Virtual Environment Training for *Dismounted* Teams – Technical Challenges

Fred Brooks, Henry Fuchs, Leonard McMillan and Mary Whitton
University of North Carolina at Chapel Hill
Chapel Hill NC 27599-3175
USA

{brooks|fuchs|mcmillan|whitton}@cs.unc.edu

Jan Cannon-Bowers
University of Central Florida
Orlando FL 32816
USA

jancb@dm.ucf.edu

I hear and I forget. I see and I remember.
I do and I understand.
Confucius

1. SUMMARY

Vision. Just as flight simulators enable pilots to safely practice responses to emergencies, the challenge now is to develop virtual environment technology for the training together of small teams on foot—military squads, Coast Guard boarding parties, police, EMTs, emergency room trauma teams, hazmat teams, etc. Such training allows repeated, varied practice. The goal is **you are there; you learn by doing with feedback; you jell as a team by doing together.** First, we must clearly envision what is wanted. This we will call the Immersive Team Trainer (ITT*)*.

The successes of flight and ship bridge simulators encourage us. Their use for training mounted teams is a well understood and trusted accomplishment. Decades of development have brought flight simulators to mature excellence; simulators for training other vehicle crews are rapidly approaching this maturity. Such simulators have been proven to be not only more cost-effective per hour than live vehicle training, but more effective as well, since VEs can provide a higher density of experiences and the chance to practice rare and dangerous scenarios safely. The vision is to extend VEs for training dismounted teams effectively and cost-effectively. This has not been done yet because it is technically much more difficult than immersive training for vehicle crews.

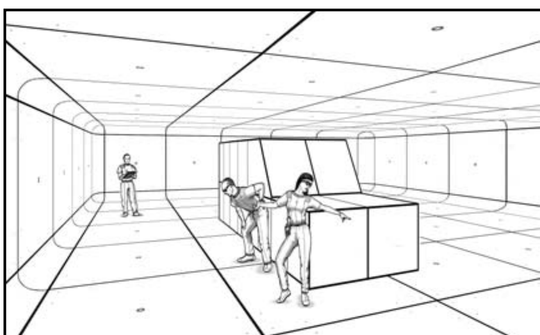
Brooks, F.; Fuchs, H.; McMillan, L.; Whitton, M.; Cannon-Bowers, J. (2006) Virtual Environment Training for *Dismounted* Teams – Technical Challenges. In *Virtual Media for Military Applications* (pp. 22-1 – 22-10). Meeting Proceedings RTO-MP-HFM-136, Paper 22. Neuilly-sur-Seine, France: RTO. Available from: http://www.rto.nato.int/abstracts.asp.

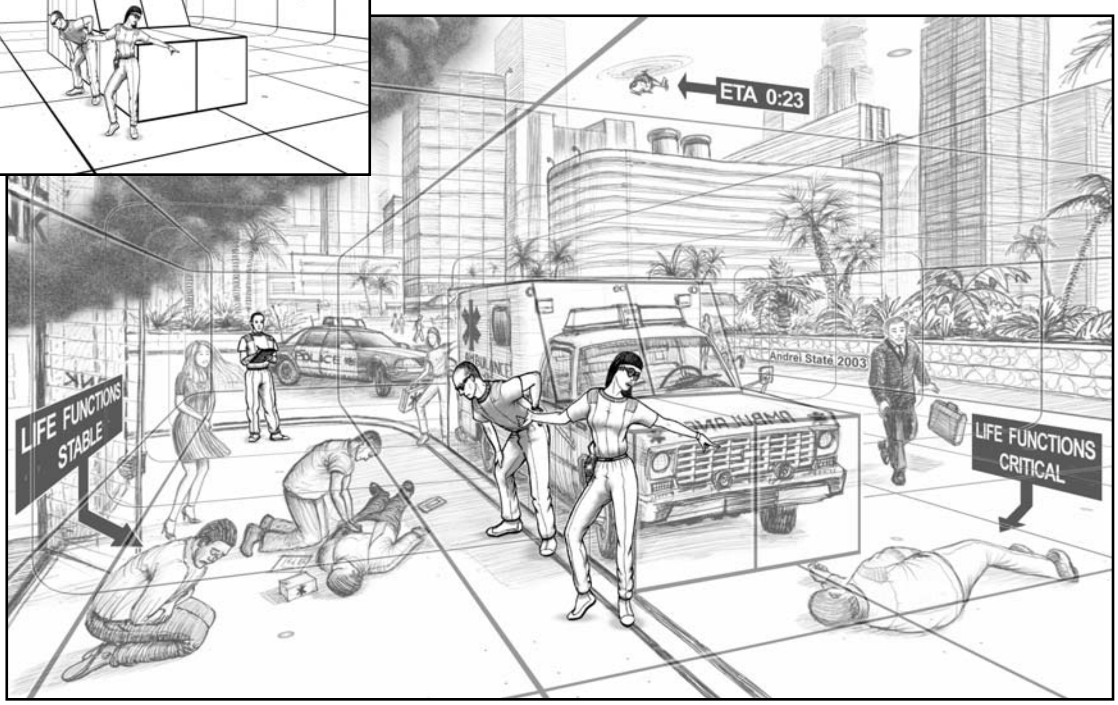
Technical Challenges. The technical challenges of such an extension abound:

	Mounted	Dismounted
Display	Unified for all observers	Distinct for each
Tracking	At most, head position; rarely that	Heads and limbs of each team member
Locomotion	Not required	Extensive, separate per trainee [Usoh, '99]
Tools and Instruments	Physical mockups	Both physical and virtual
Interaction with each other	Speech	Speech, touch, facial expression
Interaction with world	Only vehicle controls	Various manipulanda [Insko, '01]
Inanimate objects	Real touch	Manipulation of physical, virtual objects
People	Speech	Speech, body language
Physics	Well understood vehicle physics	Much more complex, moving objects
Sound	Unified for all	Distinct localization for each

2. THE VISION

Figure 1. The image on page 1 shows the emergency situation as it really happened. The inset image shows the unlit training space with an approximate physical model of the ambulance. The larger image shows the full virtual environment used by the trainees and instructors (Dwg:State).





2.1 The Virtual Environment

The virtual environment itself is easy to envision (Figure 1). Six image-bearing surfaces enclose a small team, such as an EMT team or an infantry squad. Each trainee sees, in stereo, images customized for his dynamically changing viewpoint. Thus the enclosing world looks real—providing realistic occlusion, kinetic depth effect, illumination, stereopsis.

Each trainee can see, hear, and touch his buddies. Every object he must **handle** to perform his task is real or physically mocked up [Insko, 2001]; objects the team need only **behold** are virtual, but as they move, their physics are realistic.

Sounds from the environment are properly localized for each trainee. When the team needs to move beyond the physical space available, the squad leader's walking-in-place translates the real space in the virtual world.

2.2 The Training—Critiqued Practice with Branched Scenarios

The very tracking necessary for individualizing the visual displays yields a superabundance of individual and team performance data. Real-time analysis identifies behavior patterns in trainee speech and head and limb trajectories, compares them to expert and erroneous behavior patterns, and produces diagnostics.

The trainer can use these diagnostics for after-action review and/or can branch the scenario to provide more practice for flubbed responses. Ideally, the behavior patterns can be used to generate automatic after-action reviews and "pinball scores," so that teams can practice and compete without trainers.

An automatic scenario generator takes specified knowledge, skills, and attitudes and generates a rich variety of event-branched scenarios, so that in a compressed time the trainees experience many varied situations that never repeat.

2.3 The VE System

The system itself is transportable, easy to calibrate, built with commercial off-the-shelf computers, projectors, and trackers, and demands only the space of the training volume itself.

Latency. The ITT achieves 50 ms end-to-end latency for complex virtual environments, whereas most CAVEs have 100-250ms latencies. Low latency prevents most simulator sickness.

Model Acquisition. An advance in the ITT is the economical generation of the 3-D models of virtual objects. The ITT developers acquire such models from images taken with video cameras and laser-rangefinders.

3. THE VE TECHNICAL CHALLENGES IN REALIZING THE VISION

3.1 Display

In flight simulators, the entire cabin crew shares one display. In the ITT, each team member needs a personalized stereo view. Yet a critical part of the training concept is that all team members must be physically collocated, share physical objects as well as visual environment, talk to each other, see each other's faces, touch each other. This combination of collocation and individualized displays is a major challenge.

The proposed system must go far beyond today's CAVE®-like immersive environments:

- Front-surface projection radically shrinks the required floorspace. One must, and can, solve the keystoning, depth-of-field, and shadowing problems that have inhibited front-surface projection [Jaynes. '01; Low, '01; Sukthankar,'01].
- It must have much higher resolution imagery.
- It must deliver customized stereo views for at least four persons, whereas only one can get a proper perspective view in today's CAVEs.
- It must track not only heads, but hands, feet, and tools, so as to simulate interactions between trainees and virtual objects.

Individualized Display for Each Trainee. Ceiling-mounted projectors will paint imagery not only on the walls but also on all the other surfaces, even moving real objects [Raskar, '99; Lok, '03]. The system will generate individualized head-tracked stereo imagery for each team member, and time-shuttered glasses will filter out the other trainees' views.

The farther-out technology vision augments the projectors with lenticular autostereoscopic displays that show the proper views from every viewpoint simultaneously. These will appear first on real objects in the ITT, such as the ambulance in Figure 1, later on the walls.

3.2 Tracking

One must track the position, but not the orientation, of the eyes of each trainee in order to render his imagery from the proper viewpoint. Inertial or outward-looking optical trackers yield the high required angular precision [Welch, '01]. It is also necessary to track the hands, feet, and tools of each trainee, so that one can compute interactions with virtual objects. There is promise of new hybrid tracking systems, using spread-spectrum sound and imperceptible structured light [Fuchs, '05], as well as skeletal models fitted to visual hulls from video silhouettes [Matusik, '00], and models generated from reduced marker sets [Liu 06].

3.3 After-Action Review

The very tracking necessary for calculating the proper behavior of objects also yields massive amounts of data about the actions of the trainees. The system should capture this data for after-action review. This data should be reduced to behavioral descriptors that can cue instructors for after-action review. The current path-characterization work on the ONR Virté Project is a first step [Whitton, '05].

In the longer run, these behavioral descriptors should be automatically harnessed to yield feedback to help team members train themselves and each other. The cost and logistical complexity of having instructors is one of today's inhibitors to frequent team-training practice. Even a single-variable "pinball score" (and a list of high scores to date) provides feedback and motivation that enhances training.

3.4 Model Acquisition

A substantial cost in building immersive scenarios is the development of models of the environment and objects in it. Today this is done essentially by computer-assisted-design (CAD) modeling: tedious, time-consuming, and costly. For environments that exist, the models can today be economically acquired from laser-rangefinder images combined with color photography, and from moving video cameras [Pollefeys, '02]. These techniques can also be used to acquire models of separate objects, which are then manually combined into virtual environments.

3.5 Scenario Acquisition

Besides the instructor-guided automatic scenario generation described in Section 4.2, we believe it would be very useful to capture acted-out scenarios by dynamic computer-vision techniques. For scenario generation, these need to be segmented and techniques developed for piecing segments together in new ways to create new scenarios.

3.6 Sound

Hodges and others have found that realistic sound may be more important than visual fidelity in creating the illusion of presence [Rothbaum, '97]. For effective training, we believe sound needs to be generated in proper synchrony with visual images, and displayed properly localized for trainees. Our experiments indicate that localization in azimuth only, as opposed to 3-D localization, may quite suffice for trainees who are free to move their heads. We observed experimental subjects consistently cocking their heads to determine sound-source elevation, rather than relying on their head transfer functions.

3.7 Physics for Virtual Objects

Virtual objects in the environment should display realistic physics in their interactions with each other [Ehrmann, '00; Hirota, '99; Hoff, '01]. When real objects interact with virtual ones, the physics should be as plausible as possible, even though virtual objects inherently cannot impose forces on real ones.

3.8 Computer-Generated People

Often it is the people in the environment who create much of the stress. For economical team training, these need to be virtual, computer-generated people, with realistic behaviors. The work of the Institute for Creative Technologies is a great first step [Rickel, '02].

3.9 Speech Recognition and Generation

The trainees' utterances are an important behavior. Recognizing them and doing communication analysis on them is at least as important as knowing physical behavior. This requires recognizing not only words and larger syntactic patterns, but also distinguishing the voices of the several trainees.

The trainees also will interact with computer-generated people, so real-time speaker-independent speech analysis and generation is a crucial component.

3.10 System Integration

Each individual technological challenge is great. **Making any of them work** will be an important technical contribution. **Integrating them all into a single working system** is an equal challenge in system engineering that must not be overlooked or minimized in planning funding and schedules [Brooks, '05].

4. TEAM-TRAINING SCIENCE CHALLENGES

Team-training science is fundamental to this concept: it provides training strategies; it provides guidelines and tools for developing content. Mounting evidence suggests that a set of team competencies can be identified and that these competencies are **generalizable across team task situations** [Cannon-Bowers, '98a,c].

4.1 Why Scenario-Based Training for Teams?

Team training develops shared mental models via shared experiences; all team members participate together—seeing, talking, touching, and interacting with each other. *Scenario-based training* (SBT) relies on controlled exercises or vignettes, where the trainee is presented with cues found in the actual task environment, performs, is evaluated, and is then given feedback. SBT differs from more traditional training in that there is no separate formal curriculum; instead, *the scenarios themselves are the curriculum* [Cannon-Bowers, '98a,b]. Hence scenarios must be crafted, and training executed, so that it accomplishes specified training objectives. Effective SBT requires expert scenario authors and instructors, and they are in short supply. Homeland security has substantially increased the demand for team training, the skills to be trained, and the uniformity and quality required.

Scenario-Based Training under Stress. Cannon-Bowers and others have demonstrated the power of SBT in several complex operational environments and completed a large-scale, multi-year study of decision making under stress by high performance combat teams [Cannon-Bowers, '98b; Salas, '00]. They found that exposure to many varied task instances helps develop decision makers who respond quickly and maintain situational awareness while dealing with ambiguity. Immersive simulation can do this in a controlled, cost-effective manner.

4.2 A Big Challenge –Generating Scenarios

Scenarios make or break scenario-based training. Today they are costly and time-consuming. Therefore an important task is to create a scenario generation software tool that:

- Generates valid stories and scenarios
- Is easily used by scenario authors, via a point-and-click interface
- Is flexible—accommodates many variables and input conditions
- Allows manipulation of scenario difficulty and stressors
- Is scalable for individuals, teams, and teams-plus-autonomous agents
- Is usable for many training levels—live drills, VE, tabletop exercises, classroom instruction, computer-based training, handheld computer games, and mobile telephones
- Connects to the performance-data capture tool, the feedback tool, and instructor aids
- Is easily updated and exports scripts to word processors for further customization

Build on an Existing Tool. The developers should leverage Bowers' team's experience and their public-domain work-product from aviation training. Through a Navy-FAA partnership, the Rapidly Reconfigurable Line-Oriented Evaluation (RRLOE) scenario-generation tool was developed for testing aviation proficiency in commercial pilots [Bowers, '97]. RRLOE has been delivered to over 50 aviation concerns, including every major U.S. airline. Because the RRLOE tool is generic in structure and platform, the software core can assemble event sets for scenario-based training in other domains. RRLOE strengths include its relational database that directly ties-in qualification standards, task analyses, and cross-environment standardization of terms.

But some challenging new requirements arise for scenarios in the proposed environments:

- **Branching.** RRLOE can create scenarios that target specific events and skills, but it currently does not readily handle branching on data sensed from the ongoing simulation.

- **Real-time scenario generation.** The existing software does not yet run in real time, which will be necessary to avoid combinatorial explosion, once rich branching is incorporated.
- **Interaction with autonomous agents.** The RRLOE engine should be expanded to incorporate descriptions and behavior models of autonomous agents.

Other Tools Needed. Training Management System, Instructor Interface Tools, Automatic Performance Assessment tools, Post-exercise Analysis and Feedback tools; Asset Versioning and Management System.

4.3 Research in SBT for Teams.

To maximize the usefulness of the ITT, its development should include research in basic learning science to generate new knowledge of team competencies, team performance, and team training. The ITT system to be developed under this concept should provide a high-fidelity environment for repeated, controlled and rigorous investigations. Research questions include instructional strategies, dynamic assessment, feedback, and validating training effectiveness [Tannenbaum, '98]. A research plan could be:

- **Test New Learning Strategies in Scenario-Based Training.** One should test recently introduced training strategies against the particular training tasks of a real partner who is engaged in team training: stress exposure training, cross-training to acquaint team members with each others' jobs, team self-correction, and communication training.
- **Improve the reliability of measurement and of feedback** by standard tools.
- **Broaden and accelerate adoption of immersive SBT** by enabling it to be performed with less-experienced instructors, via help for the hard parts—evaluation and feedback.

5. HOW TO DEVELOP SUCH AN IMMERSIVE TEAM TRAINER?

5.1 Development with Training Partners

Our UNC experience as tool-builders is that any useful new system must be developed with real users on real problems. For 37 years the UNC computer science team has done this in no-money-changes-hands collaborations with protein chemists, biologists, physicists, surgeons, radiologists, oncologists [Brooks, '96]. It also helps to have **two** users of a new system, with different applications, so that the system is not too specialized. But it has to be really **useful** to each of the users, or they won't keep collaborating.

5.2. U-Try-It Facility

From the beginning the project should maintain **two immersive environment set-ups**. One should be the **Research Laboratory,** where new technologies should be continually integrated into the system. The other, the **Facility for Immersive Team Training (FITT)** should be a state-of-the-art but **stable** hardware-software system, a "U-Try-It" facility, with its own staff experts in scenario-based training.

A major impediment to the acceptance and adoption of immersive system technology into everyday training is the capital cost (high, today, but coming down fast) and the need for expert staff to even try out the technology. The project's "U-Try-It" FITT should be an international asset where team trainers can go to develop and test training scenarios, to do training feasibility demonstrations, and to pilot for-real training, with the staff experts' help. Providing a means for serious inquirers to try the methodology without these up-front investments should radically accelerate adoption.

ACKNOWLEDGEMENTS

This paper is built on our earlier paper [Brooks, 2005]. The analysis of the problem and this vision for addressing it is the work of many people, especially:

- UNC CS colleagues*:* Gary Bishop, Anselmo Lastra, Ming Lin, Dinesh Manocha, Marc Pollefeys, Andrei State, Herman Towles, Leandra Vicci, Gregory Welch
- UCF colleagues*:* Clint Bowers, Florian Jentsch
- Potential training partners with whom we talked at length*:*
  Morehouse School of Medicine*:* Martha Elks, Gregory Strayhorn
  U.S. Coast Guard*:* CDR Tim Quiram, Don Robinson, LCDR Justin Ward
  UNC School of Medicine*:* Bruce Cairnes, Gregory Mears, Anthony Meyer, Judith Tintinalli

Our research in these various topics has been funded by many U.S. agencies*:* Army, DARPA, DOE, FAA, NIH/NIBIB, NIH/NLM, NSF, ONR.

REFERENCES

Bowers, C., Jentsch,F., Baker, D., Prince, C., & Salas, E. (1997). Rapidly reconfigurable event-set based line operational evaluation scenarios. *Proc. of the 41st Annual Meeting of the Human Factors and Ergonomics Soc*.

Brooks, Jr., F. P. (1996, March). The computer scientist as toolsmith II, keynote/Newell Award address at SIGGRAPH 1994. *Communications of the ACM* 39(3), 61-68.

Brooks, F.P., J. Cannon-Bowers, H. Fuchs, L. McMillan, M. Whitton (2005): "A new VE challenge: immersive experiences for team training," *Proceedings of HCI International 2005*, v. 9, (Las Vegas, Nevada, July 2005), CD-ROM.

Cannon-Bowers, J., & Salas, E. (1998a). *Making decisions under stress: implications for individual and team training*. APA Press, Washington, D.C.

Cannon-Bowers, J., Burns, J. J., Salas, E., & Pruitt, J.S. (1998b). Advanced technology in decision-making training*:* The case of shipboard embedded training. in (Cannon-Bowers 1998a) 365-374.

Cannon-Bowers, J., & Salas, E. (1998c). Theoretical underpinnings. In (Cannon-Bowers 1998a).

Ehmann, S. & Lin, M. (2000). Accelerated proximity queries between convex polyhedra by multi-level Voronoi marching. *IEEE/RSJ International Conference on Intelligent Robots and Systems.*

Fuchs, H., Cotting, D., Naef, M., & Gross, M. (in review) VISLAND*:* Visualization with imperceptible structured light for acquisition and non-planar display. (Preprints available by request).

Hirota, G., Maheshwari, R. & Lin, M. (1999). Fast volume-preserving free-form deformation using multi-level optimization. ACM *Symposium on Solid Modeling and Applications*.

Hoff, K., Zaferakis, A., Lin, M., & D. Manocha, D. (2001, March). Fast and simple 2D geometric proximity queries using graphics hardware. *Proceedings of the 2001 ACM SIGGRAPH Symposium on Interactive 3D Graphics*, 145-148.

Insko, B., Meehan, M., Whitton, M., & Brooks Jr., F.P. (2001). Passive haptics significantly enhances virtual environments. *Presence Workshop*.

Jaynes, C., Webb, S., Steele, M., Brown, M., & B. Seales, B. (2001). Dynamic shadow removal from front projection displays. *Proceedings of IEEE Visualization 2001*.

Lui, G, Zhang, J., Wang, W., McMillan, L. (2006). Human motion estimation from a reduced marker set. *Proceedings of the 2006 Symposium on interactive 3D graphics and games,* 36-42. ACM Press.

Lok, B., Naik, S., Whitton, M. & Brooks, F. (2003). Incorporating dynamic real objects into virtual environments. *Proceedings of the 2003 Symposium on Interactive 3D Graphics* (Monterey, CA), 31-40.

Low, K.-L., Welch, G., Lastra, A., & H. Fuchs, H. (2001).Life-sized projector-based dioramas. *ACM Symposium on Virtual Reality Software and Technology*.

Matusik, W., Buehler, C, Raskar, R., Gortler, S., & McMillan, L. (2000). Image-based visual hulls. *Proceedings of ACM SIGGRAPH.*

Pollefeys, M. & Gool, L.V. (2002). From images to 3D models. *Communications of ACM* 45(7), 50-55.

Raskar, R., Brown, M. S., Yang, R., Chen, W.-C., Welch, G., Towles, H., Seales, B., & Fuchs, H. (1999). Multi-projector displays using camera-based registration. *Proceedings of IEEE Visualization 99*.

Rickel, J., Marsella, S., Gratch, J., Hill, R., Traum, D., Swartout, W. Towards a New Generation of Virtual Humans for Interactive Experiences. *IEEE Intelligent Systems* (2002), 32-38.

Rothbaum, B.O., Hodges, L.F. and Kooper, R. (1997). *Virtual reality exposure therapy*. The Journal of Psychotherapy and Research. In press.

Salas, E. & Cannon-Bowers, J.A. (2000).*The anatomy of team training. Training and retraining: A handbook for business, industry, government, and the military*. S. Tobias & J.D. Fletcher. New York, Macmillan*:* 312-335.

Sukthankar, R., Cham, T.-J., & Sukthankar,G. (2001). Dynamic shadow elimination for multi-projector displays. *IEEE CVPR*.

Tannenbaum, S., Smith-Jentsch, K., & Behson,S. (1998). Training team leaders to facilitate team learning and performance. In (Cannon-Bowers 1998a) 247-270.

Usoh, M., Arthur, K., Whitton, M., Bastos, R., Steed, A., Slater, M., & Brooks, Jr., F. P. (1999). Walking>Walking-in-Place > Flying, in virtual environments. *Proceedings of SIGGRAPH 99,* 359-364.

Welch, G., Bishop, B., Vicci, L., Brumback, S., Keller, K., & Colucci, D. N. (2001). High-performance wide-area optical tracking: The HiBall Tracking System. *Presence, Teleoperators and Virtual. Environments* **10**(1):1-21.