

2007

## Optimizing The Design Of Multimodal User Interfaces

Leah Reeves

University of Central Florida



Part of the [Industrial Engineering Commons](#)

Find similar works at: <https://stars.library.ucf.edu/etd>

University of Central Florida Libraries <http://library.ucf.edu>

This Doctoral Dissertation (Open Access) is brought to you for free and open access by STARS. It has been accepted for inclusion in Electronic Theses and Dissertations, 2004-2019 by an authorized administrator of STARS. For more information, please contact [STARS@ucf.edu](mailto:STARS@ucf.edu).

---

### STARS Citation

Reeves, Leah, "Optimizing The Design Of Multimodal User Interfaces" (2007). *Electronic Theses and Dissertations, 2004-2019*. 3314.

<https://stars.library.ucf.edu/etd/3314>



# OPTIMIZING THE DESIGN OF MULTIMODAL USER INTERFACES

by

LEAH MECHELLE REEVES  
B.S. University of Central Florida, 1997  
M.S. University of Central Florida, 1999

A dissertation submitted in partial fulfillment of the requirements  
for the degree of Doctor of Philosophy  
in the Department of Industrial Engineering and Management Systems  
in the College of Engineering and Computer Science  
at the University of Central Florida  
Orlando, Florida

Spring Term  
2007

Major Professor: Kay M. Stanney

© 2007 Leah Reeves

## **ABSTRACT**

Due to a current lack of principle-driven multimodal user interface design guidelines, designers may encounter difficulties when choosing the most appropriate display modality for given users or specific tasks (e.g., verbal versus spatial tasks). The development of multimodal display guidelines from both a user and task domain perspective is thus critical to the achievement of successful human-system interaction. Specifically, there is a need to determine how to design task information presentation (e.g., via which modalities) to capitalize on an individual operator's information processing capabilities and the inherent efficiencies associated with redundant sensory information, thereby alleviating information overload. The present effort addresses this issue by proposing a theoretical framework (Architecture for Multi-Modal Optimization, AMMO) from which multimodal display design guidelines and adaptive automation strategies may be derived. The foundation of the proposed framework is based on extending, at a functional working memory (WM) level, existing information processing theories and models with the latest findings in cognitive psychology, neuroscience, and other allied sciences. The utility of AMMO lies in its ability to provide designers with strategies for directing system design, as well as dynamic adaptation strategies (i.e., multimodal mitigation strategies) in support of real-time operations. In an effort to validate specific components of AMMO, a subset of AMMO-derived multimodal design guidelines was evaluated with a simulated weapons control system multitasking environment. The results of this study demonstrated significant performance improvements in user response time and accuracy when multimodal display cues were used (i.e., auditory and tactile, individually and in combination) to augment the visual display of information, thereby distributing human information processing

resources across multiple sensory and WM resources. These results provide initial empirical support for validation of the overall AMMO model and a sub-set of the principle-driven multimodal design guidelines derived from it. The empirically-validated multimodal design guidelines may be applicable to a wide range of information-intensive computer-based multitasking environments.

*Don't just settle for moments of potential;  
seize your opportunities, no matter how seemingly small,  
and create moments of greatness.*

*--Anonymous--*

## **ACKNOWLEDGMENTS**

Over the years there have been many creative and interesting individuals who have influenced both my personal and professional goals and inspired me to pursue and finish my Ph.D. Without a doubt, Dr. Kay Stanney, has been a key mentor in all regards and someone who I am fortunate enough to call a dear friend. I have always been fond of this quote by Fyodor Dostoyevsky: "It is not the brains that matter most, but that which guides them--the character, the heart, generous qualities, progressive ideas." Kay is someone who most represents such qualities to me and who I not only value for her great technical mind and career advice but more so who I cherish for her heart and strength of character that are firmly embedded within every decision she makes or piece of advice she gives. She is an exemplary role model in all respects, and I am truly grateful for the major role she has played in my personal and professional life.

I'm now a firm believer that in addition to self-motivation and perseverance, it definitely "takes a village to raise" a quality dissertation product. In this regard there are numerous others I would like to acknowledge for their guidance and support (and patience!) throughout my graduate career, particularly my UCF committee members: Dr. Linda Malone, Dr. Mansoor Mollaghasemi, Dr. Dennis McBride, and Dr. Denise Nicholson. I would also like to thank the various DARPA, ONR, and UCF personnel and sub-contractors who I have had the privilege to collaborate with in various capacities these past 5 years or so, including: CDR Dylan Schmorrow, LCDR Jim Patrey, Laura Worcester, Dr. Robert S. Kenney, Dr. Kelly Hale, Dr. Pete McAlindon, Dr. Taha Sidani, Daniel Compton, Ali Ahmad, Dr. Clint Bowers, Dr. Glenn Martin, Dr. Brian Goldiez, Dr. Jose Sepulveda, and Joy Tatlonghari. I could not have completed my final experiments and data analysis had it not been for the testbed R&D support received from

the Lockheed Martin Advanced Technology Laboratories team of Dr. Polly Tremoulet, Dr. Susan Harkness-Regli, Dr. Helen Hastie, Michael Thomas, Dan Miksch, and Nick Morizio, the additional software engineering support received from Edward Guldemon, or the data collection support from Sean Stanney and Nick Davies.

Finally, to exploit a trite but true and most appropriate expression, I would like to “last but certainly not least” thank my mom, dad, brother, Stephanie, and dearest friends for their unconditional love and support and their ever-present encouragement that abounded when each new obstacle seemed to land squarely in the middle of my path to degree completion. “Serenity now!” is exactly what they have always provided me--exactly when I needed it most. I love and appreciate each of them dearly.

## TABLE OF CONTENTS

LIST OF FIGURES .....	xii
LIST OF TABLES .....	xiii
CHAPTER ONE: GENERAL INTRODUCTION .....	1
CHAPTER TWO: GUIDELINES FOR MULTIMODAL USER INTERFACE DESIGN .....	5
Introduction.....	5
Requirements Specification .....	5
Design for broadest range of users and contexts of use.....	6
Address privacy/security issues .....	6
Designing Multimodal Input and Output.....	6
Maximize human cognitive/physical abilities .....	7
Coherently integrate modalities, accounting for user preferences, context, and system functionality .....	7
Adaptivity .....	8
Consistency .....	9
Feedback .....	9
Error Prevention/Handling.....	10
Conclusion .....	10
Acknowledgment .....	11
CHAPTER THREE: DEVELOPING AN ARCHITECTURE FOR MULTIMODAL OPTIMIZATION (AMMO) .....	12
Introduction.....	13

Theoretical Rationale Supporting the Development of AMMO.....	15
AMMO: Information Format-to-Modality Mappings (A).....	19
Behavioral Evidence for AMMO, Component A.....	23
Neural Evidence for AMMO, Component A.....	26
The Impact of AMMO, Component A, on HSI Design.....	27
AMMO: User Attributes Most Pertinent to HIP Capabilities (B) .....	30
Research Evidence Supporting AMMO, Component B .....	30
Impact of AMMO, Component B, on HSI Design .....	32
AMMO: Ideal Mappings (C) .....	34
Information Flow and Interruption Management .....	35
Impact of AMMO, Component C, on HSI Design .....	44
AMMO: Non-Ideal Mappings (D).....	50
Future directions .....	51
Conclusion .....	53
Acknowledgment .....	54
 CHAPTER FOUR: EMPIRICALLY VALIDATING MULTIMODAL MITIGATION	
STRATEGIES DERIVED WITH AMMO.....	55
Introduction.....	56
Method .....	61
Participants.....	61
Equipment .....	62
Tasks .....	63
Baseline TTWCS task environment (unmitigated/not augmented) .....	63

Verbal-Visual information augmented with Verbal-Auditory (tonal) cue .....	67
Verbal-Visual information augmented with Verbal-Tactile cue .....	67
Verbal-Visual information augmented with both Verbal-Auditory (tonal) and Verbal-Tactile cues .....	68
Verbal-Visual information augmented with Verbal-Auditory (redundant speech) cue....	68
Verbal-Visual information augmented with both Verbal-Auditory (redundant speech) cue and VT cue.....	69
Spatial-Visual information augmented with Spatial-Auditory cue.....	69
Spatial-Visual information augmented with Spatial-Tactile cue .....	70
Spatial-Visual information augmented with both Spatial-Auditory cue and Spatial-Tactile cue .....	70
Experimental Design.....	71
Procedure .....	72
Results.....	74
Objective Performance.....	74
Subjective Performance (Workload Ratings) .....	78
Discussion .....	78
Augmenting VV interruption task information with a VA (tonal) cue.....	81
Augmenting VV interruption task information with a redundant VA speech cue .....	81
Augmenting VV interruption task information with a VT cue.....	82
Augmenting VV interruption task information with VA (tonal or redundant speech) and VT cues .....	83
Augmenting SV interruption task information with a SA cue.....	84

Augmenting SV information with a spatial-tactile (ST) cue .....	85
Augmenting SV interruption task information with SA and ST cues .....	85
Conclusions.....	86
Acknowledgment .....	88
CHAPTER FIVE: GENERAL DISCUSSION .....	89
Effects of Individual WM Capabilities .....	96
CHAPTER SIX: CONCLUSION AND FUTURE RESEARCH.....	99
LIST OF REFERENCES .....	101

## LIST OF FIGURES

Figure 1 Proposed Architecture for Multi-Modal Optimization (AMMO) .....	18
Figure 2 Extended S-C-R model.....	20
Figure 3 Conceptual model integrating AMMO's Component C with the interruption management framework presented in Table 4 (adapted from Latorella 1996, 1999).....	41
Figure 4 Illustration of TTWCS visual interface components.....	65
Figure 5 Illustration of implementation strategies for SA and ST cues.....	71

## LIST OF TABLES

Table 1	AMMO’s Information-to-Modality Suitability Mappings for Conveying Various Types of Information .....	22
Table 2	Potential Suitability of Sensory Modalities for Conveying Specific Types of C <sup>2</sup> Task Information .....	28
Table 3	Information flow moderators due to varying WM processing abilities.....	32
Table 4	Information Flow and Interruption Management Strategies (adapted from Latorella 1996, 1999; McFarlane & Latorella, 2002) .....	38
Table 5	AMMO Component C: Theorized Adaptive Design Strategies and General Implementation Guidelines.....	46
Table 6	Subset of Theorized Multimodal Design Strategies Derived from AMMO for Improved Interruption Management (adapted from Reeves & Stanney, 2007) .....	58
Table 7	Alert Task Performance for Each Treatment Condition (mitigation strategy) .....	74
Table 8	Retarget Task Performance for Each Treatment Condition (mitigation strategy).....	74
Table 9	Performance Pairwise Comparisons between Baseline and Each Mitigation Strategy ..	77
Table 10	CH Subjective Workload Rating Correlations—Perceived Mental Demand for Each Mitigation Strategy as Compared to Baseline .....	78

## **CHAPTER ONE: GENERAL INTRODUCTION**

Explosive improvements in the speed and robustness of computing technology in the 1990s and early 2000s has created an information revolution that has exponentially increased the amount and type of information available to any individual with access to a computer or the internet. The National Science Foundation (NSF) and the National Research Council (NRC) have affirmed society's increasing dependence on vast amounts of multi-dimensional data and the range of systems from which these data are conveyed (e.g., computers, personal digital assistants [PDAs], cell phones, interactive television) (Durlach & Mavor, 1995; NRC, 1997). The NRC suggest that this increasing dependence on data and technology products to complete daily tasks (e.g., personal, occupational, educational, training) may lead to information overload and, in the case of poor system design, to user misunderstanding and frustration. Consequently, there is a definite need for information assimilation and management techniques to support users in effectively managing massive data (i.e., processing an optimal amount of data in a timely manner).

Multimodal displays may provide a means of supporting information assimilation and management. With multimodal displays, rather than inundating users with mostly visual data, they could be provided with a wide variety of sensory cues, thereby leveraging more of their information processing capacity. "Well-designed multimodal systems integrate complementary modalities to yield a highly synergistic blend in which the strengths of each mode are capitalized upon and used to overcome weaknesses in the other(s)" [Oviatt, 1999, p. 74]. Recent evidence from both brain and behavioral studies has shown that improved performance (e.g., reaction time, dynamic decision making capabilities) via presentation of redundant information in an

alternate modality (i.e., augmentation with visual, auditory, and tactile modalities) may be the result of parallel processing occurring across unimodal channels, enhanced human information processing (HIP) at the sensory processing and WM stages, and enhanced sensory facilitation effects when modal stimuli are combined (Stein & Meredith, 1993; Miyake & Shah; 1999; Wickens, 2002; Ho, Spence & Tan, 2005; Calvert, Spence & Stein, 2004; Schmorrow, 2005; Schmorrow, Stanney & Reeves, 2006; Reeves & Stanney, 2007 [see Chapter Three]). Unfortunately, due to a current lack of principle-driven multimodal design guidelines, designers may encounter difficulties when choosing the most appropriate modal display and interaction techniques for given users or specific tasks (e.g., verbal versus spatial tasks). The development of multimodal display guidelines from both a user and task domain perspective is thus critical to the achievement of successful Human Systems Integration (HSI). Specifically, there is a need to determine how to *design* task information presentation (e.g., via which modalities) to capitalize on an individual operator's information processing capabilities and the inherent efficiencies associated with redundant sensory information, thereby alleviating information overload.

The objective of this effort is to address this concern by providing HSI designers with practical guidance for optimizing a multimodal interface design's effectiveness in terms of cognitive workload and subsequent human performance effects. The current work involved three studies. The first study, *Guidelines for Multimodal User Interface Design* by Reeves et al. (2004), was the outcome of a CHI'03 workshop held in Ft. Lauderdale, Florida on multimodal interaction and interface design principles. This article discusses six main categories of design guidelines (i.e., Requirements Specification, Designing Multimodal Input and Output, Adaptivity, Consistency, Feedback, Error Prevention/Handling) and represents a preliminary effort in establishing principles for multimodal interaction design. It was concluded in this study

that to develop both innovative and optimal future multimodal interfaces, additional empirical studies are needed to derive principles that specify the most intuitive and effective combinations of modalities for different users, applications, and usage contexts, as well as how and when to best integrate those modalities. The second and third studies were aimed at filling this gap, at least in part, by establishing a functional framework that could be used to provide designers with guidelines regarding how to adapt information display modalities to meet varying user and task demands.

The second study, *Developing an Architecture for Multi-Modal Optimization (AMMO)* (Reeves & Stanney, 2007) extends the preliminary research of the first study and proposes a theoretical framework (Architecture for Multi-Modal Optimization, AMMO) from which principle-driven multimodal display design guidelines may be derived. The foundation of the proposed framework is based on extending, at a functional working memory (WM) level, existing information processing theories and models with the latest findings in cognitive psychology, neuroscience, and other allied sciences. The utility of such an architecture lies in its ability to provide HSI designers with a priori strategies for directing system design, as well as dynamic adaptation strategies (i.e., multimodal mitigation strategies) in support of real-time operations. Specifically, AMMO aims to support HSI designers in the efficient and effective design of today's information-intensive, multi-tasking systems (e.g., air-traffic control, military command and control watchstations, intelligence analysis, etc.).

The third study, *Empirically validating multimodal mitigation strategies derived with AMMO* (Reeves, Stanney, Ahmad & Malone, 2007) focused on validating specific components of the AMMO model proposed in the preceding study (Reeves & Stanney, 2007). Specifically, a sub-set of guidelines derived from AMMO were evaluated with a simulated weapons control

system multitasking environment. The results of this study demonstrated significant performance improvements when multimodal display augmentation cues were used (i.e., auditory and tactile, individually and in combination) to augment the visual display of information, thereby distributing information processing resources across multiple sensory and WM resources. These results provide empirical support for validation of a sub-set of principle-driven multimodal design guidelines derived from AMMO, which may be applicable to a wide range of information-intensive computer-based task environments. This study further represents an initial step in validating the overall AMMO model.

It is envisioned that once fully empirically validated, the AMMO model will empower HSI designers with principle-driven and practical guidance regarding how to effectively distribute information across display modalities (e.g., auditory or tactile), in addition to (i.e., augment with redundancy) or instead of (i.e., augment via substitution) an overtaxed visual modality, in order to mitigate existing or potential WM overload situations in complex and information-intensive, multitasking environments. The implication is that such design guidance could result in system designs that enable massive volumes of data to be conveyed with greater versatility, mobility, and efficiency. The implication to human performance is that by intelligently and strategically using multiple modalities when designing human-computer task environments to facilitate more parallel as opposed to serial information processing by users, greater performance benefits may be realized (e.g., more efficient task coordination, attention switching, and dynamic decision making). The vision is for such performance benefits to enable a single operator to do a job normally required of two or more operators.

## CHAPTER TWO<sup>1</sup>: GUIDELINES FOR MULTIMODAL USER INTERFACE DESIGN

### Introduction

In today's pursuit of more transparent, flexible and efficient human-computer interaction, a growing interest in multimodal interface design has emerged (Oviatt, 2003). The goals are twofold: first to achieve an interaction that is closer to natural human-human communication, and secondly to increase the robustness of the interaction by using redundant or complementary information. New interaction paradigms and guidelines are necessary to facilitate the design of multimodal systems from the ground up (see McGee, Cohen & Oviatt, 1998; Pieraccini et. al., this issue). This article discusses six main categories of guidelines and represents a preliminary effort in establishing principles for multimodal interaction design. A more detailed discussion of these guidelines will be available in a forthcoming issue of the International Journal of Human Computer Interaction (Reeves, Lai, Larson, Oviatt, Baljai, Buisine, Collings, Cohen, Kraal, Martin, McTear, Raman, Stanney, Su & Wang, 2003).

### *Requirements Specification*

Critical to the design of any application are the user requirements and system capabilities for the given domain. This section provides some general considerations for multimodal system requirements specification.

---

<sup>1</sup> This chapter was originally published as: Reeves, L.M., Lai, J.C., Larson, J.A., Oviatt, S.L., Balaji, T.S., Buisine, S., Collings, P., Cohen, P.R., Kraal, B., Martin, J.C., McTear, M.F., Raman, T.V., Stanney, K.M., Su, H. & Wang, Q.Y. (2004). Guidelines for multimodal user interface design. *Communications of the ACM*, 47(1), pp. 57 59

### Design for broadest range of users and contexts of use

Designers should become familiar with users' psychological characteristics (e.g., cognitive abilities, motivation), level of experience, domain and task characteristics, cultural background, as well as their physical attributes (e.g., age, vision, hearing). An application will be valued and accepted if it can be used by a wide population and in more than one manner. Thus, multimodal designs can aid in extending the range of potential users and uses, such as when redundancy of speech and keypad input enables an application to be used in dark and/or noisy environments. Designers need to account for the best modality or combination of modalities in changing environments (e.g., private office vs. driving a car).

### Address privacy/security issues

Users should be recognized by an interface only according to their explicit preference and not be remembered by default. In situations where users wish to maintain privacy by avoiding speech input or output, multimodal interfaces that use speech should also provide a non-speech mode to prohibit others from overhearing private conversations. Non-speech alternatives should also be provided when users enter personal identification numbers, passwords (e.g., automatic bank teller), or when they might be uncomfortable if certain private information is overheard by others. For example, to reduce the likelihood of others being aware of a user's mistakes, it may be preferable to provide error messages in a visual form instead of audible speech.

### *Designing Multimodal Input and Output*

The cognitive science literature on intersensory perception and intermodal coordination has provided a foundation for determining multimodal design principles (ETSI, 2002; Oviatt, 2003 and Stanney, Reeves, Hale, Samman & Buff, 2003). To optimize human performance in

multimodal systems, such principles can be used to direct the design of information presented to users, specifically regarding how to integrate multiple modalities or how to support multiple user inputs (e.g., voice and gesture). This section provides a brief summary of general guiding principles essential to the design of effective multimodal interaction.

#### Maximize human cognitive/physical abilities

Designers need to determine how to support intuitive, streamlined interactions based on users' human information processing abilities (including attention, working memory, and decision-making) for example:

- Avoid unnecessarily presenting information in two different modalities, where the user has to simultaneously attend to both sources in order to comprehend the material being presented (Cooper, 1997; Kalyuga, Chandler & Sweller, 1999); such redundancy increases cognitive load at the cost of learning the material (Cooper, 1997).
- Maximize advantages of each modality to reduce user's memory load in certain tasks and situations, as illustrated by these modality combinations (Stanney et al., 2003; Wickens, 1992):
  - Visual presentation coupled with manual input for spatial information and parallel processing;
  - Auditory presentation coupled with speech input for state information, serial processing, attention alerting, or issuing commands.

#### Coherently integrate modalities, accounting for user preferences, context, and system functionality

Additional modalities should only be added to the system if they improve performance, satisfaction, and/or efficiency for a given user/context. When using multiple modalities:

- Match output to accepted input (e.g., not allowing visual agents to use spoken natural language if the user cannot);
- Use multimodal cues to improve collaborative speech, such as allowing gaze direction or gesture interactions to indicate turn-taking;
- If using combined modalities (e.g., speech summary combined with visual details) ensure the two presentations are synchronized;
- Ensure the current system interaction state is shared across modalities and that appropriate information is displayed in order to support:
  - users in choosing alternative interaction modalities;
  - multi-device and distributed interaction;
  - system capture of users' interaction history.

### *Adaptivity*

Multimodal interfaces should adapt to the needs and abilities of different users, as well as different contexts of use. Dynamic adaptivity enables the interface to degrade gracefully by leveraging complementary and supplementary modalities according to changes in task and context. Individual differences (e.g., age, preferences, skill, sensory or motor impairment) can be captured in a user profile and used to determine interface settings such as:

- Allowing gestures to augment or replace speech input in noisy environments or for users with speech impairments;
- Overcoming bandwidth constraints (e.g., local direct manipulation replaces gaze input that is analyzed remotely);

- Adapting the quantity and method of information presentation to both the user and the display device.

### *Consistency*

Presentation and prompts should share common features as much as possible and should refer to a common task (e.g., use the same terminology across modalities). Additional guidelines include providing consistent:

- System output independent of varying input modalities (e.g., search by typing or speaking the same keyword provides identical results);
- Interactions of combined modalities across applications (e.g., consistently enable shortcuts);

System-initiated or user-initiated state switching (e.g., mode changing) by ensuring users' interaction choices are seamlessly detected and that the system appropriately provides feedback when it initiates a modality change.

### *Feedback*

Users should be aware of their current connectivity and know which modalities are available to them. They should be made aware of alternative interaction options without being overloaded by lengthy instructions that distract from the task. Specific examples include: use descriptive icons such as microphone and speech bubble to denote click-to-talk buttons; notify users to begin speaking if speech recognition automatically starts. Also, do not confirm interpretations of input from each modality in isolation, but rather from a whole multimodal interpretation after fusion has taken place (McGee et al., 1998).

### *Error Prevention/Handling*

User errors can be minimized and error handling improved by providing clearly marked exits from a task, modality or entire system and by easily allowing users to undo a previous action or command. To further prevent users from guessing at functionality and making mistakes, designers should provide concise and effective help in the form of task-relevant, easily accessible assistance. Some specific examples include:

- Integrate complementary modalities in order to improve overall robustness during multimodal fusion, thereby enabling the strengths of each to overcome weaknesses in others.
- Give users control over modality selection so they can use a less error-prone modality (i.e., most intuitive/predictable) for given lexical content; if an error does occur, permit users to switch to a different modality.

Fuse information from multiple heterogeneous sources of information (i.e., cast a broader information net), incorporating modalities capable of conveying rich semantic information and developing multimodal processing techniques that retain information.

### Conclusion

The guiding principles presented above represent initial strategies to aid in the development of principle-driven multimodal interface guidelines. In order to develop both innovative and optimal future multimodal interfaces, additional empirical studies will be needed to determine the most intuitive and effective combinations of input and output modalities for different users, applications, and usage contexts, as well as how and when to best integrate those modalities. To fully capitalize on the robustness and flexibility of multimodal interfaces, further work also needs to explore new techniques for error handling and adaptive processing, and then to translate

these findings into viable and increasingly specific multimodal interface guidelines for the broader community.

#### Acknowledgment

This article originally was drafted among the co-authors as part of a CHI'03 workshop held in Ft. Lauderdale, Florida on multimodal interaction and interface design principles. The workshop was organized by Jim A. Larson of Intel Corporation (jim@larson-tech.com) and Sharon Oviatt of the Oregon Health & Science University (oviatt@cse.ogi.edu)

## **CHAPTER THREE<sup>2</sup>: DEVELOPING AN ARCHITECTURE FOR MULTIMODAL OPTIMIZATION (AMMO)**

The present study proposes a theoretical framework (Architecture for Multi-Modal Optimization, AMMO) from which multimodal display design guidelines may be derived. The foundation of the proposed framework is based on extending, at a functional working memory (WM) level, existing information processing theories and models with the latest findings in cognitive psychology, neuroscience, and other allied sciences. The model consists of four main components. The first component extends traditional bimodal modality-assigning schema with multiple, multimodal (visual, auditory, haptic) and cross-codal information format (verbal, spatial) mappings. The second component addresses how such mappings may be affected by an individual's WM capabilities and affinity for processing cross-codal information. The third and fourth components discuss interruption management strategies for adapting systems to meet varying human performance needs based on a user's predicted (or assessed in real time) WM resource allocation and cognitive load conditions when mappings are ideal or non-ideal. Taken together, the utility of such an architecture lies in its ability to provide strategies for directing multimodal system design and dynamic adaptation strategies in support of real-time operations. AMMO aims to support interactive system designers in the efficient and effective design of today's information-intensive, multi-tasking systems.

---

<sup>2</sup> This chapter has been submitted to the TIES Journal as: Reeves, L. M. & Stanney, K. M. (2007). Developing architecture for multi-modal optimization (AMMO). Manuscript submitted to *Theoretical Issues in Ergonomics Science*.

## Introduction

*We are analog beings trapped in a digital world, and the worst part is, we did it to ourselves....We are compliant, flexible, tolerant. Yet we people have constructed a world of machines that requires us to be rigid, fixed, intolerant.... We live in a technology-centered world where the technology is not appropriate for people. No wonder we have such difficulties* (Norman, 1998, p. 135).

With a vast increase in the amount and type of information available, a main challenge to today's human-systems interaction (HSI) designers is to create flexible interfaces that allow operators to proficiently process and act upon an optimal amount of task-essential data in a manner compliant with how humans perceive, think, and act in their natural, analog settings. To meet this challenge, multimodal system technology is showing great promise because, as the technology that supports complex information systems advances, the possibility of leveraging multiple human sensory systems becomes possible (Stanney et al., 2003; Stanney, Reeves, Hale, Samman, Buff, Bowers, Goldiez, Nicholson & Lackey, 2004). The potential to use modalities beyond visual presentation and standard mouse/keyboard interactions can provide human-computer interactions that more closely resemble the way humans naturally interact (both verbally and non-verbally) with each other and with objects in the environment (Turk & Robinson, 2000). For instance, such displays could include: speech, spatial audio, variations in frequency or pitch, sound or haptic (e.g., via a tactile vest) cues to localize a point of interest, or haptic impedance (from mouse/joystick) to avert actions. Such alternate display techniques hold promise for creating rich display environments that support adaptive cross-modal (e.g., visual, auditory, haptic) mediation and attention alerting mechanisms, which incorporate alternate display strategies to invoke alternate sensory modalities. HSI designers can take advantage of such technologies when designing for today's information-intensive, multitasking work

environments (e.g., military Command and Control [C<sup>2</sup>], stock trading, air traffic control (ATC), intelligence analysis). In such environments, the completion of time-critical and quick-paced tasks in an accurate and efficient manner depends on operators being able to start and launch multiple tasks and information packages and monitor progress without having to deal with unnecessary information (e.g., manipulate, open, or close hundreds of windows and click on thousands of objects to find pertinent task information, S&T Manning Affordability, 2000). Many such computer-based task environments now involve input and output of massive data in an effort to “push” more information onto users. A critical concern is that under such circumstances *inefficient* system design may hinder, rather than improve task or mission performance as intended, by causing system operators to experience information overload and low situational awareness (SA) (e.g., where one fails to identify or locate information that is vital to the successful completion of tasks) (Endsley, 1995; Wickens & Hollands, 2000; Wickens, 2002;). Thus, there is a need to learn how to push such information *smarter*, which is the impetus for this study—to build a functional-level framework that provides HSI designers with necessary guidance for how and when to display various types of information in the most appropriate modalities. Efforts to provide HSI designers with such guidance do exist. For instance, the US Army recently developed a thorough summary of HSI research relevant to soldier warfare system design (Mulgund, Stokes, Turieo, & Devine, 2002), resulting in a set of *unimodal* design guidelines. The European Telecommunications Standards Institute (ETSI; 2002) represented the first major effort to summarize literature on intersensory perception and cross-modal coordination as a preliminary effort to aid in developing a theoretical foundation for identifying multimodal design principles. The ETSI report provided *unimodal* and *bimodal* guidelines for potential modality-to-task information mappings. Oviatt (2003) provided a review

of user-centered design issues, with a primary focus on *bimodal* guidelines specific to pen and speech interactions and user input. Practical guidance for truly multimodal system design (i.e., beyond bimodal solutions) remains elusive.

The primary objective of this effort is to address this shortcoming by identifying how best to design multimodal task information presentation (e.g., via which modalities) to capitalize on an individual operator's information processing capabilities and the inherent efficiencies associated with redundant sensory information. The results should provide HSI designers with practical guidance for optimizing a multimodal interface design's effectiveness in terms of cognitive workload and subsequent human performance effects. The implication to human performance is that by intelligently and strategically using multiple modalities when designing human-computer task environments to facilitate more parallel as opposed to serial information processing by users, greater performance benefits may be realized (e.g., more efficient task coordination, attention switching, and dynamic decision making).

#### Theoretical Rationale Supporting the Development of AMMO

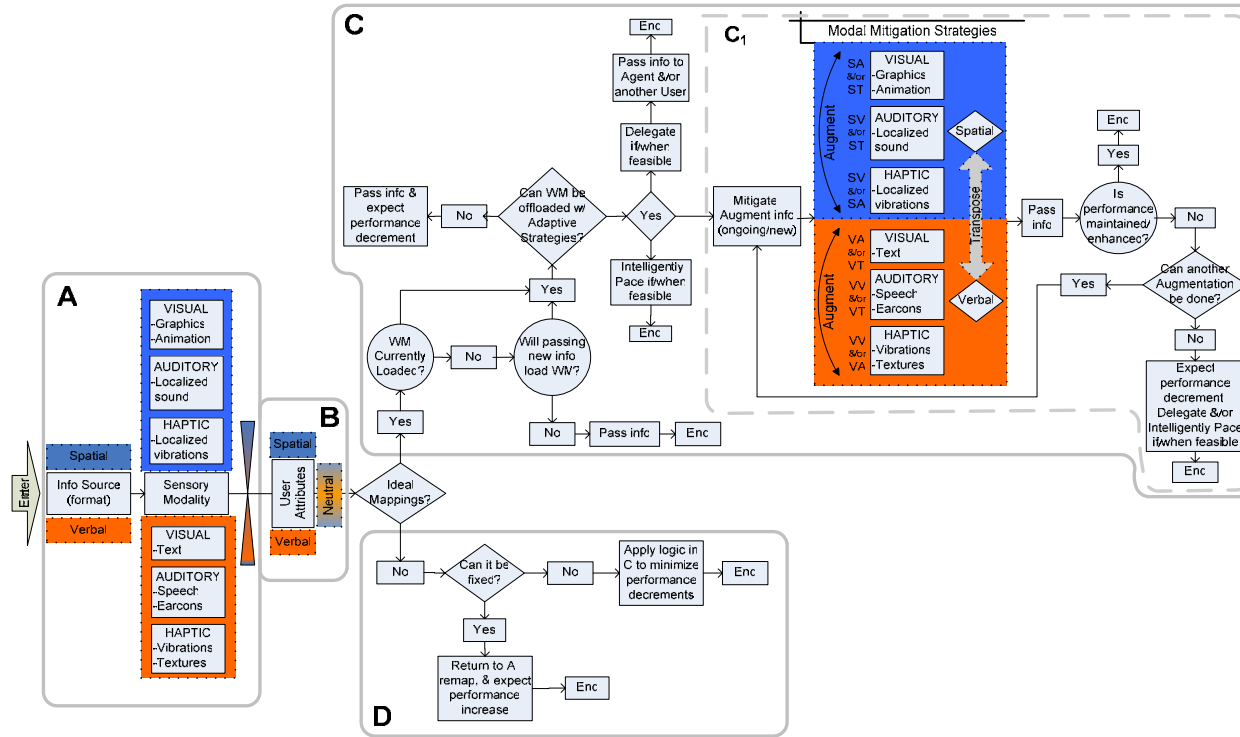
To meet the present study's objective of providing HSI designers with practical guidance for optimizing multimodal interface design, the Architecture for Multi-Modal Optimization (AMMO) model is herein proposed (see Figure 1). While the AMMO model assumes certain characteristics of a cognitive architecture to facilitate discussion in terms of well-known human information processing (HIP) theories, particularly regarding working memory (WM) and cognitive workload, it is not intended to be an all-encompassing cognitive architecture. It is intended as a framework to guide initial and 'real time' system design strategies. For instance, during initial design stages (e.g., brainstorming, prototyping, technology requirements

documentation, etc.), the flow set forth in AMMO's feedback loop (i.e., A to B to C and/or D; return to A) could be used for any or all of the following goals: to facilitate the design of task information presentation by providing guidance regarding the most effective cross-modal formats (e.g., visual, auditory, haptic) for presenting specific types of task information (e.g., verbal versus spatial cross-codal information formats); for determining how best to support task performance (i.e., via adaptive strategies, such as delegation, pacing, augmentation) when cognitive overload conditions and performance decrements are expected (i.e., as established by a task analysis or via empirical data), and; to establish manning requirements after multimodal design and adaptive strategies for a single user have been explored. As an aid to facilitating the design of real-time intelligent system models, AMMO's architecture could be integrated into system models and its flow used to drive when and how real-time adaptive design strategies are implemented to avoid cognitive overload conditions, particularly at the WM level.

While the flow in AMMO is presented serially for illustration and discussion purposes, it may be used to guide the design of multiple serial and/or parallel interactions in dynamic, information-intensive multitasking environments. For example, single or multiple information sources of varying cross-codal formats (i.e., verbal, spatial) and cross-modal formats (i.e., visual, auditory, haptic) may simultaneously enter AMMO's Component A at any given time to assess information format to display modality mappings, pass on to Component B for moderation, and then on to Component C and/or D to be assessed in terms of the most suitable and feasible (individual and/or combination of) adaptive strategies for presenting task information to the user. The structure and function of WM and its available resources are the primary driving factors chosen to guide the overall flow and adaptive logic in AMMO's feedback loop components. While numerous other factors (e.g., user goals, performance, SA, knowledge, personality,

cognitive style, task variables, context) may be used to trigger adaptive strategies (Rothrock, Koubek, Fuchs, Haas, Salvendy, 2002), WM is arguably the “hub of cognition” (Haberlandt, 1997, p. 212) and was thus considered a key factor in building the AMMO framework (Stanney, et al., 2004).

The following sections address the theoretical rationale supporting the development of AMMO. For discussion considerations, AMMO has been segmented into four main components: 1) Component A, where traditional bimodal modality-assigning schema have been further extended with multiple, multimodal (visual, auditory, and haptic) and cross-codal information format (verbal, spatial) mappings; 2) Component B, where the determined effectiveness of such mappings may need to be considered in the context of an individual’s WM capabilities and affinity for efficiently processing verbal and/or spatial information formats; 3) Component C, where strategies for adapting systems to meet varying human performance needs based on a user’s predicted (or assessed in real time) WM resource allocation and cognitive load conditions are presented, specifically for those mappings of information format to display modality that are “ideal,” and; 4) Component D, where system redesigns may need to be considered for mappings that are not ideal.



Notes:

(1) Augmenting refers to presenting interruption cues, redundant information, and/or alternate task information via (a) the same information format (i.e., spatial or verbal) but different modality (i.e., visual, auditory, or haptic), (b) a different format (i.e., transposed) but same modality, or (c) a different format and different modality. (2) Transposing refers to changing the information format from spatial to verbal (and vice versa).

Figure 1 Proposed Architecture for Multi-Modal Optimization (AMMO)

### *AMMO: Information Format-to-Modality Mappings (A)*

The purpose of AMMO's Component A is to provide a framework to aid designers in effectively mapping goal-relevant task information to the appropriate display modality according to which sensory modality might be most suited for displaying a particular type of information (i.e., spatial, verbal). As reviewed in this section, evidence from both behavioral and neural studies indicates the need for such a framework—one which extends existing bimodal HIP-based models traditionally used by human-computer interaction (HCI) designers.

When designing for short duration, simple uni- and bi- modal tasks, designers have typically been able to rely on well-established modality assigning schemas, such as Wickens' (1984; 1992) Multiple Resource Theory (MRT) (see Figure 2, gray box), to streamline cognitive load, minimize interference effects, and subsequently improve performance (e.g., response time). MRT suggests that the limited HIP resources for which tasks compete may be defined in any or all of the following dimensions: sensory input modalities (see "S" in Figure 2), which include visual, auditory, and the emerging haptic senses; in codes of WM, which represent Central processing of verbal and spatial information formats that may be integrated and stored in the episodic buffer and then controlled by the central executive (Baddeley, 2000) (see "C" in Figure 2); and in Response output (user input to a system) modalities, which include speech, manual, and the emerging brain-directed response modalities (see "R" in Figure 2). The MRT model describes parallel or separate independent processing that occurs in the various S-C-R dimensions of the model and suggests that each dimension contains limited and allocatable resources that can be distributed between and within tasks. In Wickens' S-C-R model, verbal information (e.g., tasks with words, language, or logical operations; Sanders & McCormick,

1993) is best thought to be presented auditorally and with speech as the most appropriate response; spatial information (e.g., tasks requiring moving, positioning, or orienting objects in space) is best thought to be presented visually and coupled with a manual response (see Figure 2, gray box). Further, when designing modality-to-task information mappings, MRT suggests that it is best to couple verbal and spatial information rather than loading one WM channel.

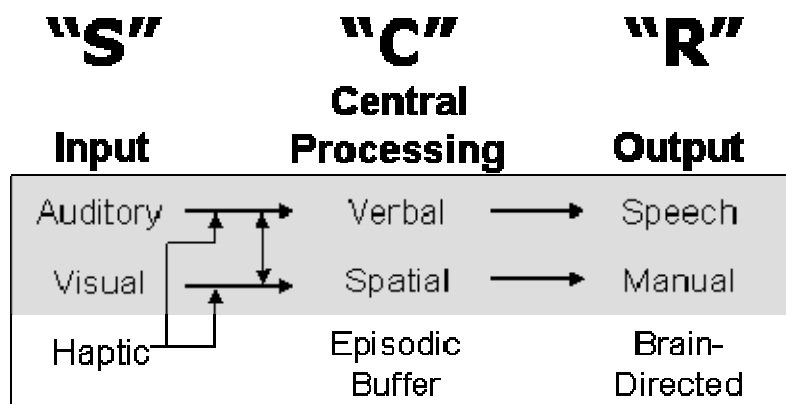


Figure 2 Extended S-C-R model.

Wickens' (1984; 1992) SCR model provides a foundation from which to build a multimodal framework. Specifically, within the traditional MRT framework, a distinction between cross-modal *types* of spatial and verbal information formats can be made, which may prove helpful in providing practical guidance for how best to coordinate and streamline a user's WM resources (Note: The model in Figure 2 also includes the latest components that have been incorporated into the SCR model - the episodic buffer, which represents a limited capacity storage system that holds information in a multimodal code awaiting binding into a unitary episodic representation [Baddeley, 2000], as well as a contemporary response modality, brain-directed responses, which use signals from the brain to direct computer interaction [Kennedy, Bakay, Moore, Adams & Goldwaihthe, 2000]). For example, verbal information can be auditory (e.g., speech, earcons),

visual (e.g., text), and haptic (e.g., textured codes, vibratory semantic patterns). Spatial information can also be auditory (e.g., spatialized sound), visual (e.g., graphics, animation), and haptic (e.g., localized vibration). As an initial step in addressing such design considerations, Stanney et al. (2004) proposed the use of a theorized modality-to-information source mapping framework (see Table 1) originally presented by ETSI (2002). Such a dichotomy is further represented in AMMO (see Figure 1, Component A), where: known modality-to-task information mappings have been extended into their respective verbal and spatial information categories, and; single or multiple stimuli (i.e., information sources) may enter at any given time and be assessed according to how appropriately they are mapped for given information formats and modalities. As discussed next, evidence from behavioral and neural studies further indicates the utility of the multimodal information format-to-modality schema as set forth in AMMO to aid designers in offloading, coordinating, and streamlining a user's WM resources via effective distribution of task information across sensory modalities and information formats.

Table 1 AMMO's Information-to-Modality Suitability Mappings for Conveying Various Types of Information

<i>Information Source</i>	<i>Possible Info Format</i>	<i>Presentation Modality</i>			<i>Practical Design Examples</i>
		<i>Visual</i>	<i>Auditory</i>	<i>Haptic</i>	
Temporal	S , V	<>	++	+	Auditory tonal cues, speech, &/or localized sound preferred; tactile vibrations &/or possibly visual text, graphics
Spatial	S	++	<>	+	Visual graphics or animation preferred; localized vibrations &/or possibly localized sound
2-dimensional Localization	S	++	+	+	Visual graphics or animation preferred; localized sound, &/or localized vibrations
3-dimensional Localization	S	<>	+	+	Localized sound, &/or localized vibrations; possibly visual graphics &/or animation
Alerts/Warnings	S , V	-	++	++	Tactile vibrations, auditory tonal cues, speech, localized sound &/or tactile vibrations preferred
Fast Reaction Time	S , V	++	+	<>	Visual graphics, animation &/or text preferred; auditory tonal cues, speech &/or localized sound; possibly tactile vibrations
Persistence	S , V	++	--	++	Visual graphics, animation, text, &/or tactile vibrations preferred
Memorability	S , V	+	++	+	Auditory cues preferred; visual graphics, animation, text, &/or tactile vibrations
Relative Quantitative	S , V	+	++	+	Auditory tonal cues, speech, &/or localized sound preferred; visual text, graphics &/or tactile vibrations
Absolute Quantitative	V	++	-	--	Visual text preferred; possibly speech
Private/Confidential	S , V	<>	-	++	Tactile vibrations preferred; possibly visual text &/or graphics

<i>Information Source</i>	<i>Possible Info Format</i>	<i>Presentation Modality</i>			<i>Practical Design Examples</i>
Outside Area of Interest (periphery)	S , V	-	++	+	Auditory tonal cues, speech, &/or localized sound preferred; tactile vibrations
Instructions	S , V	<>	+	+	Auditory tonal cues, speech, localized sound, &/or vibrations; possibly visual text &/or graphics
Object Properties	S , V	++	<>	++	Visual graphics, animation, text, &/or tactile vibrations preferred; possibly tonal cues &/or localized sound
Motion	S , V	+	<>	+	Visual graphics, animation, &/or tactile vibrations; possibly relative tonal cues or localized sound
Affective/Emotive	S , V	+	+	<>	Visual text &/or graphics, auditory tonal cues, speech, &/or localized sound; possibly tactile vibrations
Motivational	V	<>	+	<>	Tonal cues or speech; possibly visual text &/or tactile vibrations

Note: Adapted from European Telecommunications Standards Institute, 2002.

Key: ++ = best modality; + = next best; <> = neutral; - = not well suited, but possible; -- = unsuitable.

“S” indicates spatial; “V” indicates Verbal

### Behavioral Evidence for AMMO, Component A

Behavioral studies provide evidence for the importance of AMMO’s dichotomy between verbal and spatial WM processing codes and further differentiation into *types* of information, specifically to increase WM throughput and account for task interference effects when designing multimodal systems (Wickens & Liu 1988; Shah & Miyake, 1996; Chandler & Sweller, 1991, 1992; Wickens & Hollands, 2000; Wickens, 2002 Wickens & Gosney, 2003; Kobus, Brown, C., Morrison, J., Kollmorgen, G., Cornwall, R. & Schmorow, D., 2006). Such studies have reported enhanced task performance effects (e.g., improved response time and task accuracy) when distributing information across modalities in dual and multiple combinations. This may be due to increased information throughput or enhanced information organization. Specifically, while

people can store and maintain only a small amount of single modality information in WM (Miller, 1982; Card, Moran & Newell, 1984), it has been shown that they can store, maintain, and interact with considerably larger amounts of information when multiple information formats and WM modalities are invoked. Sulzen (2001), for example, illustrated that people could recall nearly three times the traditionally known WM capacity limits of  $7 \pm 2$  (Miller, 1956) or five times Cowan's (2001) predictions when presented with stimuli in non-interfering information formats and modalities (e.g., tonal, kinesthetic, tactile) along with the standard visual-verbal (i.e., written letters, words, or objects) and auditory-verbal (i.e., spoken letters, words, or objects) information historically used in WM capacity/recall studies. (Note: Capacity can be defined as "the channel capacity of absolute judgment, the capacity of working memory, or the bandwidth capacity to transmit information along a channel in bits per unit of time" [Wickens, 1992; p.381]). Though, as suggested by Cowan (2001), this likely represents a certain reasonable degree of chunking, perhaps along each modality. Thus, multimodal information presentation may help alleviate the information overload often experienced with current interactive systems, even if this is simply by facilitating a chunking structure. More empirical research is needed to determine whether such performance improvement may be due to truly separate and modally-distributed WM stores or to enhanced chunking abilities within a central WM store, which may be facilitated with modally-organized information (Cowan, 2001).

Regardless to which school of thought one may adhere to, the previously referenced and numerous other studies (Martin, 1980; Wickens, 1984; Baddeley, 1986; Mayer & Anderson, 1991; Chandler & Sweller, 1991, 1992; McKinley & Ericson, 1997; Spence & Driver, 1997, 1999, 2004; Mayer & Moreno, 1998; Bolia, D'Angelo & McKinely, 1999; Giard & Peronnet, 1999; Wickens & Hollands, 2000; Sarter 2000, 2002; Eimer, Cockburn, Smedley & Driver,

2001; Popescu, Burdea & Trefftz., 2002; Wickens & Gosney, 2003; Wickens & Seppelt, 2002; Ho et al., 2005; Ho, Tan & Spence, 2005; Hopp, Smith, Clegg & Heggstad, 2005; Hopp-Levine, Clegg, Smith and Heggstad, 2006; Kobus et al., 2006) provide support to the notion that WM throughput may be increased with effectively designed multimodal information displays. To realize such gains, however, it is essential to ensure that appropriate facilitation and depression of combined modal stimuli occur (i.e., dual-process theory of plasticity, Groves & Thompson, 1970) by avoiding incongruent modality pairings (i.e., when an information format is not appropriately mapped to its ideal presentation modality) (Wickens, 1992; Wickens & Hollands, 2000; Stanney et al., 2004) and potential subsequent task-switching costs (Arrington, Altmann & Carr, 2003; McFarlane & Latorella, 2002) or other interference effects that may occur at a user's cortical processing level (Schumacher, Seymour, Glass, Fencsik, Lauber, Kieras & Meyer 2001; Dyson & Quinlan, 2002). Consequently, when applying AMMO, the following principles should be considered to avoid incongruence and task switching costs:

- WM capacity enhancement: To enhance an operator's WM capacity, direct sensory stimuli to a multitude of sensory modalities, while avoiding extensive cross-encoding among visual and auditory percepts into linguistic terms (Baddeley, 1990, 2000; Barnard, 1999; Schneider, 1999; Stanney et al., 2004; Sulzen, 2001), which may overload HIP resources in the left hemisphere.
- Presentation of *spatial* information: When distributing spatial information among various presentation modalities to enhance WM capacity, facilitate congruency, and minimize task interference effects, it may be most effective to use multiple mapping strategies, to include graphics or animation for the visual modality, localized sounds for the auditory modality, and/or localized vibrations for the haptic modality (Stanney et al., 2004).

- Presentation of *verbal* information: When distributing verbal information among various presentation modalities to enhance WM capacity, facilitate congruency, and minimize task interference effects, it may be most effective to use multiple mapping strategies, to include text for the visual modality, speech, auditory icons, or earcons for the auditory modality, and/or vibrations for the haptic modality (Stanney et al., 2004).

#### Neural Evidence for AMMO, Component A.

Brain-imaging studies demonstrate differential cortical processing areas are involved in various forms of multimodal information processing (Bowers & LaBarba, 1991; Smith & Jonides, 1998; Springer & Deutsch, 1985; Smith & Jonides, 1998; Miyake & Shah, 1999; Thompson-Schill, Aguirre, D'Esposito & Farah, 1999; Cabeza and Nyberg, 2000; Just, Carpenter & Miyake, 2003; Calvert et al., 2004). These studies indicate physically separable and hemispheric WM systems are used for specific types of information (i.e., spatial mostly right hemisphere; verbal mostly left hemisphere), which suggests that AMMO's strategies should put more of the brain on task by fostering multimodal information processing across brain regions.

Further neural evidence in support of the design approach presented in AMMO stems from studies of the "coactivation model," which suggests that redundant sensory information (i.e., coactivation via sufficient overlap in time and space) is basically equivalent to linear neural summation of modal stimuli, where the integration of modal redundancy combinations in the superior colliculus may also result in multiplicative effects (Miller, 1982, 1986; Corballis, Hamm, Barnett, & Corballis, 2002; Roser & Corballis, 2002; Savazzi & Marzi, 2002; Iacoboni & Zaidel, 2003). For instance, Calvert and Lewis (2004) note that some studies have shown firing rates of multisensory neural coactivation at the cellular level to be up to 12x faster beyond that expected by summing impulses from unimodal stimuli, particularly when the unimodal

stimuli have the least effective sensory facilitation when presented alone (Stein & Meredith, 1993). Thus, performance gains realized through the application of AMMO may be attributed at the neural level to enhanced sensory facilitation effects from multisensory presentation, resulting in improved information processing times and subsequent human response times. The implication is that the AMMO framework may assist in determining how information processing resources should be distributed between and coordinated among multisensory information sources to facilitate coactivation and thus is an important tool for multimodal system designers.

#### The Impact of AMMO, Component A, on HSI Design

By applying AMMO Component A, designers obtain an idea of how to appropriately map multimodality sensory inputs (visual, auditory, and haptic) to both verbal and spatial information formats, thereby distributing processing across multiple sensory capacities (i.e., put more of the brain on task). Table 2 illustrates some practical examples as to how candidate modalities may be selected for specific types of  $C^2$  task information according to whether verbal or spatial (or both) HIP may be required and according to the theorized suitability for displaying various types of sensory information sources in Table 1. Considering the latest reported findings of WM capacities summarized in Schmorow, Stanney, Wilson, and Young (2005), Table 2 also considers ranges of WM capacity per modality or per central processing (i.e., for verbal and spatial information). Two schools of thought (i.e., modally separable – see channel capacity ranges under “Presentation Modality” in Table 2 [Wickens & Liu, 1988; Wickens, 1984, 1992, 2000; Sulzen, 2001] vs. a centralized WM storage area – see WM capacity ranges under “Info Format” in Table 2 [Miller, 1956; Cowan 1988; 1995; Engle, Kane, & Tuholski; 1999]) are represented because it may be the case that a central WM storage area is the constraining factor for how many pieces of information in each modality may be presented during information-

intensive, multitasking environments. While more empirical evidence is needed to substantiate either claim, the ranges in Table 2 provide system designers with bounds on the amount of information that can be readily processed by each modal system in working memory.

Table 2 Potential Suitability of Sensory Modalities for Conveying Specific Types of C<sup>2</sup> Task Information

<i>Task</i>	<i>Information Source</i>	<i>Info Format (WM Capacity Range)</i>	<i>Presentation Modality [Channel Capacity Range]</i>			<i>Practical Design Examples</i>
			<i>Visual [2-5]</i>	<i>Auditory [4-6]</i>	<i>Haptic [3-5]</i>	
ID Friendly, Enemy, Unknowns	Temporal, Alerts/Warnings, Fast Reaction Time, Memorability, Relative Qualitative, Object Properties, Motion	V [4-7]	++	+	+	Provide visual text, auditory tonal cues or speech, &/or tactile cues to aid in general identification of objects
Target Designation	Temporal, Alerts/Warnings, Fast Reaction Time, Memorability, Object Properties	V [4-7]	++	+	+	Provide visual text, auditory tonal cues or speech, &/or tactile cues to aid in target identification
Mapping, Navigation	Temporal, Spatial, 2 and 3D Localization, Fast Reaction Time, Persistence, Memorability, Object Properties, Motion, Motivational	S & V [4-7]	++	+	+	Provide combinations of visual graphics and text, localized sound and speech or earcons, &/or tactile vibrations to indicate heading, location, distance, terrain, etc.
Air Traffic Monitoring	Temporal, Spatial, 2 and 3D Localization, Persistence, Object Properties, Motion, Motivational	S [5-7]	++	+	+	Provide visual graphics or animation, localized sound, &/or possibly localized vibrations to aid in localization of self and/or others, judging axes, or perceiving motion

<i>Task</i>	<i>Information Source</i>	<i>Info Format (WM Capacity Range)</i>	<i>Presentation Modality [Channel Capacity Range]</i>			<i>Practical Design Examples</i>
Communications Monitoring	Instructions, Temporal, Persistence, Memorability	V [4-7]	+	++	=	Provide speech &/or visual text if information must be both memorable and persistent (e.g., available for later access/review)

Note: Adapted from Mulgund, et al. (2000).

Key: ++ = best modality; + = next best; <> = neutral; - = not well suited, but possible; -- = unsuitable.

“S” indicates spatial; “V” indicates Verbal. “[ ... ]” indicate WM capacity ranges; passing more than the designated amount of information at any one time is not recommended.

If the AMMO architecture is integrated into intelligent adaptive systems and used to drive real-time adaptive design strategies to avoid cognitive overload conditions, each component of the model will pass context- and task-dependent data on to the next system component. For instance, Component A of AMMO will pass the following data on to Components B and C (and/or D when appropriate):

- The type of information being presented to a user; both the information source and its format (i.e., whether verbal, spatial or both HIP WM resources are required) and the modality being used to present it;
- The amount of information being presented to a user in each modality versus the capacity of each modality (see Table 2), and;
- Determination as to whether ideal (congruent) or non-ideal (incongruent) information format-modality mappings are being used for presenting information to the user.

While the mappings derived from AMMO's Component A are theoretically well supported, their generalizability is likely to be mediated by the individual receiving the information, which will be addressed in the next section.

#### *AMMO: User Attributes Most Pertinent to HIP Capabilities (B)*

Component B of AMMO has been designed to address issues of individuals' varying HIP abilities by extending mapping strategies of Component A into another dimension that considers an individual's WM capabilities and how they may impact the effectiveness of information format-to-modality mapping strategies. Specifically, one should not make the assumption that all operators will be equally effective in interacting with multimodal technologies, as some may be unable to efficiently process various combinations of modalities and information formats simultaneously.

#### Research Evidence Supporting AMMO, Component B

Individuals may benefit to differing degrees from the S-C mappings and multimodal design strategies offered by AMMO Component A. It is thus important to consider individual factors in the AMMO model. While applied psychology studies have identified a plethora of individual attributes (e.g., age, sex, handedness, etc.) affecting various aspects of human performance, the current study focuses on attributes that both map well to the AMMO model and have been shown to be particularly relevant to human-systems interaction—a user's individual capabilities and limitations in spatial and verbal WM processing (Endsley & Bolstad, 1994; Bowers & LaBarba, 1991; Miyake & Shah, 1999; Ackerman, Beier & Boyle, 2002; Gonzalez, 2005; Stanney et al., 2004; Hale, Axelsson, Fuchs, Baskin & Stanney, 2005; Hale, Reeves, Samman, Axelsson, Milham & Stanney, 2006; Doan, 2002; Lathan & Tracy, 2002; Reeves, Ahmad & Stanney,

2005). Such differences in WM capabilities have been shown to be relatively enduring traits of an individual (Mayer & Moreno, 1998; Tindal-Ford, Chandler, & Sweller, 1997). These differences tend to affect an individual's efficiency in WM resource utilization, as well as their ability to maintain focused attention on pertinent task information while ignoring irrelevant, distracting information (Miyake & Shah, 1999; Givens & Smith, 2000; Kane & Engle, 2002). Specifically, high ability individuals tend to process more information, respond faster, and are better able to focus and sustain attention on performance relevant information. Thus, once a modality-to-task information mapping has been decided in Component A, for low ability individuals it may be necessary to present less information, slow the schedule and/or pace of information flow, and present attentional cues to direct attention (see Table 3).

Further differentiation is found with regard to the type of high ability, with high verbal individuals using more of the left hemisphere and high spatial individuals using more of the right hemisphere (Bowers & LaBarba, 1991; Miyake and Shah, 1999). Thus, an individual's capabilities and limitations for processing spatial and/or verbal information may result in them using alternative, potentially inefficient processing strategies to compensate for structural or neurological inefficiencies in hemispheric processing, where low spatial individuals would engage verbal information processing resources (left hemisphere) even when spatial processing strategies would be most appropriate and high spatial individuals may engage spatial information processing resources (right hemisphere) even when verbal processing strategies may be most appropriate. To alleviate these inefficient processing strategies, it may be necessary to augment information presentation with a redundant modality that is a more efficient format for the given individual (e.g., for a low spatial individual, augment a visual-spatial map with visual-verbal directions; for a low verbal individual, augment visual-verbal descriptions with a visual-spatial

graphic), particularly during high workload conditions (see Table 3). Such cross-codal redundancy (i.e., presenting same information via multiple information formats [i.e., verbal and spatial]; Wickens & Gosney, 2003) has been shown to lead to performance gains (Wickens and Seppelt, 2002).

The considerations in Component B of the AMMO framework may assist in determining how to appropriately tailor the presentation of task information to meet an individual user's information processing needs during information-intensive, multitasking conditions.

Table 3 Information flow moderators due to varying WM processing abilities

<i>Ability</i>	<i>Moderators</i>
Low ability	Present less information, slow the schedule and/or pace of information flow, and present attentional cues to direct attention
Low verbal	Augment with redundant cross-codal (i.e., spatial) information that is then mapped to an ideal (congruent) presentation modality
Low spatial	Augment with redundant cross-codal (i.e, verbal) information that is then mapped to an ideal (congruent) presentation modality

#### Impact of AMMO, Component B, on HSI Design

Understanding how effective an individual may be at processing particular types and amounts of information at the WM-level is an important consideration when determining the most appropriate S-C mapping schema and information quantities for a given user in a given task domain. AMMO's Component B has thus been structured to parallel the spatial/verbal WM dichotomy in Component A to account for potential user attributes at the WM-level that may

impact the effectiveness of S-C mapping strategies chosen for particular users. This AMMO module should thus serve as a moderator on the *amount* of information to be passed and how this information is to be passed (e.g., slow the schedule/pace, add an attentional cue, augment with redundant information).

Conceptualizing AMMO as a data processing loop that supports real-time adaptive human-system interaction, Component B of AMMO will receive inputs from A and then pass the following data on to Component C if mappings are ideal and/or on to D for mappings that are not ideal:

- Recommendations on the amount of information to pass, specifically, information load should be maintained at the lower bound of the WM capacity ranges presented in Table 2 if an individual has low ability to process information in the modality being passed.
- Recommendations on the rate of information conveyance, specifically, for low ability individuals, if less information cannot be passed or if performance decrements are found, then slow the schedule or rate of information presentation.
- Recommendations on the use of attention cues to assist low ability individuals with focusing and sustaining attention (details on modal attentional cuing can be found in Tables 4 and 5).
- Recommendations on augmenting with redundant cross-modal information in a format conducive to the abilities of an individual, specifically, augment with verbal information for low spatial individuals and spatial information for low verbal individuals.

The remaining AMMO discussion sections are focused on WM load and interruption management issues, congruency effects, and mapping strategies when mappings are either ideal (with regard to the guidelines set forth in Component A of AMMO) or non-ideal.

### *AMMO: Ideal Mappings (C)*

The purpose of AMMO's Component C is to provide an information management framework to aid designers in determining the most effective adaptation strategies when information format-to-modality mappings are ideal. Specifically, if mappings are identified as ideal in AMMO's Component A, then information to be presented to users will follow through AMMO's Component B, with possible moderation on the amount, pace, and format of information being passed, and then on to AMMO's Component C. Component C addresses issues regarding how to handle the information intended for the user, while minimizing potential deleterious interruption effects. Specifically, Component C's flow suggests when and what to do when WM is either already overloaded or will soon be overload if more of the same information (i.e., format and modality) continues to be passed to the user. In such cases, one or more of the following adaptive strategies can be invoked (Schmorrow et al., 2005): 1) WM load could be reduced by augmenting both existing (ongoing) information and/or new incoming cross-modal information, particularly for information critical to ongoing task performance; 2) intelligent pacing strategies could be used to decrease the presentation rate of critical information and/or hold non-critical information in queue and schedule for later presentation when the user is less overloaded and better apt to attend to it and/or; 3) information could be delegated to another user or system agent to immediately relieve a currently overloaded user (see Figure 1). From a general perspective, the design of any of these adaptive strategies in AMMO's Component C (i.e., multimodal augmentation, intelligent pacing, delegation) may be considered as task information flow and interruption management design problems (Latorella 1996, 1999; McFarlane, 2002; McFarlane & Latorella, 2002; Ho, Nikolic, Waters & Sarter, 2004; Speier,

Vessey, & Valacich, 2003; Hopp et al., 2005; Hopp-Levine et al., 2006). It is therefore essential that such issues be factored into the AMMO approach.

### Information Flow and Interruption Management

When an individual is processing data and AMMO is invoked to enhance information processing, it is important to understand the general characteristics of the information flow that transpires. For example, consider an individual operator performing a primary task (e.g., monitoring air space) who, at the same time, is supporting secondary tasks (e.g., monitoring vehicle health status or various communication channels). As the operator is engaged in these ongoing tasks (i.e., the ongoing procedure), an interruption may occur at any point as an incoming disjointed activity (e.g., new planes just entered the airspace; urgent incoming communications must be transmitted) (Speier et al., 2003). The operator would thus have to contend with the arrival of an annunciation (i.e., interrupting) stimulus that indicates the presence of the interruption (Latorella, 1999). Once detected, the operator must choose when to attend to interrupting information. Evidence to date suggests that such interruption is generally associated with a cost to human performance (e.g., decision making and response errors, task switching costs, loss of situational awareness, and increased task completion times) (Cohen, 1980; Latorella, 1996, 1999; McFarlane, 2002; McFarlane & Latorella, 2002; Speier et al., 2003; Dorneich, Whitlow, Ververs, Mathan, Raj, Muth, Hoover, DuRousseau, Parra & Sajda, 2004). Thus, understanding how an interruption affects a user's HIP resources and workload, as described in Table 4, will be of critical importance in determining how to design and apply AMMO's interruption management strategies to ensure that: a user is only minimally distracted from ongoing task performance, mismatches do not occur between the user's mental model of the system and the actual system state, and any task information conveyed via adaptive strategies

is easily attended to, interpreted by, and acted upon by the user (Latorella, 1996, 1998, 1999; McFarlane & Latorella, 2002; Ho et al., 2004; Hopp et al., 2005; Hopp-Levine et al., 2006).

To achieve these objectives, first AMMO's information format-to-modality suitability mappings (see Table 1) could be applied to the design of the overall task environment, which should minimize user interpretation times and task switching costs (e.g., by avoiding inefficient information presentation design strategies that cause a user to devote unnecessary time and WM resources to interpret non-ideally presented information) (Altmann, 2004; Arrington et al., 2003). Second, the information flow and interruption management strategies in Table 4 can be used to further mitigate costs (deleterious HIP and performance effects) that may occur at each stage of an interruption (i.e., detection and interpretation of the interruption, integration of additional performance requirements with those of the ongoing procedure's, and ongoing procedure's resumption, see Figure 3; Latorella, 1999) and thus improve overall task performance by minimizing: annunciation (interruption) lag (i.e., the time between a user receiving an interruption cue and beginning to complete the interrupting task), resumption lag (i.e., the time between leaving the interrupting task and resuming the ongoing procedures), and overall task completion time (i.e., due to combined effects gained with optimized annunciation and resumption lag times) (Trafton, Altmann, Brock & Mintz, 2003). The extent to which an interruption's deleterious effects (i.e., distraction, disturbance, and disruption effects) affect HIP workload and task performance is dependent upon how the interruption is handled by both the user (i.e., interruption management behavior, such as oblivious dismissal, unintentional dismissal, intentional dismissal, pre-emptive integration, or intentional integration) and by the system (i.e., adaptive strategies to support the intended interruption management behavior) (Latorella, 1996, 1999). AMMO's Component C may be integrated with the guidelines

summarized in Table 4 to aid in controlling for and optimizing a user's HIP workload levels via the system's presentation of information (e.g. via modality selection and timing rules/constraints and/or delegation strategies) at various stages of an interruption. Figure 3 illustrates this integration concept and depicts a process in which particular components of AMMO may be applied to the user's ongoing procedure and at specific interruption stages from Table 4 (i.e., detection, interpretation, integration, resumption).

Table 4 Information Flow and Interruption Management Strategies (adapted from Latorella 1996, 1999; McFarlane & Latorella, 2002)

<i>Stages of Interruption (See Figure 3)</i>	<i>Practical Example</i>	<i>HIP Demands &amp; Deleterious Performance Effects</i>	<i>Information Flow &amp; Interruption Management Strategies</i>
1 <i>Detection</i>	An annunciation stimulus (e.g., visual, auditory, or tactile alerting cue) of sufficient strength for sensory processing must be presented to user to facilitate detection (“grab user’s attention” away from ongoing procedure).	User’s attention is directed away from their ongoing procedure (i.e., a <i>diversion</i> ) resulting in reduced attentional resources available to maintain ongoing task performance.	Enhance detection of an interruption and ease diversion effects by improving an operator’s attention allocation and task switching capabilities via alerting cues (e.g., in another modality) (Ho et al., 2004; Hopp et al., 2005; Roda & Thomas, 2006; Trafton et al., 2003): <ul style="list-style-type: none"> <li>• <i>An alerting cue (or combination of cues) should occur in a modality that is most appropriate for the information source type (see Tables 1 and 2 and AMMO C<sub>1</sub>) and one that makes the cue dissimilar enough to the previous and current tasks to allow timely detection (Roda &amp; Thomas, 2006).</i></li> <li>• <i>A combination of modal cues could increase an alerting cue’s sensory facilitation (coactivation) effects (Miller, 1982, 1986; Iacoboni &amp; Zaidel, 2003).</i> <ul style="list-style-type: none"> <li>◦ <i>The intensity of an attention-getting cue could be mapped to the importance of an interruption (Obermayer &amp; Nugent, 2000).</i></li> </ul> </li> <li>• <i>An alerting cue should occur a few seconds before the interrupting task (i.e., empirically identify an appropriate lag, such as 0-3 seconds), where the length of this lag is not as important as its constancy or its predictability because users will learn and adapt to a consistent cueing strategy (Roda &amp; Thomas, 2006; Trafton et al., 2003).</i></li> </ul>
2 <i>Interpretation</i>	User’s attention must be maintained on the information source (i.e., alerting cue) long enough to allow translation into the associated interrupting task performance requirements (e.g., user determines if the interrupting information is vital to the ongoing procedure and must be attended to now or can be postponed until later, as with a secondary task not critical to the ongoing procedure).	Requires: attentional resources to retrieve memory representations of interrupting task from long-term memory, WM resources to instantiate a representation, and attentional resources to maintain WM representation of interrupting task. Attention and WM capacity limitations and coordination of these resources may cause deleterious performance effects (i.e., <i>distractions</i> ) to the ongoing procedure, potentially resulting in errors and increased response times.	Enhance interpretation of an interruption and mitigate distraction effects (e.g., task switching costs) by improving an operator’s ability to effectively maintain attention and WM resources on the interruption long enough to create and maintain a WM representation of the interrupting task by (Latorella 1996, 1999; McFarlane & Latorella, 2002): <ul style="list-style-type: none"> <li>• <i>Providing interrupting task information in a format and modality that is most consistent with the presentation of previous interrupting tasks requiring similar HIP resources for interpretation (Arrington et al. 2003).</i> <ul style="list-style-type: none"> <li>◦ <i>Where possible, interrupting task information should be in the appropriate modality for the information source type (see Tables 1 and 2 and AMMO C<sub>1</sub>).</i></li> </ul> </li> </ul>

<i>Stages of Interruption (See Figure 3)</i>	<i>Practical Example</i>	<i>HIP Demands &amp; Deleterious Performance Effects</i>	<i>Information Flow &amp; Interruption Management Strategies</i>
3 <i>Integration</i>	Integrating the interruption into an ongoing task set, which a user may do right after interpreting the interruption (i.e., immediate integration) or later (i.e., scheduled integration).	Integration creates <i>disturbances</i> (i.e., effects localized to preemption of the ongoing procedure such as increased annunciation and/or resumption lag) to task performance because attentional and WM resources are needed for: preemption and resumption of the interrupted position; formulation and execution of plans for performing the interruption, and; scheduling when the interruption will be performed.	Enhance integration of an interruption and mitigate disturbance effects it may have on an ongoing task (e.g., errors due to inattention, situational awareness loss of interrupted position) by facilitating effective/efficient coordination of WM resources necessary for 1) preemption and resumption of the ongoing task, and 2) formulation and execution of plans for performing the interruption immediately or scheduling it for a later time by: <ul style="list-style-type: none"> <li>• <i>Providing interrupting task information in the appropriate modality for the interrupting task's information source type (see Tables 1 and 2 and AMMO C<sub>1</sub>).</i></li> <li>• <i>Using ongoing task context to: (a) present an interruption at a cognitive break point in the ongoing task (e.g., when WM load is not at its highest, such as at higher level goal formulation or after a sub-goal is completed) (Burton &amp; Brown, 1979; Galdes &amp; Smith, 1990; Latorella 1999; McFarlane 2002), and (b) interrupt ongoing spatial tasks with verbal information (and vice versa) when respective cognitive resources for the ongoing task are loaded (Wickens, 1984, 1992).</i></li> <li>• <i>Providing external markers (e.g., modal cues as placeholders) at a point where/when an ongoing task is interrupted to facilitate later resumption of that task in a timely manner (i.e., reduce resumption lag; Latorella 1998, 1999; Trafton et al., 2003).</i></li> <li>• <i>Facilitating interruption task performance planning and execution, by determining rules for when it may be appropriate to (a) allow the user to control the timing of the interruption lag (e.g., perform task immediately; explicitly schedule until later); (b) allow the system to control the timing of the interruption lag (e.g., perform immediately; implicitly schedule until later); or (c) provide a mixed-initiative (negotiated) approach (i.e., system announces need for interruption and then supports a negotiation with the user for when/how to perform the interruption)(Zijlstra, Roe, Leonora &amp; Krediet, 1999; Cutrell, Czerwinski &amp; Horvitz, 2001; McFarlane &amp; Latorella, 2002; McFarlane 2002; Trafton et al., 2003); among others, McFarlane (2002) has established the following guidelines:</i> <ul style="list-style-type: none"> <li>○ <i>when accuracy and efficiency on the ongoing task are more important, use a negotiated approach;</i></li> <li>○ <i>when promptness and completeness on the interrupting task are more important, have the system require the user to perform the interruption immediately;</i></li> <li>○ <i>to minimize task switching, have the system schedule interruptions with consistent interruption (and resumption) lags throughout the ongoing procedure.</i></li> </ul> </li> </ul>

<i>Stages of Interruption (See Figure 3)</i>	<i>Practical Example</i>	<i>HIP Demands &amp; Deleterious Performance Effects</i>	<i>Information Flow &amp; Interruption Management Strategies</i>
4 <i>Resumption of Interrupted Task</i>	Completing the interrupting task and returning to the ongoing procedure (e.g., return to monitoring the airspace in a C <sup>2</sup> task after receiving an update on a potential new target)	Resumption of the ongoing task set is considered a <i>disruption</i> because previous interruption effects (from diversions, distractions, disturbances) propagate to disrupt future performance on the ongoing procedure once resumed; HIP requirements and disturbance effects similar to Stage 3.	Enhance resumption of the ongoing procedure (interrupted task) and mitigate overall disruptions by following the above information flow and interruption management strategies for stages 1-3 (detection, interpretation, integration), as appropriate (Latorella 1999; McFarlane & Latorella, 2002).

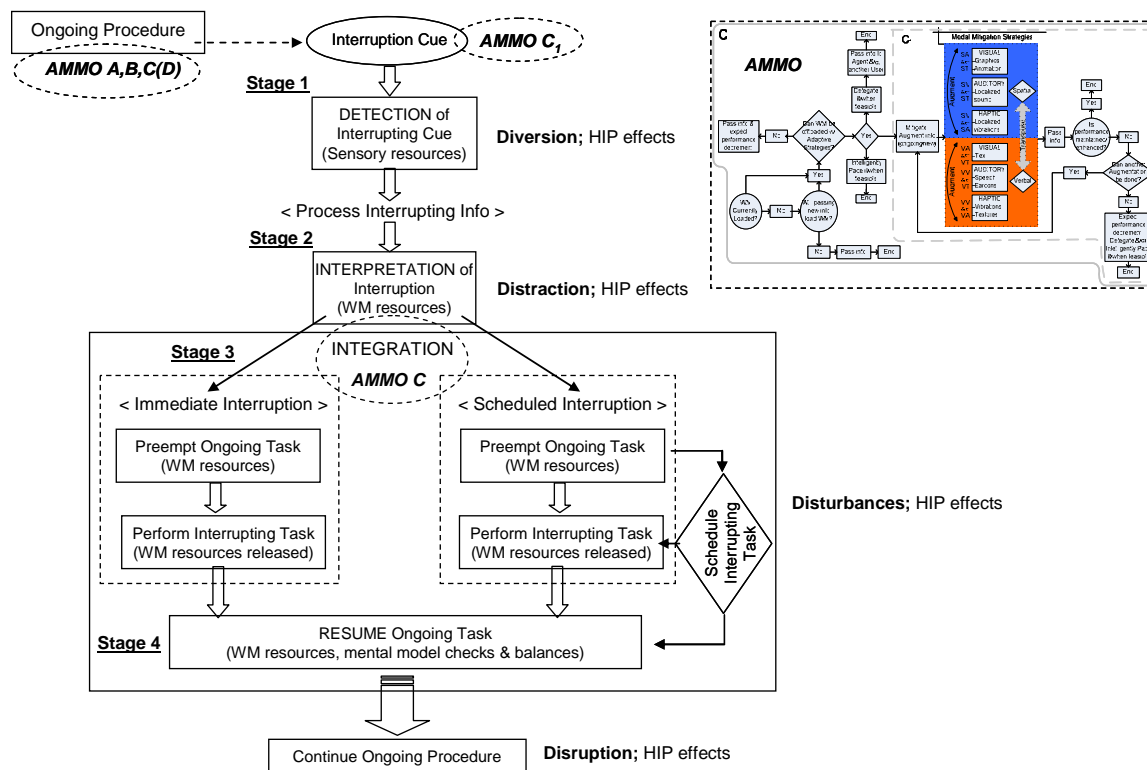


Figure 3 Conceptual model integrating AMMO's Component C with the interruption management framework presented in Table 4 (adapted from Latorella 1996, 1999).

After AMMO's A, B, C and/or D Components have been applied to optimize the design of the ongoing procedure, Figure 3 indicates how the modal mitigation strategies component of AMMO (C<sub>1</sub>) could be applied to establish the most effective cross-modal and cross-codal interruption cues to improve detection of an annunciation stimulus and minimize diversions from the ongoing procedure (see Table 4's description of Stage 1, *Detection*). AMMO's C<sub>1</sub> component is considered a recursive loop, which may be used to guide selection of the most appropriate modality or combination of modalities with which to augment ongoing and/or new task information. The top (i.e., spatial) and bottom (i.e., verbal) halves of the Modal Mitigation Strategies box each follow the information format-to-modality mapping structure set forth in Component A (see Section 2.1 and Figure 1). The main Visual, Auditory, and Haptic boxes

within in each half identify the format and modality of the ongoing or new task information that is overloading the user; the text to the left of each of these boxes (i.e., SV [spatial-visual], SA [spatial-auditory], ST [spatial-tactile], VV [verbal-visual], VA [verbal-auditory], VT [verbal-tactile]) represents potential formats (i.e., spatial vs. verbal) and modalities (i.e., visual vs. auditory vs. haptic) with which to augment the ongoing or new task information. The order in which a modal mitigation strategy is selected, whether for an ongoing or interrupting task, will likely depend on both the overall task context and WM load levels, as well as the user's WM capabilities as addressed in AMMO's Components A and B. More empirical research is needed to determine the most effective and next best modal mitigation strategies for various task and user contexts and to establish empirically-validated modal mitigation strategy parameters that could be integrated into AMMO's component C1 and applied where appropriate at each interruption stage in Table 4 (and Figure 3).

A practical example of how AMMO's Component C and sub-component C<sub>1</sub> may be used to direct the design of a multimodal mitigation strategy would be to consider the ATC task example. For instance, an operator currently being presented with SV information (e.g., monitoring ten planes on an ATC radar screen [i.e., the ongoing procedure in Figure 3]) is assessed to not yet be overloaded. However, five more planes are about to enter the operator's monitored airspace, and these planes need to be presented to the user because this is new critical information (i.e., the interruption in Figure 3) needing to be attended to immediately to avoid collisions. If presenting this new information in the same SV mapping used for the ongoing monitoring task would overload the user's WM resources (as predicted or assessed in real time), then this new information could be cued (e.g., augment SV interrupting task information with a SA, ST, VV, VA, or VT cue or combination of cues) to enhance *detection* (see Table 4 and

Figure 3, Stage 1). Then, the actual interrupting task information could be presented via one of AMMO C<sub>1</sub>'s modal mitigation strategies to offload SV WM resources (e.g., present the interrupting information in a suitably congruent format-modality combination not currently overtaxing the user's WM resources, such as SA, ST, VV, VA, or VT) and improve *interpretation* (see Table 4 and Figure 3, Stage 2) and *integration* (see Table 4 and Figure 3, Stage 3). Should the dynamically changing task conditions cause the user's WM to again become taxed to where they begin missing pertinent task information associated with the ongoing procedure, similar cueing strategies could also be used to augment the ATC's ongoing procedure information (e.g., use SA, ST, VV, VA, VT cue or combination of cues to direct user's attention back to the original 10 planes being monitored) and thus ease *resumption of the ongoing (interrupted) task* (See Table 4 and Figure 3, Stage 4). Unfortunately, there may be situations at the *interpretation* and *integration* interruption stages in which augmenting with AMMO's C<sub>1</sub> modal mitigation strategies is not feasible (e.g., due to system technology constraints and/or task environment conditions, such as when too noisy for auditory mitigations) or sufficient (e.g., user continues to be overloaded after modally mitigating and performance continues to suffer). In these circumstances, the flow in AMMO's Component C suggests implementing alternative adaptive strategies (i.e., intelligent pacing, delegating) when possible in order to minimize deleterious interruption effects on task performance and to keep WM load levels within acceptable ranges.

Consequently, Figure 3 further illustrates where the general adaptive structure in AMMO's overall C Component may be applied at the *interpretation* (see Table 4, Stage 2) and *integration* (see Table 4, Stage 3) stages to determine when and how AMMO's adaptive strategies (i.e., modal augmentation, intelligent pacing, and/or delegation) could be implemented

to minimize distraction, disturbance, and disruption effects caused by interruptions. As with AMMO's modal mitigations, intelligent pacing and delegation adaptive strategies may be applied in different combinations (in serial and/or parallel) and different orders (e.g., modally mitigate first then delegate and/or intelligently pace if/when necessary or vice versa) when establishing *integration* rules for interruptions based on a user's assessed (or predicted) WM load and task performance levels. Such rules would be used to determine whether the user should immediately perform the interrupting task and preempt the ongoing task or instead schedule when/how the interrupting task could be performed (e.g., at a different time, at a different pace, or by another user or system agent [i.e., delegation]). Task analyses, predictive modeling, and/or empirical validation via user studies would be needed to determine optimal combinations and orderings for all of AMMO Component C's adaptive strategies at each interruption stage for given user/task contexts. Although more empirical research is needed to validate such implementation guidelines as derived from Table 4 and Figure 3, the next section presents some general implementation guidance HSI designers may currently rely on when applying any of AMMO's adaptive strategies.

#### Impact of AMMO, Component C, on HSI Design

When information passes into AMMO Component C, a decision about the action(s) to be taken based on that information is made. For example, in instances when a user is performing an ongoing procedure, while WM resource requirements are not exceeding capacity, AMMO suggests both ongoing and interrupting task information continue to be passed "as is" to the user (Wickens & Hollands, 2000). On the other hand, when WM resources are already overloaded or will eventually become overloaded when information continues to be passed to the user, AMMO suggests using adaptive automation techniques (Parasuraman, Mouloua, & Hilburn, 1999) to: (a)

make certain elements of a task simpler and thus easier to perform the task (i.e., via adaptive aiding, such as by offloading WM with modal mitigation strategies [see Figure 1 and Figure 3, component C1] and/or intelligent pacing strategies if/when feasible ), and/or (b) to offload or automate an entire task from within a larger multitask context (i.e., via adaptive task allocation, such as delegating to another user or system agent if/when feasible). More empirical studies in information-intensive, operationally-relevant settings are needed to establish validated adaptive automation rules and constraints for both (a) and (b), which can then be integrated within AMMO Component C's architecture and its WM/workload-based 'if-then' decision parameters. However, the adaptive strategies and guidelines presented in Table 5 may be used by HSI designers as general guidance for when (at what stages of an ongoing procedure or interruption, see Table 4 and Figure 3) and how (combinations and orderings) to implement adaptive strategies and for designing future empirical studies for assessing how certain implementation strategies may affect a user's HIP and task performance for a given operational domain.

Table 5 AMMO Component C: Theorized Adaptive Design Strategies and General Implementation Guidelines

<i>WM Loaded</i>	<i>Adaptive Design Strategy</i>	<i>Implementation Guidelines</i>	<i>Expected Performance Benefits/Costs</i>
No	Pass info to user as is.	N/A	Maintained RT and accuracy until WM becomes loaded, and then expect performance decrements
Yes	When continuing to pass information (ongoing and/or new) in same format/modality loads WM, offload with <i>modal mitigation strategies</i> to improve information detection, interpretation, and integration ( <i>see AMMO C<sub>1</sub></i> ).	<p>Enhance <i>detection</i> of <i>ongoing</i> and <i>new</i> information via <i>cueing modal mitigation strategies</i> (see Figure 3 and Table 4, Stage 1, Detection) to direct a user's attention to the information by:</p> <ul style="list-style-type: none"> <li>• Consistently presenting attentional directing cues 0-3 seconds before the task information must be attended to by the user (Roda &amp; Thomas, 2006; Trafton et al., 2003).</li> <li>• Use free modal resources to present augmentation cues in the <i>same information format but different modality</i> (see Note 2) to facilitate congruent information format-to-modality mappings, while not unnecessarily conflicting with or additionally overloading format/modal resources being used for the ongoing or new task (Latorella, 1996, 1999; Wickens &amp; Hollands, 2000) (see AMMO's Component A and Tables 1 and 2). <ul style="list-style-type: none"> <li>○ E.g., if the user is SV loaded, augment with SA cues (Begault, 1993; McKinley &amp; Ericson, 1997; Bertolotti &amp; Strybel, 2005; Vu, Strybel &amp; Proctor, 2006; Rudmann &amp; Strybel, 1999; Bolia et al., 1999) or ST cues (Eimer et al., 2001; Kennet, Eimer, Spence &amp; Driver, 2001; Ho et al., 2005; Hopp et al., 2005; Hopp-Levine et al., 2006).</li> <li>○ If the user continues to miss cues and performance is not at acceptable levels after augmenting with a single mitigation, then augment with an additional modal mitigation (e.g., SV with both SA &amp; ST cues) to increase sensory facilitation effects of the cue and subsequent detection of task information (Spence &amp; Driver, 2004; Calvert et al., 2004).</li> </ul> </li> <li>• When cueing with the same information format but different modality is not appropriate (e.g., verbal or spatial format resources are overloaded; low spatial ability individuals may perform better with verbal cueing formats [see Component B and Table 3]), it may be effective to augment with cueing strategies in a <i>different information format but same modality</i> or a <i>different information format and different modality</i> (see Notes 3 and 4) (Wickens &amp; Hollands, 2000; Stanney et al., 2004).</li> <li>• Once optimum cueing strategies (timing and format/modality) have been determined for ongoing and/or new task information source types and user capabilities, consistently implement them (Latorella, 1996, 1999; Wickens &amp; Hollands, 2000; McFarlane 2002; Trafton et al., 2003).</li> </ul>	<p>Increased RT and reduced accuracy until WM load brought back to acceptable levels from mitigating, then expect performance improvements (Latorella, 1996, 1999; McKinley &amp; Ericson, 1997; Bolia, et al., 1999; Spence &amp; Driver, 1997, 1999, 2004; Wickens &amp; Hollands, 2000; Sarter 2000; Eimer, Cockburn, Smedley &amp; Driver, 2001; Popescu et al., 2002; Ho &amp; Spence, 2005; Ho, Spence &amp; Tan, 2005; Ho, Tan &amp; Spence, 2005; Hopp et al., 2005; Kobus et al., 2006; Schmorow, 2005; Schmorow et al., 2006; Hopp-Levine et al., 2006):</p> <ul style="list-style-type: none"> <li>• Enhanced task performance effects (e.g., improved RT and task accuracy) via attentional cueing by improving detection of both ongoing and new task information when augmenting visual information with auditory and/or haptic cues.</li> <li>• Enhanced task performance effects when using alternative task information presentation to improve interpretation and integration (i.e., distributing task information in more than one modality and/or format).</li> <li>• Enhanced task performance effects for low ability individuals when augmenting ongoing task information in a more appropriate format/modality for their needs (e.g., providing low spatial individuals with verbal- auditory information instead of or in addition to spatial-visual).</li> </ul>

<i>WM Loaded</i>	<i>Adaptive Design Strategy</i>	<i>Implementation Guidelines</i>	<i>Expected Performance Benefits/Costs</i>
Yes		<p>When cueing strategies are not sufficient to keep task performance at acceptable levels, it may be necessary to enhance <i>interpretation</i> (see Figure 3 and Table 4, Stage 2) and <i>integration</i> (see Figure 3 and Table 4, Stage 3) of <i>ongoing</i> (and/or new) task information by <i>augmenting with an alternative task information presentation</i> to more effectively distribute information across available resources, for example (Latorella, 1996, 1999; Wickens &amp; Hollands, 2000; McFarlane 2002):</p> <ul style="list-style-type: none"> <li>• If the user's spatial (or verbal) WM resources are not completely loaded but the visual (or auditory or haptic) modality resources are overloaded, then provide redundant task information in <i>the same format but different modality</i> (e.g., SV augmented with SA or ST; SA augmented with SV or ST and so forth as in Note 2) to tap other sensory resources and enhance coactivation effects, while minimizing task switching costs (e.g., by consistently providing task information in the appropriately congruent information format-to-modality mappings). <ul style="list-style-type: none"> <li>○ If augmenting with a single, additional modality is not sufficient to enhance interpretation and integration, then augment SV (VV) with both SA and ST (VA and VT) information to further increase coactivation effects and distribute workload across multiple modal resources.</li> </ul> </li> <li>• If the user's spatial (or verbal) WM resources are overloaded but visual (or auditory or haptic) modality resources are not, then provide task information in <i>a different format but same modality</i> (e.g., SV augmented with VV, such as presenting a visual text alert in a chat window to let user know the status change of one of the planes currently being monitored during an ATC task) to tap alternative WM resources and redistribute workload. <ul style="list-style-type: none"> <li>○ If the format-to-modality mapping when using the same modality is not sufficiently congruent as set forth in Component A, then provide task information in <i>a different format and different modality</i> as appropriate (e.g., SV augmented with VA &amp;/or VT or so forth as in Note 2).</li> </ul> </li> </ul> <p>Once optimum alternative task information presentation augmentation strategies have been determined for ongoing and/or new task information source types and user capabilities, consistently implement them.</p>	

<i>WM Loaded</i>	<i>Adaptive Design Strategy</i>	<i>Implementation Guidelines</i>	<i>Expected Performance Benefits/Costs</i>
Yes	<p>When modal mitigation strategies are not feasible (or sufficient, such as when WM overload continues after implementing optimized modal mitigations) and information must continue to be passed to the user (e.g., it is critical to task performance; task allocation via delegation is not available), and/or when low ability individuals are the target users (see Section 2.2), <i>intelligently pace</i> information with appropriate presentation rates and schedules for presenting information formats/modalities.</p>	<p>Employ an intelligent pacing mitigation strategy that allows the system to mitigate WM overload effects by controlling when (i.e., immediately or scheduled; see Figure 3 and Table 4, Stage 3, <i>Integration</i>) and how <i>non-critical</i> and <i>critical</i> (i.e., essential to ongoing task performance; requiring immediate attention) information is presented to the user (Latorella, 1996, 1999; Czerwinski, Cutrell, &amp; Horvitz, 2000; Mamykina, Mynatt, &amp; Terry, 2001; McFarlane, 2002; Monk, Boehm-Davis, &amp; Trafton, 2002; Hildebrandt &amp; Harrison, 2003; Schmorow et al., 2005; Kobus et al., 2006; Berka, Levendowski, Davis, Lumicao, Ramsey, Stanney, Reeves, Tremoulet &amp; Harkness-Regli, 2005; Thomas, Tremoulet &amp; Morizio, 2005):</p> <ul style="list-style-type: none"> <li>• This strategy is considered ‘intelligent’ as opposed to a simple pacing strategy because it involves an automated system not only coordinating the timing of task presentation rates during overload conditions but also prioritizing and coordinating when and how queued information is later presented to a user as an effectively designed interruption.</li> <li>• When presenting previously queued task information to a user, follow ideal information format-to-modality mapping strategies presented in AMMO’s Component A and in the above modal mitigation strategy guidelines.</li> <li>• The timing (scheduling) of information presentation should be determined by the priority of the task information, with higher priority task information (ongoing and new) being presented before lower priority information. <ul style="list-style-type: none"> <li>◦ While it may generally not be ideal to hold <i>critical</i> (high priority) task information in queue, in some circumstances (e.g., when there are more pieces of critical information than can be handled at a given time by a single available operator) it may be feasible to intelligently pace <i>critical task</i> information.</li> </ul> </li> <li>• In terms of information presentation rate, provide external pacing cues to the user (e.g., indicate the user’s task completion progress and overall time available for the task) to improve their internal pacing capabilities.</li> <li>• Implement pace recovery strategies (e.g., provide ‘window of opportunity’ timelines or status indicators for pending tasks) and warn the user of potential consequences when off the appropriate schedule and/or pace.</li> <li>• For optimized performance benefits, when available, combine neuro- and physiological measures of cognitive workload with measures of performance to aid in identifying when task information may need to be held in queue and when/how it may be presented to the user (e.g., when cognitive state gauge indicates an overload of WM resources and performance is suboptimal, instantiate intelligent pacing strategies).</li> </ul>	<p>Intelligent pacing strategies can minimize deleterious interruption effects, but expect performance decrements on the ongoing and/or interrupting task at anytime non-critical or critical task information is immediately passed to an overloaded user without mitigation (Sanders &amp; McCormick, 1992; Wickens &amp; Hollands, 2000):</p> <ul style="list-style-type: none"> <li>• Intelligent pacing adaptive strategies are appropriate during stressful, overloaded, multitasking conditions because under such task conditions users are sub-optimal at scheduling when and how long a task should take (Latorella 1996; 1999).</li> <li>• The benefits of intelligent pacing strategies have been demonstrated by Lockheed Martin Advanced Technology Lab researchers (Kobus et al., 2006; Berka et al., 2005; Thomas et al., 2005) who found greater than 100% improvement in WM throughput when using an effectively designed intelligent pacing strategy during information-intensive, multimodal C<sup>2</sup> watchstation tasks (i.e., presenting a verbal interrupting task during a primarily spatial ongoing task when spatial WM resources were detected in real time to be overloaded; presenting previously queued information once WM was detected to not be overloaded).</li> <li>• The effective timing and design of interrupting information can positively affect a user’s performance, while <i>ineffective</i> timing (e.g., information removed or presented at inappropriate cognitive peaks or valleys; system-directed presentation rates/schedules not appropriately aligned with user capabilities) may instead cause unintended deleterious interruption effects (Latorella 1996, 1999; Czerwinski, et al., 2000; McFarlane, 2002; Monk et al., 2002; Kobus et al., 2006; Berka et al., 2005).</li> </ul>

<i>WM Loaded</i>	<i>Adaptive Design Strategy</i>	<i>Implementation Guidelines</i>	<i>Expected Performance Benefits/Costs</i>
Yes	When continuing to pass information (ongoing and/or new) in same format/modality loads WM, and modal mitigation and/or intelligent pacing strategies are not feasible or sufficient, <i>delegate</i> information to another user or system agent as appropriate.	<p>Delegation is feasible (Parasuraman, Bahri, Deaton, Morrison &amp; Barnes, 1992; Parasuraman, Mouloua &amp; Molloy, 1996; Parasuraman et al., 1999; Wickens &amp; Hollands, 2000; Sheridan &amp; Parasuraman, 2006): when a critical event occurs and must be attended to immediately; assessment of operator performance levels (e.g., reaction time, false alarms, hit rate, omissions, etc.) and/or cognitive state (i.e., via predictive models or real-time physiological and neurophysiological sensors) indicates system intervention is necessary, and; when adaptive automation techniques are integrated within the human-computer environment to allow task allocation to another user or system agent [For instance: if the workspace environment is a co-operative workgroup, meaning more than one individual is involved as a team (e.g., two or more C<sup>2</sup> operators monitoring the same airspace), then certain task functions could be dynamically allocated to the next most feasible operator (e.g., the operator who has WM resources available to attend to another critical task) or to an intelligent agent when another operator is not available].</p> <p>Ensuring accuracy of individual and shared mental models of the system state and of the dynamic allocation processes is critical when implementing delegation strategies (Byrne &amp; Parasuraman, 1996; Hoc &amp; Lemoine 1998; Hoc 2001; Hoc &amp; Debernard, 2002; Prinzel et al., 2003; Sheridan &amp; Parasuraman, 2006):</p> <ul style="list-style-type: none"> <li>• Delegate entire functions and not just sub-task information; if sub-tasks must be delegated, ensure they are as independent as possible.</li> <li>• When delegating tasks or functions, use implicit delegation (i.e., system directed) when performance improvement is the primary constraint, use explicit delegation (i.e., user directed) when user control and acceptance are primary constraints, and use assisted explicit delegation (i.e., system proposes delegation strategies to the user, who then has the control to accept or reject) when possible to capitalize on strengths of both methods while avoiding complacency and trust issues.</li> </ul> <p>For optimized performance benefits, when available, <i>combine</i> neuro- and physiological measures of cognitive workload with measures of performance to aid in identifying when a user's workload needs to be offloaded and when it no longer needs to be offloaded.</p>	<ul style="list-style-type: none"> <li>• Increased RT and reduced accuracy until WM load brought back to acceptable levels after delegating.</li> <li>• Delegation strategies have been shown to alleviate a human operator's existing workload peaks and to subsequently improve task performance by reducing the amount of information needing attending to (i.e., number and costs of mental operations required) by the currently overloaded operator (Wickens &amp; Hollands, 2000; Crévits, Debernard, &amp; Denecker, 2002; Hoc &amp; Debernard, 2002; Prinzel, Parasuraman, Freeman, Scerbo, Mikulka &amp; Pope, 2003).</li> <li>• As summarized in Prinzel et al. (2003), the benefits of adaptive automation techniques, such as delegation, may include: regulated user workload, bolstered situational awareness, enhanced vigilance, maintenance of manual skill levels, increased task involvement, and overall improved operator performance (Endsley, 1996; Parasuraman et al., 1992; Parasuraman, et al., 1996; Scerbo, 1994, 1996, 2001).</li> <li>• Conversely, Prinzel et al. (2003) address how adaptive automation has not yet fully matured and more empirical evidence is needed to determine when and how adaptive aiding should take place in order to reduce potential negative effects of automation (e.g., loss of situational awareness, user trust, system reliability/stability, etc. when relinquishing some control to another operator or system agent) (Billings &amp; Woods 1994; Wickens &amp; Hollands, 2000).</li> </ul>

Notes: (1) RT refers to reaction or response time.

(2) Augmenting with *the same information format but different modality* includes: SV augmented w/ SA &/or ST; SA augmented w/ SV &/or ST; ST augmented w/ SV &/or SA; VV augmented w/ VA &/or VT; VA augmented w/ VV &/or VT; VT augmented w/ VV & VA.

(3) Augmenting with *a different information format but same modality* includes: SV augmented w/ VV; SA augmented w/ VA; ST augmented w/ VT; VV augmented w/ SV; VA augmented w/ SA; VT augmented w/ ST.

(4) Augmenting with *a different format & different modality* includes: SV augmented w/ VA &/or VT; SA augmented w/ VV &/or VT; ST augmented w/ VV &/or VA; VV augmented w/ SA &/or ST; VA augmented w/ SV &/or ST; VT augmented w/ SV &/or SA

Once processing in Component C is completed and task information is passed on to the user, the AMMO processing cycle would begin again at Component A. As discussed in the next section, during the initial pass of this processing loop, there may be instances when outputs from A are not ideally mapped and would thus pass on to Component D of AMMO instead of C for sub-component processing.

#### *AMMO: Non-Ideal Mappings (D)*

The purpose of AMMO's Component D is to address design considerations for when the mapping of information format-to-modality may not be ideal (i.e., are incongruent) and thus induce unnecessary strain on an individual's WM processing resources (e.g., invoke inefficient information processing strategies; inhibit otherwise advantageous effects of sensory facilitation). Specifically, if mappings are identified as non-ideal in AMMO's Component A, then information to be presented to users will follow through AMMO's Component B, with possible moderation on the amount of information being passed, and then on to AMMO's Component D. Component D suggests determining whether information could be redesigned by recoding (i.e., transposing) it into the appropriate format (i.e., verbal into spatial; spatial into verbal) and then using Component A to remap to suitable display modalities. If non-ideal mappings can be fixed with such a system redesign strategy, it is expected that positive performance effects will be realized, particularly with respect to response time and decision making capabilities (Sanders & McCormick, 1993; Wickens 1984, 1992; Wickens & Hollands, 2000). Such performance gains are expected due to congruently mapped display modalities and task information formats improving a user's ability to efficiently process incoming information and thus effectively

manage both ongoing and interrupting tasks. Both the neurological and behavioral evidence presented thus far supports this notion.

Unfortunately, system designers may face situations where such redesign is not possible, whether due to budget restrictions or simply interface standards that are instantiated and cannot be changed. For these conditions, other options for mitigating potentially negative cognitive workload and performance effects may need to be explored (Schmorrow et al., 2005). Currently, AMMO suggests using the adaptive strategies in Component C (i.e., modal mitigation, intelligent pacing, delegation) to optimize WM processing capabilities and minimize deleterious interruption effects. More empirical research is needed to determine how effective such strategies may be with non-ideal mapping conditions and whether such conditions would affect the existing ‘if-then’ logic and associated adaptive strategy guidelines in Component C.

### Future Directions

Although the theorized guidelines derived from AMMO (see Table 5) represent a good starting point for proactively directing the design of multimodal and other adaptation strategies, more empirical studies are needed to validate whether such guidelines hold true under various task and user conditions. For, not all design attempts to enhance performance may actually reap the benefits expected, and sometimes negative effects may be seen, particularly when multiple forms of sensory integration are involved (e.g., unwanted cross-modal effects). Furthermore, evidence from the dual-process theory of plasticity (Groves & Thompson, 1970), regarding how depression and facilitation compete to determine the final strength of a signal (whether uni-, bi-, or multimodal), also addresses how prolonged exposure to such a signal may lead to eventual habituation. The implication to general multimodal display design, and to multimodal mitigation

strategies in particular, is that certain augmentation strategies may initially work to provide attention-alerting mechanisms for users during high workload conditions but may lose their effectiveness over time. Consequently, once multimodal mitigation strategies are initially validated, additional efforts could focus on longitudinal studies to investigate potential deleterious habituation effects and alternate adaptive strategies to overcome them. Such future studies could also involve examining user populations with known significant variances in WM capabilities (e.g., low/high spatial/verbal processors) to examine specific effects of such capabilities on the selection of appropriate mitigation strategies for users with particular WM capabilities and limitations.

AMMO has been developed as both an a priori design framework for extracting multimodal display guidelines, as well as with a flow that could potentially be used to build simulation-based predictive or real-time models for directing adaptive automation. Thus, future research could also investigate the combination of AMMO's current MRT-based logic with known or predicted values of HIP parameters (e.g., WM capacity, decay) and with appropriate task context modeling in order to provide quantitative predictions (or real-time estimations) of workload and human performance effects. Both approaches (a priori design; predictive or real-time modeling) may be used to design appropriate empirical studies necessary to validate derived guidelines, as well as identify both inter- and intra-adaptive strategy (modal mitigation, intelligent pacing, delegation) rankings (or weightings) in terms of their measured effectiveness for reducing WM load and improving performance in various user and task settings (i.e., prioritized best and 'next best' strategies). For instance, when unaware of the conflicts that may occur between modalities, (e.g., incongruency, sensory conflict, capture), the modality-assigning design stage would begin by presenting information in the modality that is most appropriate or

beneficial (see Tables 1, 2, 3, 4, and 5), and then these modal mitigations could be prototyped and evaluated for effectiveness or potential unwanted sensory conflicts (ETSI, 2002; Stanney et al., 2003, 2004). A similar implementation and evaluation approach could be used to validate orderings and combinations of intelligent pacing and delegation adaptive strategies—individually or in combination and with modal mitigation strategies.

It is envisioned that through validation and implementation of AMMO in information-intensive operational environments (e.g., military C<sup>2</sup>), more robust augmented cognition may be achieved, whereby a real-time intelligence model effectively directs what type and when an adaptive strategy should be invoked once real-time cognitive sensors detect when and how a person's WM resources are overloaded. Then, conclusions may start to be drawn regarding the generalizability of particular strategies across multiple information-intensive task domains and for various types of users (e.g., low/high verbal/spatial processors).

### Conclusion

The AMMO model and associated interruption management guidelines presented in this study have been developed based on a multimodal extension of existing HIP theories and models, at a functional WM level, with the latest findings in cognitive psychology, neuroscience, and other allied sciences. It is proposed that AMMO may be used to guide HSI researchers and designers as to what cognitive workload mitigation strategies (i.e., information flow and interruption management adaptive strategies) may be most appropriate for given users and task contexts. Once empirically validated with experiments in various applied task settings and with users of varying WM abilities, the utility of AMMO would lie in its ability to provide HSI designers with both a priori design strategies and adaptive automation strategies (i.e., multimodal mitigation,

intelligent pacing, and delegation) in real-time operational settings, as well as aid in establishing manning requirements once designs are optimized. Appropriate application of AMMO could thus facilitate performance improvements (e.g., improved response time and accuracy) via a reduction in potential information processing bottlenecks and task switching costs and minimized effects of subsequent information overload conditions (i.e., where users fail to detect, interpret, integrate, and successfully act on pertinent task information). Such an architecture may provide HSI designers with the proper ammunition ('AMMO') necessary to efficiently and effectively design most any of today's information-intensive, multi-tasking systems (e.g., air-traffic control, military command and control watchstations, intelligence analysis). There is no reason for us to continue to "*live in a technology-centered world where the technology is not appropriate for people* (Norman, 1998, p. 135).

#### Acknowledgment

This work was made possible with the support of an Office of Naval Research (ONR) funding under the leadership of CDR Dylan Schmorow. The authors wish to thank ONR and CDR Schmorow for their support. The views expressed in this article are those of the authors and do not necessarily reflect those of their affiliated organizations or sponsoring organizations.

## **CHAPTER FOUR<sup>3</sup>: EMPIRICALLY VALIDATING MULTIMODAL MITIGATION STRATEGIES DERIVED WITH AMMO**

The present study focused on empirically validating a set of multimodal design guidelines in a simulated weapons control system multitasking environment. The guidelines direct when and how to interrupt users by implementing multimodal cueing strategies that use combinations of visual, auditory, and/or haptic information augmentation strategies. To validate the guidelines participants were involved in a multitasking scenario that consisted of a primary spatial-visual and verbal-visual combination task (i.e., monitoring and retargeting missiles to emergent targets) and a visual-verbal interrupting task (i.e., visual chat questions regarding the current retargeting task or other system state questions). The results of this study showed significant human performance improvements when multimodal cues were used to augment the spatial-visual and verbal-visual information display. Specifically, the average response time for the verbal-visual interrupting task showed: 13% (s.d.= 1.2) improvement when augmented with verbal-auditory (redundant speech) cues, 7% (s.d. = 1.5) improvement when augmented with verbal-tactile cues, 14% (s.d.=1.3) improvement when augmented with a combination of verbal-auditory (redundant speech) and verbal-tactile cues, 9% (s.d. = 1.5) improvement when augmented with verbal-auditory (tonal) cues, and 9% (s.d. = 1.3) improvement when augmented with a combination of verbal-auditory (tonal) and verbal-tactile cues. In addition, when spatial-visual information in the primary task was augmented with spatial-tactile and a combination of spatial-auditory and spatial-tactile cues, the average response time improved by 8% (s.d.= 1.3), and 10% (s.d. = 1.3),

---

<sup>3</sup> This chapter has been submitted to the TIES Journal as: Reeves, L. M. & Stanney, K. M., Ahmad, A., & Malone, L. (2007). Empirically validating multimodal mitigation strategies derived with AMMO. Manuscript submitted to *Ergonomics*.

respectively. These results provide empirical validation of a set of principle-driven multimodal design guidelines, which may be used as effective interruption management strategies applicable to a wide range of information-intensive computer-based task environments.

### Introduction

As today's human-computer systems are increasingly able to provide more information than a single human operator can efficiently and effectively process and act on, a challenge for designers is to create interfaces that allow operators to process an optimal amount of data in a timely manner. It has been proposed that this might be accomplished by creating multimodal display systems that augment display modalities to maximize user's information processing capabilities, particularly at the working memory (WM) level (Miyake & Shah, 1999; Calvert et al., 2004; Stanney et al., 2004; Oviatt, Coulston & Lunsford, 2004; Reeves & Stanney, 2007). When information is distributed across multiple sensory modalities (i.e., visual, auditory, haptic) and WM codes (i.e., verbal and spatial), improved WM capacity limits are theorized to occur through the use of non-interfering modalities (e.g., tonal, kinesthetic, tactile) and cross-codal information formats (i.e., verbal and spatial) (Wickens, 1984, 1992, 2002; Sulzen 2001; Wickens & Hollands, 2000; Stanney et al., 2004). Reeves and Stanney (2007) advanced these and other research findings to develop an Architecture for Multi-Modal Optimization (AMMO), which was used to derive a theorized set of guidelines that can be used to direct multimodal display design (see Reeves & Stanney, 2007, for a detailed description of AMMO and a complete list of derived guidelines). The current study focuses on a subset of these guidelines (see Table 6), those aimed at improving interruption management in information-intensive, computer-based multitasking environments (e.g., Air Traffic Control [ATC], military Command and Control [C<sup>2</sup>], stock

trading, intelligence analysis) by augmenting incoming primary task and interrupting task information with multimodal cues to minimize deleterious interruption effects (e.g., decision making and response errors, task switching costs, loss of situational awareness, and increased task completion times) (Cohen, 1980; Latorella, 1996, 1999; McFarlane, 2002; McFarlane & Latorella, 2002; Speier et al., 2003; Dorneich et al., 2004; Hopp et al., 2005; Hopp-Levine et al., 2006).

Table 6 Subset of Theorized Multimodal Design Strategies Derived from AMMO for Improved Interruption Management (adapted from Reeves & Stanney, 2007)

Adaptive Design Strategy	Implementation Guidelines	Expected Performance Benefits
When continuing to pass information (ongoing and/or new) in same format/modality loads WM, <i>offload with modal mitigation strategies</i> to improve information detection, interpretation, integration, and subsequent action.	<p>Enhance <i>detection</i> of ongoing and new information via <i>cueing modal mitigation strategies</i> to direct a user's attention to the information by:</p> <ul style="list-style-type: none"> <li>• Consistently presenting attentional directing cues 0-3 seconds before the task information must be attended to by the user (Roda &amp; Thomas, 2006; Trafton et al., 2003).</li> <li>• Use free modal resources to present augmentation cues in the <i>same information format but different modality</i> (see Note 1) to facilitate congruent information format-to-modality mappings, while not unnecessarily conflicting with or additionally overloading format/modal resources being used for the ongoing or new task (Latorella, 1996, 1999; Wickens &amp; Hollands, 2000). <ul style="list-style-type: none"> <li>◦ E.g., if the user is SV loaded, augment with SA cues (Begault, 1993; McKinley &amp; Ericson, 1997; Bertolotti &amp; Strybel, 2005; Vu et al., 2006; Rudmann &amp; Strybel, 1999; Bolia et al., 1999) or ST cues (Eimer et al., 2001; Kennet et al., 2001, Ho et al., 2005; Hopp et al., 2005; Hopp-Levine et al., 2006)</li> <li>◦ If the user continues to miss cues and performance is not at acceptable levels after augmenting with a single mitigation, then augment with an additional modal mitigation (e.g., SV with both SA &amp; ST cues) to increase sensory facilitation effects of the cue and subsequent detection of task information (Spence &amp; Driver, 2004; Calvert et al., 2004).</li> </ul> </li> <li>• When cueing with the same information format but different modality is not appropriate (e.g., verbal or spatial format resources are overloaded; low spatial ability individuals may perform better with verbal cueing formats), it may be effective to augment with cueing strategies in a <i>different information format but same modality</i> or a <i>different information format and different modality</i> (see Notes 2 and 3) (Wickens &amp; Hollands, 2000; Stanney et al., 2004; Reeves &amp; Stanney, 2007).</li> <li>• Once optimum cueing strategies (timing and format/modality) have been determined for ongoing and/or new task information source types and user capabilities, consistently implement them (Latorella, 1996, 1999; Wickens &amp; Hollands, 2000; McFarlane 2002; Trafton et al., 2003).</li> </ul>	<p>Increased response time and reduced accuracy until WM load brought back to acceptable levels from mitigating, then expect performance improvements (Latorella, 1996, 1999; McKinley &amp; Ericson, 1997; Bolia, et al., 1999; Spence &amp; Driver, 1997, 1999, 2004; Wickens &amp; Hollands, 2000; Sarter 2000; Eimer et al., 2001; Popescu et al., 2002; Ho &amp; Spence, 2005; Ho et al., 2005; Ho, Tan &amp; Spence, 2005; Hopp et al., 2005; Kobus et al., 2006; Schmorow, 2005; Schmorow et al., 2006; Hopp-Levine et al., 2006):</p> <ul style="list-style-type: none"> <li>• Enhanced task performance effects (e.g., improved response time and task accuracy) via attentional cueing by improving detection of both ongoing and new task information when augmenting visual information with auditory and/or haptic cues. <ul style="list-style-type: none"> <li>◦ While the auditory modality is typically used as an effective cueing modality and for alerts and warnings, the haptic modality is showing great promise as an alternate cueing modality, particularly when an auditory approach may not be ideal or sufficient (e.g., too noisy of an environment to detect auditory cues; the auditory information is too intrusive to ignore and disrupts performance on another task, or; the auditory information is too difficult to interpret and integrate when visual and auditory resources are overtaxing the same verbal or spatial WM resources) (Stanney et al., 2004).</li> </ul> </li> </ul>

Notes:

(1) Augmenting with *the same information format but different modality* includes: Spatial-Visual (SV) augmented w/ Spatial-Auditory (SA) (e.g., graphics w/ localized sounds) &/or Spatial-Tactile (ST) (e.g., graphics w/ localized vibrations); SA augmented w/ SV &/or ST; ST augmented w/ SV &/or SA; Verbal-Visual (VV) augmented w/ Verbal-Auditory (VA) (e.g., visual text w/ speech or earcons) &/or Verbal-Tactile (VT) (e.g., visual text w/ tactile vibrations or textures); VA augmented w/ VV &/or VT; VT augmented w/ VV & VA.

(2) Augmenting with *a different information format but same modality* would include: SV augmented w/ VV (e.g., graphics w/ text); SA augmented w/ VA (e.g., localized sound w/ speech, earcons); ST augmented w/ VT (e.g., localized vibrations w/ vibrations or textures); VV augmented w/ SV; VA augmented w/ SA; VT augmented w/ ST.

(3) Augmenting with *a different format & different modality* would include: SV augmented w/ VA (e.g., graphics with speech or earcons) &/or VT; SA augmented w/ VV &/or VT; ST augmented w/ VV &/or VA; VV augmented w/ SA &/or ST; VA augmented w/ SV &/or ST; VT augmented w/ SV &/or SA.

As presented in Table 6, AMMO's modal mitigation strategies examined in the present study involve augmenting ongoing task and interrupting task information with multimodal cues to direct attention to pertinent task information, where augmenting is done via *the same information format* (i.e., spatial or verbal) *but different modality* (i.e., visual, auditory, or haptic). By providing cueing (and redundant) task information in alternate sensory display modalities, it is suggested that WM processing at the cortical level may be effectively distributed both within and across multiple sensory modalities (i.e., visual, auditory, haptic) and WM codes (i.e., verbal and spatial) to ease deleterious interruption effects and WM overload conditions (Latorella, 1996, 1999; Wickens & Hollands, 2002; Wickens, 2002; Cellier and Eyrolle, 1992; Wickens, Goh, Helleberg, Horrey & Talleur, 2003). Further, effectively designed multimodal cueing strategies can aid operators of visually busy environments by more efficiently directing their attention to where/when it is most critically needed and thus improve their interruption management capabilities and overall task performance (Cellier & Eyrolle, 1992; Hopp et al., 2005; Hopp-Levine et al., 2006).

The benefits of modal augmentation strategies have been reported in numerous studies (Latorella, 1996, 1999; McKinley & Ericson, 1997; Bolia, et al., 1999; Spence & Driver, 1997, 1999, 2004; Rudmann & Strybel, 1999; Sklar & Sarter, 1999; Wickens & Hollands, 2000; Sarter 2000; Eimer et al., 2001; Popescu et al., 2002; Ho et al., 2004; Bertolotti & Strybel, 2005; Ho & Spence, 2005; Ho et al., 2005; Hopp et al., 2005; Kobus et al., 2006; Schmorrow, 2005; Vu et al., 2006; Schmorrow et al., 2006; Hopp-Levine et al., 2006), where enhanced task performance effects (e.g., improved response time and task accuracy) have been realized via attentional cueing and information presentation redundancy (i.e., distributing task information in more than

one modality and/or format) to minimize cognitive overload conditions. However, existing behavioral studies, which demonstrate such performance effects, are historically bimodal in nature, where modal distribution effects are investigated by augmenting with a single, additional modality (e.g., visual information augmented with auditory information; visual augmented with haptic) (Stanney et al. 2004; Reeves & Stanney, 2007). Recent neuroimaging-based studies indicate that augmenting with more than one modality (i.e., *multimodal* strategies, such as augmenting visual information with auditory *and* haptic information) may reap even greater human performance gains than bimodal strategies due to additional increases in sensory facilitation and potential cross-modal coactivation effects (Stein & Meredith, 1993; Eimer et al., 2001; Kennet et al., 2001; Dyson & Quinlan, 2002; Calvert et al., 2004; Spence & Driver, 2004). Consequently, the present study is focused on empirically validating, in an operationally-relevant simulated weapons control system multitasking environment, a subset of the bimodal *and* multimodal cueing augmentation strategies derived from AMMO and presented in Table 6. The specific *same format/different modality* hypotheses examined include:

- H1:** Augmenting verbal-visual information (VV) with a verbal-auditory (VA) (tonal) cue will have a positive effect on performance as compared to no mitigation (no augmentation).
- H2:** Augmenting VV information with a redundant VA speech cue will have a positive effect on performance as compared to no mitigation.
- H3:** Augmenting VV information with a verbal-tactile (VT) cue will have a positive effect on performance as compared to no mitigation.
- H4:** Augmenting VV information with VA (tonal or redundant speech) and VT cues will have a positive effect on performance as compared to no mitigation, where performance gains may be greater than augmenting with either VA or VT cues alone.

**H5:** Augmenting spatial-visual (SV) information with a spatial-auditory (SA) cue will have a positive effect on performance as compared to no mitigation (i.e., augmentation).

**H6:** Augmenting SV information with a spatial-tactile (ST) cue will have a positive effect on performance as compared to no mitigation.

**H7:** Augmenting SV information with SA and ST cues will have a positive effect on performance as compared to no mitigation, where performance gains may be greater than augmenting with either SA or ST cues alone.

It is envisioned that once empirically validated, multimodal display guidelines derived from AMMO, such as those presented in Table 6 and hypothesized above, could empower human-systems interaction designers with principle-driven and practical design guidance. Such guidance could aid designers regarding how to effectively distribute information across display modalities other than visual (e.g., auditory or tactile) to improve the detection, interpretation, and integration of ongoing and interrupting task information during information-intensive multitasking situations.

## Method

### *Participants*

A total of 32 (25 males, 7 females) participants were recruited. All participants except two males were right-handed, and one of those two reported being ambidextrous. Each participant was paid \$10/hr, and each experimental run took on average 4 hours. The average age for participants was 22.5 years (s.d.= 5.1 years; with a range of 16 - 35 years). All participants used computers for 9.6 years on average (s.d.= 4.1 years). The participants had an average of 7.9 years of gaming experience (s.d.= 5.43 years). Only four of the participants had experience with spatial audio,

which mostly involved first-person shooter games or similar. Only three of the participants had any tactile interaction experience, which mostly involved force feedback from a vest or gun while playing games. High school students who participated were all in advanced placement classes and college bound (participants under 18 years old needed a parental consent form signed by their parent or legal guardian). All other participants were undergraduate and graduate level college students or college graduates from both the schools of engineering and psychology.

### *Equipment*

All computer-based tasks were performed on a 3.0 GHz Intel P4 processor computer with an MSI K7N2G-ILSR NF2 AGP 8X motherboard, GEFORCE-4 TI 4600 8x AGP video card, two CORSAIR 512 Mb PC3200 PC400 DDR memory chips, and a Creative SB Audigy 2 Platinum 6.1 sound card. The operating system was Linux Red Hat “Strike” version. The interface was presented on a 19” Viewsonic 0.22 dot pitch flat screen monitor at 85 Hz refresh rate and 1024x768 screen resolution, with audio presented through Creative THX 550 speakers. Tactile cues were presented via a tactile vest made of neoprene material and developed by the University of Central Florida’s (UCF) Institute for Simulation and Training (IST). The vest’s tactors were created with standard cell phone batteries (i.e., approximately 1.5 V DC, model 6CL-5472A from VibratorMotor.com) at average frequencies (approx 50 Hz each). When activated, the tactors created a buzzing sensation to the participants, similar to a cell phone set to ‘vibrate mode.’

A simulated Tactical Tomahawk Weapons Control System (TTWCS) task interface was programmed in java, using OpenAL for spatial audio and tonal cues. Synthetic text-to-speech was presented via Lockheed-Martin Advanced Technologies Laboratory’s (LMATL) proprietary speech engine. All participants’ input was performed via a standard keyboard and mouse.

## *Tasks*

### Baseline TTWCS task environment (unmitigated/not augmented)

The simulated TTWCS task environment used for this study entailed participants performing the role of a Tactical Strike Coordinator (TSC), whose overall objective was to monitor and adjust (i.e., retarget missiles to emergent targets) an in-progress missile strike package for 90 seconds while also tending to various other task environment demands (e.g., chat information, questions from a CO, target updates, etc.) that may interrupt performance on the monitoring and retargeting tasks. A strike package consisted of a set of missiles following individual pre-set missions, where each missile was assigned to service a specific default target. There were also emergent (newly appearing) targets that had to be serviced, which were high priority targets with a limited window of opportunity (timeframe within which they had to be serviced). Emergents randomly appeared during the 90 second trial and were not part of the original pre-assigned strike package. Although multiple task components were dynamically changing during the entire task trial and needed to be monitored (e.g., system update information presented in chat window, windows of opportunity changing as reassign missiles, etc.), the TSC's performance scores were based on two main tasks, which occurred simultaneously during all task scenario conditions and constantly competed for the TSC's available WM and attentional resources:

- Retarget Task-- retargeting missiles based on emerging targets in the Tactical Situation Display (Tacsit) window, while also maintaining maximum coverage on as many high and medium default targets as possible; this task was a combination of verbal and spatial WM and executive functioning.

- Alert Task-- responding to Alert questions presented in a visual chat interface window, which may interrupt performance on the Retargeting task at any time; this task was predominantly a verbal WM and executive function task.

Figure 4 illustrates the main components of the TSC's visual interface. For this simulated TTWCS task platform, there were three types of missiles available: *Unitary* missiles are traditional high explosive devices with approximately 50% explosives by weight; *Penetrating* devices have hardened casings, which allow them to punch through bunkers or earth, and have approximately 25-30% explosives by weight, and; *Submunition* devices are cluster bombs which consist of grenade-like balls encased in plastic impregnated with ball bearings or metal darts designed to shower the target area. Every target required a specific missile type, so the TSC had to be sure each default and emergent target was appropriately mapped to the specific type and amount of missiles required to service it (i.e., to successfully destroy the target). Targets (red diamonds; filled red for emergent) and missiles each had specific icons to represent them, with alphanumeric codes printed below them to aid in their identification (ID) by the TSC. For example, T033S-EH indicated (T)arget # 0333 requiring a (S)ubmunition missile, and the target was (E)mergent and (H)igh priority. The same naming convention was used for missiles, where the ID would start with an L/R/F (Loiter/Retarget/Fire and Forget) followed by an M for missile. For example, LM032P-DL would indicate (L)oitering (M)issile #032, which is a (P)enetrating missile currently assigned to a (D)efault and (L)ow priority target. Above each target icon was a symbol, indicating to the TSC the number of missiles required to fully service it (ranging from 1 to 3 missiles).



“yes/no” dialogue box to confirm they wanted to finalize the current retarget change. To aid in retargeting strategies, the TSC had the opportunity to use the missile timeline window to see time-on-target (TOT) for each missile (i.e., the time a missile would impact its current target represented with a black rectangle or a potential emergent target represented with a red rectangle).

For the *Alert Task*, the TSC responded to visual questions presented in the response window (see Figure 4, lower center), which interrupted the ongoing Retargeting Task. A standard system beep prompted the participant each time an Alert Task question appeared. Operators had approximately 15 seconds to answer and click the “Done” button in the response window before the question disappeared and could no longer be answered. Types of questions asked related to task information presented in either the Tacsit or the visual Chat window and included the following examples:

- Which missile will reach its target [last/first] if all go directly to their default target?
- How many missiles are you monitoring right now?
- How many targets are you monitoring right now?
- What is your communications channel?
- What is your heading?

For incentive, operators were told they would be scored on both response time and accuracy, with a correct answer gaining them 100 points and bonus points awarded for answering before the 15 second deadline (although for data analysis, participants’ performance was assessed with raw scores of correct/incorrect and pure response time).

As discussed next, each modally augmented (mitigation) task scenario condition involved the baseline TTWCS interface described above but with some form of a multimodal (auditory and/or

tactile) cueing strategy added to direct attention to either (1) a new emergent target that had just appeared (Retarget Task) or (2) to answer a chat question (Alert Task) in a timely manner. The design of each modal cue was based on established design guidelines necessary to meet perceptual thresholds (Sherrick & Cholewiak, 1986; Sanders & McCormick, 1993; see Stanney et al., 2004 for a summary of modal design guidelines).

#### Verbal-Visual information augmented with Verbal-Auditory (tonal) cue

For this treatment condition, the visual chat (Alert question) was augmented with a three-ping auditory cue from two front speakers (on either side of the computer monitor) alerting the participant to answer a question they might have missed. This three-ping warning sound cue was given when there were approximately eight seconds left to answer the current Alert Task question. The three-ping auditory cue was chosen because its auditory properties differed in both time and frequency to the standard system beep (i.e., the one used to indicate when a new visual Alert Task question first appeared on screen). Similarly with all the modal cue conditions, participants were instructed and trained to know that the cue indicated only approximately 7-8 seconds remained to answer a pending Alert question.

#### Verbal-Visual information augmented with Verbal-Tactile cue

In this scenario condition, the visual chat question was augmented with a vibratory cue from all eight tactile vest quadrants (four front, four back; see bottom right of Figure 5). As with the tonal auditory three-ping cue, this vibratory cue was given when there were approximately eight seconds left to answer the current Alert Task question. Each tactor was 50 Hz with a 1 second duration sufficient for appropriate torso detection. Although the two-point threshold has generally been found to be much smaller at the hand and fingertip than at the torso (Sherrick &

Cholewiak, 1986), in recent studies (Cholewiak & Collins, 2000; Erp & Veen, 2003; Tan, Gray, Young & Taylor, 2003; Cholewiak, Brill & Schwab, 2004; Lindeman, Page, Yanagida & Sibert, 2004; Hopp et al., 2005; Hopp-Levine et al, 2006) the torso has proven an effective body area for general vibrotactile cueing in operationally-relevant task domains.

#### Verbal-Visual information augmented with both Verbal-Auditory (tonal) and Verbal-Tactile cues

This task scenario involved the combination of cues described in 2.3.2 and 2.3.3, with the visual Alert Task question being augmented with both the tonal and vibratory cues when there were approximately eight seconds left to answer a question. To increase sensory facilitation effects in the redundancy of the VA and VT cues occurring together, the cues were designed to occur temporally close, starting within 150ms of each other and ending within 500ms of each other (the final third ping of the three-ping tonal cue finished just after the vest tactors stopped vibrating).

#### Verbal-Visual information augmented with Verbal-Auditory (redundant speech) cue

To examine another form of redundancy and its potential effectiveness in improving interruption management capabilities, an additional type of verbal-auditory cue was implemented—redundant speech. For this scenario condition, when a visual Alert Task question first appeared, it was augmented with a redundant synthetic speech cue from the two front speakers located on either side of the computer monitor. The synthetic speech cue asked the participant the same exact Alert question that was shown to them visually. The redundant combination of visual and auditory information was meant to help participants to multitask in the visually busy TTWCS environment via more efficient and effective timesharing of modal and WM resources (Wickens & Hollands, 2000; Wickens & Gosney, 2003).

### Verbal-Visual information augmented with both Verbal-Auditory (redundant speech) cue and VT cue

This scenario combined the cues described in 2.3.3 and 2.3.5. In this treatment condition, the redundant speech cue occurred immediately when the visual Alert Task appeared, and then the VT cue occurred when there were about eight seconds remaining to answer the question.

### Spatial-Visual information augmented with Spatial-Auditory cue

This SA cueing strategy scenario was focused on cueing the participant when each new emergent target appeared in the Tacsit window for the Retargeting Task. The SA cue occurred simultaneously (within 100ms) when a new emergent appeared in the Tacsit window. Figure 5 illustrates the placement of the localized speakers at 45 degrees elevation (above and below the horizontal plane of each participant's ears) along the median plane and in the same vertical plane as the computer monitor, where the upper speaker emitted the cue for a northward appearing emergent target, and the lower speaker emitted the cue for a southward appearing emergent. To create effective SA cueing strategies that designate visual targets, SA cues specifying the exact target location are generally not essential and have shown to increase search time (Rudmann & Strybel, 1999; Bertolotti & Strybel, 2005; Vu et al., 2006). Thus, 45 degrees was chosen to provide cues specifying the 'local' target area in the north or south direction and to avoid potential up-down/front-back discrimination issues that may have occurred with larger angles of elevation in the median plane (Blauert, 1983; Middlebrooks, 1997; Marentakis, 2006). To further ensure no sound localization/discrimination issues occurred, the north and south cues were coded with additional contextual information (Melara & O'Brian, 1987). That is, a higher pitch tone was used for the north cue (i.e., 200 Hz for one second) and a lower pitch tone was used for the south cue (i.e., 50 Hz for one second).

#### Spatial-Visual information augmented with Spatial-Tactile cue

Figure 5 illustrates the location of the coded north and south vibratory cues on the tactile vest, which were used to indicate when an emergent target appeared to the north or south in the Tacsit window for the Retargeting Task. As with the SA cue, the ST cue occurred within 100ms of an emergent target appearing in the Tacsit window. It should be noted that for this ST cue, and the SA cue condition that was previously described in section 2.3.7, the east and west directions were not necessary to implement given the nature of the design of the TTWCS simulation in this study. As shown in Figure 5, all targets (emergent and regular) appeared to the east in the Tacsit window in all simulation runs.

#### Spatial-Visual information augmented with both Spatial-Auditory cue and Spatial-Tactile cue

This task scenario condition combined the cues described in 2.3.7 and 2.3.8, where the visual emergent targets were simultaneously augmented with both the SA and ST cues within 100ms of when the emergent target appeared in the Tacsit window. To increase sensory facilitation effects in the redundancy of these 2 cues occurring together, the SA and ST cues were designed to start and end within 150ms of each other.

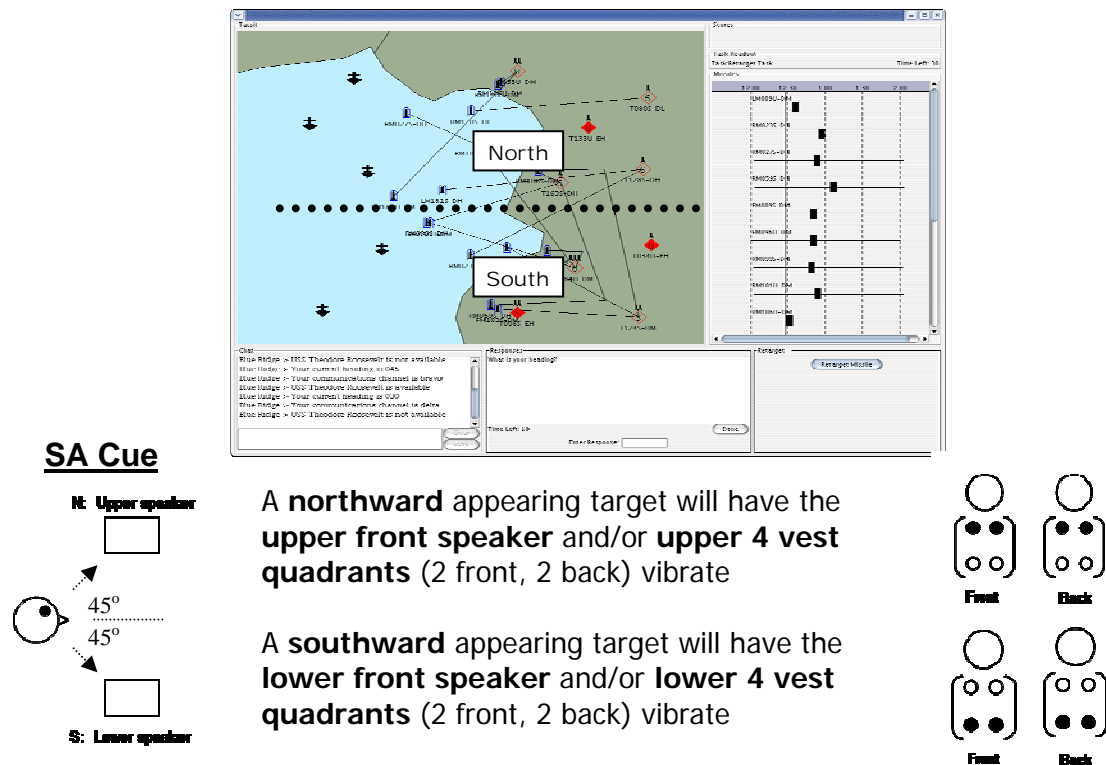


Figure 5 Illustration of implementation strategies for SA and ST cues

### Experimental Design

The study employed a within-subjects repeated measure design to test the effects of the mitigation strategies factor (9 levels). The nine mitigation design conditions included: one baseline (unmitigated; not augmented) and the eight multimodal mitigation (augmented) conditions described in Section 2.3. Based on previous neurophysiological experimental work with the same task environment (Berka et al., 2005), the WM load across each task scenario was considered average to high based on the following task parameter values: 4-6 emergent targets, 6 default targets, 15 missiles, 15 second question interval for Alert Task questions with a 5 second pause between questions).

The dependent variables used to assess objective performance effects for the TTWCS Alert Task questions task included:

- # questions attempted divided by # asked,
- # correct (and incorrect) user responses to questions divided by the # attempted,
- overall average response time per question, and
- average response time for correctly answered questions.

The TTWCS Retarget Task was assessed with an average retarget performance dependent variable based on user's score out of a possible optimum retargeting (expert) score for the particular scenario. Subjective performance dependent variables included workload ratings for both the Retargeting and Alert Tasks via the Modified Cooper-Harper (CH) Scale (Cooper & Harper, 1969; Wierwille & Casali, 1983). The Modified CH uses a 10 point Likert-type scale ranging from 0 (completely undemanding; very relaxed and comfortable; i.e., chewing gum) to 9 (completely demanding; i.e., time-pressured physics exam) to assess participant's perceived mental demand level for each task scenario. A repeated measures ANOVA was used to compare the various dependent variables to the mitigation design treatment factor (modal augmentation cueing strategy), with post-hoc comparisons performed using the Bonferroni adjustment for multiple comparisons when significance was found for treatment factors.

### *Procedure*

Before the start of the experiment, participants completed an informed consent, demographics and other questionnaires. Participants were then assigned to a particular experimental condition based on a randomized order of multimodal augmentation strategies (treatment), with each mitigation strategy testing condition being performed twice within the random ordering. Regardless of mitigation task scenario order, every participant performed training and testing on the baseline condition (i.e., no augmentations) first. Every training and testing condition

consisted of 2-90sec trials for a total of three minutes. Participants read written task instructions and then completed as many training sessions as necessary on the baseline task to ensure they were at least at an 80% performance level before beginning any actual testing scenarios. Participants' scores were presented to them on the computer monitor at the end of each and every task session, whether training or testing. Baseline task training and assessment took approximately 45 minutes.

Before beginning any multimodal augmentation task conditions, participants were given a brief demo and training on the audio and tactile technologies and associated cues they would be interacting with during the remaining task conditions. Participants did not move on to the actual simulation training and testing modal augmentation scenarios until they could accurately identify during the general demo and training what each modal cue represented (i.e., the cues to attend to an emergent target and where; the cue to answer an Alert Task question). No participant required longer than 10 min to learn all the cues during this general training and demo session, providing some evidence of their intuitiveness. Before each multimodal augmentation task condition, participants read a brief written overview of the multimodal cues they would be receiving before they completed the respective 3 minute training scenario and then the immediately following 3 minute test condition. To avoid practice and learning confounds in later statistical analyses, each person was also retested on the baseline condition after all randomized modal augmentation conditions had been completed. It took approximately 2 hours for each participant to complete all testing task scenarios.

At the end of each test session (baseline and augmented), users completed a Modified CH questionnaire to rate their perceived mental demands of the Retargeting and Alert Tasks for that particular test condition. Participants were paid cash immediately upon completing the study.

## Results

### *Objective Performance*

The mean and standard deviations of all performance variables for both the Alert Task and Retarget Task are given for each treatment condition (mitigation/augmentation strategy) in Tables 7 and 8, respectively.

Table 7 Alert Task Performance for Each Treatment Condition (mitigation strategy)

Treatment	Attempted Over Asked			Correct Over Attempted			Overall Avg. Response Time			Correct Avg. Response Time		
	Mean	SD	% Imprv.	Mean	SD	% Imprv.	Mean	SD	% Imprv.	Mean	SD	% Imprv.
Baseline	0.934	0.062		0.815	0.136		7.415	1.369		6.832	1.160	
VV w/ VA	0.973	0.057	4.2	0.869	0.136	6.6	6.742	1.459	9.1	6.402	1.329	6.3
VV w/ VT	0.967	0.057	3.5	0.863	0.124	5.9	6.871	1.527	7.3	6.65	1.403	2.7
VV w/ VA & VT	0.972	0.045	4.1	0.865	0.113	6.2	6.747	1.284	9.0	6.373	1.216	6.7
VV w/ VA (S)	0.979	0.034	4.8	0.862	0.141	5.8	6.45	1.211	13.0	6.18	1.182	9.5
VV w/ VA (S) & VT	0.985	0.028	<b>5.5</b>	0.876	0.136	<b>7.5</b>	6.368	1.290	<b>14.1</b>	6.157	1.290	<b>9.9</b>
SV w/ SA	0.963	0.051	3.1	0.85	0.13	4.3	7.009	1.499	5.5	6.599	1.324	3.4
SV w/ ST	0.966	0.068	3.4	0.848	0.141	4.1	6.802	1.341	8.3	6.38	1.199	6.6
SV w/ SA & ST	0.966	0.051	3.4	0.85	0.136	4.3	6.668	1.375	10.1	6.193	1.194	9.4

Notes: SD = standard deviation. **Bold** indicates the mitigation strategy with the greatest % improvement in performance over baseline per dependent variable.

Table 8 Retarget Task Performance for Each Treatment Condition (mitigation strategy)

Measure	Average Retarget Performance	
TREATMENT	Mean	S.D.
Baseline	0.665	0.068
VV w/ VA	0.678	0.068
VV w/ VT	0.679	0.085
VV w/ VA & VT	0.671	0.085
VV w/ VA (S)	0.684	0.079
VV w/ VA (S) & VT	0.671	0.079
SV w/ SA	0.682	0.085
SV w/ ST	0.66	0.096
SV w/ SA & ST	0.674	0.091

The repeated measures multivariate ANOVA results and Bonferroni-adjusted pairwise comparisons allowed comparisons of the Alert Task and Retarget Task performance for each mitigation strategy condition as compared to baseline (no mitigation), as well as between each mitigation strategy. The mitigation strategy treatment effect was significant ( $F(40, 926) = 2.607$ ,  $p < .000$ ), and Table 9 provides the repeated measures pairwise comparisons for each mitigation strategy to baseline. Except for the SV w/ SA strategy, there were significant improvements in performance for each mitigation strategy as compared to baseline.

As shown in Table 9 and in support of the first four hypothesis, each of the verbal cueing strategies resulted in significantly faster response times (i.e., ‘overall average response time’ dependent variable) with respect to baseline for the Alert Task performance. Regarding accuracy (i.e., the ‘correct over attempted’ dependent variable), only the redundant speech verbal cueing strategy failed to show significant improvement over baseline, although this strategy did show significant improvement with regards to the overall number of questions attempted for the Alert Task. Regarding the last three hypotheses (i.e., the spatial cueing mitigation strategies) in terms of overall average response time for the Alert Task, only the SV w/ SA mitigation strategy failed to show significant improvement over baseline. Thus Hypotheses H6 and H7 were supported but not H5. In terms of accuracy on the Alert Task, none of the spatial mitigation strategies significantly improved performance over baseline. This is most likely due to the spatial cueing strategies being designed to improve emergent target location and reaction time for spatial components of the Retargeting Task and thus overall multitasking performance, and not designed to directly improve accuracy on the verbal components of the Alert task. Thus, as indicated by SV w/ ST and SV w/ SA & ST average response time results for the Alert Task, the spatial mitigation strategies were effective in improving overall multitasking performance.

No mitigation strategy showed significance with respect to baseline or the other mitigation strategies for the primary Retarget Task--indicating Retarget Task performance did not improve significantly but, more importantly, that it did not deteriorate significantly either. Thus, performance was maintained on the primary task, while the user was able to attend to more of the secondary Alert Task questions and correctly answer them. This suggests improved task switching and overall interruption management capabilities were facilitated by implementing bimodal and multimodal verbal cueing strategies for the predominantly verbal Alert Task.

The present results could not confirm parts of H4 and H7 regarding combined modal cueing strategies significantly improving performance over their respective individual cueing strategies. However, future studies looking at increased workload levels between spatial and verbal tasks within multitasking conditions, as well as individual difference factors (e.g., user's spatial/verbal WM capabilities), may provide additional support for combining modal cueing strategies to significantly improve response time and/or accuracy over single modal cueing strategies.

Table 9 Performance Pairwise Comparisons between Baseline and Each Mitigation Strategy

Measure	Treatment	Mean Difference (BL-Treatment)	P-value
Attempted Over Asked	VV w/ VA	-.039(*)	.001
	VV w/ VT	-.033	.379
	VV w/ VA & VT	-.038(*)	.037
	VV w/ VA (S)	-.044(*)	.006
	VV w/ VA (S) & VT	-.051(*)	.001
	SV w/ SA	-.029	.439
	SV w/ ST	-.032	.311
	SV w/ SA & ST	-.032	.124
Correct Over Attempted	VV w/ VA	-.055(*)	.001
	VV w/ VT	-.048(*)	.000
	VV w/ VA & VT	-.050(*)	.042
	VV w/ VA (S)	-.047	.074
	VV w/ VA (S) & VT	-.061(*)	.003
	SV w/ SA	-.035	.312
	SV w/ ST	-.033	1.000
	SV w/ SA & ST	-.035	.822
Overall Average Response Time	VV w/ VA	.673(*)	.000
	VV w/ VT	.544(*)	.031
	VV w/ VA & VT	.668(*)	.004
	VV w/ VA (S)	.965(*)	.000
	VV w/ VA (S) & VT	1.047(*)	.000
	SV w/ SA	.405	.530
	SV w/ ST	.612(*)	.002
	SV w/ SA & ST	.746(*)	.003
Correct Average Response Time	VV w/ VA	.430	.134
	VV w/ VT	.182	1.000
	VV w/ VA & VT	.458	.282
	VV w/ VA (S)	.652(*)	.004
	VV w/ VA (S) & VT	.675(*)	.008
	SV w/ SA	.233	1.000
	SV w/ ST	.452	.217
	SV w/ SA & ST	.639	.035
Average Retarget Performance	VV w/ VA	-.013	1.000
	VV w/ VT	-.014	1.000
	VV w/ VA & VT	-.006	1.000
	VV w/ VA (S)	-.019	1.000
	VV w/ VA (S) & VT	-.006	1.000
	SV w/ SA	-.017	1.000
	SV w/ ST	.005	1.000
	SV w/ SA & ST	-.008	1.000

\* The mean difference is significant at the .05 level.

Note: BL = baseline.

### *Subjective Performance (Workload Ratings)*

Spearman correlation coefficient was used to compare the participant's subjective workload ratings (i.e., their perceived mental demand based on the CH ratings) for both the Alert Task and Retarget Task for each mitigation strategy as compared to the baseline condition (Table 10). When comparing perceived mental demand for each mitigation strategy as compared to baseline, high and significant correlations were found for both the Alert Task and Retarget Task subjective workload ratings for all mitigations. Furthermore, the Friedman non-parametric test determined that the workload ratings for each mitigation were not significantly different from baseline ratings for either the Alert Task ( $X^2 = 10.778$ ,  $p = 0.215$ ) or the Retargeting Task ( $X^2 = 6.621$ ,  $p = 0.578$ ). This suggests participants did not perceive the multimodally-enhanced TTWCS task environment as any more mentally demanding than the baseline TTWCS environment.

Table 10 CH Subjective Workload Rating Correlations—Perceived Mental Demand for Each Mitigation Strategy as Compared to Baseline

	Perceived Mental Demand for the Alert Task		Perceived Mental Demand for the Retarget Task	
	Spearman Corr.	P-value	Spearman Corr.	P-value
VV w/ VA	0.791	0.000	0.806	0.000
VV w/ VT	0.876	0.000	0.756	0.000
VV w/ VA & VT	0.788	0.000	0.683	0.000
VV w/ VA (S)	0.681	0.000	0.537	0.002
VV w/ VA (S) & VT	0.812	0.000	0.702	0.000
SV w/ SA	0.726	0.000	0.747	0.000
SV w/ ST	0.781	0.000	0.788	0.000
SV w/ SA & ST	0.885	0.000	0.657	0.000

### Discussion

The overall objective performance results of this study support all hypotheses (except for the SV w/ SA strategy; hypothesis H5), regarding expected performance benefits of individual and

combined modal cueing strategies when compared to baseline (no cueing strategies). Additionally, subjective performance results revealed that users did not perceive the Alert or Retargeting Task during any mitigation strategy condition to be more mentally demanding than the baseline (unmitigated) condition. These findings illustrate the effectiveness of implementing multimodal cueing (augmentation) strategies to improve the interruption management capabilities of users in computer-based multitasking environments.

The use of verbal-auditory and verbal-tactile cues and combined verbal auditory with tactile cues enabled users to respond more quickly (from 7.3% to 14.1% quicker) to interruptions (the Alert Task) than users in the uncued (baseline) condition. Verbally cued participants also produced a significantly greater number of correct answers (from 5.8% to 7.5% greater) and thus fewer errors on the interrupting Alert Task. Although neither spatial or verbal cueing strategies significantly improved performance on the Retargeting (interrupted) Task, the fact that performance was maintained on the primary task for all mitigation strategy conditions when secondary Alert (interrupting) Task questions were being answered more quickly and with less errors is indication of successful interruption management strategies being facilitated and is in line with similar implications reported in recent interruption management and bimodal cueing studies (Latorella 1996, 1999; Sklar & Sarter, 1999; van Erp & van Veen, 2004; Hopp et al., 2005; Hopp-Levine et al, 2006). Further, that the auditory and tactile verbal cues did not interfere with performance on the visual primary Retargeting Task is in line with multiple resource theory (MRT) (Wickens, 1984) expectations, as cueing with free modal resources appears to have sufficiently directed attention to critical, pending Alert Task questions and did not additionally load format/modal resources being used for the ongoing Retargeting Task.

While performance on the primary task was maintained, performance on the secondary task was improved with single and combined spatial cueing strategies. Specifically, adding ST or the combined SA and ST cues to the primary Retargeting task facilitated significant performance improvements in response time on the secondary verbal Alert Task. The fact that the spatial mitigation strategies did not also show significance for improving Retargeting Task performance is not surprising because the spatial cueing strategies were aimed at improving a user's ability to locate emergent targets and thus overall performance in the Retargeting Task, and not at determining exactly how long it took a user to respond to an emergent target once it appeared on screen and was cued. Future studies may examine such response time effects for the spatial mitigation strategies in alternate multitasking environments.

Taken together, the present results suggest the cognitive mechanisms by which the multimodal cueing strategies investigated in this study improve the effectiveness and efficiency of users' attentional and WM processing, task switching capabilities, and thus their overall interruption management and task performance (Hopp et al., 2005; Hopp-Levine, et al., 2006). Similar to Ho et al.'s (2004) findings, for instance, the present study's findings suggest the auditory tonal and vibrotactile cueing strategies may have allowed users to perform 'negotiated interruption' techniques (i.e., when a user has some control over when to attend to and complete a task) (Latorella, 1999; McFarlane & Latorella, 2002; McFarlane, 2002) and avoid unintentional dismissals of the Alert Task questions. That is, users were trained to know that if they were busy with the Retargeting Task they could wait for the cue to occur at the '8 seconds remaining' mark to remind them to answer a pending Alert Task question before it disappeared from the computer screen. This is in line with other memory research (Kliegel, Martin, McDaniel, & Einstein, 2002) indicating that tonal and vibrotactile cueing strategies tend to transform user's task-

switching processing from a more memory-intensive, time-based task to a more resource-efficient, event-based task, where users can rely on cues to know when they more urgently need to shift attention between tasks (e.g., from the Retargeting Task to the Alert Task in this study). Hopp et al. (2005) made similar speculations based on the results of their tactile cueing study. The next sections address further research findings and implications for each mitigation strategy investigated in this study.

#### *Augmenting VV interruption task information with a VA (tonal) cue*

In support of H1, this cueing strategy significantly improved over the baseline (unmitigated/unaugmented) condition the overall average time to respond to Alert Task questions by approximately (~) 9% (s.d. = 1.46). The proportion of correctly answered questions improved by approximately ~ 7% (s.d. = 0.14) and the time to answer them by > 6 % (s.d. = 1.33). The performance improvement results from this VA tonal cueing strategy support existing research that auditory tonal cues may be used: for rapid cueing of critical information, such as alerts (Sanders & McCormick, 1993; Wickens & Hollands, 2000); as effective warnings for time-relevant events (Welch & Warren, 1986; Ho & Spence, 2005), and for attention directing in multitasking situations to facilitate improved resource allocation and thus overall interruption management capabilities (Latorella, 1996; 1999).

#### *Augmenting VV interruption task information with a redundant VA speech cue*

In support of H2, this cueing strategy significantly improved over the baseline condition the proportion of Alert Task questions attempted by > 5% (s.d. = 0.03) and the overall average time to respond to them by ~ 13% (s.d. = 1.21). These performance gains over baseline were expected, particularly with regards to response time improvements, because the redundant

auditory (speech) was presented simultaneously when the visual Alert Task question appeared on screen. Existing empirical research reveals that providing visual-auditory redundancy facilitates faster and more accurate responses than visual only presentation in visually loaded multitasking environments (e.g., C<sup>2</sup> type tasks; demanding driving simulation tasks) (Spence & Read, 2003; Wickens & Gosney, 2003). Similar performance improvements have been found when redundantly using visual and auditory information presentation for verbal task information to improve user response times to critical or high priority alerts and warnings (Belz, Robinson, & Casali, 1999; Wickens & Hollands, 2000; Ho et al., 2004). The performance improvement effects in the present study may be attributed to MRT and the premise that users in the redundant modality condition were able to process information with both visual and auditory attention and WM resources and thus avoid overloading either modality's resources (Mayer & Moreno, 1998; Moreno & Mayer 2002) when trying to manage interruptions (Alert Task questions) and maintain performance on the Retargeting Task.

#### *Augmenting VV interruption task information with a VT cue*

In support of H3, this cueing strategy significantly improved over the baseline condition the overall average time to respond to Alert Task questions by > 7% (s.d. = 1.52). The proportion of correctly answered questions improved by ~ 6% (s.d. = 0.12) and the time to answer them by ~ 3% (s.d. = 1.40) (not significant). The performance improvement results from this VT cueing strategy support existing research that vibrotactile cues can be effectively used in visually busy C<sup>2</sup> type multitasking environments to direct attention to important interruption tasks as needed, enabling users to more efficiently allocate information processing resources to the ongoing task and reduce user reaction time and dependence on the visual modality (Hopp et al., 2005; Ho et

al., 2005; Hopp-Levine et al., 2006). Such performance enhancing effects may be attributed to fewer attentional switching costs, given the complementarities between the visual and tactile senses for effectively conveying verbal task information, (e.g., alerts, warnings) (Stanney et al, 2004).

*Augmenting VV interruption task information with VA (tonal or redundant speech) and VT cues*

In support of H4, both of these multimodal cueing strategies (VV w/ VA & VT; VV w/ VA (S) & VT) significantly improved performance as compared to the baseline condition. The VV w/ VA & VT strategy improved the proportion of Alert Task questions attempted by > 4% (s.d. = 0.05) and the overall average time to respond to them by 9% (s.d. = 1.28); the proportion of correctly answered questions improved by > 6% (s.d. = 0.11) and the time to answer them by ~ 7 % (s.d. = 1.22) (not significant). The VV w/ VA (S) & VT mitigation strategy significantly improved the proportion of Alert Task questions attempted by ~ 6% (s.d. = 0.03) and the overall average time to respond to them by ~ 14% (s.d. = 1.29); the proportion of correctly answered questions improved by ~ 8% (s.d. = 0.14) and the time to answer them by ~ 10 % (s.d. = 1.29). These combined modal strategies, however, did not support H4 regarding significantly improving performance over either the individual VA (tonal or redundant speech) or VT strategy. This finding is in contrast to what was expected based on recent evidence from neuroimaging-based studies demonstrating improved response times based on increases in sensory facilitation and potential cross-modal coactivation effects (Stein & Meredith, 1993; Eimer et al., 2001; Kennet et al., 2001; Dyson & Quinlan, 2002; Calvert et al., 2004; Spence & Driver, 2004). It is also contrary to the expected resource allocation benefits of using the tactile modality to offload visual, linguistic-visual, and linguistic-auditory modalities, which may utilize

the same WM resources (Sulzen, 2001). This lack of significance may be due to the strength of the unimodal cues being sufficiently strong enough (i.e., met appropriate threshold levels) to direct users attention to the cued information, with the additional cue modality not providing any significant added benefits (c.f. Stein & Meredith, 1993; Calvert et al., 2004). Future empirical studies should examine potential means of achieving effective augmentation with combined cueing strategies when the unimodal cueing strategies alone are not sufficient to significantly improve performance. Of particular interest will be how varied task workload levels and individual differences in users' verbal/spatial WM capabilities may affect the significance in performance improvements when comparing mitigation strategies.

#### *Augmenting SV interruption task information with a SA cue*

Although H5 could not be confirmed regarding significance levels for any of the dependent variables, this SV w/ SA mitigation strategy improved over the baseline condition the proportion of Alert Task questions attempted by > 3% (s.d. = 0.05) and the overall average time to respond to them by ~ 6% (s.d. = 1.50); the proportion of correctly answered questions improved by > 4% (s.d. = 0.13) and the time to answer them by > 3 % (s.d. = 1.32) (not significant). The lack of significance was unexpected in that existing SA cueing strategy research has shown that spatial auditory cues may be used with visual target detection tasks (e.g., cockpit applications; general C<sup>2</sup> tasks) to decrease general visual search times and improve traffic detection and avoidance via reduced visual workload (Begault, 1993; McKinley & Ericson, 1997; Bolia et al., 1999; Ho et al., 2004; Ho & Spence, 2005; Ho et al., 2005). As discussed previously, the limitation regarding the Retargeting Task performance assessment capabilities due to how the TTWCS simulation environment was programmed may have limited the amount of significant findings for all the

spatial cueing strategies. Future research will implement more sensitive measures for assessing response time and accuracy performance effects when spatial cueing strategies are implemented.

#### *Augmenting SV information with a spatial-tactile (ST) cue*

In support of H6, the SV w/ ST cueing strategy significantly improved over the baseline condition the overall average time to respond to Alert Task questions by > 8% (s.d. = 1.34). This performance improvement result supports existing research evidence that spatial vibrotactile cues can be effectively used in visually cluttered C<sup>2</sup> type multitasking environments to direct attention to important interrupting task information as needed and thus reduce a user's reaction time and dependence on the visual modality (ETSI, 2002; Ho et al., 2004; Hopp et al., 2005; Hopp-Levine et al., 2006). The VT performance enhancing effects may be attributed to fewer attentional switching costs, given the complementarities between the visual and tactile senses for effectively conveying spatial task information (e.g., spatial orientation, object identification) (Stanney et al., 2004). The implication from the present findings is that the reduced visual workload and WM resources that the ST cue provided in the Retargeting Task could have provided free resources for the user to allocate to performance on the Alert Task, thereby potentially leading to enhanced interruption management capabilities on the overall TTWCS task.

#### *Augmenting SV interruption task information with SA and ST cues*

In support of H7, the SV w/ SA & ST mitigation strategy significantly improved the overall average time to respond to Alert Task questions by > 10% (s.d. = 1.38). However, as with the verbal combined cueing strategy results, this combined spatial multimodal mitigation strategy did not realize significantly greater performance gains than the individual SA or ST mitigation strategy and is in contrast to what was expected. Thus, the lack of significance for the combined

spatial strategies may also have been due to the strength of the unimodal spatial cues being sufficiently strong enough to direct users attention to the cued information, with the additional cue modality not providing any significant added benefits (c.f. Stein & Meredith, 1993; Calvert et al., 2004). However, more sensitive measures to assess the response time effects of the spatial cueing strategies on detection of the spatial information are needed to make such inferences. Future empirical studies will employ such measures, as well as investigate how varied task workload levels and individual differences in users' verbal/spatial WM capabilities may affect the significance in performance improvements when comparing mitigation strategies.

### Conclusions

This study provides a source of empirical validation of the *same information format/different modality* bimodal and multimodal theoretical design guidelines presented in Table 6. The results of this study both support and extend existing research evidence that demonstrates the effectiveness of implementing multimodal cueing (augmentation) strategies to improve the interruption management capabilities of users in operationally-relevant, computer-based multitasking environments. The overall implication of this study is that the following design guidelines, which were herein empirically validated, can enhance human-system performance during multi-tasking:

- When passing information (ongoing and/or new) in the same format/modality loads WM in a visually busy task environment, offload with mitigation cueing strategies in alternative modalities.

- Offload WM and improve response time and accuracy by using free modal resources to present augmentation cues in the *same information format but different modality*:
  - Augment verbal-visual information with verbal-auditory tonal cues and/or verbal-tactile cues to remind users of a pending critical verbal task needing attention.
  - Augment verbal-visual information with verbal-auditory redundant speech cues to facilitate efficient distribution of attention and WM resources.
  - Augment verbal-visual information with verbal-auditory redundant speech cues to improve attention and WM resource allocation, and verbal-tactile cues to remind users of pending critical verbal tasks needing attention.
  - Augment spatial-visual information with spatial tactile cues to facilitate target detection and reduce users' reaction time and dependence on the visual modality.
  - Augment spatial-visual information with both spatial-auditory and spatial tactile cues to realize significantly improved performance gains as compared to uncued conditions.

These guidelines may be applicable to a wide range of information-intensive computer-based task environments. Future studies should focus on: examining performance effects and implications when both spatial and verbal cueing strategies are implemented in a single mitigation condition; investigating and validating the remaining theoretical guidelines in Table 6 (e.g., the *different format/same modality* and *different format/different modality* modal mitigation strategies; individual difference effects on mitigation strategy selection/implementation), and; determining how to implement effectively designed modal mitigation strategies in similar  $C^2$  task environments based on real-time performance monitoring (e.g., determining when users have not yet attended to a critical task and implementing attention-directing cues) and real-time cognitive

state assessment (e.g., using physio- and neurophysiological sensors to monitor and assess cognitive workload and implementing cueing strategies based on this real-time assessment).

#### Acknowledgment

This work was made possible with the support of an Office of Naval Research (ONR) Ph.D. fellowship grant under the leadership of CDR Dylan Schmorow. The authors wish to thank ONR and CDR Schmorow for their support. The authors would also like to thank Polly Tremoulet, Helen Hastie, Michael Thomas, and particularly Nick Morizio of Lockheed-Martin Advanced Technologies Laboratory (LMATL) for their testbed development support in conjunction with DARPA's Augmented Cognition Phase 2 and 3 efforts. The views expressed in this article are those of the authors and do not necessarily reflect those of their affiliated organizations or those of LMATL, DARPA or ONR.

## CHAPTER FIVE: GENERAL DISCUSSION

Today's 21<sup>st</sup> Century human-computer systems are increasingly able to provide humans with more information than can effectively and efficiently be processed via single modalities or information format codes. Thus, multimodal display systems that augment display modalities to maximize user's information processing and interruption management capabilities in information-intensive, multitasking computer-based environments hold great promise (Miyake & Shah, 1999; Calvert et al., 2004; Stanney et al., 2004; Oviatt et al., 2005; Reeves & Stanney, 2007). Unfortunately, a lack of principle-driven multimodal design guidelines regarding how to choose the most appropriate display modalities and information formats for given users and applications could prevent multimodal systems from realizing their true potential.

To address this issue, Chapter Three introduces the Architecture for Multimodal Optimization (AMMO) model, the utility of which lies in its ability to extend existing bimodal S-C mapping guidance and provide principle-driven strategies for directing multimodal system design and dynamic adaptation strategies in support of real-time operations. Figure 3 of Chapter Three depicts a conceptual model that illustrates how particular components of AMMO and guidelines derived from it can be integrated with an interruption management framework (adapted from Latorella 1996, 1999) to aid in controlling for and optimizing a user's HIP workload levels via the system's presentation of information (e.g. via modality selection and timing rules/constraints and/or delegation strategies) at various stages of an interruption (i.e., detection, interpretation, integration, resumption). Based on this integration framework a set of theorized adaptive design strategies and general implementation guidelines is proposed (see Chapter Three, Table 5), which may be used by human-systems interaction designers as general

guidance for when (e.g., at what stages of an ongoing procedure or interruption) and how (e.g., combinations and orderings) to implement adaptive strategies and for designing future empirical studies for assessing how certain implementation strategies may affect a user's HIP and task performance for a given operational domain. The current empirical work focused on evaluating, in a simulated weapons control system multitasking environment, a subset of AMMO's guidelines, those associated with the *same format/different modality* bimodal and multimodal *cueing* augmentation strategies (i.e., augmenting verbal-visual information with verbal-auditory and/or verbal-tactile cues; augmenting spatial-visual with spatial-auditory and/or spatial-tactile). Additional empirical study is needed to validate the other AMMO components and the theorized guidelines presented in Table 5 of Chapter Three.

As the AMMO model and implementation guidelines suggest, when continuing to pass information (ongoing and/or new) in same format/modality loads WM, the detection of ongoing and new information may be enhanced via cueing modal mitigation (augmentation) strategies to direct a user's attention to pertinent task information. Using free modal resources (as predicted or assessed in real time) to present augmentation cues in the same information format but different modality can facilitate congruent information format-to-modality mappings, while not unnecessarily conflicting with or additionally overloading format/modal resources being used for the ongoing or new task (Latorella, 1996, 1999; Wickens & Hollands, 2000). If a user continues to miss cues or performance is not at acceptable levels after augmenting with a single mitigation, then augmenting with an additional modal mitigation (e.g., SV with both SA & ST cues) may increase sensory facilitation effects of the cue and subsequent detection of task information (Spence & Driver, 2004; Calvert et al., 2004).

As addressed in Chapter Four, the present results show significant improvements in response time and accuracy performance when *same format/different modality* individual (bimodal condition) or combined (multimodal condition) cueing strategies are implemented in a simulated TTWCS multitasking environment. The enhanced human performance benefits seen here support the supposition that such cueing strategies may be used as operational interruption management strategies applicable to a wide range of information-intensive, computer-based multitasking environments. The verbal-visual with verbal-auditory (tonal) bimodal cueing strategy results support previous findings (Welch & Warren, 1986; Sanders & McCormick, 1993; Latorella, 1996; 1999; Spence & Driver, 2000; Wickens & Hollands, 2000; Ho & Spence, 2005) that auditory cueing strategies may be used: for rapid cueing of critical information, such as alerts, as effective warnings for time-relevant events, and for attention directing in multitasking situations to facilitate improved resource allocation, response time and accuracy, and thus overall interruption management capabilities in visually loaded multitasking environments (e.g., C<sup>2</sup> type tasks). The results of the verbal-visual with verbal-auditory (redundant speech) strategy in the present study also support existing research regarding improving response time, but this strategy did not *significantly* improve accuracy too as expected from findings in other studies investigating visual-auditory redundancy (for a review of such studies, see Wickens & Seppelt, 2002). The present results may be due to issues with users not always being able to clearly interpret the synthetic speech voice, which many participants noted as an issue at the end of the study in their free response comments. Another issue could be that the present study implemented redundant speech to present the Alert Task question to improve its detection, instead of redundantly presenting the actual task information itself to improve its interpretation and integration when presented in the visual chat window. Most all the visual-

auditory redundancy studies summarized by Wickens and Seppelt (2002) were more inline with the latter strategy. Future studies examining modal mitigation strategies for improving interpretation and integration of interrupting information, as well as detection, should consider such information presentation issues when designing strategies with modal redundancy.

The results of the spatial auditory cueing strategy also did not support existing findings that spatial auditory cues may be used as effective design strategies for significantly improving performance in visual target detection tasks (e.g., cockpit applications; general  $C^2$  tasks) by decreasing general visual search times and improving traffic detection and avoidance via reduced visual workload (Begault, 1993; McKinley & Ericson, 1997; Bolia et al., 1999; Ho et al., 2004; Ho et al., 2005). The present results are likely due to the TTWCS simulation tested environment not having sensitive enough measures for assessing Retargeting Task performance effects like the previous spatial-auditory cueing research studies did. These previous studies employed specific dependent measures to assess how long it took users to locate targets as soon as they appeared and after a spatial-auditory cueing strategy was implemented. The present study, on the other hand, was designed to assess overall Retargeting Task performance and not specifically examine how fast or accurately a user responded to the spatial-auditory cue. Future studies may employ such dependent measures as an additional assessment technique regarding the effectiveness of the spatial-auditory cues.

Previous vibrotactile study findings (Hopp et al., 2005; Ho et al., 2005; Hopp-Levine et al., 2006) are supported by both the verbal-visual with verbal tactile and spatial-visual with spatial-tactile bimodal cueing strategy results. As with the verbal-auditory cues, the verbal-tactile cues proved effective in the visually busy  $C^2$  type multitasking environment to direct

attention to important interruption tasks as needed, enabling users to more efficiently allocate information processing resources to the ongoing task and reduce user reaction time and dependence on the visual modality. As with the limitations in assessing the specific response time and accuracy improvements in target detection/localization due to the spatial-auditory cue, the spatial-tactile cueing strategy may not have employed sensitive enough dependent measures for direct assessment. However, when implementing the spatial-tactile cueing strategies, performance was maintained on the primary Retargeting Task while users were able to more quickly respond to Alert Task questions. The implication is that the reduced visual workload and WM resources that the spatial-tactile cue provided in the Retargeting Task could have provided free resources for the user to allocate to performance on the Alert Task, thereby leading to enhanced interruption management capabilities on the overall TTWCS task.

The results of the tactile-enabled *multimodal* cueing strategies (i.e., verbal-visual w/ verbal-auditory *and* verbal –tactile; spatial-visual with spatial-auditory *and* spatial-tactile) extend the findings from existing behavioral bimodal cueing studies and provide support for the implication from recent neurological evidence that multimodal cueing strategies may aid in distributing a user’s cortical processing requirements across multiple, modally-designated resource areas in WM (Stanney et al., 2004; Reeves & Stanney, 2007). For instance, adding the spatial-auditory cue alone increased performance overall, but it did not significantly increase response time performance over the uncued baseline condition until the tactile modality was added. A similar result was seen when the verbal-auditory (redundant speech) was implemented alone and then when the verbal-tactile cueing strategy was added to it. The redundant speech cue did not significantly improve over baseline the accuracy on the Alert Task—only response time significantly improved. However, both response time *and* accuracy were significantly

improved over baseline when the verbal-tactile cue was added to the redundant speech cue, further indicating the effectiveness of implementing the tactile modality to aid in distributing a user's cortical processing requirements across more than one modality. These findings extend the existing theories to date (see Chapter Three, Table 1 and ETSI, 2002) regarding the suitability of the tactile modality as similar to or better than the auditory modality for displaying task information relying on fast reaction times and memorability in information-intensive, multitasking environments.

The results of the present study did not meet expectations regarding the combined multimodal cueing strategies (i.e., verbal-visual w/ verbal-auditory *and* verbal –tactile; spatial-visual with spatial-auditory *and* spatial-tactile) significantly improving performance time over their respective individual auditory or tactile cueing strategies. Recent evidence from neuroimaging-based studies demonstrate improved response times based on increases in sensory facilitation and potential cross-modal coactivation effects (Stein & Meredith, 1993; Eimer et al., 2001; Kennet et al., 2001; Dyson & Quinlan, 2002; Calvert et al., 2004; Spence & Driver, 2004). Although the combined strategies did generally improve performance over the individual strategies (see Chapter Four, Table 7), the bimodal and multimodal strategy results may not have been significantly different from each other because the stimulus intensity of the individual cueing strategies were sufficient to meet necessary threshold levels to facilitate detection. To see the potential additive coactivation effects (i.e., from the linear neural summation of redundant modal stimuli sufficiently overlapped in time and space) reported by Corballis et al. (2002) and others (Miller, 1982, 1986; Roser & Corballis, 2002; Savazzi & Marzi, 2002; Iacoboni & Zaidel, 2003) or the potential multiplicative effects as noted by Calvert and Lewis (2004) (i.e., up to 12x faster beyond that expected from summing impulses from unimodal stimuli), the unimodal

stimuli would need to have the least effective sensory facilitation when presented alone (Stein & Meredith, 1993). As the present study was not a basic research effort but rather an applied effort that would aim to enhance performance with the lowest cost solution, individual modalities were each presented in the most effective manner possible. Thus, this criterion of “least effective sensory facilitation” was not met in the present study, as indicated by the individual modal strategies’ significance in improving performance. Future empirical studies should examine potential means of achieving effective augmentation with combined cueing strategies when the unimodal cueing strategies alone are not sufficient to significantly improve performance. Of particular interest will be how varied task workload levels and individual differences in users’ verbal/spatial WM capabilities may affect the significance in performance improvements when comparing future individual and combined mitigation strategies.

Both the bimodal and multimodal cueing strategy results, which showed significant improvements over the baseline multitasking performance, provide a source of validation for Component C and Sub-Component C<sub>1</sub> of AMMO and the adaptive design strategies derived from them (see Chapter Three, Table 5; Chapter Four, Table 6). These results extend the existing interruption management framework of Latorella (1996; McFarlane & Latorella, 2002) by providing principle-driven design guidance regarding how to effectively interrupt users and improve detection, interpretation and integration of this interrupting information. Particularly, the results suggest that the cognitive mechanisms underlying the *same format/different modality* modal mitigation adaptive strategies can enable significant improvements in the effectiveness and efficiency of users’ attentional and WM processing, task switching capabilities, and thus overall interruption management and task performance in information-intensive, multitasking environments. In addition to examining the effects of varying workload levels on different

mitigation strategies' effectiveness, future studies should investigate additional combinations of the multimodal mitigation strategies derived from AMMO C<sub>1</sub>, including *different format/same modality* and *different format/different modality* implementation guidelines.

### Effects of Individual WM Capabilities

The information format-to-modality mappings in AMMO's Component A (See Chapter Two, Section 2.1) were used to develop the principle-driven framework for AMMO's Modal Mitigation Strategies Sub-Component (C<sub>1</sub>), which was used to derive the (now partially validated) modal mitigation strategy implementation guidelines in Chapter Three's Table 5. While the derived mappings and associated guidelines are theoretically well supported as discussed in Chapter Three, their generalizability is likely to be mediated by the individual user receiving the information. Although not addressed in Chapter Four's empirical study but discussed in Chapter Three's Section 2.1 and represented by AMMO's Component B, a user's individual capabilities and limitations in spatial and verbal WM processing may be significant factors in determining how a person performs in an information-intensive, time-critical multitask environments, such as a military C<sup>2</sup> system (Miyake & Shah, 1999; Gonzalez, 2005; Stanney et al., 2004; Hale et al., 2005, 2006; Lathan & Tracy, 2002; Reeves et al., 2005; Kane & Engle, 2002). For instance, high ability individuals tend to process more information, respond faster, and are better able to focus and sustain attention on performance-relevant information. Thus, once a task information format-to-modality mapping has been decided in AMMO's Component A, Component B suggests propagating the potential effects of a user's spatial/verbal WM capability through the rest of the AMMO C components to aid in tailoring the modal mitigation strategies or other adaptive strategies to meet the individual user's needs. For instance, for low

ability individuals it may be necessary to present less information, slow the schedule and/or pace of information flow, and/or present specific attentional cues or redundant information in a format/modality that is more efficient for the given individual (e.g., for a low spatial individual, augment a visual-spatial map with visual-verbal directions; for a low verbal individual, augment visual-verbal descriptions with a visual-spatial graphic).

The importance of being able to assess and account for such individual difference factors in multimodal system design has further been elucidated during recent conversations with Dr. Robert S. Kennedy (December 2005). Dr. Kennedy discussed historical findings from his involvement with a nine year study on flight simulators and the results of subsequent meta-analyses of various significant factors affecting flight performance. Of particular note is the finding that individual differences had the biggest effects on performance, accounting for ~ 60-65% of the variance. Practice/Training accounted for ~ 20-25% of the variance, and system factors (e.g., equipment) accounted for ~ 15%. Dr. Kennedy also anecdotally noted that knowing what pilot was on board provided him with more indication of expected performance than any other factors being evaluated in a given experiment. Consequently, an attempt was made in the present study to investigate potential individual difference effects. The following ability tests and questionnaire were chosen for their well-established construct validity and based on previous studies investigating potential performance prediction capabilities in multimodal, multitask environments (Gonzalez, 2005; Reeves et al., 2005; Hale et al., 2005, 2006; Reeves & Stanney, 2007): Raven's Standard Progressive Matrices (SPM) Plus version (estimate of  $g_F$ ; assesses spatial ability on various scales and relatively free of cultural bias) (1998, 2000) and Mill Hill Vocabulary (verbal) (Raven, Raven & Court, 1998); ETS Surface Development (spatial visualization) (1976a); ETS Map Planning (spatial scanning) (1976b), and; VVLSR (Visual-

Verbal Learning Style Rating; self-reported learning style preference) (Mayer & Massa, 2003). Participants completed the ability tests and questionnaire in approximately one hour and fifteen minutes and before they began the empirical study described in Chapter Four. Unfortunately, the standard deviations in the sample population were too small based on participant's assessed test scores and questionnaire ratings for the spatial/verbal ability factors to be analyzed any further. To investigate potential correlations between spatial/verbal ability tests and user's performance with various modal mitigation strategies in future studies, population samples should be grouped into between subjects factors according to a priori assessment of spatial and verbal WM capabilities.

While the current study only validated a portion of the AMMO model and its derived guidelines, the framework may be used to guide the design of numerous future empirical studies. It is envisioned that through continued validation and implementation of AMMO and its derived guidelines, conclusions may start to be drawn regarding the generalizability of particular adaptive strategies (i.e., multimodal mitigation, intelligent pacing, and delegation) across multiple information-intensive task domains and for various types of users (e.g., low/high verbal/spatial processors).

## CHAPTER SIX: CONCLUSION AND FUTURE RESEARCH

Empirical results from this study support the use of *same format/different* modality cueing strategies in information-intensive, multitasking environments to improve users' attentional and WM processing resource allocation abilities, task switching capabilities, and thus overall interruption management and task performance. These results provide validation for aspects of Components A and C of the proposed AMMO model. The AMMO model's overall framework and its derived implementation guidelines can be used to guide the design of future empirical studies in various operational and training system environments. These studies should be aimed at validating the overall AMMO model and improving the sensitivity and diagnosticity of its *if/then* logical parameters for varying task and user requirements. Future studies should also focus on investigating potential verbal/spatial WM individual difference effects and how such effects should be integrated and represented in AMMO's Component B, its overall framework, and its derived implementation guidelines.

Future studies should also explore the option of using real-time cognitive monitors (e.g., a physio- or neuro-physiological-based WM index) to direct the AMMO's mitigation strategies. The cognitive monitor could identify periods of spatial and/or verbal WM overload and thus trigger when a system should provide appropriate adaptive aiding (e.g., invoking multimodal cues, switching from verbal or spatial presentation formats, invoking intelligent pacing or delegation strategies to meet operator requirements in real time) (Stanney et al., 2004; Kobus et al., 2006; Schmorow et al., 2005). A real-time performance monitor (e.g., determining when users have not yet attended to a critical task) could also be integrated with the cognitive monitor to determine when and how certain workload conditions (overload or under load) effect task

performance and vice versa. The integrated neurophysiological-based (e.g., EEG, fNIR, physiological sensors) WM index and performance monitor could then be used as a real-time gauge to ensure adaptive strategies are only invoked when necessary (e.g., when performance is not at a required level) and in the proper form (e.g., modal mitigation strategies in the proper format/modality for the given spatial/verbal WM load conditions and user WM abilities). This integrated approach could aid in avoiding potential “costs” associated with unnecessary and inefficient switching between adaptive strategies (e.g., modality switches, changes in user/system control).

Once such approaches are empirically validated with experiments in various applied task settings and with users of varying WM abilities, the utility of AMMO would lie in its ability to provide HSI designers with *both* a priori design strategies and adaptive automation strategies (i.e., multimodal mitigation, intelligent pacing, and delegation) in real-time operational settings, as well as aid in establishing manning requirements once designs are optimized. The ultimate objective is to leverage AMMO to facilitate performance improvements (e.g., improved response time and accuracy) in computer-based multitasking systems via a reduction in potential information processing bottlenecks, task switching costs, and minimized effects of information overload conditions (i.e., where users fail to detect, interpret, integrate, and successfully act on pertinent task information).

## LIST OF REFERENCES

- Ackerman, P. L., Beier, M. E., & Boyle, M. O. (2002). Individual differences in working memory within a nomological network of cognitive and perceptual speed abilities. *Journal of Experimental Psychology: General*, 131(4), 567-589.
- Altmann, E. M. (2004). Advance preparation in task switching: What work is being done? *Psychological Science*, 15, 616-622.
- Arrington, C. M., Altmann, E. M., & Carr T. H. (2003). Tasks of a feather flock together: similarity effects in task switching. *Memory & Cognition*, 31, 781-789.
- Baddeley, A.D. (1986). *Working memory*. Oxford: Clarendon Press.
- Baddeley, A. (1990). *Human memory: Theory and practice*. Boston: Allyn & Bacon.
- Baddeley, A. (2000). The episodic buffer: a new component of working memory? *Trends in Cognitive Sciences*, 4(11), 417-423.
- Barnard, P. J. (1999). Interacting cognitive subsystems: Modeling working memory phenomena within a multiprocessor architecture. In A. Miyake & P. Shah (Eds.), *Models of working memory* (pp. 298–339). New York: Cambridge University Press.
- Begault, D. R. (1993). Head-up auditory displays for traffic collision avoidance system advisories: a preliminary investigation. *Human Factors*, 35, 707-717.
- Belz, S. M., Robinson, G.S., & Casali, J. G. (1999). A new class of auditory warning signals for complex systems: auditory icons. *Human Factors*, 41(4), 608-618.

- Berka, C., Levendowski, D. J., Davis, G., Lumicao, M. N., Ramsey, C. K., Stanney, K. M., Reeves, L. M., Tremoulet, P. D., & Harkness-Regli, S. (2005). EEG Indices Distinguish Spatial and Verbal Working Memory Processing: Implications for Real-Time Monitoring in a Closed-Loop Tactical Tomahawk Weapons Simulation. In *CD-ROM Proceedings of the Human-Computer Interaction International Conference*, Mahwah, NJ: Lawrence Erlbaum Associates, Las Vegas, Nevada, July 22-27.
- Bertolotti, H., & Strybel, T. Z. (2005). The effects of audio spatial cue precision on target detection and identification. In *CD-ROM Proceedings of the Human-Computer Interaction International Conference*, Mahwah, NJ: Lawrence Erlbaum Associates, Las Vegas, Nevada, July 22-27.
- Blauert, J. (1983). *Spatial hearing: the psychophysics of human sound localization*. Cambridge: MIT Press.
- Bolia, R. S., D'Angelo, W. R., & McKinley, R. L. (1999). Aurally aided visual search in three-dimensional space. *Human Factors*, 41(4), 664–669.
- Billings, C. E., & Woods, D. D. (1994). Concerns about adaptive automation in aviation systems. In M. Mouloua and R. Parasuraman (Eds.), *Human performance in automated systems: Current research and trends* (pp. 264-269). Hillsdale, NJ: Lawrence Erlbaum Associates.
- Bowers, C. A., & LaBarba, R. C. (1991). Secondary-task assessment of workload: Neuropsychological considerations for applied psychology. *Perceptual and Motor Skills*, 73, 487-496.

- Burton, R. R., & Brown, J. S. (1979). An investigation of computer coaching for informal learning activities. *International Journal of Man-Machine Studies*, 11, 5-24.
- Byrne, E. A., & Parasuraman, R. (1996). Psychophysiology and adaptive automation. *Biological Psychology*, 42, 249-268.
- Cabeza, R., & Nyberg, L. (2000). Imaging cognition II: an empirical review of 275 PET and fMRI studies. *Journal of Cognitive Neuroscience*, 12, 1-47.
- Calvert, G. A., & Lewis, J.W. (2004). Hemodynamic studies of audio-visual interactions. In G. A. Calvert, C. Spence, & B. Stein (Eds.). *Handbook of Multisensory Processing* (pp. 483-502). MIT Press: Cambridge, MA.
- Calvert, G. C., Spence, C., & Stein, B. E. (Eds.) (2004). *The handbook of multisensory processing*. MIT Press: Cambridge, MA.
- Card, S., Moran, T. and Newell, A. (1984). *The psychology of human-computer interaction*. Hillsdale, NJ: Lawrence Erlbaum Associates.
- Cellier J-M, & Eyrolle, H. (1992). Interference between switched tasks. *Ergonomics*, 35, 25-36.
- Chandler, P. and Sweller, J. (1991). Cognitive load theory and the format of instruction. *Cognition and Instruction*, 8(4), 293-332.
- Chandler, P. and Sweller, J. (1992). The split-attention effect as a factor in the design of instruction. *British Journal of Educational Psychology*, 62, 233-246.
- Cholewiak, R. W., Brill, J. C., Schwab, A. (2004). Vibrotactile localization on the abdomen: effects of place and space. *Perception & Psychophysics*, 66, 970-987.
- Cholewiak, R. W., & Collins, A. A. (2000). The generation of vibrotactile patterns on a linear array: influences of body site, time, and presentation mode. *Perception and Psychophysics*, 62, 1220-1235.

- Cohen, S. (1980). Aftereffects of stress on human performance and social behavior: A review of research and theory. *Psychological Bulletin*, 88, 82–108.
- Cooper, G., (1997). Research in to Cognitive Load Theory and Instructional Design at UNSW, [http://www.arts.unsw.edu.au/education/CLT\\_NET\\_Aug\\_97.HTML](http://www.arts.unsw.edu.au/education/CLT_NET_Aug_97.HTML).
- Cooper, G. E., & Harper, R. P. (1969). The use of pilot ratings in the evaluation of aircraft handling characteristics. Washington, DC: NASA, Report No. TN-D-5153.
- Corballis, M.C., Hamm, J.P., Barnett, K.J. & Corballis, P.M. (2002) Paradoxical interhemispheric summation in the split brain. *Journal of Cognitive Neuroscience*, 14, 1151–1157.
- Cowan, N. (1988). Evolving conceptions of memory storage, selective attention, and their mutual constraints within the human information processing system. *Psychological Bulletin*, 104, 163-191.
- Cowan, N. (2001). The magical number 4 in short-term memory: A reconsideration of mental storage capacity. *Behavioral and Brain Sciences*, 24(1), 87–185.
- Crevitz , I., Debernard, S., & Denecker, P. (2002). Model building for air-traffic controllers' workload regulation. *European Journal of Operational Research*, 136, 24-332.
- Cutrell, E., Czerwinski, M., Horvitz, E. (2001). Notification, disruption, and memory: effects of messaging interruptions on memory and performance. Interact 2001: IFIP Conference on Human–Computer Interaction, Tokyo, Japan.
- Czerwinski, M., Cutrell, E., & Horvitz, E. (2000). Instant Messaging: Effects of Relevance and Time, In S. Turner, P. Turner (Eds), *People and Computers XIV: Proceedings of HCI 2000*, Vol. 2 (pp. 71-76), September 2000, British Computer Society.

- Doane, S. (2002). *New Measures of Complex Spatial Processing Abilities: Relating Spatial Abilities to Learning and Performance*. Mississippi State, MS: Mississippi State University Department of Psychology. (DTIC No. ADA413799, found on-line at: <http://handle.dtic.mil/100.2/ADA413799>).
- Dorneich, M., Whitlow, S., Ververs, P.M., Mathan, S., Raj, A., Muth, E., Hoover, A., DuRousseau, D., Parra, L., & Sajda, P. (2004). *DARPA Improving Warfighter Information Intake under Stress - Augmented Cognition Concept Validation Experiment (CVE) Analysis Report for the Honeywell Team*. DARPA/IPTO technical report submitted to CDR Dylan Schmorrow, December 31, 2004.
- Durlach, N. I. & Mavor, A. S. (1995). *Virtual reality: scientific and technological challenges*. Washington, DC: Academic Press.
- Dyson, B. J., & Quinlan, P. T. (2002). Within- and between-dimensional processing in the auditory modality. *Journal of Experimental Psychology: Human Perception and Performance*, 28, 1483-1498.
- Eimer, M., Cockburn, D., Smedley, B., & Driver, J. (2001) Cross-modal links in endogenous spatial attention are mediated by common external locations: evidence from event-related brain potentials. *Experimental Brain Research*, 139(4), 398–411.
- Endsley, M. (1995). Toward a Theory of Situational Awareness in Dynamic Systems *Human Factors* 37(1), 32–64.
- Endsley, M. R. (1996). Automation and situational awareness. In R. Parasuraman & M. Mouloua (Eds.), *Automation and human performance: Theory and applications* (pp. 163-181). Mahwah, NJ: Lawrence Erlbaum, Associates.

- Endsley, M. R. and Bolstad, C. A. (1994). Individual differences in pilot situation awareness. *International Journal of Aviation Psychology*, 4(3), 241-264
- Engle, R.W., Kane, M.J., & Tuholski, S.W. (1999). Individual differences in working memory capacity and what they tell us about controlled attention, general fluid intelligence, and functions of the prefrontal cortex. In A. Miyake and P. Shah (Ed.) *Models of working memory: Mechanisms of active maintenance and executive control* (pp. 102-134). Cambridge, UK: Cambridge University Press.
- Erp, J. B. F. Van, & Veen, H. A. H. C. Van (2003). A multi-purpose tactile vest for astronauts in the international space station. *Proceedings of Eurohaptics*, 405-408.
- ETS (1976a). Surface Development Test. *Kit of Factor-Referenced Cognitive Tests*. Princeton, NJ: Educational Testing Service.
- ETS (1976b). Map Planning Test. *Kit of Factor-Referenced Cognitive Tests*. Princeton, NJ: Educational Testing Service.
- European Telecommunications Standards Institute (ETSI) (2002). *Human factors: Guidelines on the multimodality of icons, symbols, and pictograms*. (Report No. ETSI EG 202 048 v 1.1.1 (2002-08). Sophia Antipolis, France: ETSI
- Giard, M. H., & Peronnet, F. (1999). Auditory-visual integration during multimodal object recognition in humans: A behavioral and electrophysiological study. *Journal of Cognitive Neuroscience*, 11(5), 473–490.
- Givens, A., & Smith, M.E. (2000). Neurophysiological measures of working memory and individual differences in cognitive ability and cognitive style. *Cerebral Cortex*, 10(9), 829-839.

- Galdes, D. K., & Smith, P. J. (1990). Building an intelligent tutoring system: Some guidelines from a study of human tutors. *Proceedings of the Human Factors Society 34<sup>th</sup> Annual Meeting*, 1407–1411.
- Gonzalez, C. (2005). Task workload and cognitive abilities in dynamic decision making. *Human Factors*, 47(1), 91-101.
- Groves, P. M., & Thompson, R. F. (1970). Habituation: a dual-process theory. *Psychological Review*, 77, 419-50.
- Haberlandt, K. (1997). *Cognitive psychology* (2nd ed.). Boston: Allyn & Bacon.
- Hale, K.S., Axelsson, P., Fuchs, S., Baskin, A. & Stanney, K. (2006). Multimodal Information Perceptualization for C4ISR Systems, Scientific and Technical Report: Final Report. Contract No. N61339-04-C-0037, Orlando, FL: NAVAIR Orlando.
- Hale, K.S., Reeves, L., Samman, S., Axelsson, P., Milham, L.L., & Stanney, K. (2005). Multimodal Information Perceptualization for C4ISR Systems: Scientific and Technical Report Summary Report and Validation Plan. Contract No. N61339-04-C-0037, Orlando, FL: NAVAIR Orlando.
- Hildebrandt, M. & Harrison, M.D. (2003). Putting time (back) into Dynamic Function Allocation. *Proceedings of the 47th Annual Meeting of the Human Factors and Ergonomics Society* (pp.488-492).
- Ho, C., & Spence, C. (2005). Assessing the effectiveness of various auditory cues in capturing a driver's visual attention. *Journal of Experimental Psychology: Applied*, 11, 157-174.

- Ho, C., Spence, C., & Tan, H. Z. (2005). Warning signals go multisensory. In CD *Proceedings of the 11th International Conference on Human-Computer Interaction* (Vol. 9 - Advances in Virtual Environments Technology: Musings on Design, Evaluation, & Applications), Mahwah, NJ: Lawrence Erlbaum Associates, Las Vegas, Nevada, July 22-27.
- Ho C.-Y., Nikolic M. I., Waters M. J. & Sarter N. B. (2004). Not now! Supporting interruption management by indicating the modality and urgency of pending tasks. *Human Factors*, 46 (3), 399-409.
- Ho, C., Tan, H. Z., & Spence, C. (2005). Using spatial vibrotactile cues to direct visual attention in driving scenes. *Transportation Research Part F: Traffic Psychology and Behavior*, Vol. 8, pp. 397-412.
- Hoc, J.M. (2001). Towards a cognitive approach to human-machine cooperation in dynamic situations. *International Journal of Human-Computer Studies*, 54, 509-540.
- Hoc, J. M., & Debernard, S (2002). Respective demands of task and function allocation on human-machine co-operation design: a psychological approach. *Connection Science*, 14(4), 283-295.
- Hoc, J. M., & Lemoine, M. P. (1998). Cognitive evaluation of human-human and human-machine cooperation modes in air traffic control. *International Journal of Aviation Psychology*, 8, 1-32.
- Hopp-Levine, P. J., Smith, C. A. P., Clegg, B. A., & Heggstad, E. D. (2006). Tactile interruption management: tactile cues as task-switching reminders. *Cognition, Technology & Work*, 8(2), 137-145).
- Hopp, P. J., Smith, C. A. P., Clegg, B. A., & Heggstad, E. D. (2005). Interruption management: The use of attention-directing tactile cues. *Human Factors*, 47(1), 1-11.

- Iacoboni, M., & Zaidel, E. (2003). Interhemispheric visuo-motor integration in humans: the effect of redundant targets. *European Journal of Neuroscience*, 17, 1981-1986.
- Just, M. A., Carpenter, P. A., & Miyake, A. (2003). Neuroindices of cognitive workload: Neuroimaging, pupillometric, and event-related potential studies of brain work. *Theoretical Issues in Ergonomics*, 4, 56-88.
- Kalyuga, S., Chandler, P., & Sweller, J., (1999). Managing Split-attention and Redundancy in Multimedia Instruction, *Applied Cognitive Psychology*, 13, 351-371.
- Kane, M. J., & Engle, R. W. (2002). The role of prefrontal cortex in working-memory capacity, executive attention, and general fluid intelligence : an individual-differences perspective. *Psychonomic Bulletin & Review*, 9, 637–671.
- Kennedy P.R., Bakay R.A., Moore M.M., Adams K., & Goldwaithe J. (2000). Direct control of a computer from the human central nervous system. *IEEE Transactions on Rehabilitation Engineering*, 8(2),:198-202.
- Kennett, S., Eimer, M., Spence, C., & Driver, J. (2001). Tactile-visual links in exogenous spatial attention under different postures: Convergent evidence from psychophysics and ERPs. *Journal of Cognitive Neuroscience*, 13(4), 462–478.
- Kobus, D., Brown, C., Morrison, J., Kollmorgen, G., Cornwall, R. & Schmorow, D. (2006). *DARPA Improving Warfighter Information Intake under Stress - Augmented Cognition Phase II: The Concept Validation Experiment (CVE)*. Submitted for DTIC publication.
- Lathan, C. E. & Tracy, M. R. (2002). The effects of operator spatial perception and sensory feedback on human-robot teleoperation performance. *Presence: Teleoperators and Virtual Environments*, 11(4), 368-377.

- Latorella, K. A. (1996). Investigating interruptions: implications for flight deck performance. Doctoral dissertation, State University of New York at Buffalo.
- Latorella, K. A. (1998). Effects of modality on interrupted flight deck performance: implications for data link. *Proceedings of the Human Factors and Ergonomics Society 42nd Annual Meeting*.
- Latorella, K. A. (1999). Investigating interruptions: implications for flight deck performance. NASA Technical Memorandum 209707. Washington, DC: National Aviation & Space Administration.
- Lindeman, R. W., Page, R., Yanagida, Y, & Sibert, J. L. (2004) Towards full-body haptic feedback: the design and deployment of a spatialized vibrotactile feedback system. *Proceedings of the ACM symposium on Virtual reality software and technology*, 146-149.
- Mamykina, L., Mynatt, E., & Terry, M. A. (2001). Time aura: interfaces for pacing. *In Proceedings of ACM CHI*, 3(1), 144-151.
- Marentakis, G. (2006). Deictic Spatial Audio Target Acquisition in the Frontal Horizontal Plane. Doctoral dissertation, University of Glasgow, United Kingdom. Available online at: [http://www.dcs.gla.ac.uk/~stephen/papers/theses/georgios\\_marentakis\\_thesis.pdf](http://www.dcs.gla.ac.uk/~stephen/papers/theses/georgios_marentakis_thesis.pdf)
- Martin, M. (1980). Attention to words in different modalities: Four-channel presentation with physical and semantic selection. *Acta Psychologica*, 44, 99-115.
- Mayer, R.E. & Anderson, R.B. (1991). Animation needs narrations: An experiment test of a dual-coding hypothesis. *Journal of Educational Psychology*, 83(4), 484-490.
- Mayer, R. E., & Massa, L. J. (2003). Three facets of visual and verbal learners: Cognitive ability, cognitive style, and learning preference. *Journal of Educational Psychology*, 95(4), 833-846.

- Mayer, R.E. & Moreno, R. (1998). A split-attention effect in multimedia learning: evidence for dual processing system in working memory. *Journal of Educational Psychology*, 90(2), 312-320.
- Melara, R. D., & O'Brian, T. (1987). Interaction between synesthetically corresponding dimensions. *Journal of Experimental Psychology:General*, 116, 323-336.
- McFarlane, D.C. (2002). Comparison of four primary methods for coordinating the interruption of people in human-computer interaction. *Human-Computer Interaction*, 17(3), 63-139.
- McFarlane, D. C., & Latorella, K. A. (2002). Scope and importance of interruption in human-computer interaction design. *Human-Computer Interaction*, 17(3), 1-61.
- McGee, D. R., Cohen, P.R., Oviatt, S. (1998). Confirmation in Multimodal Systems, *Proceedings of the International Joint Conference of the Association for Computational Linguistics and the International Committee on Computational Linguistics (COLING-ACL '98)*: Montreal, Quebec, Canada.
- McKinley, R. L., & Ericson, M. A. (1997). Flight demonstrations of a 3-D auditory display. In R. H. Gilkey & T. R. Anderson (Eds.) *Binaural and Spatial Hearing in Real and Virtual Environments* (pp. 683-699). Mahwah, NJ: Lawrence Erlbaum Associates.
- Middlebrooks, J. (1997). Spectral Cues for Sound Localization. In H.R. Gilkey and T. Anderson R., (Eds), *Binaural and Spatial Listening in Real and Virtual Environments* (pp. 77-98). Mahwah, NJ: Lawrence Erlbaum Associates.
- Miller, G.A. (1956). The magical number seven, plus or minus two: Some limits on our capacity for processing information. *Psychological Review*, 63, 81-97. Found online Dec 2006 at: <http://cogprints.org/730/00/miller.html>

- Miller, J. (1982). Divided attention: Evidence for coactivation with redundant signals. *Cognitive Psychology*, 14, 247-279.
- Miller, J. (1986). Time course of coactivation in bimodal divided attention. *Perceptual Psychophys.* 40: 331–343.
- Miyake, A., & Shah, P. (Eds.). (1999). *Models of working memory: Mechanisms of active maintenance and executive control*. New York: Cambridge University Press.
- Monk, C., Boehm-Davis, D., & Trafton, J. G. (2002). The Attentional Costs of Interrupting Task Performance at Various Stages. In the *Proceedings of the 2002 Human Factors and Ergonomics Society Annual Meeting*. Santa Monica, CA: HFES.
- Moreno, R. & Mayer, R. E. (2002). Learning science in virtual reality multimedia environments: role of methods and media. *Journal of Educational Psychology*, 94(3), 598-610.
- Mulgund, S, Stokes, J., Turieo, M., & Devine, M. (2002). *Human/Machine interface modalities for solidier systems technologies*. Technical report submitted to the U.S. Army Natick Solider Center.
- National Research Council (1997). Computer Science and Telecommunications Board. *More than screen deep: toward every-citizen interfaces to the nation's information infrastructure*. Washington, DC: National Academy Press.
- Norman, D. (1998). The invisible computer. Cambridge, MA: MIT Press.
- Obermayer, R.W., & Nugent, W. A. (2000). *Human-computer interaction for alert warning and attention allocation systems of the multi-modal watchstation*. SPIE 2000, SPIE–The International Society for Optical Engineering, Bellingham, WA.

- Oviatt, S., Coulston, R., and Lunsford, R. 2004. When do we interact multimodally?: cognitive load and multimodal communication patterns. In *Proceedings of the 6th international Conference on Multimodal interfaces* (State College, PA, USA, October 13 - 15, 2004). ICMI '04. ACM Press, New York, NY, 129-136.
- Oviatt, S.L. (2003). Multimodal interfaces. In *The Human-Computer Interaction Handbook: Fundamentals, Evolving Technologies and Emerging Applications*, (ed. by J. Jacko and A. Sears), Lawrence Erlbaum Assoc., Mahwah, NJ, 2003, chap.14, 286-304.
- Oviatt, S. L. (1999). Ten myths of multimodal interaction. *Communications of the ACM*, 42(11), 74-81.
- Parasuraman, R., Bahri, T., Deaton, J. E., Morrison, J. G., & Barnes, M. (1992). *Theory and design of adaptive automation in aviation systems* (Technical Report, Code 6021). Warminster, PA: Naval Air Development Center.
- Parasuraman, R., Mouloua, M., & Hilburn, B. (1999). Adaptive aiding and adaptive task allocation enhance human-machine interaction. In M. W. Scerbo & M. Mouloua (Eds.) *Automation Technology and Human Performance: Current Research and Trends*. (pp. 119-123). Mahwah, NJ: Erlbaum.
- Parasuraman, R., Mouloua, M., & Molloy, R. (1996). Effects of adaptive task allocation on monitoring of automated systems. *Human Factors*, 38, 665-679.
- Popescu, G. V., Burdea, G. C., & Trefftz, H. (2002). Multimodal interaction modeling. In K. M. Stanney (Ed.), *Handbook of virtual environments: design, implementation, and applications* (pp. 435-454). Mahwah, NJ: Lawrence Erlbaum Associates, Inc.

- Prinzel, L. J., Parasuraman, R., Freeman, F. G., Scerbo, M. W., Mikulka, P. J., & Pope, A. T. (2003). Three experiments examining the use of electroencephalogram, event-related potentials, and heart rate variability for real-time human-centered adaptive automation design (Technical Report, Code 212442). Hanover, MD: NASA Center for AeroSpace Information
- Raven, J., Raven, J. C., & Court, J. H. (1998). *General Overview: Standard Progressive Matrices*. Oxford: OPP Limited.
- Raven, J., Raven, J. C., & Court, J. H. (2000). *Mill Hill Vocabulary Scales*. Oxford: OPP Limited.
- Raven, J., Raven, J. C., & Court, J. H. (2000). *Standard Progressive Matrices*. Oxford: OPP Limited.
- Reeves, L. M. & Stanney, K. M. (2007). Developing an architecture for multi-modal optimization (AMMO). Manuscript submitted to *Theoretical Issues in Ergonomics Science*.
- Reeves, L. M. & Stanney, K. M., Ahmad, A., & Malone, L. (2007). Empirically validating multimodal mitigation strategies derived with AMMO. Manuscript submitted to *Ergonomics*.
- Reeves, L.M., Ahmad, A., & Stanney, K.M., (2005). Assessing influences of verbal and spatial ability on multimodal C2 task performance. *In Proceedings of the 11th International Conference on Human-Computer Interaction* (Vol. 11 - Foundations of Augmented Cognition), Mahwah, NJ: Lawrence Erlbaum Associates, Las Vegas, Nevada, July 22-27.
- Reeves, L.M., Lai, J.C., Larson, J.A., Oviatt, S.L., Balaji, T.S., Buisine, S., Collings, P., Cohen, P.R., Kraal, B., Martin, J.C., McTear, M.F., Raman, T.V., Stanney, K.M., Su, H., & Wang, Q.Y. (2004). Guidelines for multimodal user interface design. *Communications of the ACM*, 47(1), 57-59.

- Roda, C., & Thomas, J. (2006). Attention aware systems: theories, applications, and research agenda. *Computers in Human Behavior*, 22(4), 557-587.
- Roser, M., & Corballis, M.C. (2002). Interhemispheric neural summation in the split brain with symmetrical and asymmetrical displays. *Neuropsychologia*, 40, 1300–1312.
- Rothrock, L., Koubek, R., Fuchs, F., Haas, M., & Salvendy, G. (2002). Review and reappraisal of adaptive interfaces: toward biologically-inspired paradigms. *Theoretical Issues in Ergonomics Science*, 3(1), 47-84.
- Rudmann, D. R. & Strybel T.Z. (1999). Auditory spatial facilitation of visual search performance: effect of local vs global visual distractor density. *Human Factors*, 41, 141-160.
- Sanders, M. S., & McCormick, E. J. (1993). *Human factors in engineering and design* (7th ed.). New York: McGraw–Hill.
- Sarter, N. B. (2000). The need for multisensory interfaces in support of effective attention allocation in highly dynamic event-driven domains: the case of cockpit automation. *International Journal of Aviation Psychology*, 10, 231–245.
- Sarter, N.B. (2002). Multimodal information presentation in support of human-automation communication and coordination. In Eduardo Salas (Ed.), *Advances in human performance and cognitive engineering research* (pp. 13-36). New York, NY: JAI Press.
- Savazzi, S. & Marzi, C.A. (2002) Speeding up reaction time with invisible stimuli. *Current Biology*, 12, 403–407.
- Scerbo, M. W. (1994). Implementing adaptive automation in aviation: the pilot-cockpit team. In M. Mouloua & R. Parasuraman (Eds.), *Human performance in automated systems: current research and trends* (pp. 249-255). Hillsdale, NJ: Lawrence Erlbaum Associates.

- Scerbo, M. W. (1996). Theoretical perspectives on adaptive automation. In R. Parasuraman & M. Mouloua (Eds.), *Automation and human performance: Theory and applications* (pp. 37-64). Mahwah, New Jersey: Lawrence Erlbaum Assoc.
- Scerbo, M. W. (2001). Adaptive automation. In W. Karwowski (Ed.) *International encyclopedia of ergonomics and human factors*. London: Taylor and Francis, Inc.
- Schmorrow, D. D. (Ed.) (2005). *Foundations of Augmented Cognition*. Mahwah, NJ: Lawrence Erlbaum Associates.
- Schmorrow, D. D., Stanney, K. M., & Reeves, L. M. (Eds.) (2006). *Foundations of Augmented Cognition (2<sup>nd</sup> Ed.): Augmented Cognition: Past, Present, and Future*. Arlington, VA: Strategic Analysis, Inc.
- Schmorrow, D., Stanney, K.M., Wilson, G., & Young, P. (2005, in press). Augmented cognition in human-system interaction. In G. Salvendy (Ed.), *Handbook of human factors and ergonomics* (3<sup>rd</sup> edition). New York: John Wiley.
- Schneider, W. (1999). Working memory in a multilevel hybrid connectionist control architecture (CAP2). In A. Miyake & P. Shah (Eds.), *Models of working memory* (pp. 340– 374). New York: Cambridge University Press.
- Schumacher, E. H., Seymour, T. L., Glass, J. M., Fencsik, D. E., Lauber, E. J., Kieras, D. E., & Meyer, D. E. (2001). \_Virtually perfect time sharing in dual-task performance: Uncorking the central cognitive bottleneck. *Psychological Science*, 12(2), 101-108.
- Shah, P., & Miyake, A. (1996). The separability of working memory resources for spatial thinking and language processing: An individual differences approach. *Journal of Experimental Psychology: General*, 125, 4-27.

- Sheridan, T., & Parasuraman, R. (2006). Human-automation interaction. *Reviews of Human Factors and Ergonomics*, 1, 89-129.
- Sherrick, C. E., & Cholewiak, R. W. (1986). Cutaneous sensitivity. In K. Boff, L. Kaufman, & J.L. Thomas (Eds.), *Handbook of perception and human performance* (pp. 12-1 - 12-58). New York: Wiley.
- Sklar, A. E., & Sarter, N. B. (1999). Good vibrations: tactile feedback in support of attention allocation and human–automation coordination in event-driven domains. *Human Factors*, 41, 543–552.
- Smith, E. E., & Jonides, J. (1998). Neuroimaging analyses of human working memory. *Proceedings of the National Academy of Sciences, USA*, 95, 12061–12068.
- Spence, C., & Driver, J. (1997). Cross-modal links in attention between audition, vision, and touch: implications for interface design. *International Journal of Cognitive Ergonomics*, 1, 352-373.
- Spence, C., & Driver, J. (1997). Audiovisual links in attention: implications for interface design. In D. Harris (Ed.), *Engineering Psychology and Cognitive Ergonomics, Vol.2., Job Design and Product Design* (pp. 185-192). Ashgate Publishing: Hampshire.
- Spence, C., & Driver, J. (1999). A new approach to the design of multimodal warning signals. In D. Harris (Ed.) *Engineering psychology and cognitive ergonomics, Vol. 4: job design, product design and human-computer interaction* (pp. 455-461). Aldershot, England: Ashgate.
- Spence, C., & Driver, J. (Eds.) (2004). *Crossmodal Space and Crossmodal Attention*. Oxford, England: Oxford University Press.

- Spence, C., & Read, L. (2003). Speech shadowing while driving: on the difficulty of splitting attention between eye and ear. *Psychological Science*, 14, 251-256.
- Speier, C., Vessey, I., & Valacich, J. S. (2003). The effects of interruptions, task complexity, and information presentation on computer-supported decision-making performance. *Decision Sciences*, 34(4), 771-797.
- Springer, S. P., & Deutsch, G. (1985). *Left Brain, Right Brain*. New York: W. H. Freeman and Company.
- S&T Manning Affordability (2000). Multi-modal watchstation. Available from: <http://www.manningaffordability.com/S&tweb/index.htm>.
- Stanney, K., Reeves, L., Hale, K., Samman, S., Buff, W. (2003). OSD02-CR14: Multimodal Information Perceptualization (MIP) for C4ISR Systems. Navy Technical Report TR-203-002.
- Stanney, K., Samman, S., Reeves, L., Hale, K.S., Buff, W., Bowers, C., Goldiez, B., Nicholson, D., & Lackey, S. (2004). A paradigm shift in interactive computing: Deriving multimodal design principles from behavioral and neurological foundations, *International Journal of Human-Computer Interaction*, 17(2), p. 229-257.
- Stein, B. E. & Meredith, M. A. (1993). The merging of the senses. Cambridge, MA: MIT.
- Sulzen, J. (2001). Modality based working memory. School of Education, Stanford University. Retrieved, February 5 2003, from <http://ldt.stanford.edu/~jsulzen/james-sulzen-portfolio/classes/PSY205/modality-project/paper/modality-expt-paper.PDF>.
- Tan, H. Z., Gray, R., Young, J. J., & Taylor, R. (2003). A haptic back display for attentional and directional cueing. *Haptics-e: The Electronic Journal of Haptics Research*, 3: <http://www.haptics-e.org/>.

- Thomas, M., Tremoulet, P., Morizio, N. (2005). Performance augmentation through cognitive enhancement. In *CD-ROM Proceedings of the Human-Computer Interaction International Conference*, Mahwah, NJ: Lawrence Erlbaum Associates, Las Vegas, Nevada, July 22-27.
- Thompson-Schill SL, Aguirre GK, D'Esposito M, Farah MJ (1999). A neural basis for category and modality specificity of semantic knowledge. *Neuropsychologia*, 37, 671-676.
- Tindal-Ford, S., Chandler, P., & Sweller, J. (1997). When two sensory modes are better than one. *Journal of Experimental Psychology: Applied* 3(4), 257-287.
- Trafton, J. G., Altmann, E. M., Brock, D. P., & Mintz, F. E. (2003). Preparing to resume an interrupted task: effects of prospective goal encoding and retrospective rehearsal. *International Journal of Human-Computer Studies*, 58, 583–603.
- Turk, M., & Robinson, G. (2000). Perceptual user interfaces. *Communications of the ACM*, 43(3) 33-34.
- Van Erp, J. B. F., & Van Veen, H. A. H. C. (2004). Vibrotactile in-vehicle navigation system. *Transportation Research Part F* 7, 247–256.
- Vu, K. L., Strybel, T. Z., & Proctor, R. (2006). Effects of displacement magnitude and direction of auditory cues on auditory spatial facilitation of visual search. *Human Factors*, 48(3), 587-599.
- Welch, R. and D. H. Warren (1986). Intersensory interactions. In K.R. Boff, L. Kaufman, & J.P. Thomas (Eds.), *Handbook of perception and human performance. Vol. I: Sensory processes and perception* (pp. 25.1-25.36). New York: John Wiley.
- Wickens, C.D. (2002). Multiple resources and performance prediction. *Theoretical Issues in Ergonomic Sciences*, 4, 1-19.

- Wickens, C.D. (1992). *Engineering psychology and human performance* (2<sup>nd</sup> ed.). New York: Harper Collins Publishers.
- Wickens, C. D. (1984). *Engineering Psychology and Human Performance*. Columbus, OH: Charles E. Merrill Publishing Co.
- Wickens, C. D., Goh, J., Helleberg, J., Horrey, W., & Talleur, D. A. (2003). Attentional models of multitask pilot performance using advanced display technology. *Human Factors*, 45(3), 360-380.
- Wickens, C. D., & Gosney, J. L. (2003). Redundancy, Modality, Priority and Instructions in Dual Task Interference Between Simulated Vehicle Control and In-Vehicle Technology. Technical Report AHFD-03-18/GM-03-3. Savoy, IL: University of Illinois, Aviation Human Factors Division. Available from:  
<http://www.humanfactors.uiuc.edu/Reports&PapersPDFs/TechReport/03-18.pdf>
- Wickens, C. D. & Hollands, J. G. (2000). *Engineering Psychology and Human Performance* (3rd Ed.). Upper Saddle River, NJ: Prentice Hall.
- Wickens, C. D., & Seppelt, B. (2002). Interference with driving or in-vehicle task information: The effects of auditory versus visual delivery. Technical Report AHFD-02-18/GM-02-03. Savoy, IL: University of Illinois, Aviation Human Factors Division. Available from:  
<http://www.humanfactors.uiuc.edu/Reports&PapersPDFs/TechReport/02-18.pdf>
- Wickens, C. D., and Liu, Y. (1988). Codes and modalities in multiple resources: A success and a qualification. *Human Factors*, 30, 599- 616.
- Wierwille, W. W., & Casali, J. G. (1983). A Validated Rating Scale for Global Mental Workload Measurement Applications. Proceedings of the human factors Society 27th Annual Meeting, 1983 (pp. 129-133). Santa Monica, CA: Human Factors Society.

Zijlstra, F. R. H., Roe, R. A., Leonora, A. B., & Krediet, I. (1999). Temporal factors in mental work: effects of interrupted activities. *Journal of Occupational and Organizational Psychology*, 72, 163–185.