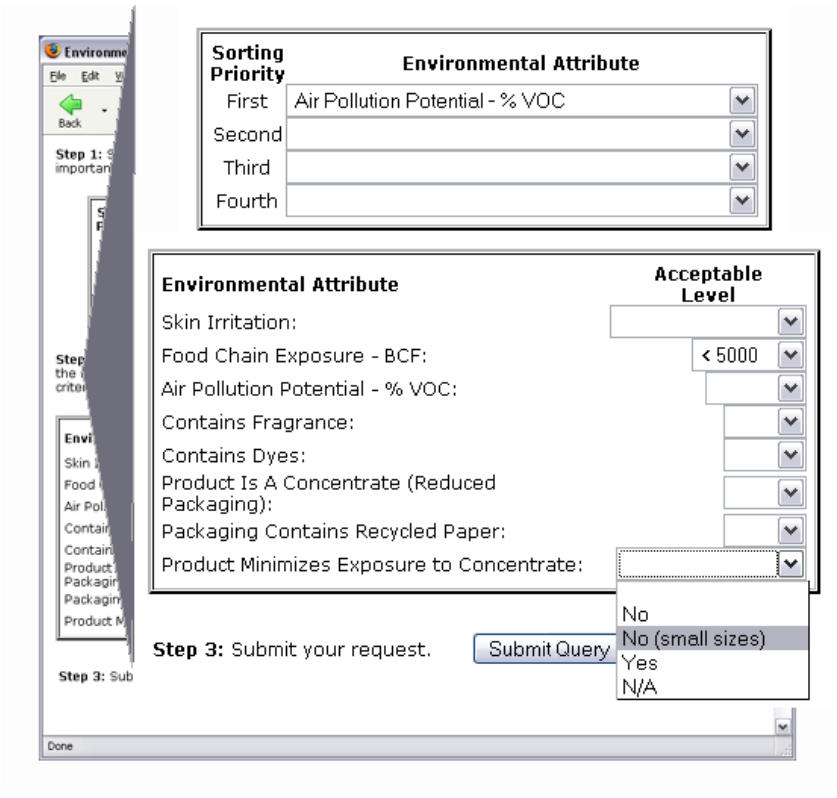

Figure 1: A screen capture of the US-EPA MART. The main interface is magnified. Users may filter cleaning products by setting priority of up to four attributes. Users may also chose acceptable attribute value levels.

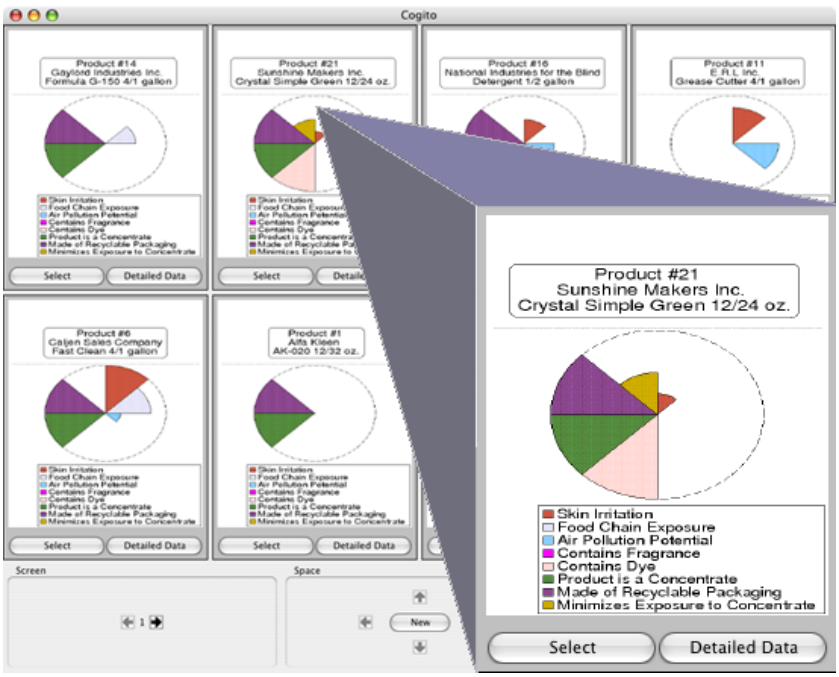

Figure 2: Screen capture of the graphical *cogito* interface. A cell on the interface is magnified.

In each experiment, 28 participants (11 female in the first experiment and 13 in the second) were recruited from the University of Regina Computer Science Participant Pool.[2]. In both experiments participants were asked to complete a pre-task questionnaire, answer a series of task questions, and complete a post-task questionnaire. The task questions were designed to evaluate the quality of the tool interfaces in terms of task complexity. People who use these tools are interested in finding products that match their values. The question of how well people can decide based on values is a more difficult task, because of the potential mismatch between user's model and the categories available. To some users, only one attribute might be important, while other users might need two, three, or more attributes to make their decisions. These preferences may also change over time. The authors created question sets to represent this variable complexity. There were a total of three sets, each with questions dealing with 1, 2, and 3 attributes.

To analyze user satisfaction, we recorded user response times and computed user task scores. Task scores were computed (Eqn 1) by determining the absolute and relative values of participant scores in relation to their responses. Eqn 2 was used to measure the efficiency of users. Based on the equation, participants who performed well on the interfaces in both task score and response time obtained a higher efficiency score. Task score results were squared to account for the occurrence of a high user task score matched with a *lengthy* response time and vice versa. It is also important to note that in order to discredit any bias in our final results, the interfaces were counterbalanced accordingly.

$$TaskScore = (1 - (\frac{\#CorrectAnswers - \#CorrectGiven}{\#CorrectAnswers})) \times \frac{\#CorrectGiven}{\#AnswersGiven} \quad (1)$$

$$Efficiency = \frac{TaskScore^{\,2}}{UserResponseTime} \quad (2)$$

4 Evaluation and Results

Results from both experiments are described in Table 1.[3] When observing results from the first experiment for tasks involving one attribute, users performed more adequately using SART whereas for tasks involving two and three, users performed more adequately using MART and WART. This also coincided with the participant preferences indicated in the results obtained in the post-task questionnaire analysis, where we asked participants to reveal their preferences in terms of interface. Based on their responses, for tasks involving one attribute SART was more preferred, whereas for tasks involving more than one attribute MART was more preferred. Based on these results, for our second evaluation incorporating the *cogito* tools, we developed a comparison structure of SART versus the textual *cogito* interface and MART versus the graphical *cogito* interface.

Observing the results in Table 1, it is clear that the users performed more adequately using the *cogito* interfaces. Figures 3 and 4 illustrate the user efficiency metric. Again, there is a clear indication that users performed more adequately using the *cogito* tools with respect to performing the prescribed tasks.

We analyzed results of our experiments using a 2×3 ANOVA (analysis of variance), comparing SART with the textual *cogito* interface and MART with the graphical *cogito* interface. F and p-values are reported to illustrate any significant patterns, interactions, and/or trends within the data.[4]

[2]The University of Regina Computer Science Participant Pool enables undergraduate students enrolled in computer science courses at the University of Regina to participate in evaluations similar to that described in this paper. Students receive up to 2% course credit as reward for their participation.
[3]Analysis conducted using SPSS.
[4]As F-values increase, p-values decrease. P-values less than 0.05 show, with 95% confidence, that there are significant differences between the estimated marginal mean values of the analysis data

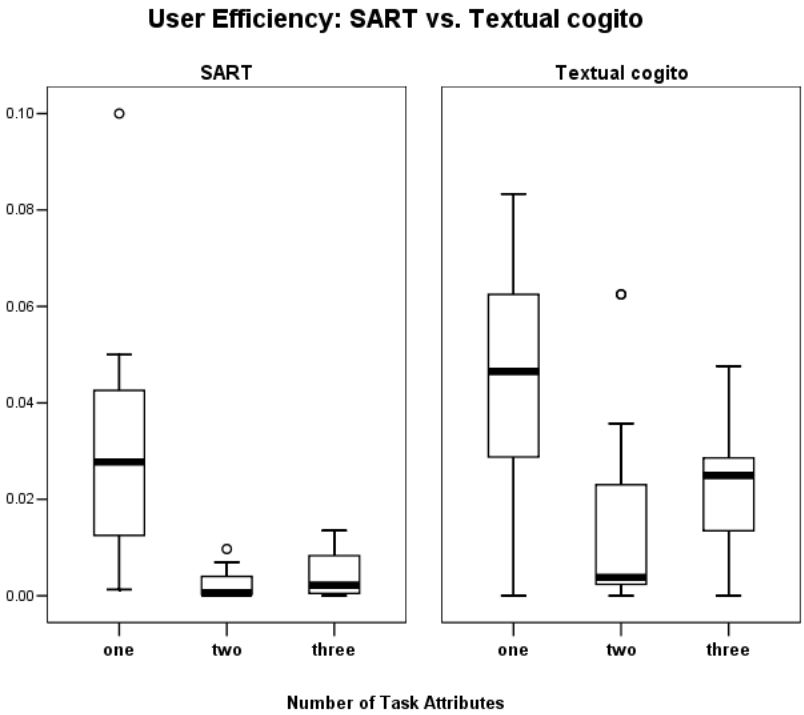

Figure 3: Illustration of user efficiency for the SART and textual *cogito* interface. The illustration provides a comparison of the data accumulated from both experiments. Here, both interfaces were assigned the same question set.

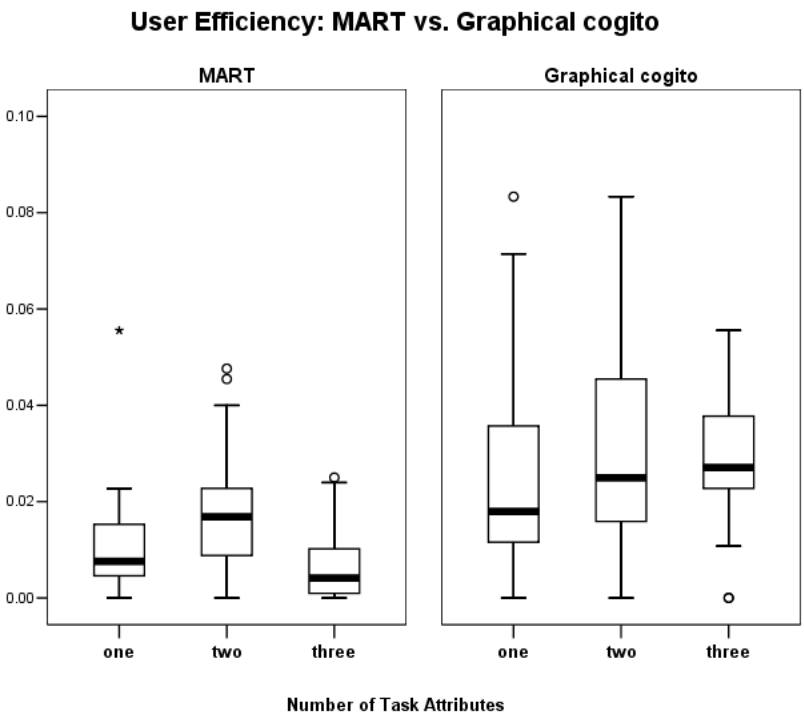

Figure 4: Illustration of user efficiency for the MART and graphical *cogito* interface. The illustration provides a comparison of the data accumulated from both experiments. Here, both interfaces were assigned the same question set.

Table 1: Experiment Results: Mean and Standard Deviations for User Response Time and Task Score for the US-EPA tools and the *cogito* interfaces (Colours are to indicate the same question set and are used to aid comparison)

	Time (sec)		Score (0-1)			Time (sec)		Score (0-1)	
Int.(Attr.)	$\bar{x}$	σ	$\bar{x}$	σ	Int.(Attr.)	$\bar{x}$	σ	$\bar{x}$	σ
First Experiment Results, US-EPA suite					Second Experiment Results, *cogito* Interfaces				
SART(1)	32.58	24.03	0.81	0.33	Text.(1)	24.00	11.58	0.96	0.19
MART(1)	83.18	50.65	0.74	0.35	Graph.(1)	50.89	33.66	0.86	0.30
WART(1)	61.68	36.80	0.61	0.49					
SART(2)	50.39	23.07	0.26	0.22	Text.(2)	34.04	29.66	0.51	0.38
MART(2)	58.36	35.83	0.79	0.42	Graph.(2)	35.11	15.04	0.79	0.42
WART(2)	60.21	29.38	0.48	0.42					
SART(3)	123.50	98.05	0.54	0.40	Text.(3)	42.12	15.85	0.88	0.28
MART(3)	103.57	81.64	0.60	0.39	Graph.(3)	37.93	16.32	0.92	0.40
WART(3)	68.07	44.27	0.65	0.48					

4.1 SART vs. Textual Cogito

Comparing the results from Table 1 for SART and the textual *cogito* interface, it appears the textual *cogito* interface provides users with better support. In terms of user response times for SART versus the textual *cogito* interface, there were significant interactions for both the task complexity (F = 35.13, $p \leq 0.000$), with linear relationships, and interface (F = 19.28, $p \leq 0.000$). In terms of user task scores for SART versus the textual *cogito* interface, there were also significant interactions for both task complexity (F = 68.29, $p \leq 0.000$), with non-linear relationships, and interface (F = 20.60, $p \leq 0.000$). Thus, results indicating the textual *cogito* interface being superior to SART were verified

Apparent from results in Table 1 and Figure 3 is a sharp drop in both SART and the textual *cogito* interface for the two-attribute question set. Upon further analysis, it became clear that users had to perform additional tasks in order to fully answer this question. By the definition provided by the GSA of *air pollution potential*, an air pollution value of *N/A* means that the product is exempt from materials and ingredients that have potential to harm the air. Thus, a value of *N/A* would be the same as having 0% air pollution potential. When observing the data, it was shown that 21/28 (75%) of participants who used SART did not include these products. As well, it was shown that 17/28 (61%) participants who used the textual *cogito* did not include these products. Here, we observe the crucial role between attribute terminology and end-user comprehension and how it reflects upon interface quality.

4.2 MART vs. Graphical Cogito

In terms of user response times for MART versus the graphical *cogito*, there were significant interactions for both the task complexity (F = 22.46, $p \leq 0.000$), with non-linear relationship, and interface (F 20.64, $p \leq 0.000$). However, in terms of user task score, there were no significant interactions in terms of task complexity (F = 0.01, p = 0.91) and only marginal significance in terms of the interfaces (F = 5.81, p = 0.02). Although this was observed, given the overall results described in Table 1 and the depiction of user efficiency in Figure 4, the graphical *cogito* interface provides the users with marginally better support.

Furthermore, with respect to participant preferences for interface design, the US-EPA versus *cogito* interface, the *cogito* interfaces were the most preferred consistent across all tasks, as described in [Maciag and Hepting, 2008].

5 CONCLUSION

This paper described how user performance could be used to measure the quality of user interfaces for decision support tools. Two experiments were conducted and results of each were matched and analyzed. The evaluation highlighted the need to provide users with a high quality interface for information seeking and searching tasks. The experiments provided indication that the *cogito* interfaces provided better support to participants. In terms of user performance, the *cogito* interfaces were shown to be successful. Some of the issues that remain include system terminology, e.g. attributes with values as *N/A*. Future work will include researching ways to provide personalized user interfaces for decision support tools. Some work in this direction has already been done [Maciag and Hepting, 2005; Maciag et al., 2005]. More adequately understanding user preferences will help personalize the user's search experience, highlighting those features, attributes, and values that the user prefers thus, improving their overall user experience.

6 ACKNOWLEDGMENTS

We wish to thank Don Sharpe at the University of Regina for his helpful comments on statistics. The authors also gratefully acknowledge the support of the Centre for Sustainable Communities at the University of Regina and the Natural Sciences and Engineering Research Council (NSERC) of Canada.

REFERENCES

Fischer, G. User modeling in human-computer interaction. *User Modeling and User-Adapted Interaction*, 11:65–86, 2001.

Grinstein, G., A. Kobsa, C. Plaisant, B. Shneiderman, and J. T. Stasko. Which comes first, usability or utility. *In Proc. IEEE Visualization Conference*, pages 605–606, 2003.

Hamming, R. *Numerical Methods for Scientists and Engineers*. McGraw-Hill, 1962.

Hepting, D. H. Interactive evolution for systematic exploration of a parameter space. In Dagli, C. H., editor, *Proceedings of ANNIE 2003*, pages 125–131. American Society of Mechanical Engineers, 2003.

Kochhar, S., J. Marks, and M. Friedell. Interaction paradigms for human-computer cooperation in graphical-object modeling. *In Proc. GI*, pages 180–191, 1991.

Maciag, T. and D. H. Hepting. Analysis of user classifiers for personalization of environmental decision support system interfaces. *In Proc. ANNIE*, 2005.

Maciag, T. and D. H. Hepting. On the role of reflection and representation in environmental decision support systems. *In Proc. International Environmental Modelling and Software Society (iEMSs)*, 2008.

Maciag, T., D. H. Hepting, and D. Slezak. Personalizing user interfaces for environmental decision support systems. *In Proc. Rough Sets and Soft Computing in Intelligent Agent and Web Technology*, 2005.

Phillips-Wren, G. E. and G. A. Forgionne. Aided search strategy enabled by decision support. *Information Processing and Management*, 42:503–518, 2006.

Pu, P. and L. Chen. Integrating tradeoff support in product search tools for e-commerce sites. *In Proc. EC*, pages 269–278, 2005.

Raskin, J. *The Humane Interface: New Directions for Designing Interactive Systems*. Addison-Wesley, 2000.

Rosson, M. B. and J. M. Carroll. *Usability Engineering, Scenario-based Development of Human-Computer Interaction*. Morgan Kaufmann Publishers, 2002.