Development of a Naval Air Warfare Decision Support Interface Using Rapid Prototyping Techniques[1]

Mark St. John, Daniel I. Manes, and Ronald A. Moore
Pacific Science & Engineering Group, Inc.
6310 Greenwich Drive, Suite 200
San Diego, CA 92122
(619) 535-1661
stjohn@pacific-science.com
dmanes@pacific-science.com
ramoore@pacific-science.com

C.A.P. Smith
Space and Naval Warfare Systems Center
53560 Hull Street, Building A33, D44210
San Diego, CA 92152-5000
(619) 553-5411
cap@spawar.navy.mil

Abstract

In this paper, we describe our development of a Naval air warfare decision support interface called the Response Manager using rapid prototyping and Decsion-Centered Design. Specifically, we took eight steps to ensure that our design would meet the needs of its intended users and possess a high degree of usability: 1) recruit expert informants, 2) assemble a comprehensive set of possible features to be used in the design, 3) have experts evaluate and rank these features for development priority, 4) generate alternative display concepts for several higher ranked features, 5) have experts evaluate these alternative designs through interviews and surveys, 6) generate display designs that integrate the knowledge gained from the previous steps and incorporate human factors interface design guidelines, 7) have experts evaluate these display designs through interviews and surveys, and 8) repeat steps 4-7 as resources allow.

1. Introduction

Decision-Centered Design (e.g. Wagoner, Gilmour, Durante, Smith, & Castellano 1997) consists of several phases. It begins by attaining an understanding of the cognitive and procedural characteristics of the human task that the interface is intended to support. Once these characteristics are reasonably understood, prototyping can begin. Prototyping involves generating, developing, and refining display concepts, and then developing those concepts into actual display designs. The goal is to produce displays that suit the cognitive and procedural characteristics of the task and maximize usability. Following the prototyping phase, a formal

[1] This work was sponsored by the Office of Naval Research and managed by SPAWAR System Center, San Diego, code 44.

usability study may be conducted to certify the design's usability or to find additional areas for improvement. The depth of each phase depends on many factors that vary from project to project.

While prototyping in general is a standard engineering practice, the focus and added value of Decision-Centered Design comes from what is called *rapid* rototyping. Rapid prototyping consists of fast and frequent iterations between low fidelity designs and concepts on the one hand and user interviews and feedback on the other hand. The hope is that this iterative procedure will lead to better designs more quickly and for less cost than traditional high fidelity prototyping. "Rapid (throw-it-away) prototyping aims to collect information on requirements and the adequacy of possible designs [from users, and it] recognizes that requirements are likely to be inaccurate when first specified" (Preece, 1994). Thus, rapid prototyping allows designers to refine requirements, concepts, and designs based on feedback from users. Critically, because of prototypes, even low fidelity prototypes (such as storyboards), users can see concrete illustrations of concepts. This concreteness makes for much more useful and accurate user feedback.

We recently implemented a rapid prototyping procedure to aid our design of a decision support tool for Naval air warfare. The tool, called the Response Manager, supports the evaluation and selection of responses that can be taken by a single Aegis-class ship toward threatening air contacts. The goal for the tool is to remind operators of appropriate responses for evaluation and which responses have already been taken. When an operator must concurrently monitor several contacts, these reminders can be quite valuable.

The rapid prototyping procedure we implemented consisted of the following eight steps which helped to manage and control the process: 1) recruit expert informants, 2) assemble a comprehensive set of possible features to be used in the design, 3) have experts evaluate and rank these features for development priority, 4) generate alternative display concepts for several higher ranked features, 5) have experts evaluate these alternative designs through interviews and surveys, 6) generate display designs that integrate the knowledge gained from the previous steps and incorporate human factors interface design guidelines, 7) have experts evaluate these display designs through interviews and surveys, and 8) repeat steps 4-7 as resources allow. We illustrate these steps with examples from our development of the Response Manager.

2. Response Manager Background

In fast-paced, busy environments, such as single ship, littoral, naval air warfare, it is difficult to 1) generate alternative plans and options for consideration and 2) keep track of each concurrently developing situation and which actions have already been taken. Currently, Commanding Officers (COs) and Tactical Action Officers (TAOs) use such *ad hoc* tools as grease pencils and note cards to aid memory and keep track of each situation, but these involve bookkeeping that consumes valuable time and mental resources, making these methods subject to error.

A solution to these problems would be to provide the CO/TAO with a tool called a *Response Manager*. A Response Manager is a display that supports tactical decision making by presenting plans and options for handling various situations as well as history information regarding actions already taken. We call these plans and options *response sets*, and there may be several different sets to fit different situations. Because the Response Manager is computerized, it also supports rapid modification of the response sets as conditions and official Rules of Engagement change, and it ensures that the latest response sets are displayed consistently throughout the Combat Information Center and across watches.

Work conducted in the Tactical Decision Making Under Stress (TADMUS) program at the Space and Naval Warfare Systems Center (SSC) in San Diego and sponsored by the Office of Naval Research served as the impetus for the Response Planner and Manager (RPM) project. The TADMUS project developed a Decision Support System for single-ship air warfare based on principles of naturalistic decision making. It was designed as a series of distinct display modules that could augment existing geo-plot and text displays in current ship combat information centers (Moore, Quinn, & Morrison, 1996). One of these modules presented the preliminary concept for a Response Manager (see Figure 1). In this design, response options are displayed on a Gantt chart showing the range from own-ship along the x-axis. The white vertical line indicates the current range of one aircraft of interest, and responses turn green once they have been executed.

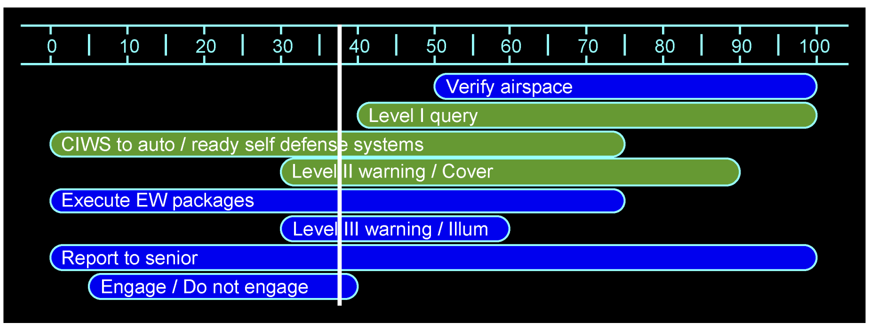

Figure 1. Early design of a Response Manager module from the TADMUS project.

Naturalistic decision making theories (e.g., Klein, 1989, 1993; Klein, Orasanu, Calderwood, & Zsambok, 1993), provided a framework in which to develop the original design and use of the Response Manager as well as our current design efforts. According to these theories, tactical decision makers interpret the current situation to the extent that it is similar to another situation with which they are already familiar. The remembered situation and its associated plan of action are retrieved from memory for use in the current situation.

The Response Manager supports this process by: 1) displaying relevant data about each air contact being assessed, so that relevant memories can be retrieved, 2) displaying a set of relevant responses (and their timing) that could be taken for each contact so that appropriate responses can be considered, and 3) providing reminders about which responses have been taken and which have yet to be taken in conjunction with each contact, so that the histories for multiple concurrent contacts can be remembered without overloading short-term memory.

The Response Manager display shown in Figure 1 was the result of careful analysis of the operator's tasks and decisions. Operator's tasks and decisions were identified through extensive, in-depth interviews. Then, the displays were created by expert designers to maximize operators' efficiency for performing those tasks and making those decisions. The RPM project used this design as a foundation and used rapid prototyping (a stage in Decision-Centered Design) to develop the concept. Specifically, our goal was to clarify a number of remaining open questions regarding Response Manager functionality and to refine the concepts and display designs of the Response Manager.

3. Rapid Prototyping Steps

3.1 Recruit Expert Informants

All stages of prototyping depend on access to Subject Matter Experts (SMEs). This is especially true for rapid prototyping, where the short time between iterations requires frequent input from experts. Ideally, several SMEs should be available for frequent and in-depth discussions about the content domain and display concepts and designs. Additionally, it is valuable to have access to a larger number of SMEs for one-time interviews and surveys regarding each prototype iteration.

The RPM project team consisted of five very experienced SMEs who we called upon for frequent discussions. We also scheduled one-time interviews with a variety of naval officers as opportunities arose.

3.2 Assemble a Comprehensive Feature List

Complex interfaces, such as decision support tools, can contain a tremendous number of specialized features, and each feature can potentially be represented in a variety of ways. Hence, a useful first step for organizing the research effort is to develop a comprehensive list of potential features and design questions. Typically, many different types of information might impact a decision, so it is important to determine whether each piece of information should be immediately available to the user, available only upon request, or not available at all. Other issues include how information should be organized and represented and what user interaction should be possible.

In designing the Response Manager, we began assembling a list of potential design features by identifying the different components of the TADMUS Response Manager. We then extended this list based on extensive interviews with project SMEs and a review of debriefing interviews with naval officers who had participated in TADMUS experiments (Heacox, 1996; Kelly, 1996; Moore, 1996; Moore & Fehér, 1997). This work culminated in a comprehensive set of 39 HCI design features, organized into the following five categories: 1) response organization and grouping, 2) response and track status, 3) constraints on responses, 4) range and time, and 5) miscellaneous. For each feature, we wrote a short description and developed a set of potential advantages and disadvantages. Table 1 contains a sample of five of the features from this list, one from each category.

Table 1. Sample items from the comprehensive list of 39 potential Response Manager features.

Category	Feature	Potential Advantages	Potential Disadvantages
Response organiza-tion and grouping	Display all available air responses for each contact regardless of current threat assessment (*vice* display only appropriate responses)	• Support a less biased interpretation • Support considering more responses	• Lead to overly passive or aggressive responses • Increase display complexity
Response and track status	Indicate when a response was ordered (requested) and completed	• Aid memory • Increase the salience of unordered and uncompleted responses	• Increase display complexity
Con-straints on responses	Indicate equipment failures that would effect the execution of responses	• Increase awareness of ownship's state • Reduce interaction with other personnel	• Increase display complexity • Increase user interaction
Range and time	Provide range windows (bars whose endpoints define "windows of opportunity" for completing a response)	• Support proper timing of range-dependent responses • Reduce time to find a given response	• Emphasize timing at the expense of good judgement • Increase display complexity
Miscell-aneous	Provide means of entering and displaying comments	• Increase flexibility • Aid memory • Reduce interaction with other personnel	• Increase display complexity • Increase user interaction

3.3 Establish Feature Rankings

Once a list of potential features and questions has been prepared, the features should be ranked according to functional importance and development priority. Researching features in rank order is important for project success because all projects are limited by time and money. Giving research priority to the most important features and questions ensures that adequate time can be given to developing the foundation of the interface. This will allow for an overall design that is well integrated and performs the minimum required functions, even if project resources are restricted or insufficient. Of course, establishing this ranking is difficult and requires substantial expertise, both in the content domain and in human factors. Experts in the content domain and experts in human factors should determine what is important and what is achievable. But since there are few people who are experts in both fields, there is substantial art in developing a ranking. Discussion among both SMEs and human factors researchers can significantly benefit this process.

In the case of the Response Manager, we asked five project SMEs and two human factors experts who were very familiar with air warfare to rank the 39 design features in terms of their operational utility and to provide justification for their ranking by commenting on each feature. The participants determined their rankings individually to ensure that any differences in opinion would be recorded. This process was valuable for organizing our development efforts and generating discussion about the value of each feature. We also conducted a focused interview of three project SMEs regarding these features. The three participants were interviewed simultaneously to encourage discussion and consensus building. This helped to further our understanding of the advantages and disadvantages of the features and the issues surrounding them.

3.4 Generate Alternative Display Concepts

The next step is to focus on groups of two or three of the highest ranking features, and develop alternative display concepts for each. These storyboards may demonstrate different variations on a piece of functionality (e.g., whether to show time remaining or distance remaining), alternative ways to represent a piece of information (e.g., graphically versus verbally), or how much detail to represent. The storyboards for this phase can often be highly simplistic, focusing solely on the key questions that need to be answered. Experts will generally be able to understand and discuss the concept following an explanation of the project's background and the purpose of the interface.

One question we faced in designing the Response Manager was how the responses should be laid out on the display. Project team members debated between a cascading chart with extensive range information or a straight, vertical list containing only limited range information. In an effort to resolve our differing opinions, we developed a storyboard for each design, keeping elements of the design unrelated to our key questions the same on both displays (see Figure 2). The major difference between the two storyboards was that one employed cascading range windows (bars whose endpoints define "windows of opportunity" for completing a response), while the other contained bars that extended across the length of the display and whose labels lined up along their left edge. Note that the underlying issues of the physical and procedural constraints on response ranges is very complex and open to debate among experts.

Figure 2. Storyboards showing two alternative arrangements for the layout of responses.

Another question was how the responses we wanted operators to consider should be organized on the display. The first alternative was borrowed from the TADMUS display, which was derived from Navy practice and documents called "[standing] battle orders". This alternative involved arranging responses into a single list from the first response to consider when an aircraft is detected to the last response to consider before making the decision whether to shoot down the aircraft. This list was an integration of four categories of responses (communication, aircraft identification, ship defense, and weapon preparation). The alternative arrangement under consideration was to separate these different categories of responses into four separate lists. These categories were established through discussions with project SMEs. Both arrangements, therefore, were sensible, operationally valid, and preferred by at least some experts. We next created simple storyboards showing these two different arrangements.

Many other features, both major and minor, were addressed in a similar fashion.

3.5 Have Experts Evaluate these Alternative Designs

Questionnaires and interviews are very effective in eliciting information from experts. While the answers to the questions themselves are often highly informative, the explanations for those answers often give valuable insight into the operator's decision making processes. When conducting in-depth interviews, five to 10 SMEs is usually sufficient to evaluate most sets of design alternatives. Formal preference ratings may require more users to establish statistically reliable effects. Such a small number of participants is acceptable because experience has shown that additional users provide mostly redundant information. Nielsen (1997), writing in the *Handbook of Human Factors and Ergonomics*, estimates that the probability of additional users finding new problems with an interface drops exponentially, meaning that more than a handful of users buys little extra information about a prototype design.

In designing the Response Manger, once we had generated a number of storyboards representing alternative design concepts, we would recruit between six and nine SMEs unfamiliar with the project to evaluate them. In order to answer the question of whether it is preferable to use a cascading chart or a vertical list, we showed the two alternative storyboards to six different SMEs, and had them fill out a questionnaire. We asked them to choose which display they preferred and to rate the potential value of range windows on a five-point scale. They also provided explanations for their answers. The results indicated a strong preference for the cascading chart display (five to one), and the average rating for the range windows was midway between "very useful" and "somewhat useful." The comments indicated that cascading range windows could be a valuable memory aid, provide useful visual cues, and help time operator actions. The SMEs also provided suggestions for improving the display, such as making the range windows adjustable.

Similarly, in order to answer the question of displaying four separate lists versus a single integrated list, we created storyboards for each alternative and showed them to SMEs. We found that even though the four separated lists made sense to all of the SMEs, most of them preferred the single integrated list arrangement.

3.6 Generate Display Designs that Integrate Knowledge Gained

Once most of the higher ranked potential features and issues have been evaluated, the findings and accumulated insights can be used to generate more complete interface storyboards. These prototypes should represent a careful integration of the feature designs developed throughout the design process. It is often helpful to establish a style guide for colors, shapes, and data access methods, so that similar data can be retrieved by similar actions. Additionally, display elements should be arranged in an organized manner so that related information is placed close together. Human factors interface design guidelines, both generic to all interfaces and specifically developed for military displays, are also invaluable.

Several integrated Response Manager storyboards were developed toward the later stages of the RPM project. These combined a number of concepts and lessons learned from the previous steps of the process. The final version (see Figure 3) was a high fidelity prototype that allowed visualization of all the major features as well as our vision for the look and feel of the final system. These designs served to refine our implementation of the cascading range windows and numerous other features, such as status indicators, Rules of Engagement indicators, and a threat-response guide. The status indicators, which appear on the right sides of the bars, indicate how many times a response has been ordered or completed and whether there were any problems in executing the response. The ROE indicators, which appear as small triangular cutouts in the edges of the bars, show the range specified by ROE for executing particular responses. Finally, the threat-response guide indicates the responses that are appropriate for the current threat assessment (i.e., unknown, friend, assumed commercial, or assumed enemy).

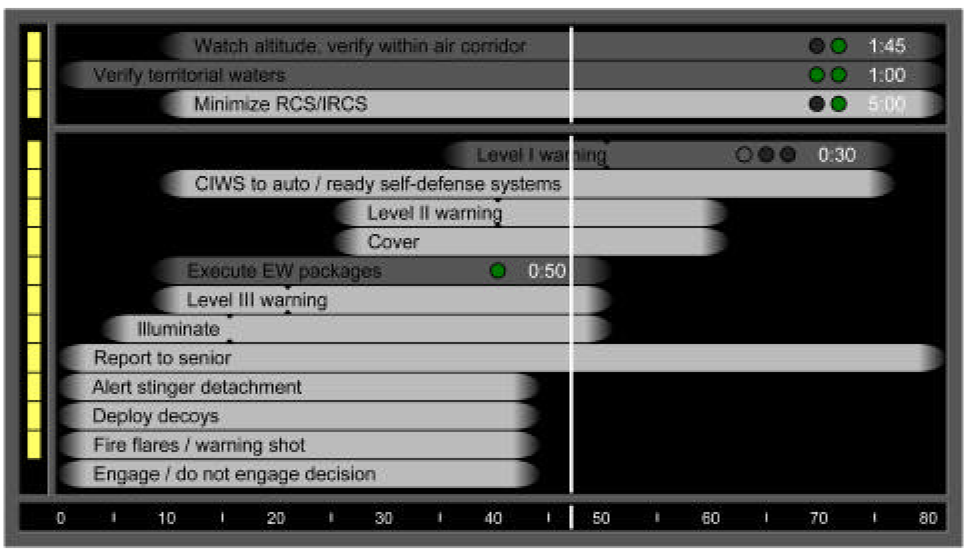

Figure 3. The RPM Response Manager.

3.7 Have Experts Evaluate these Display Designs

The next step is to have SMEs evaluate the integrated display designs. As always, it works well to orient the SMEs to the display and how it would be used in practice. When describing the display, it is helpful to begin with a discussion of the major modules and their purpose, then to describe the details. Building some structure into the interview helps ensure that major points get covered and the discussion doesn't go far off topic, but a loose structure gives the SMEs more freedom to discover issues and problems that might otherwise go unnoticed. It is often useful to interview SMEs in groups of two or three so that they can react to one another, thereby deepening the discussion. Finally, a handful of SMEs is often sufficient to identify all but the more idiosyncratic problems.

After one of the design iterations, we conducted a focused interview with three of our project SMEs to elicit ideas on how to improve the design. They provided valuable input on improving the sequencing of responses and adding a visual separator between monitoring-type responses and action-type responses. They also helped clarify the threat assessments represented on the display, both in terms of their labels and the colors used to designate them. Finally, they provided further validation that the cascading range windows, status indicators, and ROE indicators were indeed valuable features.

3.8 Repeat Steps 3.4 to 3.7 as Resources Allow

There is generally no specific point to stop the process and consider the design complete, but there is a point where a further iteration would fall victim to the law of diminishing returns. If the latest iteration doesn't result in a worthwhile improvement over the previous one, it may be time to move on to a formal evaluation. Furthermore, because most projects are constrained by time and resources, there may be no need to debate when to discontinue the design process.

4. Discussion

We used a rapid prototyping procedure to design a Response Manager decision support tool for Air Warfare. Starting from a foundation of extensive interviews and a preliminary design from the TADMUS project, we implemented eight steps to govern the rapid prototyping process. The steps began with collating, organizing, and ranking a large number of potential features and issues that expert users had identified from the preliminary design. Then we progressed through the ranked list, iteratively investigating alternative concepts and display designs. The frequent interaction with users who could see and experiment with concrete prototypes of our design ideas resulted in valuable feedback.

A common problem with rapid prototyping is that the process can be difficult to manage. We believe that the eight steps described here will be useful in minimizing this problem. The steps provide structure to the iterative development process. The most important features and issues are identified early (though revisions are certainly possible as more is learned), and development begins with these features. Periodic integration into a coherent display also helps to maintain control over the development process. There is always a "state of the art" integrated display to document and illustrate current project design and thinking.

The value of this Decision-Centered Design procedure is that initial ideas, even when based on substantial cognitive task analyses, can be incomplete and askew. The multiple iterations involved in rapid prototyping allow opportunity for concepts and designs to be refined, effectiveness to be increased, and opportunity for missed requirements to be discovered. Importantly, the concreteness of the prototypes allows researchers to convey design ideas in a way that users can understand, think about, and even simulate using.

But, of course, rapid prototyping works best in conjunction with other research methods. For instance, the design of a new interface best begins with a firm understanding of the cognitive and procedural tasks of the user. This understanding is best achieved through cognitive task analyses and intensive “knowledge engineering” interviews with users. Rapid prototyping then proceeds from this foundation.

Also, some design questions are better answered through formal behavioral experiments rather than user interviews and preferences. It is not uncommon to find cases where users claim to prefer interface features that actually impede performance. Further, documentation of actual performance enhancement is crucial for promoting the final interface design. For example, formal empirical evaluation was required to help resolve one of our most serious design issues—how operators might be influenced by the specific responses shown in the Response Manager display. Our concern was that the decision makers would tend to choose only among the responses presented to them, even though other responses might be advised if the situation were not as originally assumed (a tendency we referred to as a “response presentation effect”). Presenting responses applicable to several situations, however, might result in confusion, clutter, or information overload. Operator preferences were decidedly mixed.

To help resolve this issue, we conducted two controlled experiments to examine the effects of different Response Manager designs on actual decision making performance (St. John, 1998; St. John, Kelly, and Seymour, 1997). We found that even highly experienced users tended to draw responses mostly from the response set they were shown. Importantly, this presentation effect did not influence the threat assessments of aircraft—only responses/choices. This finding means that presenting only one of the two sets did not influence the way users interpreted threats and non-threats, but it did influence which responses they issued to deal with those threats and non-threats. We concluded that showing more options—both response sets—was preferable because it would remind operators of their full array of choices.

In summary, rapid prototyping plays an integral, but not solitary, role in the Decision-Centered Design process. In conjunction with extensive up-front cognitive task analyses, formal experimentation where appropriate, and careful attention to transition toward the end, it can dramatically improve the usability and efficiency of an interface.

5. References

Heacox, N. J. (1996). *RPM display design considerations*. Technical Memorandum; December 5, 1996.

Kelly, R. T. (1996). *RPM HCI issues and design considerations*. Technical Memorandum; November 22, 1996.

Klein, G. A. (1989). Recognition-primed decisions. In W. R. Rouse (Ed.), *Advances in man-machine systems research*, vol. 5 (pp. 47-92). JAI Press, Inc.

Klein, G. A. (1993). A recognition-primed decision (RPD) model of rapid decision making. In G. A. Klein, J. Orasanu, R. Calderwood, & C. E. Zsambok (Eds.) *Decision making in action: models and methods* (pp. 138-147). Norwood, NJ: Ablex Publishing Corporation.

Klein, G. A., Orasanu, J., Calderwood, R., & Zsambok, C. E. (Eds.). (1993). *Decision making in action: models and methods*. Norwood, NJ: Ablex Publishing Corporation.

Moore, R. A. (1996). *RPM notional display screens*. Static Display Mock-Ups.

Moore, R. A. & Fehér, B. (1997). *Review of early HCI issues*. Technical Memorandum; April 2, 1997.

Moore, R. A., Quinn, M. L., & Morrison, J. G. (1996). A tactical decision support system based on naturalistic cognitive processes. In *Proceedings of the Human Factors and Ergonomics Society 40th Annual Meeting*, (p. 868). Santa Monica, CA: Human factors and ergonomics society.

Nielsen, J. (1997). Usability testing. In G. Salvendy (Ed.) *Handbook of human factors and ergonomics, 2nd edition.* New York: John Wiley & Sons.

Preece, J. (Ed.) (1994). *Human-computer interaction* (Ch. 27). New York: Addison-Wesley.

St. John, M. (1998). The impact of a response management tool on air warfare tactical decision making. In *Proceedings of the 20th Annual Conference of the Cognitive Science Society*, (pp. 1013-1018). Mahway, NJ: Lawrence Erlbaum Associates.

St. John, M., Kelly, R. T, & Seymour, G. E. (1997). *The impact of response management tools on air warfare tactical decision making*. Technical report. San Diego, CA: Pacific Science and Engineering Group.

Wagoner, R. C., Gilmour, T. H., Durante, G. G., Smith, R., & Castellano, F. X. (1997). Command 21: Speed of command for the 21st century. Research report of the Command 21 concept generation team. Newport, RI: Naval War College.