

## Verbal Coaching During a Real-time Task

Bruce Roberts, Nicholas Pioch, William Ferguson

► **To cite this version:**

Bruce Roberts, Nicholas Pioch, William Ferguson. Verbal Coaching During a Real-time Task. International Journal of Artificial Intelligence in Education (IJAIED), 2000, 11, pp.377-388. hal-00197334

**HAL Id: hal-00197334**

**<https://telearn.archives-ouvertes.fr/hal-00197334>**

Submitted on 14 Dec 2007

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

## **Verbal Coaching During a Real-time Task**

**Bruce Roberts, Nicholas Pioch, and William Ferguson**, *BBN Technologies, 10 Moulton Street, Cambridge, MA 02138*

*e-mail: {broberts, npioch, wferguson}@bbn.com*

**Abstract** TRANSoM is a collaborative effort among university and industry researchers aimed at producing an intelligent tutoring system for training pilots of remotely operated vehicles (ROVs). ROVs are unmanned, tethered, underwater vehicles used in a range of applications such as inspection, search and salvage, and mine countermeasures. Pilots have to learn to maneuver the ROV, keeping track of its tether and its surroundings, using little more than a video camera and sonar. To minimize workload while a trainee is practicing, the primary mode of feedback during a mission is verbal. The design of this verbal coaching was modeled closely on techniques observed in human instructors. We report on several of the issues faced in implementing this coaching paradigm that make it effective yet unobtrusive.

### **INTRODUCTION**

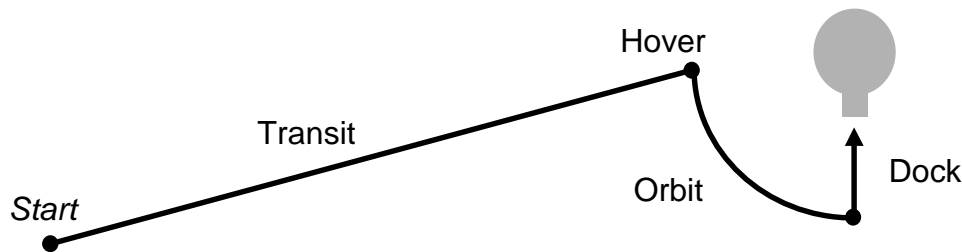
The TRANSoM (Training for Remote Sensing and Manipulation) project is a combined industry and university effort applying the technologies of intelligent tutoring systems and virtual environments to the task of training pilots of remotely operated underwater vehicles (ROVs). ROVs are used widely for scientific exploration, construction, inspection, search and salvage, survey, and other activities in both military and commercial settings. They are controlled by an operator on board ship, which communicates with the ROV via a tether. The pilot controls thrusters on the ROV with a joystick and receives information from the ROV by means of a video camera and sonar mounted on it. A navigation plot, which tracks the ROV with respect to a global reference frame, is sometimes available.

Piloting an ROV is a challenging task despite the aid of video, sonar and guidance systems. As one pilot described it, "You can get so disoriented." In the underwater environment visual capabilities are limited by the natural turbidity of the water and backscatter from the lights needed to illuminate the scene. Unseen currents and the drag of the tether have significant influence on the dynamics of the ROV. Nevertheless, ROV operators must maintain a mental model of the vehicle, its tether, and its physical surroundings based on often incomplete and confusing data.

We are addressing these piloting skills in our training system: first by trying to understand how good pilots construct and maintain their world model, and second by devising a set of apprenticeship learning activities which foster this type of reasoning in novice pilots. As part of this effort, we have built a tutor to develop the maneuvering skills that form the foundation of good piloting. These predominately sensorimotor skills present an instructional challenge to the ITS, not the least of which is the inherently high workload associated with the task itself. We have been exploring the use of verbal feedback as the chief modality for real-time critique. This paper describes our experience integrating verbal coaching within an ITS for ROV maneuvering skills.

### **PRACTICE ENVIRONMENT**

Several component maneuvering tasks have been identified. They differ in the control strategies and information sources a pilot uses during their execution. We assembled these tasks as phases of a single "mission" to serve as the context for practicing maneuvering skills (Figure 1).



**Figure 1.** Maneuvering Task phases (Plan View)

Taken together, these phases of the mission exercise all the basic maneuvering skills identified by our expert pilots. To transit, you must drive in a straight path from your initial position to your destination. To hover, you must keep the vehicle stationary, compensating for any external forces from the current or tether. To orbit, you must circle around a target at a constant distance while maintaining the target in the center of your field of view. To dock, you must move forward along a constrained path and carefully align the vehicle with an external object, in this case, a protruding portal on the target.

Each phase has its own information gathering and control strategies. For example, at the beginning of the transit, the vehicle is too far away to see the target in the video camera so the pilot must rely on the navigation plot's view of the ROV. In later phases such as hover and dock, when the target is nearby, the video camera's view of the target is the primary focus of attention. These different strategies lead to distinct classes of mistakes, and coaching comments must be chosen appropriately to point out problems and suggest control actions that would alleviate them.

We have created a virtual environment (VE) that incorporates the salient dynamical, visual, and aural features of the ROV's underwater environment (e.g., obstacles, strength and direction of current, visibility, particulate flow, and lighting) and makes these features easily reconfigurable by the ITS. The ITS guides a student's progress through a curriculum of increasingly challenging situations. The VE allows us to augment the world with additional features chosen for their instructional value and to populate the world with agents responsible for critiquing, guiding, and reviewing the student's progress. The role these agents play in instruction is borrowed from research on the value of coached practice and apprenticeship learning and the successful application of this approach in intelligent tutoring systems (Collins, Brown, & Newman, 1987; Lesgold et al, 1988).

The general coaching framework is to demonstrate the mission, allow the student to practice the mission, and then review the student's performance.

- The *demonstration* features an expert flying the mission accompanied by a talk-through of the important aspects of the mission. It is an opportunity to present to the student explicit strategies for control (e.g., dealing with currents) and information assessment (e.g., how to use the displays).
- During *practice* the student can solicit help in the form of reminders about maneuvering goals and how to achieve them. Various "bird's eye" views of the ROV and its path are available to reorient the pilot. Unsolicited coaching is triggered by monitors that assess the student's progress and deviations from expert performance, such as drifting too far from the prescribed path. This feedback is delivered verbally in order to minimize the workload impact on the student, whose visual attention is already divided among several displays.
- During the *review* phase of instruction, the student can replay portions of the mission, see a summary of measures compared with expert performance, see a view of the actual path and an expert's path, and re-fly segments of the mission where errors occurred.

The practice environment includes several visual and aural features that act as instructional scaffolding and invite the student's own interpretation of what went wrong and how to improve his or her performance. In general they provide extra information to improve awareness of a) the vehicle in its environment, b) the consequences of pilot actions, c) the effects of external forces, and d) information available in the typical ROV displays.

Figure 2 shows the simulated ROV running on a Silicon Graphics workstation.<sup>1</sup> A joystick controller borrowed from a physical ROV is used for input. The screen has a collection of displays used by an ROV operator: a video camera, a sonar, a map, and assorted instruments showing, for example, compass heading, depth, and tether length. The dynamic response of the vehicle matches that of a physical ROV used by the project for task analysis and evaluation. The TRANSoM simulation has been used as a testbed to explore applications of VE in training, exploring spatialized audio and the fully immersive visual experience afforded by head-mounted displays. For more details on the implementation of the VE, see Pioch, Roberts, and Zeltzer (1997).



**Figure 2.** Simulated ROV operator station

## VERBAL COACHING

Task analysis led us to investigate ways of incorporating verbal coaching in addition to visual coaching aids as part of the maneuvering task training. The fact that a pilot's hands and eyes are fully occupied with flying the ROV makes it problematic to introduce additional visual features and practically impossible to present the student with something he or she would have to read. On the other hand, the speech modality is relatively underutilized in normal ROV operations, where it is used to coordinate with other crew members, such as the sonar operator or tether handler. Although the demonstration and review phases of the session permit extended textual and graphical interactions, the student is not actually *doing* the task at that point in the instruction. We feel that timely feedback and suggestions during the mission itself are an extremely important form of coaching. This presented a unique opportunity to build on the

<sup>1</sup> OpenInventor™ was used to model the 3-D environment in TRANSoM.

relatively few precedents that exist in applying ITS technology to real-time tasks (Ritter & Feurzeig, 1988; Eliot & Woolf, 1994).<sup>2</sup>

Several challenges arose in the course of creating verbal coaching for the maneuvering task.

- Simplistic schemes that trigger commentary based on exceeding a performance threshold rapidly become intrusive and distracting. To be acceptable to a student, the tutor must offer a dialogue, not a collection of repetitive or disconnected utterances.
- We wanted to structure dialogues in the tutor to reflect those often used by instructors that we had observed. One common technique was to give the student just enough information to allow him or her to make a correction without giving too much away. For example, “Check your depth” allows the student to diagnose the problem and correct it, while “Apply upward thrust” tells the student to correct the deviation without realizing what was wrong or how the situation arose.
- Multiple deviations can occur simultaneously or in rapid succession. Coaching responses must be coordinated to avoid a cacophony of suggestions and feedback.
- Behavior that is questionable at one moment may be entirely appropriate in other circumstances. An obvious example is the difference in the primary maneuvering goals during hover (stay motionless) and transit (make steady progress toward the destination). The mix of the coaches must vary as the task progresses.

The verbal coaching implementation must allow for the authoring and tuning of tutor behavior in light of these considerations. We have developed a theoretical framework and implementation that monitors violations and the student's response to corrective feedback. The framework defines a collection of coaching agents focused on particular performance features, imposes a common dialogue structure on the feedback generated by these agents, and incorporates a multi-threaded control structure to coordinate the activity of the agents. Follow-up feedback to uncorrected problems becomes increasingly specific, and multiple simultaneous violations are managed by establishing priorities among the violations so that feedback triggered by multiple problems is not interleaved.

## SENSORS AND COACHES

Sensors provide a mechanism for monitoring student performance during a practice session. They act as instrumentation on the simulation to provide information needed by the coaches. They are also the basis for curriculum decisions that control progress through a sequence of scenarios and the source for the student assessment provided during the review stage of a trial. Sensors track properties of the mission scenario as it unfolds and detect conditions that trigger coaching responses to situations as they arise. Sensors have direct access to control movements and other control panel inputs, the status of the ROV (*x*, *y*, *z*, heading, and velocity), simulated time, collisions, and inputs from the student in the form of button presses or simple speech commands. The output of sensors can be recorded for later analysis, monitored with respect to a threshold, and have other properties derived from it.

The following sensors are currently used by the tutor:

- *Path Sensor*—horizontal deviation from a prescribed path
- *Depth Sensor*—vertical deviation from a prescribed path
- *Bearing Sensor*—angular deviation between ROV heading and bearing to target
- *Hover Sensor*—distance to a prescribed hover point in 3-D space

---

<sup>2</sup> Of related interest is STEVE (Johnson et al, 1998), a VE based pedagogical agent that includes speech output. STEVE manifests itself as a simulated human figure in the visual scene to act as a guide for the student. The focus of this work to date has been on the operation and repair of machinery, not real-time tasks.

- *Progress Sensor*—rate of motion with respect to a prescribed path
- *Control (Smoothness) Sensor*—variance of the commanded thrust
- *Collision Sensor*—collisions with the sea floor or other objects

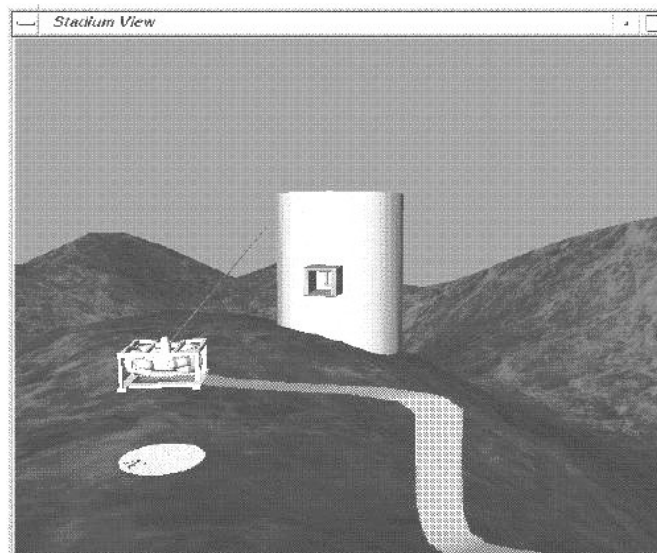
Similarly, a coach uses one or more sensors to monitor a particular aspect of the student's maneuvering. The coach's interaction with the student follows a "correction" template, which captures a common type of interaction observed between human tutors and their students (Freedman, 1996).

A coach is defined as a linear sequence of actions. It first alerts the student to the type of the error ("What's your heading?"). If, after a time, the error goes uncorrected, the coach tells the student the specific nature of the error ("Your heading is drifting to the right."). If, after another interval, the error continues unabated, the coach finally tells the student exactly how to remedy the error ("Twist left 5 degrees."). All the while, the coach is checking to see if the student corrects the problem. If this occurs, the coach first acknowledges the student's response and then restarts itself after a refractory period during which it is disabled.

Coaches are tied into the phase structure of the mission; each phase has its own possible student deviations and so has its own set of coaches to help the student correct them. The tutor manages transitions between the phases by shutting down one set of coaches and activating another. The tutor author specifies how these phase transitions are to be detected and which coaches should be active for each phase.

Although our primary emphasis has been on how to provide unsolicited coaching for a student in trouble, we also allow the student to ask for help directly. Currently, the student can ask these four general questions:

- *Where am I?* Show a "Stadium View" of the scene for the student to check the location of the ROV relative to the target (Figure 3). The tutor's response to this question is purely visual. We have not found it necessary to supplement the information in the Stadium View verbally. To guard against a student becoming overly reliant on the Stadium View, the tutor removes the view after a brief time.



**Figure 3.** Stadium View showing the ROV, its path, its direction of gaze, a compass, and the target.

- *What should I be doing?* Describe the current goal. For example, during the transit phase the tutor tells the student to head directly toward the target, keeping it constantly in view.

- *How do I do it?* Describe what actions to take and what information is needed to achieve the current goal. For example, during the transit phase the tutor tells the student to use the navigation plot to monitor the ROV's track and to use side-stick to compensate for any current.
- *Tell me more?* Elaborate the most recent unsolicited coaching message. It is available whenever a coach has been triggered. Currently, the clarification consists of the next most specific message that would be given if the student failed to correct the situation.

## DEFINING A COACHING AGENT

To provide versatility in defining verbal coaching dialogues, a multi-threaded control mechanism was implemented along with a set of process primitives to allow parallel execution and non-local returns from within executing process threads. Using these primitives allowed coaching agents to be defined separately while at the same time establishing interdependencies to be maintained during their execution. In addition, coaching agents can be nested; that is, agents can spawn sub-agents and still retain control of the overall flow of the coaching activity.

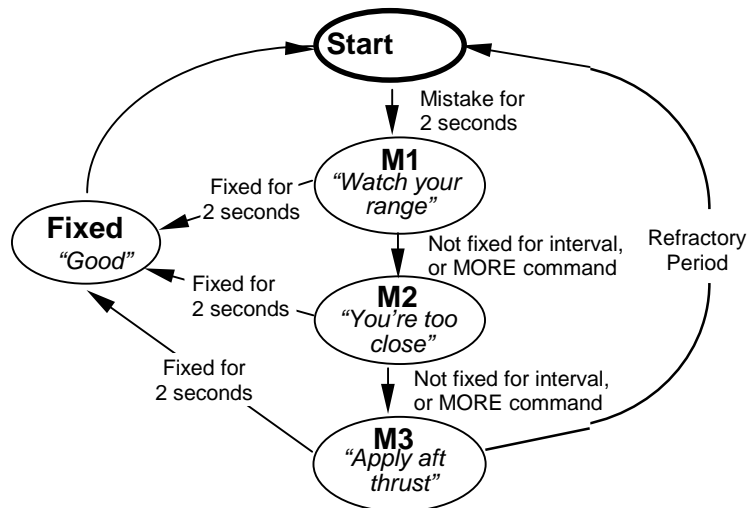
Defining a coach consists of writing a sequence of actions or condition-action pairs that execute in sequence. Two special types of primitive actions mediate the simultaneous execution of multiple coaches. The primitive action *race*, borrowed from the ACTOR language (Agha, 1986), spawns multiple threads of execution and terminates when any one of its spawned threads terminates:

```
race <coach1> <coach2> ...
```

The following family of primitive actions pause execution of their thread until a condition is satisfied:

```
wait <duration>
wait-for <condition>
wait-until <duration> <condition>
```

The action *wait* simply pauses for the specified <duration>; the action *wait-for* pauses until the specified <condition> becomes true; the action *wait-until* pauses until the specified <condition> remains true for the <duration>.



**Figure 4.** Finite State Machine representation of the “Maintain Range” coach, which is active during the orbit phase.

A typical coach definition is shown in Figure 4 as a Finite State Machine. This coach is used to monitor the range of the ROV to the target during the orbit phase. When activated (**Start**) it will "wait-until" a triggering condition holds, in this case, the range sensor going above the sensor threshold for a minimum time. Once triggered, it outputs the first level coaching message (**M1**), and then enters a "race" to detect either a corrective action, an interval during which the triggering condition remains unrectified, or a button press asking for more information. In the first case the coach acknowledges the correction (**Fixed**) and resumes monitoring (**Start**). In the latter two cases, it outputs the next coaching message (**M2**); and similarly the third level coaching message (**M3**). After the final message, the coach enters a refractory period before it resumes monitoring.

The same parallel control structure has proved to be valuable for managing multiple simultaneous coaches during the practice stage of a trial. For example, a portion of the definition for the orbit phase is shown in Figure 5. It instantiates three coaches, assigning each a priority, an interval to pause between coaching levels, and an associated sensor. The coaches remain active until the phase terminates; i.e., the orbit is completed or the student quits. A coach in any state will be interrupted when a higher priority coach is triggered. In this case, the lower priority coach enters a dormant state until the interrupting coach completes. Then, after a global refractory period, all coaches are reset to their initial state (**Start**).

```
(race
  (run-coaches
    (maintain-target-view-coach high 10
      (orbit-bearing-sensor))
    (maintain-range-coach medium 12
      (orbit-range-sensor))
    (maintain-depth-coach low 8
      (orbit-depth-sensor)))
  (wait-for (orbit-complete? (orbit-bearing-sensor)))
  (wait-for (quit-button-pressed)))
```

**Figure 5.** Typical Phase definition

These examples illustrate several temporal parameters that have been found helpful for adjusting the behavior of the verbal coaching to make it more effective and acceptable by students. Collectively they control the intrusiveness and sense of urgency associated with the coaching. The parameters include:

- *Interval between successive levels of coaching within a single coach.* The coach's correction template defines a progression of utterances containing increasingly specific advice to the student. A parameter determines the amount of time the coach waits for a corrective response after each utterance. For the maneuvering task, coaches allow the student about 10 seconds to respond before offering the next piece of advice.
- *Interval between successive activations of a single coach.* To avoid appearing to harp, each coach enters a refractory period after it has finished. For the maneuvering task, a coach cannot be triggered again for about 7 seconds.
- *Interval between successive utterances regardless of their source.* To avoid too many interruptions in general, a global refractory period was added that inhibits *any* coaching for 5 seconds after a coach has finished, regardless of its source.
- Other parameters are defined for each coach which convey its apparent seriousness or importance to the student:
- *Priority.* A scalar parameter assigned to each coach determines whether one coach can interrupt another. Once a coach has been triggered, it "keeps the floor" until it terminates; i.e., it becomes the sole discourse focus. Coaches with lower priorities cannot break into the sequence of coaching defined for the active coach. However, if a condition appears that triggers a higher priority coach, it effectively grabs the focus by



terminating the active coach. For example, keeping the target in view when orbiting is more important than maintaining the proper distance and depth.

- *Leniency.* In theory, the condition under which a coach is triggered can be an arbitrarily complex function of the student's performance. In practice, simpler functions tend to be preferred since they are more easily computed and more easily understood by the student. Typically a sensor associated with each coach continuously monitors the simulation and computes derived values that are compared against predefined bounds. To this is added a "grace period," the amount of time a sensor value must be out of bounds before triggering the coach. For example, during orbiting the "Maintain Range" coach is triggered when the student drifts 2 meters closer or farther from the target for more than 2 seconds.

The values for these parameters have been derived empirically, starting from a priori estimates and refined through user feedback, evaluation by expert pilots, and our ongoing experience with the tutor. User validation and software verification continue to be critical activities in developing an effective and acceptable verbal coaching capability.

Currently, all parameter values are fixed for the duration of a session. This does not preclude incorporating adaptive changes as our theory of how to manage these details of verbal coaching matures. The tutor might adjust the priority and the leniency of a coach based on the level of mastery of the student, for example, to lower the priority of well learned skills (thus shifting the training focus to weaker skills) or to lower the leniency (thus "raising the bar" and urging the student toward higher standards of performance). For novices, the tutor might turn some coaches off altogether in order to focus on a few key skills at a time.

## LESSONS LEARNED

Our experience with the verbal coaching scheme has led to the following additional qualitative suggestions for integrating verbal coaching in a tutor:

- *Discourse cues.* Two situations arise during the corrective verbal coaching that benefit from such cues:
  - a) *Acknowledgement.* Whenever the student corrects the problem, the coach utters an acknowledgement of the form "Good. You have..." This signals the end of the interaction with that activation of the coach.
  - b) *Interruption.* Whenever a higher priority coach interrupts an active coach, it precedes its first-level advice (stating the problem) with the phrase, "New problem." This signals the end of the interaction with the interrupted coach and the introduction of another.
- *Short utterances.* Brief coaching messages are less intrusive than lengthy ones. The correction template in Figure 4 provides an example of how to decompose coaching feedback into smaller statements. The student's ability to ask for elaboration or clarification (using the *Tell Me More* command) provides some control of when and how much information is delivered.
- *Recorded speech.* Despite advances in speech generation and text-to-speech technology, our subjects prefer the natural quality of pre-recorded speech. We continue to use recorded speech fragments wherever possible.
- *Simultaneous visual cues.* When a coach is triggered, the visual cues associated with being able to recognize and correct the problem are highlighted during the coach's initial utterance. For example, on hearing "What's your heading?" the student simultaneously sees the compass and target bearing indicators illuminate. This is meant to suggest that by paying increased attention to these information sources the student will be better able to detect and correct deviations.

- *Corrections.* In response to suggestions from our expert advisors, we added an additional layer of monitoring to each of the coaches in order to detect two prevalent types of novice reaction to errors: over-correction (“Too far!”) and reverse-correction (“Wrong way!”). There is a cost to this second layer of coaching, however. Adding conditions to the set of situations being monitored by a coach requires careful tuning in order to balance the timeliness of a coach’s response against its confidence in interpreting the subject’s actions. For example, waiting to see if the pilot will over-correct requires delaying the response to *any* correction, including a good one.

## EXTENSIONS

The verbal coaching machinery we have developed can be extended to include additional features of natural instructional dialogues. The correction template can be enhanced to include other parts of the discourse context:

- *Pedagogical context.* It is useful instructionally to connect feedback to prior examples of the same class of error, in order to emphasize recurring errors or situations; e.g., “You’re too low *again*.” (Moore, 1993; Rosenblum & Moore, 1993)
- *Control context.* The shift in focus required when one coach interrupts another could be made less abrupt by explicitly acknowledging the status of the other currently active coaches: e.g., “Forget your depth, you’re about to collide with that mine!”
- *User initiative.* We have added a mechanism for the student to acknowledge awareness of a coached problem. This inhibits subsequent redundant messages and provides a greater sense of control over the dialogue with the coach.

New classes of templates can also be defined. For example, a control template actively talks the student through a maneuver; e.g., “Get ready to stop...OK, ease back on the stick now...not too much...good.”

Recent development in TRANSoM has shifted focus from chiefly sensorimotor skills involved in maneuvering an ROV to the cognitive skills required for maintaining situation awareness (Fletcher et al, in review). Task analysis revealed four primary collections of skills: sensor integration, location awareness, tether awareness, and system awareness (using the status displays to recognize faulty behavior). A prototypical search task provides the context for practicing situation awareness. In this task the student is launched near a well-marked object and must use the sonar and video displays alone to detect potential targets, record their locations on a sketch map, fly out to them, record their visual attributes, and return to the startpoint. This cycle of detect, prosecute, identify, and return is repeated until all interesting targets have been mapped. Our experience with expert approaches to the search task revealed that maintaining a sketch map—using it to record discoveries and guide subsequent navigation decisions—was key to developing an awareness of the vehicle and its surroundings. We have extended the TRANSoM tutor’s suite of performance sensors and coaching interventions to address these new skills. For example, one coach detects potential tether infractions such as too many twists, excessive slackness or tautness, and snags around objects. Another coach specializes in recovering lost sonar contacts by suggesting possible causes such as depth, heading, occlusions, or sonar settings. Additionally, two new classes of verbal interventions have been found valuable for this task:

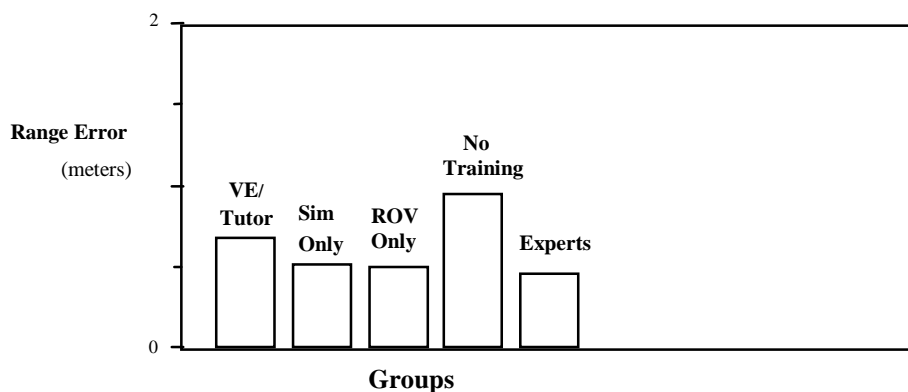
- *Prompts.* Prompts are short exchanges designed to assist and monitor a student’s progress through the task. They serve as reminders for the current goal and its associated actions. For example, if a student leaves the startpoint without selecting a target to prosecute, the tutor will prompt him to do so. The student responds by selecting a target on his sketch map. Prompts serve two roles: (1) to encourage good habits—in this case, always having a definite target in mind, and (2) to verify the tutor’s model of the student’s intention.

- *Probes.* Probes are questions asked by the tutor during the practice session. They are instructional by virtue of calling attention to features of the environment that should be constantly monitored. As quizzes, they also elicit facts about the student's current model of his environment. Because they interrupt the session and require an immediate response, the simulation is paused while the student answers. After the student responds, the tutor acknowledges (or supplies) the correct answer and resumes the simulation. For example, the tutor may ask "What object corresponds to this (highlighted) blip (on the sonar display)?" which is answered by selecting a target on the map, or "Where is (a named object)?" which is answered by choosing a rough direction and distance in a simple dialog window. Probes are restricted to periods of relative inactivity when the intrinsic task workload is low, such as during transit to or from a target.

## EVALUATION

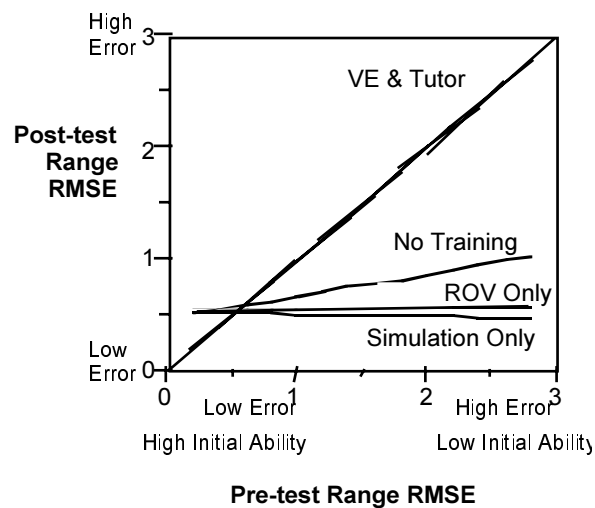
A training transfer experiment to study the effectiveness of the maneuvering skills tutor is reported briefly in (Fletcher, 1998) and in greater detail in (Fletcher & Roberts, 1998). Four groups of eight subjects each were tested: full-training (the complete VE/Tutor environment), simulation-only (no tutoring), ROV-only (driving an actual ROV, no tutoring), and no-training. Pre- and Post-tests were administered using an ROV. Training and transfer trials were modeled after the Maneuvering Task described earlier. An additional Inspection Task was introduced in the Post-test in which the subject followed a circuitous pipe, finding and recording numbers and letters spaced along it at irregular intervals. Measures of performance were the same as used in the tutor to assess skills and trigger coaching; for example, the Hover sensor, mentioned earlier, which reports root-mean-squared deviation from an ideal position.

The most significant test results were found in transferring skills learned in the Maneuvering Task to the Inspection Task (Figure 6). All training groups showed significant positive transfer, as shown for example by the range root-mean-squared error, which measured ability to control the distance from the pipe as one moves along. All training groups showed significantly improved performance over the No Training group, and reported consistently lower workloads. In fact, their performance approached that of expert pilots. However, the full-training group did not differ significantly from the simulation-only and ROV-only groups.



**Figure 6.** Training Transfer results

Further analysis revealed another anomaly. Figure 7 compares Pre-test and Post-test range measures for each group during the orbit phase of the Maneuvering Task. What is striking is the apparent lack of benefit of the tutoring for those subjects with initial low ability (high error). In this case, the training appeared to have no effect; the post-test errors were equal to the pre-test errors. Both simulation-only and ROV-only groups showed improvement in the form of reduced errors from pre-test to post-test.



**Figure 7.** Pre-test versus Post-test performance

The verbal coaching paradigm was one of several instructional features included in the version of the tutor used for the study. For example, it also included a problem progression scheme that recommended more challenging situations based on demonstrated mastery of current conditions. In fact, several of the full-training group failed to complete the curriculum in the available time and consequently were not exposed to the most extreme conditions of high current and low visibility. This is one possible explanation for the observed results. Another explanation may be that the form and timing of the verbal feedback inadvertently rewarded inappropriate behaviors, or that the low ability students became overly dependent on it. So far we have concentrated primarily on the micro-structure of natural coaching utterances, but may need to devote similar resources to the macro-structure. A cross-session mechanism for dynamically tailoring the selection of a suite of coaches and priority assignments according to the student's abilities or preferences needs to be developed and tested. For example, since we already suggest a training focus for each trial, based on prior performance, we could limit the coaching interactions to skills associated with that single focus.

The transfer study also revealed a strong and surprisingly persistent effect of visualization aptitude.<sup>3</sup> In general, one expects variability in performance to be diminished by training. However, the performance of the full-training group was the most variable of all the treatment groups and was found to be significantly related to visualization aptitude, unlike any of the other groups. One overall lesson is that despite the high praise and enthusiasm of expert subjects, our approach to coaching a real-time sensorimotor task may have shortcomings that are only revealed through the rigor of a transfer study. Additional experiments will be necessary to understand the sources of these findings.

## CONCLUSIONS

We have shown how a general verbal coaching paradigm has contributed to the instructional effectiveness and reconfigurability of an ITS in a real-time task. The style and format of the coaching was driven largely by observations of expert pilots and instructors. Further study is needed to refine the coaching parameters and policies. TRANSOM's coaching template

<sup>3</sup> Visualization ability was measured using the Paper Folding (VZ-2) and Card Rotation (S-1) tests from the Kit of Factor-Referenced Cognitive Tests published by Educational Testing Service.

framework could be adapted for use in training systems in other real-time maneuvering tasks, such as piloting unmanned aerial vehicles, submarine navigation, or flight simulators.

### Acknowledgements

The authors wish to acknowledge the contributions of the entire TRANSOM team: Richard Pew, Yvette Tenney and Jason Vantomme (BBN); Stewart Harris, Barbara Fletcher, and Jason Fritz (Imetrix); Nat Durlach, David Zeltzer, Johnny Yoon, Jonathan Pfautz and David Manowitz (MIT); Bruce Perrin, Barbara Barnett, James Curtin, and Katja Helbing (Boeing); Sandra Katz and Jenny Hwang (LRDC). This work was sponsored under the Office of Naval Research contract N00014-95-C-0181. Terry Allard and Harold Hawkins at ONR deserve special thanks for their continuing support.

### References

- Agha, G. (1986). *Actors: A Model of Concurrent Computation in Distributed Systems*. Cambridge, MA: MIT Press.
- Collins, A., Brown, J.S., and Newman, S.E. (1987). *Cognitive Apprenticeship: Teaching the Craft of Reading, Writing and Mathematics*. In: Resnick, L.B. (ed.): *Cognition and Instruction: Issues and Agendas*. Hillsdale, NJ: Lawrence Erlbaum Associates.
- Eliot, C., and Woolf, B.P. (1994). *Reasoning about the User within a Simulation-based Real-time Training System*. Proceedings of the Fourth International Conference on User Modeling. Bedford, MA: The MITRE Corporation.
- Fletcher, B. (1998). *TRANSOM ROV Pilot Training: Test Results*. Underwater Intervention '98 Conference Proceedings. Washington, D.C.: Marine Technology Society.
- Fletcher, B. and Roberts, B. (1998). *Training ROV Pilots with a VE-based Intelligent Tutoring System*. Proceedings of the International/Industry Training, Simulation and Education Conference. Orlando, FL.
- Fletcher, B., Roberts, B., Pioch, N., Perrin, B., and Barnett, B. *VE-based Training for Situation Awareness*. Submitted for publication in *PRESENCE: Teleoperators and Virtual Environments*.
- Freedman, R. (1996). *Using Tutoring Patterns to Generate More Cohesive Text in an Intelligent Tutoring System*. Proceedings of the Second International Conference on the Learning Sciences. Evanston, IL.
- Johnson, W.L., Rickel, J., Stiles, R., & Munro, A., (1998). *Integrating Pedagogical Agents into Virtual Environments*, *Presence: Teleoperators and Virtual Environments* 7(6):523-546.
- Lesgold, A., LaJoie, S., Bunzo, M., and Eggan, G. (1988). *SHERLOCK: A Coached Practice Environment for an Electronics Troubleshooting Job*. University of Pittsburgh Learning Research and Development Center, Pittsburgh, PA.
- Moore, J.D. (1993). *Indexing and Exploiting a Discourse History to Generate Context-Sensitive Explanations*. Proceedings of the DARPA Human Language Technology Workshop. Princeton, NJ: Morgan Kaufmann Publishers.
- Pioch, N.J., Roberts, B., and Zeltzer, D. (1997). *A Virtual Environment for Learning to Pilot Remotely Operated Vehicles*. Proceedings of the International Conference on Virtual Systems and Multimedia 1997. Los Alamitos, CA: IEEE Computer Society Press.
- Ritter, F. and Feurzeig, W. (1988). *Teaching Real-Time Tactical Thinking*. In: Psotka, J., Massey, L.D., Mutter, S.A. (eds.): *Intelligent Tutoring Systems: Lessons Learned*. Hillsdale, NJ: Lawrence Erlbaum Associates.
- Rosenblum, J. and Moore, J. (1993). *Participating in Instructional Dialogues: Finding and Exploiting Relevant Prior Explanations*. Proceedings of World Conference on Artificial Intelligence and Education. Edinburgh, Scotland.