

Human Performance Modeling for Command and Control of the Tactical Tomahawk

Track
C² Decision Making and Cognitive Analysis Category

Mary L. Cummings

Advisor
Stephanie Guerlain

Department of Systems & Information Engineering
University of Virginia
151 Engineer's Way
PO Box 400747
Charlottesville, VA
22904-4747
mlc3u@virginia.edu
434-293-0779
603-697-6044 (Fax)

Student paper submission

Abstract

The U.S. Navy is currently developing a new version of the Tomahawk land attack missile, called the Tactical Tomahawk, which will have the capability of redirection in-flight. This is a significant future command and control addition since current versions of the Tomahawk cannot be redirected in-flight. The implementation of the Tactical Tomahawk means that not only will battlefield commanders have more flexibility and options; it also means that a layer of human command and control will be needed where none previously existed. Requiring an operator to manage high value assets in the close-in combat arena through constant replanning requires substantial cognitive contribution, which will significantly impact the effectiveness of future Tomahawk engagements as well as the entire command and control picture for these types of engagements. This paper reports the results of the cognitive work analysis for the new human-in-the-loop aspect of the Tactical Tomahawk and how the analysis contributed to the design of both a monitoring and retargeting human computer interface. In addition, results from a human-in-the-loop simulation study in which 42 Navy personnel were tested on the two-screen user interface are discussed.

Introduction

The Tomahawk missile is the Navy's premier land attack missile, and indeed, the U.S. military has declared, "Because of its long range, lethality, and extreme accuracy, *Tomahawk*[®] has become the weapon of choice for the U.S. Department of Defense (U.S. Navy 2000)." Both the Gulf War and the recent strikes in Iraq have demonstrated the precision and strategic value of the missile. However, one of the primary drawbacks to the current Tomahawk missile is its "fire and forget" capability. The Tomahawk contains its own internal guidance and navigation system, but once it is launched, it cannot be redirected in-flight. This limitation to the system causes potential redundant demolition of targets, thus wasting the \$1,000,000 missile. In addition, since the missile's flight path, once fired, cannot be modified, the missile cannot respond to dynamic situations in which, for example, a target has moved, or a more critical target emerges.

In response to this shortcoming for what is otherwise a very effective weapon system, the U.S. Navy is in the process of designing a version of the Tomahawk that will have the capability of redirection in-flight. This new version, called the Tactical Tomahawk, will not only be able to provide battlefield commanders with the ability to redirect Tomahawk missiles in-flight, they will also have the ability to position a Tomahawk missile in a loiter pattern to await further instruction. This loiter pattern, much like a commercial aircraft holding pattern, allows missiles to be placed in strategic positions where it is highly likely that a target will emerge in a combat situation. An example of an emergent target would be surface-to-air enemy missile launch platforms, and often their positions are unknown until they actually begin electronic transmissions. Since this often occurs in the middle of a battle, Tomahawk missiles have not been used against these targets since the locations were not known in advance and thus preprogrammed. The new Tactical Tomahawk will however, be able to target this surreptitious threat.

The implementation of the Tactical Tomahawk means that not only will battlefield commanders have more flexibility and options, it also means that a layer of human control will be needed where none previously existed. Introducing the ability to control a very fast-moving tactical weapon in the close-in combat arena requires substantial cognitive contribution, which will without a doubt significantly impact the effectiveness of future Tomahawk engagements. Since no human operator interface exists for this system, one must be built from the ground up.

When previous complex human control systems like the Tactical Tomahawk interface were built, the design approach typically centered on the cognitive limitations of the human as the primary design constraints. While this cognitivist approach to interface design is important, recent research suggests that it does not take into account the impact of the work environment in design. This criticism spawned a new area of human factors called ecological interface design (EID). EID does not ignore or discount the importance of cognitive considerations in design, but it does stress the primary importance of the environmental work domain constraints. The external environment must be deconstructed and analyzed before the cognitive constraints can be effectively understood and incorporated into a design.

However, before a system can be designed to best optimize human-machine interaction using EID principles, it must be analyzed to determine the nature of the human interactions both from a user's perspective as well as an environmental perspective. To this end, the cognitive work analysis (CWA) approach has been developed to provide cognitive systems engineers with a framework for analysis that identifies not only the user's goals and constraints, but also the impact of the constraints of the environment (Vicente 1999). The CWA is a tool for understanding how people interact in a particular domain, constraints of both the domain and worker, tools required, and errors that can be made in such environments. The end result should be a product in which not only worker's interactions are understood, but also one in which otherwise hidden relationships are revealed. When developing a decision support tool for a system that is in the conceptual stages like that of the Tactical Tomahawk, the CWA is particularly useful in identifying critical environmental limitations that help to define otherwise ambiguous system boundaries.

In a cognitive work analysis, once the work domain has been analyzed to identify environmental constraints and modeled through a structural means-ends representation of the system, a control task analysis (CTA) is used to focus on the action means-ends that must be employed to achieve a desired outcome of a particular decision. Since a CWA typically focuses on previously established systems, the CTA usually focuses on expert performance to demonstrate not how a system *should* act, but a system *could* act if given a flexible problem space. Since the Tactical Tomahawk retargeting system has no experts as of yet, this CTA deviated somewhat to demonstrate what tasks are expected to take place during the operational phase.

1. Control Task Analysis & the Decision Ladder

The control task analysis should focus on what needs to be done for task completion as opposed to who should complete the task or how it should be done. In addition, the CTA must answer two fundamental questions, 1) what constraints exist for

the pursuit of goals, and 2) what information and relationships are relevant for particular classes of situations (Vicente 1999)? To illustrate how control actions relate to the decision-making sequences of a worker, Rasmussen (1994) developed the decision ladder model, which represents information processing activities and subsequent states of knowledge that result when the activities are performed. The decision ladder maps rather than models the structure of a decision-making process, and in doing so, helps to identify the requirements of the control tasks. The decision ladder should represent what information processes need to occur, independent of who will perform a task or how a particular control task will be accomplished. In the case of systems with computer-based decision support tools, the decision ladder represents the decision process and states of knowledge that must be addressed by the tool whether or not a computer or a human makes the decision. Because the primary decision-making element of the Tactical Tomahawk retargeting system is the missile selection/retargeting sequence, this is the primary focus of the CTA decision ladder analysis, and is represented in figure 1.

The left side of the decision ladder, the upward leg, represents the information processes that are required for situation analysis, the pinnacle depicts the judgment, and then the downward leg symbolizes the support systems that must exist to implement the decision. For the missile selection/retargeting sequence, the left side of the decision ladder represents notification of either a flex or emergent target situation and the data that must be gathered to make an informed decision. The top of the decision ladder represents the chosen missile, and the circular arrow depicts the iterative nature of this decision. For each retargeting scenario, the decision is affected by various factors, which are never the same. The right side of the decision ladder then represents those activities and states of knowledge that are required to actually implement the decision to retarget. The decision ladder represents the process and states of knowledge that must occur whether or not a computer or a human must make the decision.

This decision ladder represents the fundamental knowledge-based decision process for the missile selection/retargeting sequence. It is possible that within the ladder, an expert would be able to skip certain rungs of the ladder based on experience and familiarity with the system. For example, when the controller reaches the information gathering level of identifying candidate missiles, it is possible that the controller could be given orders from a higher, external authority to use a specific missile against a particular target. If this is indeed the case, then the controller no longer has to move up the ladder to evaluate all missiles weighted against various other factors. The decision has been heavily influenced by external factors and thus the controller moves from the identification stage, to the definition of missile solution phase to begin the action of retargeting. Essentially the decision-making process is removed from the controller and all that remains is to begin the actual sequence of commands to retarget. Thus the controller does not have to move up the information gathering side of the decision ladder at all and can take a shortcut to the implementation depicted in the right side.

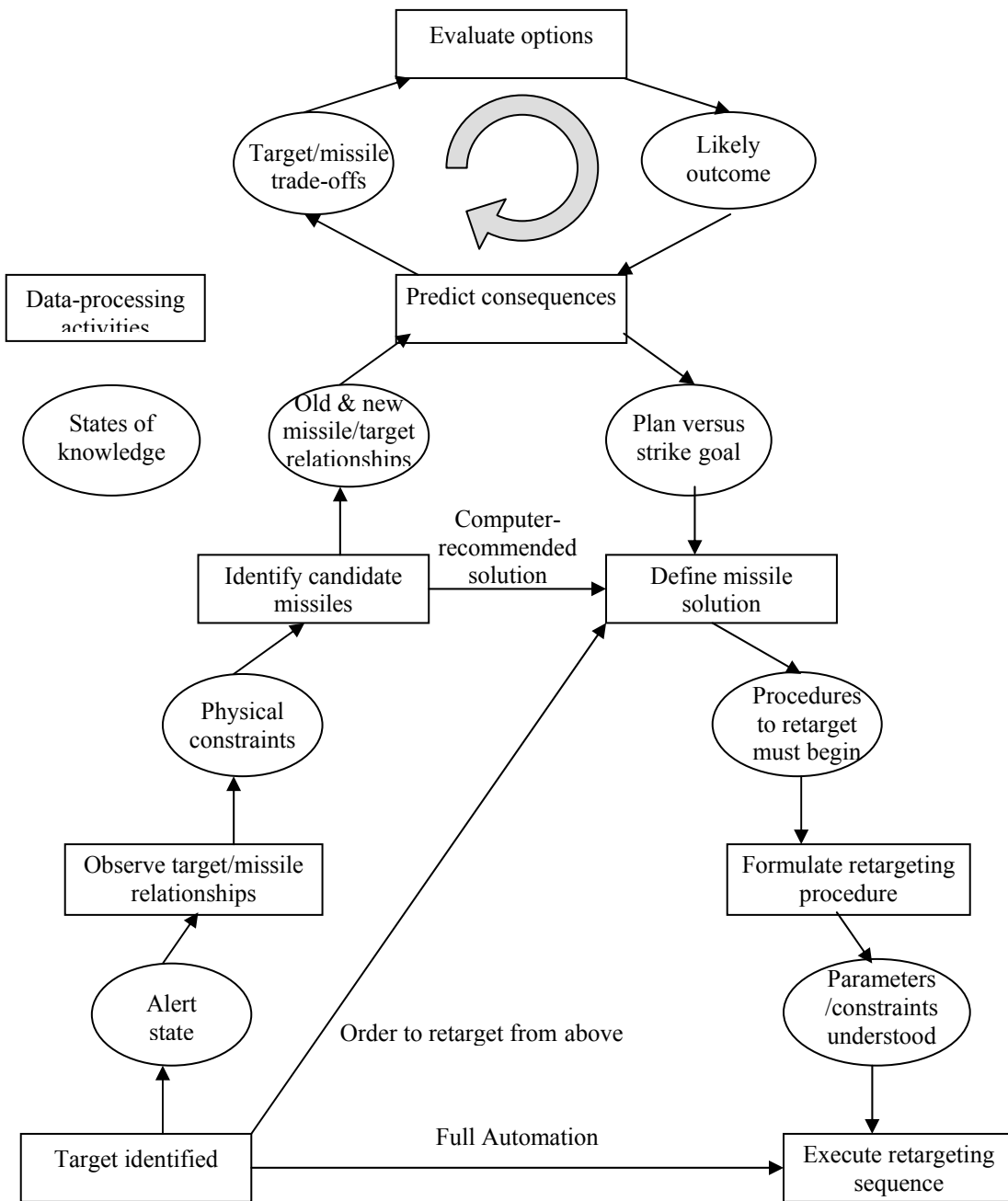


Figure 1: Tactical Tomahawk Retargeting Decision Ladder

However, human processes are not the only way shortcuts can be introduced. Another manner in which various elements of the decision ladder can be bypassed is through the introduction of automation. Since a primary question in the Tactical Tomahawk research is what decision processes should be automated, it was useful to map those decision paths that would likely be automated, and what effect that might have on human performance.

2. The TTIMR Prototype

Once the initial cognitive work analysis was completed for the Tactical Tomahawk, a human computer interface prototype was developed in order to explore workload issues, as well as possible decision strategies for controllers in this complex supervisory task. The prototype is in effect the cognitive model, which allows for both illustrations of the concepts developed in the CWA as well as further defining both the environment and user-related issues. This prototype, a dual screen interface, is called the Tactical Tomahawk Interface for Monitoring and Retargeting (TTIMR, figure 2). Developed in Macromedia Director[®], the right side of the screen provides an overview map display that represents all missiles and targets in a given strike. The left side of the display contains two primary components, the decision matrix and chat box (figure 3). Because this analysis focuses primarily on the retargeting human performance elements, just the left side of the screen will be discussed in detail.

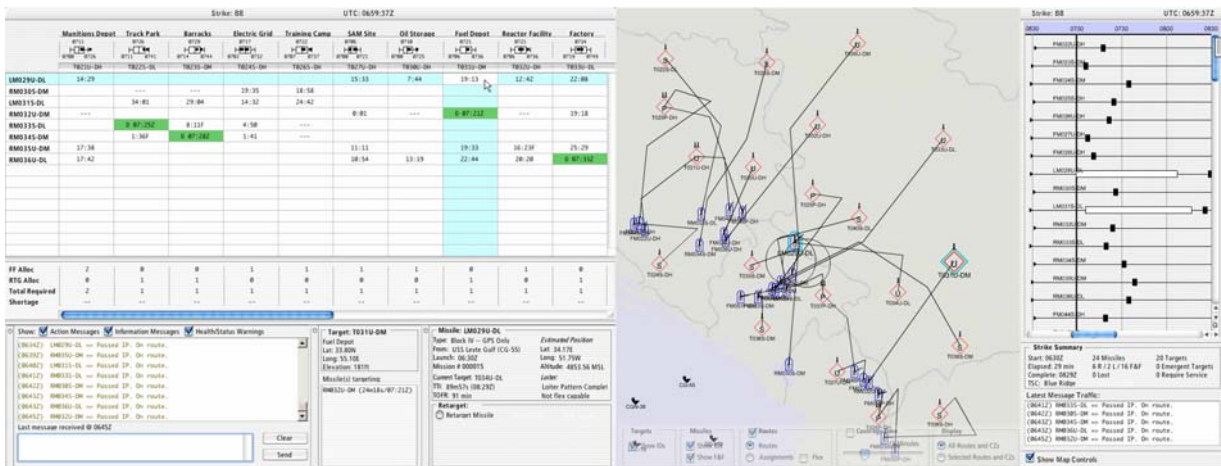


Figure 2: The Tactical Tomahawk Interface for Monitoring and Retargeting (TTIMR)

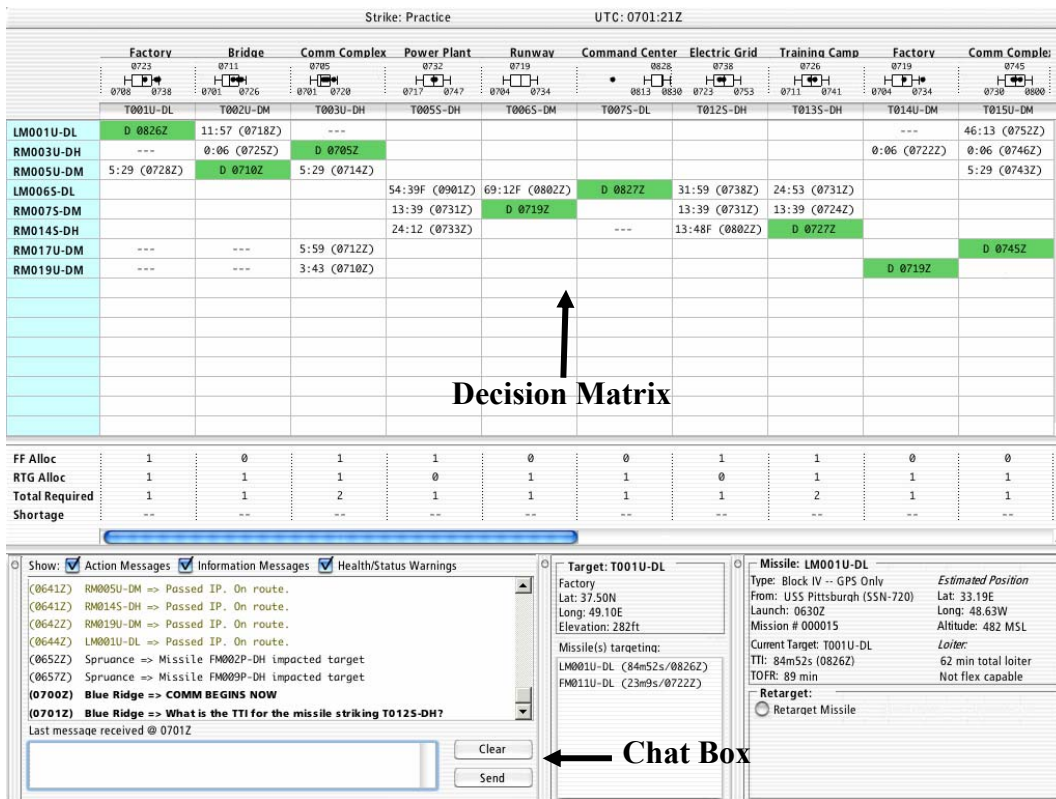


Figure 3: The TTIMR Retargeting Decision Matrix and Chat Box

The decision matrix is a tool that provides the user with the ability to see not only the current status of all missiles capable of retargeting, but all future possibilities as well. All retargetable missiles are listed in the left hand column, and all targets in a strike, including emergent targets, are represented across the top. Current missile/target pairs and their associated time on targets are highlighted, and empty cells in the matrix demonstrate that no possible relationship exists between a missile and a target. The remaining cells provide the user with information about not just whether a missile can get to a target, but equally as important, when the missile can get to the target and the time remaining left for the operator to make the decision to retarget. The dashes in the matrix are linked with the small white box at the top of every target. The small white box is called the window of opportunity (WOO).

The WOO is an interactive decision tool which gives the operator (and also battlefield commanders) greater flexibility in determining if and when a missile should be retargeted to an existing target. The development of the WOO can be directly linked with the abstraction hierarchy (figure 2), in which one of the three priority measures was to determine the time on target (TOT). The abstraction hierarchy made it clear that regardless of the situation, the TOT had to be a primary consideration, however, with the movement of missiles in flight, the TOT is not a static number and constantly changes with the evolving missile/target relationships in the lower level of general function

For example, when mission planning occurs prior to the launch of missiles, a strike planner determines the TOT for each missile going to each target. However, due to both the current system limitations and the nature of missiles flying through the air at ~500 knots, the problem of retargeting a missile in-flight is not just simply “Can a missile get to the target?” rather, the problem is “Can the missile get there by a certain time?” For example, if a missile was preplanned to hit a target at 0730Z specifically at that time because of possible friendly forces in the area either earlier or later in the strike, then it would be very important to only retarget a missile that could get to the target at its original TOT.

However, due to communication network and other missile constraints (like the fact that a loiter missile can only leave its pattern at one point), it would be very difficult to guarantee that a missile could get to a target at specifically 0730Z. However, if a window of time were put around that particular TOT, then the probability of being able to retarget a missile to the target would increase in direct relationship to how large that window of time was. The WOO allows the controller to determine for a particular TOT, can a missile get there within a time window ranging anywhere from +/- 1 min to the length of the strike. For targets that do not require missiles to arrive at or within a certain time period, the WOOs could be extended (through a click and pull mechanism) for the length of the strike. If a missile absolutely had to arrive within a 5 minute window, the WOO could be hard wired and would not be interactive. However, this would not be advisable because what if a missile could make it to a target in 7 minutes? It is possible that some sort of communication and negotiation would allow involved parties to determine if this would be the best course of action and make the most informed decision possible.

The WOO provides both the greatest flexibility possible in retargeting missiles while also presenting all possible solutions to operators (and thus battlefield commanders) in a concise format. Future versions of the WOO could be incorporated to include some form of user initiated notification.

The other significant display component for the left hand side of the display is the chat box, also referred to as the communication box. The popularity of the chat box is not just in mainstream pop culture, but is already in use today aboard many Navy vessels. The chat box is particularly useful because it provides for not just real time communication, but also archiving the message history so operators can reconstruct what has transpired if needed, thus providing for overall enhanced situational awareness. However, one of the drawbacks to chat systems, both in and out of the military, is the tendency to overuse and perhaps saturate the network because of the ease of use. In addition, it can also be very difficult to add the appropriate input at the correct times if too many or undisciplined users participate in the chat.

The chat box was included in the TTIMR design not only to provide a realistic feel for the system, but also as an embedded secondary workload measurement tool as it allows for both measurement of time delay in responses as well as accuracy of responses. Traditional secondary tasking can be intrusive and introduce an unrealistic artifact, however, embedded secondary tasks do not fundamentally change the task or task performance and provide more sensitive measurements (Shingledecker 1987; Tsang and Wilson 1997; Wickens and Hollands 2000).

3. The Human-in-the-Loop Experiment

After the TTIMR prototype was completed, an experiment was conducted to determine the effectiveness of the TTIMR design as well to begin exploration of workload and decision strategies. The initial research goal of the experiment was to determine how increasing the number of missiles that a controller assigned would affect overall performance as well as increasing the tempo of operations.

3.1 Participants

Forty-two Navy personnel, both active duty and retired, participated as subjects. Since the retargeting ability for the Tactical Tomahawk does not exist, there are no subjects that have any actual experience with a system similar to TTIMR. The active duty personnel (N=26) consisted of both officers and enlisted personnel who currently worked with the fire-and-forget versions of the Tomahawk. These personnel would most likely be the ones who would actually be the controllers of the future Tactical Tomahawk. The remaining personnel (N=16) were retired Navy personnel who either had significant experience in operating older versions of the Tomahawk or had experience in strike planning.

3.2 Procedure¹

All subjects received approximately three hours of training which included a slide presentation to explain the TTIMR prototype, two training sessions each lasting approximately 25 minutes, as well as observing an additional training session other than their own. The rules for retargeting were explained during all phases of training and reviewed before testing. After training, all subjects were tested on two separate test sessions, with approximately a 25-minute break in between.

3.3 Experimental set-up

The TTIMR prototype training and testing was conducted using two dual screen side-by-side monitor stations. Each station used a Dell personal computer with two 17-inch color monitors with a screen area of 1024x768 pixels and 16-bit high color resolution. The dual monitors for each system were supported by a MATROX[®] graphics card. Subjects were given the static brief four at time, then split into two pairs to do both the training and testing. During testing, all user actions, mouse movements, and messaging in the interface on both screens were recorded by software into a text file. In addition, using a scan converter and a VCR, the left hand screen (the retargeting display) was recorded for later data analysis.

¹ This research protocol was approved by the UVA Social Sciences Institutional Review Board.

3.4 Experimental Design

The experimental design was a 3 X 2 X 2 repeated measures mixed factorial ANOVA model. The independent variables were missiles (3 levels: 8, 12, 16), operational tempo (2 levels: low, high), and order effect (2 levels: low first, high first). The missile factor represents increasing retargetable missiles and was included to determine if and when subjects would become overloaded due to an increasing number of objects to cognitively process. The operational tempo factor represents a difference in arrival rate of problematic situations in the program. In the low tempo sessions, emergent targets and problems were spaced approximately four minutes apart and in the high tempo sessions, the arrivals were spaced only two minutes apart. Lastly, the order effect factor was included due to the fact that training was limited, and it was possible that learning would take place between the two testing sessions. By including the order effect, any significant learning would be detected in the results.

The missile and order effect factors were between-subjects, but within subjects for operational tempo. All subjects were tested on both low and high operational tempo sessions but in each, only saw one level of missiles. Subjects were assigned to both the missile and strike factors randomly, as well as whether or not they say the low or high operational tempo session first.

The primary dependent variable was a predetermined weighted average for overall performance, called the Figure of Merit (FOM). This number represented the overall performance of the subject taking into account scenario complexity, time of response, and accuracy in retargeting answers which included satisficing responses. The FOM took into account the response times and accuracy of answers for the four problems that were presented in each session. A single covariate was used which was the age of the subject. This covariate was used because the subjects were clearly broken into two groups, those that were retired Navy personnel (average age = 47), and those subjects who were active duty (average age = 32).

3.5 Results

Given the variations of number of missiles, rate of problem arrival (tempo), and order effect of scenarios, the tempo independent variable was the only significant factor influencing Figure of Merit scores ($p=.012$) (table 1). In addition, the age covariate was significant ($p = .008$). There were two significant higher order interactions: Tempo*Order, $p < .001$ ($F = 18.457$, $DOF (1,35)$, $power = .987$), and Tempo*Age, $p = .030$ ($F = 5.126$, $DOF (1,35)$, $power = .596$).

Table 1: ANOVA Results for Overall Performance

Factor	F	DOF	p
Tempo	6.943	1, 35	.012
Order	.076	1, 35	.784
Missiles	2.802	2, 35	.074
Age	8.023	1, 35	.008

3.6 Discussion

The operational tempo factor of low and high is essentially a measure of workload, since operators were required to handle problematic situations in a compressed timeframe. The significant results for the FOM dependent variable were expected since the arrival rates of the problems were simulated prior to the human experiment through a discrete event simulation. Indeed, simulation done in advance was used to determine the arrival rates needed to produce an effect. Perhaps the most interesting results from this portion of the data analysis were that both the number of missiles and the order in which the sessions were presented were not significant. From an experimental viewpoint, the lack of an order effect confirms that no significant learning took place between sessions which helps to reduce possible confounds. While additional experimentation could be used to further these results, it appears as if the number of retargetable missiles a subject is required to control, up to 16, is not significant. This means that the matrix design is an effective one in that the level of cognitive effort needed to sort through the presented data is minimized. This is a significant finding for the Navy whose operational requirement document currently dictates that each controller should only be required to control a maximum of four retargetable missiles.

The higher order interactions for the FOM variable were examined graphically to determine the possible meaning behind the interactions. For the Tempo*Order interaction, if subjects saw the low tempo session first, they did much better on the high tempo testing session. However, this difference was much smaller than if the high tempo session was seen first. This makes intuitive sense because by testing on the low tempo session first, the frustration level was lower and more time was available for decisions. This result indicates that some learning was taking place between sessions, but more for those subjects who experienced the low tempo session first.

While other covariates such as video game experience and time in the Navy were explored, only the age covariate produced significant results. For the Tempo*Age interaction, older people did worse on the high tempo session but performed similarly on the low tempo session. Both deterioration in reaction times and a general unfamiliarity with computers could be factors in the significance of both the age effect and the tempo*age interaction.

In addition to the planned experimental design, two additional factors were explored to assess their relationship to the subjects' overall performance, which were chat box fixation and parallel decision processing. One of the decision strategies taught to all the subjects during the training was that when dual targets emerged, the decision should not be made serially, i.e. the decision should be made in parallel due to possible competition among resources in an attempt to obtain the overall optimum solution. This strategy required that the WOOs be manipulated according to the pre-established rules. However, despite the fact that this decision strategy was taught and emphasized to all subjects, only about 25% actually solved dual emergent target problems with the parallel process. The remaining 75% reverted to a serial decision process.

Chat box fixation was another pattern that emerged throughout the course of testing. All subjects were told repeatedly that emergent target situations were their primary priority and that answering queries through the chat box was the least important of all tasks. Despite the heavy emphasis on attending to the chat box when nothing else

was happening, many subjects would become fixated on the communications and answer all queries before attending to the more pressing emergent target problems. From the observer's standpoint, this over-attention on the chat box degraded the overall performance of the subjects, and from an operational standpoint, could have costly consequences.

To statistically confirm any relationship between both parallel decision processes and chat box fixation on overall performance, the data for both were checked with Pearson's Correlation against the Figure of Merit dependent variable. While there was a marginally significant weak correlation for the parallel decision process (.209, $p=.057$), there was a significant moderate correlation for chat box fixation (-.323, $p=.003$). This means that those people who fixated on the chat box had lower overall performance scores.

While correlations can give some insight into a relationship between variables, they do not establish cause and effect. Given that the correlations showed that it was possible that the parallel decision process and chat box fixation factors could be significant predictors, a linear mixed repeated measures model analysis was completed using the level of missiles, operational tempo as independent variables as well as the dichotomous variables of parallel processing and chat box fixation. The results for the parallel decision process were marginally significant ($p=.070$, $DOF(1,70)$, $F=3.384$.) but even more interesting were the results for chat box fixation were significant ($p=.007$, $DOF(1,70)$, $F=7.709$). This analysis can be considered exploratory only since this was a post hoc analysis and not originally part of the intended experimental design but it does suggest that much more research needs to be completed in both areas, but especially the use (or overuse) of the chat box.

4. Conclusion

The introduction of the Tactical Tomahawk into the military's already complex command and control structure promises great flexibility and lethality on the battlefield. However, the added capability of in-flight retargeting requires significant human cognitive contribution and the need to include the human-machine interface considerations in the early design stages is absolutely critical. Through extensive domain modeling through the cognitive work analysis, a rapid prototype for the Tactical Tomahawk in-flight retargeting mission was built and tested, providing critical feedback for the direction of a future operational human computer interface. Specifically, the TTMR decision matrix design is effective in reducing cognitive workload so that the number of retargetable missiles one person should be expected to effectively control is greater than the Navy's initial estimates of 4 missiles per controller. However, regardless of the number of missiles a person is responsible for, the operational tempo significantly impacts the performance of controllers. In addition, the chat box communication tool, while effective in many aspects of command and control, can be a problem for people required to both monitor the communications network and retarget missiles.

These results highlight the need for more extensive research in other human performance and command and control areas. This research should be extended to include multiple controller interactions, more extensive research on the dynamics of chat

box interaction, and the problems with serial versus parallel decision making in this complex domain.

5. Acknowledgments

This material is based upon work supported by the Naval Surface Warfare Center, Dahlgren Division under the guidance of Alan Thomas, through a grant from the Office of Naval Research Knowledge and Superiority Future Naval Capability Program. Any opinions, findings, and conclusions or recommendations expressed in this material are those of the authors and do not necessarily reflect the views of NSWC/DD or ONR.

References

- Miller, C. A. and K. Vicente (2001). "Comparison of Display Requirements Generated Via Hierarchical Task and Abstraction-Decomposition Space Analysis Techniques." International Journal of Cognitive Ergonomics **5**(3): 335-355.
- Rasmussen, J., A. M. Pejtersen, et al. (1994). Cognitive Systems Engineering. New York, J. Wiley.
- Shingledecker, C. A. (1987). In-flight workload assessment using embedded secondary radio communication tasks. The Practical Assessment of Pilot Workload. Neuilly Sur Seine, France, Advisory group for Aerospace research and Development: 11-14.
- Tabachnick, B. G. and L. S. Fidell (2001). Using Multivariate Statistics. New York, HarperCollins.
- Tsang, P. and G. F. Wilson (1997). Mental Workload. Handbook of Human Factors and Ergonomics. G. Salvendy. New York, John Wiley & Sons, Inc.: 417-489.
- Vicente, K. (1999). Cognitive Work Analysis. Toward safe, productive, and healthy computer-based work. Mahwah, NJ, Erlbaum.
- Wickens, C. D. and J. G. Hollands (2000). Engineering Psychology and Human Performance. Upper Saddle River, NJ, Prentice-Hall Inc.