USING A VIRTUAL ENVIRONMENT TO ELICIT SHIPHANDLING KNOWLEDGE

Michael K. Martin, Ph.D., Elizabeth Sheldon, Steve Kass, Ph.D., Andrew M. Mead, LT., Ph.D., Sherrie A. Jones, Ph.D., & Robert Breaux, Ph.D.

Naval Air Warfare Center, Training Systems Division, Orlando, FL

ABSTRACT

This paper discusses initial development of Virtual Environment Training Technologies (VETT) to facilitate knowledge elicitation, task analysis, and performance measure development for naval training. Primary components included on-going hardware and software development in the VETT test-bed, iterative prototyping capabilities, and flexible performance recording and review capabilities. The customer for this effort, the Surface Warfare Officers' School (SWOS), envisions the use of virtual environments to train "seaman's eye," an expert perceptual technique currently learned through extensive apprenticeship training. The chosen target task for the R&D process was Underway Replenishment (UNREP). The process began with the use of standard Knowledge Elicitation (KE) techniques (e.g., document reviews and interviews) to obtain an UNREP task description. This description was then used to develop a prototype simulation and a generic UNREP scenario on the VETT test-bed. Subject matter experts included Merchant Marine harbor pilots, visiting line Surface Warfare Officers (SWOs), and in-house expert reservists. The SWOs performed the simulated UNREP scenario while video was recorded from their viewpoint (i.e., first-person perspective). The videos provided further KE capability by permitting SWOs to explain what they did and why they did it in a context-specific manner during a debrief session. This arrangement produced more elaborate explanations than the initial interviews, apparently because it reminded the experts of subtle, nonverbal cues used to perform UNREP. The context-specific interviews stimulated more detailed discussions of the validity of the simulation and potential performance measures. Based on initial analysis, the prototype was taken to SWOS for a simulation validity experiment. Progressive refinements will be described and the results of a simulation validation experiment will be presented.

AUTHORS

Michael Martin is a Research Psychologist at NAWCTSD. He is the Principal Investigator of the Virtual Environment Training Technologies (VETT) Training Effectiveness Evaluation program and has a Ph.D. in Experimental / Cognitive Psychology from the University of Kansas.

Elizabeth Sheldon is a graduate student at the University of Central Florida currently working on the VETT (Virtual Environment Training Technology) project at NAWCTSD as a research assistant. She holds a masters degree in counseling and human development.

Steve Kass worked on the VETT program as a Research Psychologist at NAWCTSD, but is currently an Assistant Professor in the Psychology Department at the University of West Florida. He has a Ph.D. in Human Factors Psychology from the University of Central Florida.

LT Andrew M. Mead is designated Aerospace Experimental Psychologist at the Naval Air Warfare Center, Training Systems Division in Orlando. He earned a Ph.D. in Experimental Cognitive Psychology from the University of Delaware. His interests are spatial cognition, cognitive modeling, and simulator side effects issues.

Sherrie Jones is a Research Psychologist at NAWCTSD, serving as Principal Investigator for the Virtual Environment Training Technologies (VETT) Side-Effects research program. Miss Jones earned a Ph.D. in Human Factors Psychology from the University of Central Florida.

Robert Breaux received a Ph.D. from Texas Tech University in Experimental Psychology, and serves the NAWCTSD as IPT lead for Virtual Reality Technology.

USING A VIRTUAL ENVIRONMENT TO ELICIT SHIPHANDLING KNOWLEDGE

**Michael K. Martin, Ph.D., Elizabeth Sheldon, Steve Kass, Ph.D.,
Andrew M. Mead, LT., Ph.D., Sherrie A. Jones, Ph.D.,& Robert Breaux, Ph.D.**

Naval Air Warfare Center, Training Systems Division, Orlando, FL

INTRODUCTION

This paper presents a project that uses Virtual Environment (VE) technology to facilitate Knowledge Elicitation (KE), task analysis, and the development of performance measures. The project is the Virtual Environment Training Technologies (VETT) project at the Naval Air Warfare Center Training Systems Division (NAWCTSD). VETT's current R&D efforts will be transitioned to a training system for developing, maintaining, and refreshing "seaman's eye" for students at the Surface Warfare Officers' School (SWOS) in Newport, RI.

SEAMAN'S EYE

Seaman's eye is a traditional term used to express competence at sea (Crenshaw, 1975). It connotes judgment and understanding of ships, the sea, and the physics of shiphandling. It is the skill required to drive ships much like other people drive cars -- by looking out the window, judging relative motion, and adjusting speed and course appropriately.

Due to the large amounts of inertia involved, speed and course changes occur very slowly. To accommodate the slow response of one's own ship (i.e., ownship), a shiphandler must always be thinking ahead of ownship and anticipating future course and speed adjustments. This involves exact judgment of environmental conditions (e.g., winds and currents). It involves timing of engine and rudder commands to counteract or exploit the physical forces acting on ownship). It also involves a sense of relative motion precise enough to allow accurate predictions of own and surrounding ships' future positions.

Like other complex skills, seaman's eye is acquired in three stages (cf. Anderson, 1987; VanLehn, 1996). In the first stage, the cognitive stage, an understanding of fundamental shiphandling concepts is acquired through books and in classrooms. This stage consists primarily of the accumulation of declarative (i.e., verbalizable) knowledge of what we know about objects, how events unfold, and static relationships among concepts (Gordon, 1992). With regard to seaman's eye, this includes learning *about* the physical principles underlying the forces that act on ownship, the steps to be taken during a particular maneuver, rules of thumb, and standard communication procedures. Few would say officers at this stage of skill development have seaman's eye.

In the next stage, the associative stage, officers begin to observe experienced shiphandlers applying these concepts onboard ship. During this stage, officers begin to associate environmental conditions with actions taken to control ownship. They also begin to associate the actions taken with reasons for why they were taken. In addition, they learn to anticipate what actions an experienced shiphandler may take next. During this stage, officers begin to demonstrate the rudiments of seaman's eye. The skills required to visually estimate ranges, bearings, and relative motion begin to take shape. Static relationships among fundamental concepts learned from books and in classrooms begin to give way to a more dynamic knowledge that can be applied in the ever-changing environment at sea. They begin to learn *how* to apply their shiphandling knowledge.

The third stage of skill development, the autonomous stage, is viewed as the most crucial stage for acquiring the perceptual basis of seaman's eye, as it involves direct shiphandling experience. For this reason, the acquisition of seaman's eye requires extensive apprenticeship training. During this final stage, weak associations between environmental conditions and shiphandling maneuvers gained during prior stages of skill development are strengthened. What used to require slow, deliberate attention now begins to occur automatically. Static knowledge of ships, the sea, and the physics of shiphandling becomes progressively easier to apply to dynamic, often time-critical, situations. It is during this stage that shiphandling expertise begins to express itself. In general, it is also during this stage that experts begin to lose their ability to verbally express how they perform their task (Anderson, 1987; Gordon, 1992). This loss of declarative knowledge as expertise is acquired also appears to occur in shiphandling, as indicated by Crenshaw (1975, p. 7):

“Many excellent shiphandlers … handle the ship by intuitive feel and are at a loss to explain how they decide to give a certain command.”

The perceptual and nonverbal aspects of seaman’s eye combine to create a challenging problem with regard to KE, task analysis, and performance measure development. Seaman’s eye is a complex skill that often relies on the analysis of subtle, higher-order perceptual cues from the environment and precise timing of engine and rudder commands. The spatial and temporal (i.e., perceptual) nature of seaman’s eye can make it very difficult to describe these cues. As a result, KE, task analysis, and the development of performance measures based on verbal descriptions and verbal responses is complicated.

We attempted to meet this challenge by using a combination of standard, verbally based techniques and VE-enabled techniques. The approach was iterative and relied on rapid prototyping and performance recording/replay made possible by VE technology. Generally, our approach consisted of having shiphandlers tell us what they were going to do, do it, and then tell us what they did and why they did it. An important benefit of this approach was that it kept the development of the VE simulation centered on the user. This arrangement allowed Surface Warfare Officers (SWOs) to provide constructive criticism of the VE in terms of fidelity requirements, thereby assisting in subsequent prototype development.

STANDARD APPROACH

The standard approach to KE consists of familiarization with a task or domain, interviews with Subject Matter Experts (SMEs), and field observations (McGraw & Harbison-Briggs, 1989). Familiarization with the task or domain is the first step and provides a framework for the KE process (Hoffman, Shadbolt, Burton, & Klein, 1995). It is accomplished primarily through literature review and documentation analysis. Domain familiarization is then used to guide interviews with SMEs and organize field observations.

Task

Underway Replenishment (UNREP) is a restricted maneuvering evolution that requires bringing ownship alongside a supply ship to take on fuel, supplies, or munitions while at sea. After replenishing, ownship continues on its prior course. UNREP can be divided into three phases: approach, alongside and breakaway. A typical approach starts at approximately 500 yards astern of the oiler, offset to the side where replenishment will occur. The goal of the approach is to bring ownship to the designated UNREP station position alongside the oiler. The UNREP station position can be defined as the position in which ownship attains the proper lateral separation while alongside, typically between 100 and 140 feet, and is lined up with the UNREP station on the oiler. The approach ends with passing of the Phone and Distance Line (P&D Line), a line passed between the two ships that enables communication as well as providing a tool for measuring lateral separation.

The goal of the next phase (alongside) is for ownship to maintain the UNREP position while supplies are being transferred. When replenishment is complete and all lines (e.g., the P&D line) have been retrieved, the breakaway commences. The purpose of the breakaway is for ownship to pull forward of the oiler, make slight course changes, and resume its prior course.

Unstructured Interviews

Following a document review, two SWOs were interviewed regarding fundamental shiphandling skills and seaman's eye. An unstructured interview format was chosen because it provides opportunities for domain experts to guide the context of the interview and highlight important information (Hoffman et al., 1995).

The SWOs discussed shiphandling skills and abilities such as knowledge of the engine and rudders, intricacies of the ship, communication, local knowledge, currents, weather, vessel traffic procedures, and basic maneuvers. They also described a few of the perceptual cues used, such as listening to the clicks on a gyrocompass to determine the rate of a course change and hoisting a flag to determine the direction of the wind.

As expected, the verbal probes for components of seaman’s eye yielded information similar to that found during a document review. The verbal reports generated by the interviews included, among other things, a list of the steps required in a shiphandling task, and references to salient aspects of the task environment. Thus, the interviews primarily yielded the declarative knowledge associated with seaman's eye, but in general, only touched on its perceptually based elements. Phrased differently, the interviews tended to provide information related to the first stage of skill development, whereas information about the perceptual basis of seaman's eye acquired in the second and third stages of development was limited.

Conceptual Graphing

Experts often find it difficult to communicate fundamental concepts underlying their domain of expertise (Schumacher & Czerwinski, 1992). The same appears to be true for shiphandling expertise. Therefore, in an effort to remove some of the burden of using language to describe concepts that are fundamentally spatial and temporal, a conceptual graphing technique was employed.

The premise of conceptual graphing is for the expert to select important elements from their domain and pictorially link them in a manner that reflects the relationships among the concepts they deem important (Hoffman et al., 1995). Along these lines, two SWOs were asked to construct a diagram of the basic elements of shiphandling. The concepts drawn on the graph referred to the various forces experienced when handling a ship -- again, information very similar to that obtained during a document review.

VE-ENABLED APPOACH

Unstructured interviews and conceptual graphing techniques provided little more than the information obtained with a document review. Our efforts to elicit knowledge related to the more advanced stages of skill development were unsuccessful. To get beyond the limitations of language, we required an approach to KE that would allow us to observe and analyze how an officer maneuvering the ship uses the more sophisticated perceptual cues present at sea.

Although field observation is an alternative to the verbally based techniques above, going onboard ship for observation would have been costly and time-intensive. For our purposes, we required a method of rapidly acquiring in-context observation and measurement of a large number of experts as they performed a shiphandling task. The method must allow us to observe and record information about where the officer is looking at all times, the information required from navigational instruments, and the timing of engine and rudder commands. We also required information about the control dynamics of ownship (e.g., course, speed).

In parallel with our initial KE efforts, shiphandling objects and forces were being modeled on the VETT test-bed. Due to the relatively rapid development of the VE, it became feasible to place SWOs in a rudimentary VE shiphandling task for the initial purpose of acquiring feedback on the simulation. While providing feedback on the simulation, the SWOs also began providing unsolicited feedback on how they were performing the task. At this point, it became evident that the VE itself was a useful tool for KE. This began a coordinated, iterative process of simulation development, KE, and performance measure development.

VE Test-Bed

The hardware components of the VETT test-bed for this phase consisted of an ONYX2 with four MIPS R10000 processors, Infinite Reality Graphics, and one Raster Manager by *Silicon Graphics, Inc*. Head tracking was provided by a *Polhemus Fastrak*, and the display consisted of a VR4 Head Mounted Display (HMD) by *Virtual Research.* The software components included dVISE by *Division, Inc.* and VEGA Marine by *Paradigm Simulation, Inc*.

Scenario Development

Available documentation and interviews were used to develop a basic UNREP scenario, which was then implemented on the test bed. For the purposes of KE, it was important that the scenario be realistic enough to provide an accurate record of UNREP performance. Based on our prior experience with document reviews and interviews, however, it seemed likely that our UNREP scenario would be perceptually impoverished.

To counter this, we began a KE process in which SWOs first provided a written description of how they perform an UNREP. This technique, of course, relies on how well the SWOs remember past UNREP experiences, and is, thus, likely to omit the subtleties of shiphandling.

The SWOs then performed an UNREP on the VE test-bed. If the VE provided any sense of realism at all, it should be useful as a memory aid and remind the SWOs of subtle cues they use to handle the ship.

A debrief session completed the process. During the debrief, a videotape of the SWOs' VE UNREP performance was used to provide a specific context for KE. To make the debrief even more context-specific, the videotapes of their performance were created by using the video-feed from the test-bed to record their viewpoints from a first-person perspective. Commands and requests for information were also recorded. With the videotape as a guide for the interview process, the SWOs explained what they were doing, what they were looking at, and why.

Testing with this procedure began by having two SWOs perform a VE UNREP. During the scenario, the SWOs were required to request information normally obtained from navigational instruments. This arrangement was designed to provide a record of the navigational information required during an UNREP.

As expected, the video debrief produced data that extended the information concerning the perceptual and timing aspects of shiphandling beyond that previously obtained with verbally based KE. Examples include monitoring the oiler's wake to guide lateral positioning during the approach, and comparing the distance between the bows of ownship and the oiler at the end of the approach to determine changes in lateral separation. The data also addressed problems the SWOs encountered with the VE, such as the difficulty of determining the direction of ownship's bow because there was not a reference marker on it. Other recommendations included:

- elevating the officer's viewpoint in the simulation,
- increasing the contrast between the side and stern of the oiler,
- starting the oiler's wake closer to its bow,
- quickening the response of ownship to engine and rudder inputs, and
- including hydrodynamic effects.

Many of these suggestions were implemented on the test-bed. Following this, two additional SWOs performed a VE UNREP and video debrief utilizing the improved UNREP scenario.

During the debrief, one SWO demonstrated the strategy of maintaining a constant head and body position while looking at ownship's bow, and the oiler's wake and bow to determine whether the appropriate lateral separation had been achieved. This strategy was used to control visual perspective so that small changes in relative motion (i.e., motion parallax) could be detected more accurately. It was never mentioned in any of the verbal techniques employed previously. Range and bearing information was utilized to confirm the accuracy of the perceptual estimation. The SWO also provided feedback on the fidelity of the simulation. For example, the width of the oiler's wake appeared too wide on the approach, but it seemed realistic when alongside. The other SWO calculated lateral separation during the approach, and requested information regarding range and relative bearing. This SWO also reported estimating the rate of parallax while alongside by using the port and starboard stantions on the oiler to detect fore/aft movement.

Based on the feedback received from this KE procedure, the following modifications were made to the simulation. The Venturi effect was added, the P&D line was repositioned to the forecastle, and the scenario was modified such that the oiler's wake was completely rendered at the scenario's start. Other refinements included display of the ships' pitch and roll in response to sea-state, and enhanced scenario record and replay capabilities.

In addition, the phases of UNREP were redefined from the three broad phases of approach, alongside, and breakaway to seven specific phases. These phases were defined as follows:

1. slide-out -- recording of this phase began when the first communication occurred (at approximately 1000 yards astern of the oiler) until ownship was 600 yards astern of the oiler,

2. approach – from 600 yards astern (the approximate "waiting station" position) until the bow of ownship crossed the stern of the oiler,

3. slide-in -- bow crossing until P&D line across,

4. transition to alongside station -- P&D line across until participant reported being at the alongside station,

5. alongside -- on station until P&D line removed,

6. breakaway -- P&D line removed until stern of ownship crossed the bow of the oiler, and

7. resume prior course -- stern crossing to the end of the data.

SIMULATION VALIDATION STUDY

Following scenario development, data were collected on 18 shiphandlers of varying experience. Our specific goals were: (1) analyze the information processing requirements of conning an UNREP, and (2) validate the VETT simulation of a CG-47 class cruiser, the Venturi effect, standard commands, and various perceptual cues a conning officer uses during UNREP.

Participants completed two sets of questionnaires and performed two UNREP scenarios on the VETT test-bed. The questionnaires addressed demographics, and user acceptance of the simulation. The procedure required approximately 1.5 hours for each participant to complete.

Test-Bed

In order to carry the system to SWOS, the hardware components were changed to an. Octane with two MIPS R10000 processors, MXI Graphics, Octane Channel Option, and an Indigo^2 Impact MIPS R10000 by *Silicon Graphics, Inc.* Head tracking was provided by an IS600 Inertial Tracker by *Intersense*. The display and software components remained the same.

User Acceptance

To assess whether the fidelity of the simulation was acceptable, participants were asked to rate the accuracy of the VE in terms of: visual representation, visual detail, the Venturi effect, and acceleration /deceleration (see Figure 1).

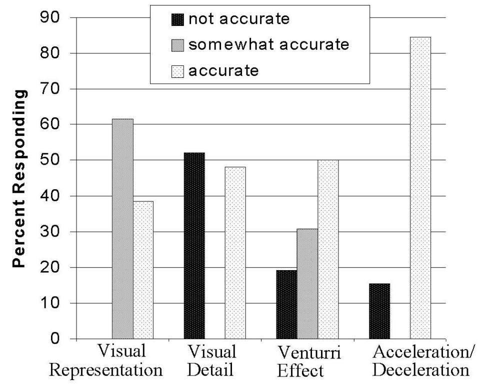

Figure 1: User Acceptance Data

Visual Representation – This referred to whether the wake of the ships, bow of ownship, Phone and Distance Line, and size of the ships were accurately represented.

Visual Detail – Subjects were asked whether there were any visual points of reference missing from the simulation needed to complete an UNREP.

Venturi Effect – This referred to whether the forces associated with the Venturi effect were accurately modeled.

Acceleration/Deceleration – This referred to whether the surge (acceleration and deceleration) of ownship was accurately modeled.

Transfer Validation

If the simulation was accurately modeled, then it should be expected that more skilled shiphandlers would perform better on the UNREP task than those who are less skilled. Skill was divided into three levels of experience: none, specific, and general. These translate as: 1) no shiphandling experience, 2) cruiser experience, and 3) shiphandling experience on platforms other than cruiser class, respectively.

Upon completion of the experiment, a graphic representation of each participant's course track was generated. These course tracks were then rated on a 100-point scale by five experienced shiphandlers. The performance measure for each course track was the mean rating of the five experts. The mean ratings for the three experience levels are shown in Figure 2.

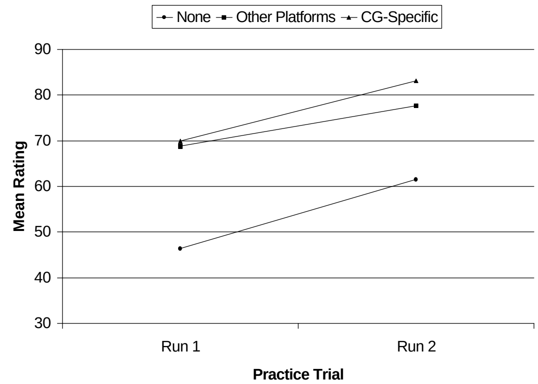

Figure 2: Mean UNREP Performance Rating.

The main effects of shiphandling experience, $\underline{F}(2, 15) = 5.50$, $\underline{p} < .016$, and practice runs; $\underline{F}(1, 15) = 7.47$, $\underline{p} < .015$, were reliable. Participants with shiphandling experience, CG-specific and other platform types, received higher ratings than those with no shiphandling experience (see Figure 2). There was no difference in performance of experienced shiphandlers. There was also an improvement from the first to the second run, indicating that learning had occurred.

CONCLUSION

VE technology, being software based, fosters rapid prototyping. Rapid prototyping and the realism provided by VEs offer training specialists new, and possibly more robust, methods for performing KE and analyzing tasks that are difficult to describe.

By using VEs for KE, simulation development and training program design can proceed in an iterative, user-centered manner. The progression from basic simulations to more sophisticated simulations synchronizes KE and simulation development efforts. It also facilitates the acquisition of information about subtle cues required for task performance. This is because information about such cues is more likely to be elicited when a SME tries to perform a task in a simulation where they are conspicuously absent. By using comments about these cues, KE efforts can contribute directly to decisions regarding simulation fidelity, and subsequent prototype development. In this way, the extra cost of fidelity higher than that required for training may be avoided.

The controlled environment provided by VEs also allows training specialists to observe dangerous tasks being performed by people of widely varying skill levels without compromising safety. This facilitates the development of objective performance measures sensitive to differences in skill.

VEs may actually alter the process by which training systems are designed. Their use as a KE tool involves end-users at the beginning of the training system design process. User-centered design, coupled with the integrated efforts of training and simulation specialists, should produce better training systems.

Information (e.g., perceptual cues) used by expert shiphandlers was revealed during our VE-enabled approach to KE. This type of information could not be obtained through more traditional, verbally based KE techniques. This information was then used to design a VE shiphandling simulation that reliably separated shiphandlers into two skill levels. This suggests that the iterative process of using VEs as KE tools was effective for obtaining information related to the latter stages of skill development (i.e., the autonomous stage).

REFERENCES

Anderson, J.R. (1987). Skill acquisition: Compilation of weak-method problem solutions. *Psychological Review, 94*, 192-210.

Crenshaw, R.S., Jr., CAPT, U.S. Navy (Retired). (1975). *Naval Shiphandling*. Annapolis, Maryland: Naval Institute Press.

Gordon, S.E. (1992). Implications of cognitive theory for knowledge acquisition. In Hoffman, R.R. (Ed.), *The Psychology of Expertise: Cognitive Research and Empirical AI* (pp. 99-120). Mahwah, N.J.: Erlbaum.

Hoffman, R.R., Shadbolt, N.R., Burton, A.M., & Klein, G. (1995). Eliciting knowledge from experts: A methodological analysis. *Organizational Behavior and Human Decision Processes*, Vol. 62 (2), 129-158.

McGraw, K.L. & Harbison-Briggs, K. (1989). *Knowledge Acquisition: Principles and Guidelines*. Englewood Cliffs, N.J.: Prentice Hall.

Schumacher, R.M. & Czerwinski, M.P. (1992). Mental models and the acquisition of expert knowledge. In Hoffman, R.R. (Ed.), *The Psychology of Expertise: Cognitive Research and Empirical AI*, (pp. 61-79). Mahwah, N.J.: Erlbaum.

VanLehn, K. (1996). Cognitive skill acquisition. *Annual Review of Psychology, 47,* 513-539.