An Initial Evaluation of the Iowa Dental Surgical Simulator

Lynn Johnson, Ph.D.; Geb Thomas, Ph.D.; Steven Dow; Clark Stanford, D.D.S., Ph.D.

Abstract: The University of Iowa colleges of dentistry and engineering are collaborating to build an effective but low-cost surgical simulator that uses force feedback to teach and assess the tactile skills of dentistry. A prototype simulator was built, and a formative evaluation examined the realism of the haptics and identified directions for future work. Using a cross-over design, twelve experienced practitioners probed two virtual teeth using two instruments attached to the force feedback device. The session was videotaped, the forces recorded, and a standardized questionnaire completed. Two analyses were conducted: an ANOVA examined practitioners' questionnaire responses, and a t-test analyzed the probing forces. Significant tooth order by instrument order interaction and instrument effects were found. Practitioners were generally satisfied. They preferred a standard joystick to an explorer, felt that two-dimensional graphics were sufficient, and emphasized that the existing vibration be eliminated. Random placement of caries should help teach generalized skills.

Dr. Johnson is Associate Professor and Dr. Stanford is Associate Professor, both at the College of Dentistry, and Dr. Thomas is Assistant Professor and Mr. Dow is Research Assistant, both at the College of Engineering, all at the University of Iowa. Direct correspondence to Dr. Johnson at the Department of Oral Pathology, Radiology and Medicine, College of Dentistry, University of Iowa, Iowa City, IA 52242; 319-335-7354 phone; 319-335-7351 fax; Lynn-Johnson@uiowa.edu e-mail.

Key words: simulation, force feedback, evaluation, fidelity, robotics

Submitted 6/30/99, returned for revision 8/27/99, accepted 6/8/00

The colleges of dentistry and engineering at the University of Iowa (UI) are collaborating on a prototype Iowa Dental Surgical Simulator (IDDS). Our goal is to build a surgical simulator that uses force feedback to teach and assess the tactile skills of dentistry. We want to develop the most efficient and cost-effective simulator possible.

The term "simulation" applies to a range of products.[1] While surgical simulators, which focus on tactile skills, are new to dental education, computer-based cognitive simulations, which emulate problem-solving skills, have been used for at least two decades[2-42] and are useful in teaching problem-solving skills to dental and dental hygiene students.[5,6]

In the simulation fidelity literature, Alessi describes the dilemma between offering realistic simulations and ensuring the transfer of learning. He theorizes that high-fidelity simulations increase transfer, but inhibit initial learning for novices, while lower fidelity simulations increase initial learning, but inhibit transfer. Alessi recommends using lower fidelity simulations for novices to increase initial learning and suggests increasing fidelity simulations as a student's skills increase, to enhance transfer.[7] Other researchers support Alessi's theory[8-13] that suggests using a simpler or lower fidelity surgical simulator with dental students and a more complex higher fidelity simulator for practicing dentists.[14,15]

In the past decade, dental simulations have become defined to include realistic mannequins with synthetic teeth and cheeks, which novice dental students use to learn dentistry's "finger skills" before practicing on live patients.[16,17] Surgical simulators[18,19] claiming to teach the tactile skills of dentistry have also been recently reported. The force feedback in surgical simulators, termed high fidelity, more closely emulates the practice of dentistry than either cognitive or mannequin simulators. Because high-fidelity simulations do not guarantee transfer, the UI colleges of dentistry and engineering launched a surgical simulator project, deliberately choosing to work in small increments while carefully evaluating their work with each step.[20] To guarantee learning, these evaluations will provide the information required to build a surgical simulator of the lowest level of fidelity.[21] This is the first in a series of required studies to develop such a surgical simulator.

The skill we have chosen for this initial phase is detecting occlusal dentine caries—a disease process that is seldom seen and only felt. This goal has two objectives: 1) to create a prototype simulator that enables a student to diagnose dentin caries; and 2) to complete a formative evaluation of the dentin caries simulator with twelve faculty members. This evaluation intends to answer the following two questions:

1. Does the simulator provide realistic forces in the force feedback device?
2. What simulator design improvements are required?

Methods

The Iowa Dental Surgical Simulator has three hardware components: a computer, monitor, and a force feedback device, as well as software. Participants viewed a monitor, which displayed cross-sections of two teeth—a tooth with an amalgam preparation and a normal tooth, both with a carious region—at 640 x 480 pixel resolution. A 200 MHz PC computer drove the simulation. Participants grasped a joystick handle or explorer handle attached to a force feedback device (Impulse Engine 2000, by Immersion Corp[22]). The force feedback device controlled the on-screen cursor movement and responded with resistive forces between 0.16 and 1.96 Newtons whenever the dentist attempted to move the on-screen cursor past the edges of the teeth.

The software consisted of two operations or threads. The first thread looped 1000 times per second while reading the x, y position of the cursor, calculated a response force, and sent force commands to the force feedback device. The second thread looped approximately thirty times per second and updated the screen graphics. Software force models developed for three tooth regions—healthy enamel, healthy dentin, and carious dentin—differed in the response forces. For healthy enamel and dentin, an upward force from the enamel surface was provided. For carious dentin, less upward force was provided. To simulate a carious lesion in healthy enamel, a small region below the surface provided no resistive force until the cursor was moving out of the region; then a small force provided a characteristic tug, which occurs when an explorer is extracted from a carious lesion. All the software was written in C and C++.

Evaluation Design

Because the surgical simulator development is in its early stages, the primary goal is to validate the forces. Thus, this formative evaluation involved only experts.[23-25] Twelve licensed dentists (volunteers) who had practiced for ten or more years provided information about the feel of the force feedback and the design of the IDSS.

The formative evaluation protocol was designed to identify flaws and suggest corrections in the design of the prototype simulator (Figure 1). The twelve participants were randomly divided into two experimental groups. Group A "explored" one tooth for caries using the joystick, then probed the second tooth with the explorer. Group B explored a tooth, first using the explorer, then probed the second tooth using the joystick. To avoid any bias occurring from the order in which the teeth were probed (order effect), each group's participants alternated which tooth was probed first. The research design was a 2 X 2 factorial design. Figure 2 illustrates the two instruments.

Each participant spent up to two minutes with the demonstration program, to fully understand how the force feedback device operated. Then each par-

Group / Probing	*Practice*	*First Experience; Questions*	*Second Experience; Questions*
Group A1 (N = 3)	✓	Joystick—Carious lesion	Explorer—Preparation
Group A2 (N = 3)	✓	Joystick—Preparation	Explorer—Carious lesion
Group B1 (N = 3)	✓	Explorer—Preparation	Joystick—Carious lesion
Group B2 (N = 3)	✓	Explorer—Carious lesion	Joystick—Preparation

Figure 1. Formative evaluation design

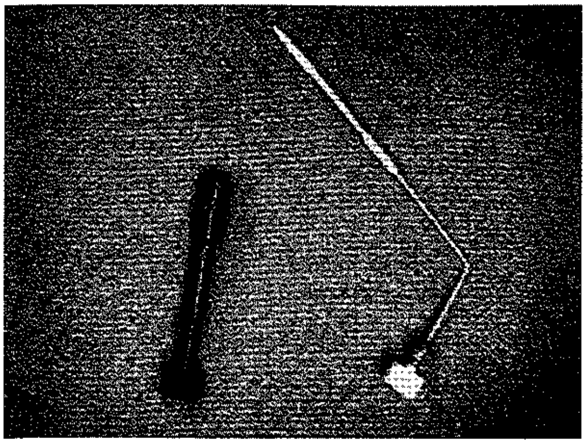

Figure 2. The two instruments attached to the force feedback device used for probing. The joystick is on the left and the explorer is on the right.

ticipant had up to five minutes to probe the assigned tooth, before answering a series of predetermined questions.

The cross sections of two teeth were presented on the screen (Figure 3). The right tooth contained a carious lesion in the central groove of a mandibular molar, while the left tooth contained a preparation for a restoration. Determining the presence of caries was the goal of probing the right tooth; detecting whether all dentinal caries had been removed was the goal of probing the left tooth.

A researcher who was present at all sessions took notes regarding participant actions and asked the survey questions. A camera also videotaped each session, recording the activity on the screen, including the system clock's display and all participants' comments. The researcher's directions instructed participants to "think out loud" during the entire session. All points were noted. During each session, the computer recorded the continuous stream of force data. Previous studies[26] indicated that the amount of force exerted by a dentist might increase when a lesion is expected. To confirm this, the dentists were asked to continuously state their current clinical findings (healthy enamel, healthy dentin, or carious dentin). These statements, combined with the videotaped and recorded forces, allowed the researchers to match practitioners' comments, such as "This feels cheesy," with the forces exerted during a post hoc review.

Figure 3. Sample screen of the tooth graphics. Note the probe cursor over the left tooth.

Analysis

Two analyses were conducted. One examined the practitioners' remarks during the session and their responses to a standard set of questions. The other examined the differences in the forces exerted by practitioners. Finally, the results of a separate study on simulation vibration are reported.

A general linear model ANOVA of each of the Likert scale questions (questions 1 through 8) determined the effect of instrument, tooth type, and orders of tooth presentation, instrument presentation, and presentation interaction. For questions 9 and 10, participants ranked the four main components of the simulator for future improvements. Components ranked as having a first priority were assigned a value of 1; components ranked as second priority a 2, etc. Separate ANOVAs were performed on participant responses for questions 9 and 10 for the effects of components, tooth, instrument, tooth order and instrument order, and their second order interactions with components.

During each trial, the practitioner's forces were recorded. Figure 4 presents these forces and their positions. The dark lines represent the forces produced by the simulator. The upper tip of each line shows the position of the cursor tip at the moment the force was delivered, while the line length represents the forces exerted. The figures indicate several aspects of the practitioners' strategy in exploring the teeth. In both examples, the practitioners concentrated their efforts in suspicious areas, but varied the amount of force exerted. Although each diagram portrays the forces exerted from an individual, the top image represents those dentists who used less force when probing for caries, and the bottom image represents practitioners who used more force.

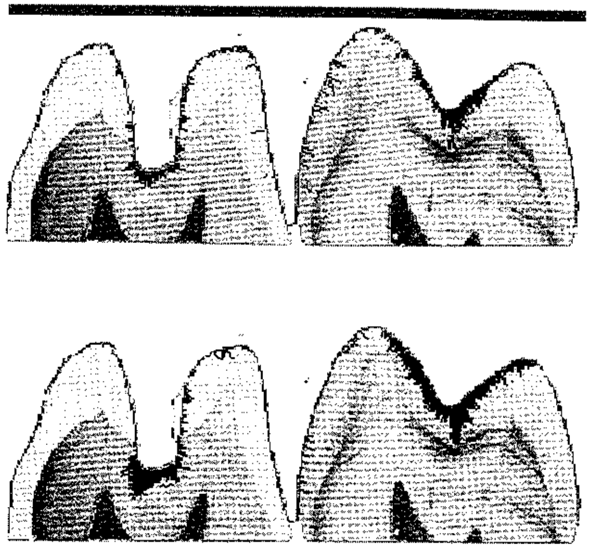

Figure 4. Illustration of the forces applied by the haptic device

T-tests ($p < 0.1$) compared the medians of the feedback forces to determine why some practitioners were unable to detect the caries on each tooth. Because this pilot study has a small sample size, the alpha value for all analyses was set at 0.01.

Results

Dentist Questionnaire

After probing a tooth, each practitioner answered a series of questions. One portion was identical for each tooth, while the remaining questions were unique to each probing experience. Table 1 summarizes the questionnaire.

The only significant effects for the first five questions were for the order of presentation interaction in question 1 ($F[1,18] = 4.90$, $p = 0.04$) and the instrument effect in question 5. For question 1, participants who initially used the explorer handle to probe the carious lesion on the occlusal surface thought the simulation was unrealistic. With a family error rate of 20 percent, Tukey's test indicated that these participants responded more negatively than all three other groups of participants, who were generally satisfied, as indicated by their mean response of 3.44.

The significant instrument effect in question 5 ($F[1, 18] = 5.5$, $p = 0.03$) suggests that the instrument attached to the force feedback device played an important role in the naturalness of the practitioners' hand positions. A one-sided t-test ($T[21] = -2.57$, $p < 0.01$) indicates that practitioners thought the hand position with the explorer handle was less natural than with the joystick handle.

Participants' comments regarding the naturalness of their hand position and the instrument use clarify the statistical findings. Table 2 contains significant comments from each treatment group.

The analysis and practitioners' comments reveal dissatisfaction with the explorer's current instrumentation, possibly because the experimental explorer, when used in a virtual environment, lacks the full functionality of an explorer used with live patients and is more frustrating than an unrealistic joystick.

Table 1. Questionnaire items and descriptive statistics (Scale: 0 = low and 5 = high)

Item	N	Mean	SD
Common Questions			
1. The "feel" of the surface enamel and healthy dentin is realistic.	24	3.0	1.5
2. I feel comfortable using this instrument as opposed to using the other instrument.	24	3.6	1.2
3. I am able to concentrate on testing for caries versus using the simulator.	24	3.8	1.3
4. The two-dimensional graphics are of sufficient quality.	24	4.2	0.8
5. The position of my hand when using the instrument is natural.	24	2.8	1.6
Carious Tooth Questions			
6. The "feel" of the enamel overlying the carious lesion is realistic.	12	2.6	1.7
7. The "feel" of the "pop" when I found the carious lesion is realistic.	12	3.9	1.7
Preparation Questions			
8. The "feel" of carious dentin is realistic.	12	3.2	1.5
Ranking Questions			
9. Rank in order of importance, with 1 being most important, the possible improvements to this simulator as a tool for teaching tactile sensitivity:	24		
❑ graphics		3.2	0.6
❑ force feedback device		2.2	1.0
❑ physical model of the patient		3.4	0.8
❑ forces delivered by the device		1.2	0.4
10. Rank in order of importance, with 1 being most important, the possible improvements to the realism of this simulator:	24		
❑ graphics		2.9	0.9
❑ force feedback device		2.5	1.0
❑ physical model of the patient		2.8	1.1
❑ forces delivered by the device		1.8	1.0

Table 2. Representative comments from question #5

Treatment Group	1st Tooth Comments	2nd Tooth Comments
	Joystick	*Explorer*
Preparation followed by Carious Lesion	I'd be better if I had a probe.	I need to push it in order to make it go down.
	I would rather have an explorer in my hand when doing this.	The hand position is not at all natural. The alignment is a problem; it is better with the joystick.
Carious Lesion followed by Preparation	It is nothing like an explorer.	I do not have the ability to angle the explorer like I want.
	Explorer	*Joystick*
Preparation followed by Carious Lesion	I cannot hold the explorer the way I want.	I like this better.
Carious Lesion followed by Preparation	The length is unreal, because the movement is at a different point. It feels realistic until I move it.	It would be better with a thinner handle.
	I cannot do the forces easily.	This is good.
	The shaft is too long. The axis of rotation is in the wrong place.	I like the joystick better. From a tactile standpoint I can grab it so that the point of rotation is more realistic.

Figure 5 presents the practitioners' responses to each of the first eight questions and suggests considerable variation in the dentists' perceptions that the simulation of the healthy enamel and dentin was realistic, with the mean of 3.0 indicating marginal satisfaction with its realism. The plot also suggests that most practitioners found the instruments comfortable and could concentrate while searching for caries with the simulator. Results for question 4 indicate that, except for one outlying point, the simulator graphics were satisfactory. Results for question 5 suggest that although participants were generally satisfied with their hand position, they varied substantially in range, from highly unsatisfied to very satisfied. For three tooth-specific questions, participants were undecided about the success of the modeling of the enamel overlying the carious region, but were generally satisfied with the tug or "pop" sensation felt when extracting the cursor from the lesion. They also disagreed about the feel of the carious dentin in the tooth with the preparation, but overall were satisfied with the simulation.

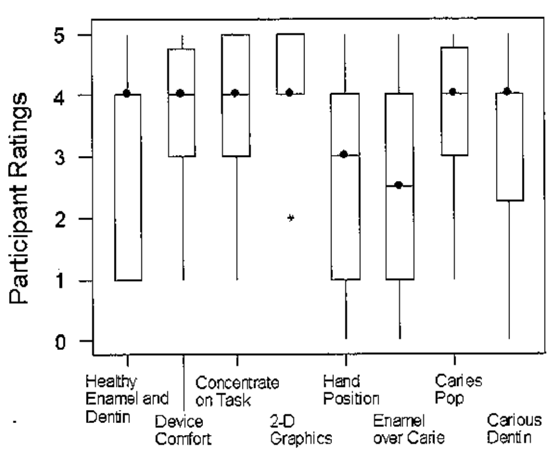

Figure 5. Boxplot of Likert scale questions. The dots indicate the medians, and the boxes represent the upper and lower boundaries of the first and third quartiles. The whiskers extend to the smallest and largest responses that are less than one quartile from the end of the box. Points outside the range are intended by an *.

Separate ANOVAs performed on questions 9 and 10 showed no significant effect except for the main effect of components (question 9—$F[3, 76] = 23.5, p < 0.001$; question 10—$F[3, 76] = 6.56, p = 0.001$). Followup Tukeys for both questions with a family error rate of 5 percent on the four category data indicated that practitioners felt that the tactile feedback was the most important system component needing improvement. The force feedback device itself (including the handles used) was the second priority. While the relative importance of the virtual patient model and graphics was unclear, both were ranked behind the other components.

Force Data

Fifty percent of the experts detected the "carious lesion," and 75 percent detected "caries" in the preparation. Three theories accounted for variations in practitioners' detection of caries: 1) the feedback loop was too slow, resulting in a vibration that caused participants to probe sideways, rather than up and down; 2) the practitioners used a drag technique versus a poking technique; and 3) the practitioners applied insufficient forces. Examining the forces exerted in the y direction suggested an explanation. A *t*-test conducted of the medians of the y direction feedback forces for the carious lesion was insignificant and ruled out vibration or practitioner technique causing the difference between the practitioners.

A *t*-test compared the medians of the simulator feedback forces for practitioners who successfully identified the caries on each tooth with unsuccessful practitioners. The results suggested that experienced practitioners unsuccessful in detecting caries applied less force (see Table 3.), indicating a need to revise the force models so that carious lesions can be more successfully detected.

Vibration

This formative evaluation was designed to identify simulator deficits. Reviewed videotapes revealed that all participants complained of vibrations with each tooth. A followup study[27] demonstrated that vibrations occurred from 1.6 to 4.8 Newtons. It also found that the vibration occurred because the force feedback loop speed was too slow (1000hz). When a practitioner "probed" a surface, the simulator exerted what appeared to be a series of discrete resistive forces, or vibrations. A faster microcomputer or improved software would update these forces faster, appear to give a continuous stream of resistance forces, and eliminate the vibration.

Discussion

The Iowa Dental Surgical Simulator's long-term goal is to teach and assess the tactile skills required of a skilled dentist, using a simulator with material and physical properties, which also balances high learning effectiveness with low cost. The goal of this simulator's first version is to teach the tactile diagnosis of dentin caries. This study's results suggest that although practitioners were generally satisfied with the simulator, specific target issues require resolution.

Table 3. Results of *t*-test on median feedback forces for successful/unsuccessful caries detection

Median Source	N	Mean	SD	T
Carious Lesion Detection				
Successful	6	1003.83	190.89	0.07*
Unsuccessful	6	696.00	325.65	
Preparation Caries Detection				
Successful	9	336.33	133.26	.37
Unsuccessful	3	252.00	140.28	

* p < .1

Originally, the IDSS was visualized as having realistic three-dimensional graphics and actual dental instrumentation. However, budget limitations forced design compromises for the prototype. The graphics were two-dimensional images, and the instruments were either the manufacturer's joystick or a modified explorer. The results from two-dimensional graphics showed that three-dimensional, high-fidelity graphics were not essential; the cognitive simulation literature supports the use of simplistic graphics unless they are essential to the task being simulated.[7,8,10] Because the primary goal of the IDSS is tactile, realistic visuals are unnecessary. This emphasis encourages the developers to focus all resources on the force feedback component of the simulator.

The practitioners' preference of the joystick over the explorer presents a paradox. The simulation fidelity literature seems to support using a simplistic joystick instead of a more realistic, higher fidelity explorer. But it also supports the presentation of the essentials of the emulated environment.[28] Our simulator's joystick was a familiar object that operated identically as in a game environment. The explorer was perceived as awkward and not operating as it would in a clinical environment with six degrees of freedom—x, y, z, roll, pitch, and yaw. Based on this feedback, a new force feedback device with an increased number of degrees of freedom is being pursued.

The researchers theorize that the wide variation in responses regarding the quality of forces may have been caused by participants' reactions to the vibration. The simulation literature clearly recommends that developers create simulations that correspond to actual systems[29] and that lack features that distract or confuse the user.[7] Thus, the vibration distracter should be eliminated. Currently, the IDSS developers are investigating low-cost hardware and software solutions to this problem.

The fidelity literature theorizes that lower fidelity simulations should be used by novices to increase initial learning and that, to enhance transfer, the level of realism should increase as students progress.[7] A simulator capability that randomly places caries in the virtual mouth versus having students always probe the same carious lesion would reflect increasing fidelity. Students would become proficient in the generalized skills of caries detection versus the specific skill of caries detection on a single tooth.

Acknowledgments

The researchers are indebted to the twelve University of Iowa faculty members who volunteered their time for this study. Without their expertise, cooperation, and candor, this study would have been impossible.

REFERENCES

1. Edens K, Gredler M. A mechanism for screening microcomputer simulations. Educ Tech 1990;30(12):46-8.
2. Tira DE. A computer-based dental diagnostic case simulation (DDS) system. Comput Biol Med 1980;10(1):23-39.
3. University of Nebraska Medical Center College of Dentistry. Computer-assisted dental simulations—final report. NIH grant no. LM 02911. October 31, 1980.
4. Finkelstein MW, Johnson LA, Lilly GE. Interactive videodisc patient simulations of oral diseases. J Dent Educ 1998;52(4):217-20.
5. Johnson LA, Cunningham M, Finkelstein MW, Hand J. Geriatric patient simulations for dental hygiene. J Dent Educ 1997;61(8):667-77.
6. Lavell N. A summative evaluation of interactive videodisc patient simulations. Unpublished manuscript, 1991.
7. Alessi SM. Fidelity in the design of instructional simulations. J Comput Instruct 1988;15(2):40-7.
8. Dwyer FM. Strategies for improving visual learning. State College, PA: Learning Services, 1987.
9. MacLachan J. Psychologically-based techniques for improving learning with computerized tutorials. J Comput Instruct 1986;13:65-70.
10. Swezy RW. Effects of instructional strategy and motion presentation conditions on the acquisition and transfer of electromechanical troubleshooting skill. Human Factors 1991;33(3):309-23.
11. Gagne RM. Training devices and simulators: some research issues. Am Psychologist 1954;9:95-107.
12. Wolfe J. The teaching effectiveness of games in collegiate business courses: 1973-1983 Update. Simulation & Game 16(3):251-88.
13. Schneider W. Training high-performance skills: fallacies and guidelines. Human Factors 1985;27:285-300.
14. Koonce JM. Effects of ground-based aircraft simulator motion conditions upon prediction of pilot proficiency. Unpublished dissertation, University of Illinois at Urbana-Champaign, 1974.
15. Hays RT. Simulator fidelity: a concept paper. Alexandria, VA: U.S. Army Research Institute, Technical Report 490, 1980.
16. Williams J. Personal conversation, 1999.
17. Caughman F. Personal conversation, 1998.
18. University of Colorado Health Sciences Center. Center for Human Simulation [Online]. www.uchsc.edu/sm/chs, December 28, 1998.
19. Lewinstein I. DentSim: a tactile simulator for teaching and assessment. A paper presented at the Annual Session of the American Association of Dental Schools, Minneapolis, MN, March 4, 1998.
20. Tessmer M. Formative evaluation alternatives. Perform Improv 1994;7(1):3-18.
21. Flagg B. Formative evaluation for educational technologies. Hillsdale, NJ: Lawrence Erlbaum Associates, 1990.
22. Immersion Corporation. Impulse Engine [Online]. www.forcefeedback.com/research/research.html, April 12, 1999.
23. Dick W, Cary L. The systematic design of instruction, 3rd ed. Glenview, IL: Scott Foresman, 1990.
24. Medley M, Weston C. A comparison of student feedback obtained from three methods of formative evaluation of instructional materials. Instructional Science 1988;17:3-27.
25. Saroyan A. Differences in expert practice: a case from formative evaluation. Instruc Sci 1993;21:451-71.
26. Gomez N. Typical forces exerted on a tooth during caries detection. Masters' thesis, University of Iowa, December 1997.
27. Dow SD, Thomas G, Johnson LA. Signal detection performance with a haptic device. A paper presented at the 1999 Annual Conference of the Human Factors and Ergonomics Society, 1999.
28. Thurman RA. Instructional simulation from a cognitive psychology viewpoint. ETR&D 1993;41(4):75-89.
29. Halff HM, Hollan JD, Hutchins EL. Cognitive science and military training. Am Psychol 1986;41:1131-9.