Human Factors: The Journal of the Human Factors and Ergonomics Society

Augmented, Pulsating Tactile Feedback Facilitates Simulator Training of Clinical Breast Examinations

Gregory J. Gerling and Geb W. Thomas

The online version of this article can be found at:

Published by:

$SAGE

http://www.sagepublications.com

On behalf of:

Human Factors and Ergonomics Society

Additional services and information for *Human Factors: The Journal of the Human Factors and Ergonomics Society* can be found at:

Email Alerts: http://hfs.sagepub.com/cgi/alerts

Subscriptions: http://hfs.sagepub.com/subscriptions

Reprints: http://www.sagepub.com/journalsReprints.nav

Permissions: http://www.sagepub.com/journalsPermissions.nav

Citations: http://hfs.sagepub.com/content/47/3/670.refs.html

>> Version of Record - Jan 1, 2005

What is This?

Augmented, Pulsating Tactile Feedback Facilitates Simulator Training of Clinical Breast Examinations

Gregory J. Gerling, University of Virginia, Charlottesville, Virginia, and **Geb W. Thomas,** University of Iowa, Iowa City, Iowa

Haptic training devices can facilitate tactile skill development by providing repeatable exposures to rare stimuli. Extant haptic training simulator research primarily emphasizes realistic stimuli representation; however, the experiments reported herein suggest that providing augmented feedback can improve training effectiveness, even when the feedback is not natural. A novel clinical breast examination training device uses inflated balloons embedded in silicone to simulate breast lumps. Oscillating the balloon water pressure makes the lumps pulsate. The pulsating lumps are easier to detect than the static lumps used in current simulators, and this manipulation seems to effectively introduce trainees to small, deep lumps that are initially difficult to perceive. A study of 48 medical students indicates that training with the dynamic breast model increased the number of lumps detected, $F(1, 47) = 9.34$, $p = .004$, decreased the number of false positives, $F(1, 47) = 5.78$, $p = .020$, and improved intersimulator skill transfer, $F(1, 47) = 26.56$, $p < .001$. The results suggest that at least in this case, augmented, tactile feedback increases training effectiveness, despite the fact that the feedback does not attempt to mimic any physical phenomenon present in the natural stimulus. Applications of this research include training techniques and tools for improved detection of palpable cancers.

INTRODUCTION

Each year breast cancer kills 40,000 women in the United States, with approximately 211,240 new cases estimated for 2005 (Jemal et al., 2005). Fortunately, if tumors are treated before reaching 2.0 cm in maximum diameter, the 5-year survival rate exceeds 98%. Because of high breast cancer mortality rates worldwide, research has focused on early detection as one possible means of saving lives.

A clinical breast exam (CBE) is a common component of many breast cancer screening protocols, often used as a complement to mammography. To conduct a CBE, a health care professional methodically palpates the patient's breast, pressing the tissue against the patient's rib cage with his or her finger pads, feeling for tissue irregularities. The critical skill in this exam is tactile perception, the elicitation and perception of nonuniform pressures across the finger pad surface resulting from the variable stiffness of the underlying material.

The potential benefit of CBE is currently limited by sensitivity ranges of 39% to 59% (Shen & Zelen, 2001). This limited sensitivity may be caused by inadequate training or training procedures. Many physicians self-report low CBE confidence and skill levels, possibly because they never received formal training or validated their CBE skills (Fletcher, O'Malley, & Bunce, 1985; Pilgrim, Lannon, Harris, Cogburn, & Fletcher, 1993; Wiecha & Gann, 1993). Forty-three percent of residents, faculty, and nurse practitioners lack confidence in their CBE skills (Wiecha & Gann, 1993), and most surveyed physicians acknowledge a need to increase their CBE competence (Freund, 2001). Practitioners may underutilize CBE if they do not feel proficient (Korn, 1998).

Training tactile skills is difficult for several reasons. Tasks such as distinguishing tumors

from normal breast tissue nodularity demand discrimination of subtle differences. Consistent task performance requires the simultaneous application of multiple skills, which include maintaining a precisely controlled pressure, moving with a consistent frequency and duration, and visualizing the three-dimensional tissue volume. Although 0.2- to 1.0-cm lumps are palpable (Bloom, Criswell, Pennypacker, Catania, & Adams, 1982; Wolfe, 1974), small, deep tumors are initially very difficult to find without gradual learning and practice, advancing from larger to smaller tumors.

Training can substantially improve performance, and effective training tools can improve training success (Bennett et al., 1990; Campbell, Fletcher, Pilgrim, Morgan, & Lin, 1991; Hall et al., 1980). Current training approaches emphasize a thorough search pattern, adequate pressure, proper finger positioning, and the ability to discriminate a solid mass from normal breast tissue, including normal, potentially confusing structures within the breast, such as milk ducts (Coleman & Heard, 2001; Pennypacker et al., 1982). These skills are typically introduced with live patient volunteers, artificial breast models, and training videos. The tactile discrimination skills can be trained with simulators, live patients with benign breast tumors, or both.

Breast model training can provide a 44% to 66% skill improvement (Bennett et al., 1990; Clarke & Savage, 1999; Hall et al., 1980) and allow trainees to detect tumors as small as 2 or 3 mm in diameter (Adams, Hall, & Pennypacker, 1976; Bloom et al., 1982). One consistent limitation of CBE breast model training research, however, is the increase in false positives after training, which suggests that breast model training may increase a health care practitioner's willingness to diagnose more breast anomalies as tumors (Bennett et al., 1990; Campbell et al., 1991; Lee, Dunlop, & Dolan, 1998).

Current CBE breast model training techniques, developed in the late 1970s and 1980s, emphasize realistic stimulus representation. Much of this work is reported in a series of nine papers published by Pennypacker, Hall, Goldstein, and colleagues (e.g., Adams et al., 1976; Bloom et al., 1982; Hall, Goldstein, & Stein, 1977), a team that founded the Mammatech Corporation to provide the medical community

with low-cost breast models resulting from the research. The team initially emphasized training procedures for discrimination, feedback, and attaining a fixed performance criterion (Hall et al., 1980) and later focused on performance proficiency and maintenance improvement (Pennypacker et al., 1982). This research led to the development of the current static CBE breast models, which are flattened silicone hemispheres embedded with five hard lumps (i.e., artificial tumors) and covered with an opaque, flexible skin. The trainee palpates the breast model to discover the lumps. After the search is completed, the gel may be turned over and a cloth backing removed to reveal the position of the lumps within the translucent silicone.

Salas and Cannon-Bowers (2001) noted the recent reverse trend toward using low-fidelity simulators, such as these silicone breast models, to train complex skills. They suggested that these somewhat less sophisticated displays do well in representing the knowledge, skills, and attitudes to be trained as well as facilitating transfer of training. Breast models made of rubber-like materials can avoid the technical limitations of mechanical resolution, update rate, and repeatability associated with haptic displays based on programmable force-feedback devices. Developers of haptic simulators seeking greater device flexibility, generality, and programmability have recently focused on developing electromechanical devices that simulate palpated tissue or the interaction between tissue and a surgical instrument. Such simulators are being developed for laproscopic surgery (Tendick et al., 2000), spinal needle biopsy (Ra et al., 2001), endoscopic sinus surgery (Yagel et al., 1996), epidural anesthesia (Stredney, Sessanna, McDonald, Hiemenz, & Rosenberg, 1996), and prostate cancer exams (Burdea, Patounakis, Popescu, & Weiss, 1999). Much of this research has focused on the realistic and efficient modeling of the mechanics and dynamics of soft tissues and tool-tissue interaction (Berkley et al., 1999; Picinbono, Lombardo, Delingette, & Ayache, 2000). Combining the realism of the low-fidelity, rubber-like materials with the flexibility of the electromechanical instrumentation could lead to training simulators that provide relatively natural haptic feedback while incorporating a wide variety of programmable stimuli conditions.

To explore this design opportunity, we developed a clinical breast exam training device that presents lumps by inflating one or more balloons embedded in a breast-shaped silicone matrix (Gerling, 2001). This approach provides the advantage of facilitating extended practice by allowing the position of active lumps to be reconfigured between trials. While developing the model, we observed that oscillating the water pressure in the balloons seemed to help trainees localize and detect subtle lumps. The pulsation, which relies on the training device's unique design, is a novel form of augmented task feedback. This feedback may simplify high-difficulty tasks by allowing the trainee to focus on the perception subtask rather than the judgment subtask.

Current training procedures with a CBE training device are easily modified to incorporate the augmented, pulsating feedback. Training with both the static and dynamic training devices begins by presenting one or more stimuli (a silicone breast with one or more lumps) to the trainee. The trainee palpates the breast model, searching for and reporting the location of each suspected lump. After each trial, the trainer provides postperformance feedback to the trainee, explaining whether or not a tumor was present at a designated location. With traditional CBE training devices, this procedure may be repeated just once or twice before the trainee memorizes the fixed positions of the tumors, after which the trainee gains little benefit from the postperformance feedback. The dynamic training device, however, allows the procedure to be repeated indefinitely because each lump may be independently activated, offering a new configuration of lumps to trainees and allowing valuable postperformance feedback to aid and speed skill acquisition. When the trainee misses a lump, the dynamic training device can also provide augmented feedback by oscillating the water pressure in the lumps. Providing this feedback directly after a trial with a missed lump helps the trainee locate the missed lump and identify the previously hidden stimulus.

Previous research suggests that this additional feedback provided on missed trials is likely to improve training effectiveness. Various researchers have demonstrated, for example, that frequent feedback quickly improves performance and consistency (Schmidt, Young, Swinnen, & Shapiro, 1989; Winstein & Schmidt, 1989; Wulf & Schmidt, 1989). Feedback that is customized to the trainee's needs, such as providing frequent feedback when a task is first introduced and then reducing its frequency as the trainee becomes more competent, also increases its effectiveness and supports long-term retention (Swinnen, 1996).

An interesting aspect of the pulsating feedback is that it is not realistic. It is a caricature of the realistic stimulus presentation, using exaggeration to direct the trainee's attention toward specific aspects of the perceptual stimuli. Nevertheless, the dynamic breast model's augmented haptic presentation may help the trainee to develop critical discrimination skills necessary in the real task environment.

The following experiment compares the training effectiveness of the feedback facilitated by the dynamic and static training devices. A between-subjects experimental design was chosen to balance the clinical relevance and benchmarking provided by the well-known static training device and the novel capabilities of the dynamic training device. To emphasize the effect of the feedback rather than the effect of the testing device (static or dynamic), care was taken to ensure that training conditions presented by the two devices were as similar as possible, although this somewhat limited the full capabilities of the dynamic simulator. For example, the testing involved only 5 of the 15 lumps available in the dynamic model, matching as well as possible the stimuli presented by the static model. Also, the experimental protocol suggested that several static models were available. To do this, the static model was discreetly rotated to change the absolute lump positions, and the model remained hidden from sight between tests. Although these measures limited the participants' opportunity to memorize the static model's lump positions, no advantage was provided to the dynamic trainees to compensate for how static trainees could have learned the tumor locations. The experimental objective was to determine whether providing the additional feedback facilitated by the dynamic training device, in particular the pulsating lump feedback, would improve clinical breast exam training effectiveness for medical students.

EXPERIMENT

The experiment compares training with the dynamic versus static training models on lump detection accuracy, false positive rate, and inter-simulator skill transfer. Two hypotheses were tested: (a) that training with the dynamic breast model leads to higher lump detection without increasing false positives, and (b) that lump detection skills learned on the dynamic breast model transfer to the static breast model.

Method

Participants. Forty-eight first- through third-year medical students at the University of Iowa, a population similar to that used in previous studies (Coleman, Coon, & Fitzgerald, 2001), participated in the experiment, which followed a protocol approved by our institutional review board. The 30 women and 18 men were between 22 and 40 years old with a mean age of 25. Using variance estimates from a pilot study, we selected this sample size so that a between-subject difference of one lump detected could be tested at the 5% confidence level.

Apparatus. Three breast models were used in the experiment: the dynamic model, the Mammacare firm model (CPM-F), and the Mammacare soft model (CPM-S; Mammatech Corp., Gainesville, FL). Participants trained with either the CPM-F model or the dynamic model.

The dynamic breast model (Figure 1) is a breast-shaped silicone matrix embedded with a series of handmade balloons that, when pressurized with water, create simulated lumps in the breast. The hardness of each lump may be independently controlled, including deflating the balloon entirely, which makes the lump undetectable.

The silicone matrix is opaque and has a hard silicone backing with embedded ribs and entry points of thin tubes leading to the balloons. The balloons and tubes are constructed by heat fusing a pair of polyethylene sheets. The silicone matrix is made with a general purpose, high-strength, tin-based silicone polymer (BJB Enterprises TC-5005 with 85% cross-linker). It is fairly homogeneous with little nodularity. The silicone formula was selected so that its stress/strain properties were similar to the properties of live breast tissue reported by Krouskop, Wheeler, Kallel, Garra, and Hall (1998) and Gerling, Thomas, and Weissman (2003). Simulated ribs fabricated with a ceramic-like sculpturing compound, Super Sculpey, lie below the silicone matrix. Below these, a hard silicone backing made of a translucent, platinum-based silicone rubber (BJB Enterprises TC-5030) supports the structure and simulates the interrib muscle structures.

The dynamic breast model includes 15 lumps that may be individually inflated to different hardness (Table 1). The inflated lumps vary in size (0.3–1.25 cm), depth of placement (shallow, medium, and deep), fixedness (fixed and mobile), and hardness (20–50 durometers). The lump sizes of 0.3 cm (4 lumps), 0.5 cm (4 lumps), 1.0 cm (5 lumps), and 1.25 cm (2 lumps) are similar to those in other CBE palpation training devices (Bennett et al., 1990; Bloom et al., 1982;
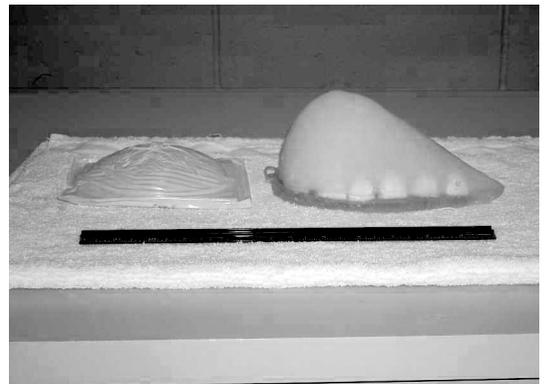

Figure 1. Breast model comparison, static (left) and dynamic (right), with a 12-inch (30.48-cm) ruler

TABLE 1: Lump Characteristics in the Dynamic and Static Breast Models

Overall Comparison	Lump ID	Size (cm)	Hardness (Durometers, Shore A)	Mobility	Depth
			Dynamic Model		
Weight: 830.8 g	1[a]	1.25	0–45 (40)	Fixed	Medium
Hardness: 3 dur	2[b]	1.25	0–45 (40)	Fixed	Deep
Dimensions	3	1.0	0–45	Fixed	Deep
L 19 cm	4	1.0	0–45	Fixed	Deep
W 11.5 cm	5[c]	1.0	0–45 (35, 30)	Mobile	Medium
H 9 cm	6	1.0	0–45	Mobile	Medium
Nodularity: no	7[a]	1.0	0–45 (40)	Fixed	Deep
	8	0.5	0–45	Mobile	Medium
	9	0.5	0–45	Fixed	Deep
	10[b]	0.5	0–45 (40)	Mobile	Shallow
	11[b]	0.5	0–45 (40)	Mobile	Medium
	12	0.3	0–45	Mobile	Shallow
	13[a]	0.3	0–45 (40)	Fixed	Deep
	14[b]	0.3	0–45 (35)	Mobile	Shallow
	15[a]	0.3	0–45 (35)	Fixed	Deep
			CPM-F Models		
Weight: 298.3 g	1	1.0	40	Fixed	Deep
Hardness: 2 dur	2	0.3	60	Mobile	Medium
Dimensions	3	0.5	60	Mobile	Deep
L, W 13 cm	4	0.7	60	Mobile	Deep
H 3.5 cm	5	1.0	40	Fixed	Deep
Nodularity: yes					
			CPM-S Model		
Weight: 298.0 g	1	1.0	40	Fixed	Deep
Hardness: 1 dur	2	0.3	60	Fixed	Shallow
Dimensions	3	0.5	60	Mobile	Deep
L, W 13 cm	4	0.7	60	Mobile	Medium
H 3.5 cm	5	1.0	40	Fixed	Deep
Nodularity: yes					

[a]Pretest and training; [b]posttest; [c]both.

Campbell et al., 1991; Pilgrim et al., 1993). The exact lump size varies with water pressure; however, the range of inflation pressures used increased the physical dimensions by less than 5%, as measured along the direction perpendicular to each lump's major axis (Gerling, 2001). These variations are small relative to the variance in clinical tumor size estimation. The depth of lump placement ranges from just under the outer silicone surface to between the ribs. Lumps arranged along the back wall of the dynamic model are fixed, whereas lumps farther out into the silicone matrix are more mobile.

Cancerous tumors have a hardness of between 0 and 60 durometers (Bloom et al., 1982); the dynamic model lumps have a hardness of between 0 and 45 durometers. Lump hardness varies linearly with water pressure over the range of interest (R^2 values ranged between .832 and .998); the exact relationship depends on lump size. A regression of water pressure (in pounds per square inch, or psi) versus hardness (Shore A durometer scale) provides separate slopes and offsets for each lump size (Gerling, 2001). Large lumps need less pressure to reach high hardness values than do small lumps. An external pressure system delivers between 20 and 45 psi (137.9–310.26 kPa) to lumps selected by opening and closing valves. Once the lump is inflated, its pressure is maintained by closing

its valve. A relief valve protects any balloons from experiencing excessive water pressures.

The two static breast models used as experimental controls (Mammatech Corp., CPM-S and CPM-F) have nearly hemispherical shape, are made of soft silicone with a tough skin, and have a flexible, square backing (Figure 1; Gerling et al., 2003). Each static model contains five lumps made of fibrous cotton wound tightly into a cylinder. The lumps vary in size (0.3–1.0 cm), hardness (20, 40, and 60 durometers), depth of placement (shallow, medium, and deep), and mobility (fixed and mobile). Lump size, hardness, position, and depth are fixed and equal in both models. The models simulate a low amount of glandular nodularity with a slightly lumpy silicone surface underneath the skin and small air pockets within the silicone matrix. Table 1 summarizes the physical differences among the CPM-S, CPM-F, and dynamic breast models. To match the five static lumps in the CPM-F model, five lumps of similar size, hardness, and depth were used in the dynamic breast model for the pretest, posttest, and training. The specific lumps used are indicated in Table 1.

Experimental design. The repeated-measures experiment included two pretests, with both the dynamic and CPM-F models; a training session; a break; and three posttests, with the dynamic and CPM-F models, followed by the CPM-S model. The 48 participants were randomly assigned to eight experimental cohorts (A_1–A_4 and B_1–B_4) balanced by gender and year in medical school, factors that reflect prior opportunities to practice breast examination skills. Cohorts A_1 through A_4 trained with the CPM-F model, whereas Cohorts B_1 through B_4 trained with the dynamic model. The subscripts indicate each variation of the four possible orderings of the pretests and posttests. For example, A_1 indicates pretesting with the CPM-F followed by the dynamic model, training with the CPM-F, postesting with the dynamic model followed by CPM-F, and then testing with the CPM-S. The between-subjects independent variables were the training device (static or dynamic), the order of the pretest (static then dynamic or dynamic then static), and the order of the posttest (static then dynamic or dynamic then static).

Six within-subject dependent variables were defined: static model detection improvement, dynamic model detection improvement, composite detection improvement, dynamic model false positive improvement, composite false positive improvement, and intersimulator skill transfer. Static and dynamic detection improvement is measured by the number of lumps found after training minus the number of lumps found on the same device before training (two measures per participant, one for CPM-F and one for the dynamic model training). Composite detection improvement is the sum of the static and dynamic detection improvement scores for each participant. Composite false positive improvement was scored as "worse" if the number of combined false positives in the dynamic and static posttests was larger than the combined pretest false positives; as "same" if the posttest and pretest false positive sums were equal; as "better" if the posttest sum was smaller; and as "NA" if there were no false positives in any of the first four tests. Intersimulator skill transfer is the within-model improvement for dynamic model tests when training with the CPM-F model or the within-model improvement for CPM-F model tests when training with the dynamic model.

Procedure. During the pretest and posttest sessions, each participant was provided a 2-min interval to examine each breast model, consistent with the typical time spent in a breast examination (Campbell et al., 1991; Fletcher et al., 1985). For the dynamic model, all five lumps were simultaneously inflated and the water pressure was kept constant during the test. The participant reported the presence or absence of identified lumps to the research assistant. The location of each lump discovered by the participant was recorded on a diagram. Immediately following the participant's set of tests, the diagrams were scored. A trial was scored as correct if the participant noted the lump's presence in the correct position, as a miss if the participant failed to note the presence of the lump, or as a false positive if the participant claimed to detect a lump where no balloon had been inflated. Neither tumor size nor depth consistency was scored.

The 15-min training session for all breast models covered search pattern, finger pressure, part and number of fingers used, finger motion, nodularity effects, breast area coverage, and

lump properties on either the dynamic or CPM-F model. The research assistant provided training according to detailed, written instructions. Of the five lumps available, three lumps of different sizes were used for the training practice, starting with the largest and moving to the smallest. With the dynamic breast model, lumps could be turned on and off and pressure could be oscillated while the participant applied finger pressure. If the participant reported difficulty in detecting the stimulus, the water pressure was oscillated until the participant reported that he or she could detect the stimulus. Water oscillation was necessary in nearly all cases for the middle- and small-sized tumors. Oscillation was induced for approximately 10-s periods while the trainee palpated the model. Up to about five such periods would occur for each lump as necessary or until the trainee found the lump.

Once the lumps were located, the assistant could covertly inflate or deflate the tumor to retest and validate the participant's ability to perceive the stimulus. With the static model, participants alternately palpated areas with and without lumps, for lump and no-lump conditions. After the 20-min rest, posttraining scores were gathered on each of the two models and the third model (CPM-S) following pretest instructions. Before the posttest, the CPM-F model was rotated 90° to change the positions of the lumps. Between experimental stages, the static models were removed from the trainee's sight and placed in a box with other static models to create the illusion that multiple static models were employed.

Results

Preliminary analysis of variance (ANOVA) of the six dependent variables showed no significant difference for pretest and posttest training order, so Cohorts A_1 through A_4 and B_1 through B_4 were collapsed in the final reported analysis. Table 2 provides a summary of the significant results for the six dependent variables.

Training had a significant effect on composite detection improvement, $F(1, 47) = 9.34$, $p = .004$. Training with the dynamic model resulted in an average composite detection improvement of 1.35 lumps ($SD = 0.92$), as compared with 0.60 lumps ($SD = 0.96$) for the static model (Figure 2).

Training also had a significant effect on the number of lumps detected in the CPM-S posttest, $F(1, 47) = 2.94$, $p = .093$, with an average of 3.04 lumps found ($SD = 1.12$) after dynamic training, as compared with 2.54 ($SD = 0.88$) after static training.

The dynamic model training improved lump detection on both the dynamic and static models, whereas static model training improved lump detection performance only for the static model (Figure 3). Both types of training increased the number of lumps detected on the static training device with approximately the same effectiveness: 1.04 ($SD = 1.4$) and 1.17 ($SD = 1.09$) for the static and dynamic models, respectively, $F(1, 47) = 0.12$, $p = .731$. Only the dynamic model

TABLE 2: Summary of Results for the Six Dependent Variables

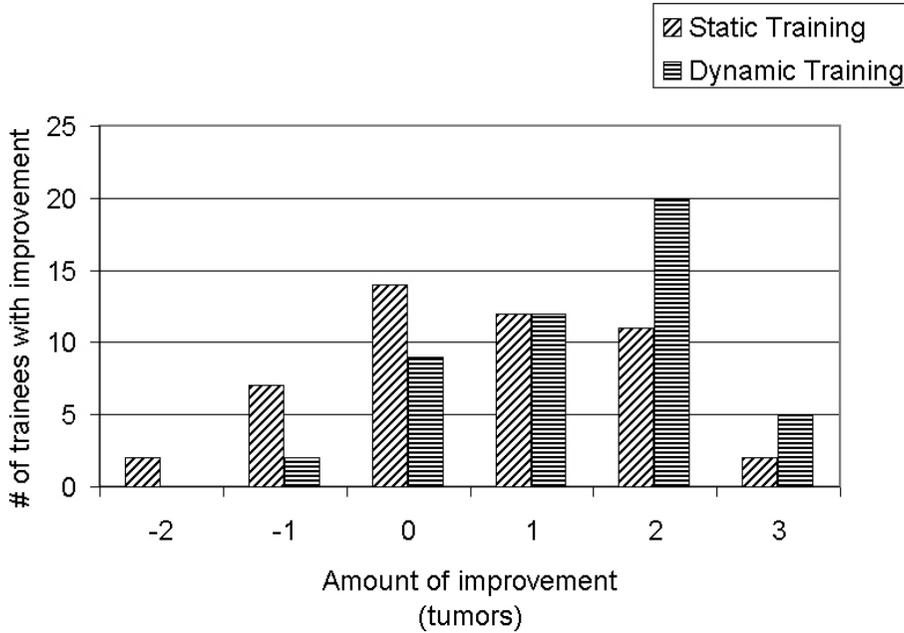|	Static Training (CPM-F)	Dynamic Model Training	df	F	p
Composite detection improvement	0.60 lumps	1.39 lumps	1, 47	9.34	.004***
Lumps detected in CPM-S posttest	2.54 of 5.00 lumps	3.04 of 5.00 lumps	1, 47	2.94	0.093*
Static model detection improvement	1.04 lumps	1.17 lumps	1, 47	*ns*	*ns*
Dynamic model detection improvement	0.17 lumps	1.54 lumps	1, 47	26.56	<.001***
Composite false positive improvement	0.42 false lumps	–0.70 false lumps	1, 47	5.78	.020**
False positives in CPM-S posttest	0.625 false lumps	0.125 false lumps	1, 47	7.56	.009***

*$p < .1$, **$p < .05$, ***$p < .01$.

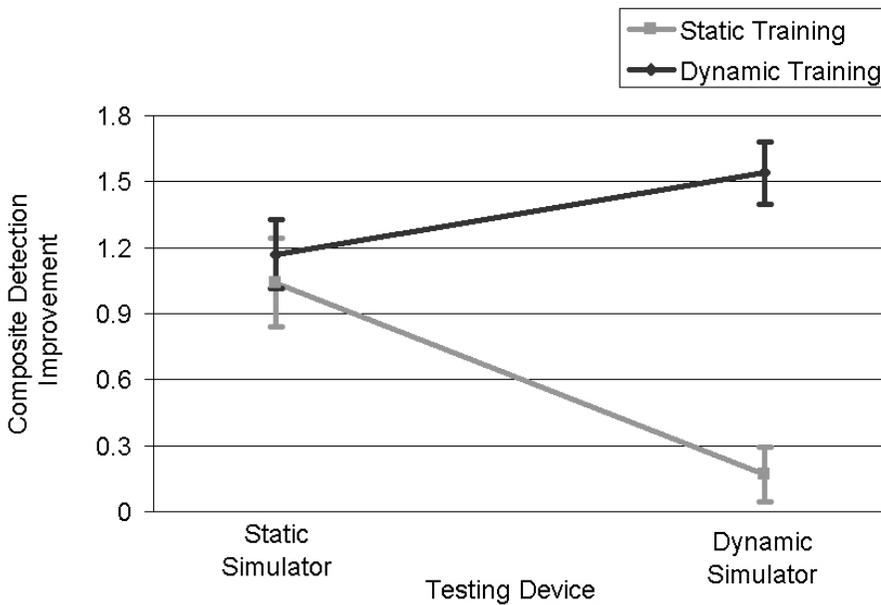*Figure 2.* Composite lump detection improvement.

Figure 3. Intersimulator skill transfer showing composite detection improvement from pretest to posttest within the same device. Lines differentiate dynamic versus static training. Dynamic training has significant composite detection improvement on both devices, but the static model trainees improved detection only when tested with the static model. Error bars show the standard error of mean.

training improved lump detection on the dynamic model (an improvement of 1.54 lumps, $SD = 0.98$, as compared with an improvement of 0.17 lumps, $SD = 0.87$, after training with the static model).

Composite false positive reports decreased (-0.70 false lumps, $SD = 1.22$) following dynamic breast model training but increased ($+0.42$ false lumps, $SD = 1.9$) following static model training, $F(1, 47) = 5.78$, $p = .020$ (Figure 4). Also, in the third-model false positives, participants who trained with the static model reported an average of 0.625 false positives ($SD = 0.82$), as compared with 0.125 reported after training on the dynamic model ($SD = 0.34$), also statistically significant, $F(1, 47) = 7.56$, $p = .009$.

DISCUSSION

The results support both experimental hypotheses: (a) Training with the dynamic breast model training leads to higher lump detection without increasing false positives, and (b) lump detection skills learned with dynamic model training transfer to the static model.

Training on the dynamic model improved composite lump detection by 1.35 lumps and decreased false positives by 0.70 lumps. Training on the static model improved composite lump detection by only 0.60 lumps and *increased* the false positive rate by 0.42 lumps. This suggests that dynamic model training improves participants' ability to distinguish between lumps and potentially confusing regions without lumps, whereas static model training may simply adjust a trainee's willingness to report a suspicious sensation as a lump. This result contradicts a potential bias that would inaccurately favor the static model. The lumps in the static model always occur in the same relative position, whereas the lumps in the dynamic model occur in different positions during training and testing. If the trainee remembered the position of the lumps in the static model, partially revealed during the training session, he or she could use this information when performing the posttest. The positions of the lumps in the dynamic model, however, changed between training and the posttest, which eliminated this potentially biasing strategy.

The increase in the number of lumps detected is particularly striking when compared with the results of previous CBE training studies that reported an increase or lack of decrease in the rate of false positives (Bennett et al., 1990; Campbell et al., 1991; Lee et al., 1998). Signal detection
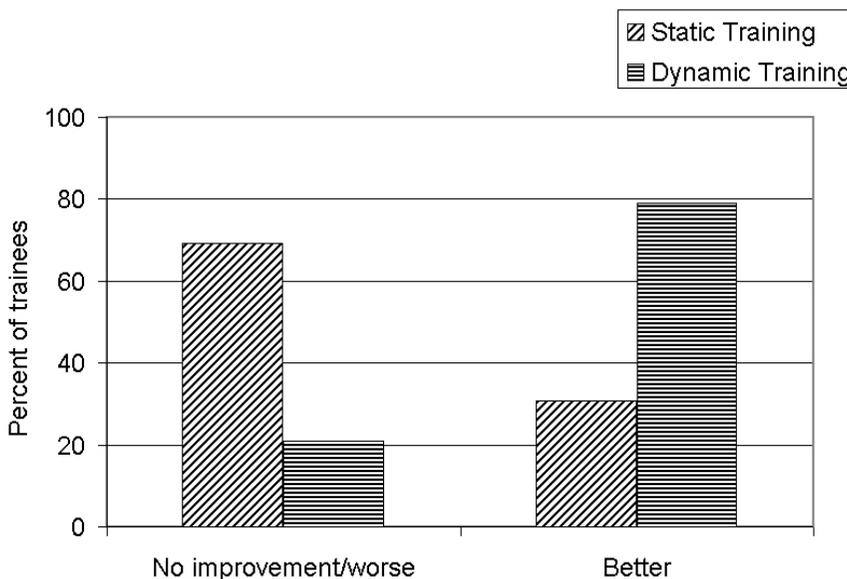

Figure 4. Composite false positive improvement by breast model (includes 13 of 24 static trainees and 19 of 24 dynamic breast model trainees who reported false positives on either pretest or posttest and excludes those who never reported a false positive).

theory suggests that such results may indicate a change in the selection criterion rather than an increase in lumps detected, although this could not be tested statistically because of the low number of repetitions available in this experimental protocol. The results reported here are consistent with the premise that the dynamic breast model enhances discrimination performance rather than shifting the selection criterion. Both training approaches and training devices are similar, except for the possibility of reconfiguring lumps for multiple, unique test scenarios and the pulsating tumor feedback in the dynamic model; consequently, the improved discrimination effect is likely a result of these differences.

With respect to the second hypothesis, intersimulator skill transfer was evident in two statistical results: (a) the intersimulator skill transfer variable and (b) lump detection on the final (third) posttest. Skills developed with static model training did not improve performance on the dynamic model, whereas skills developed with dynamic model training improved performance on the static model. This suggests that dynamic training may be more robust than static training. This might be caused by static model trainees learning specific characteristics (e.g., shape or texture) of static model lumps that are not present with the dynamic lumps, some memory effect, or a search or palpation behavior that is successful with the static model but is not helpful in detecting the dynamic model's lumps. Dynamic model training, however, appears to develop skills that are also successful in detecting lumps in the static model. In fact, when participants were detecting lumps in the static model, the dynamic model training was just as effective as the static model training.

The generality of the dynamic training is reinforced by the results of the test with the third model. The third posttest model, CPM-S, is very similar to the static model, CPM-F. This similarity might have provided a significant advantage for the static model trainees because lumps with identical properties were situated in exactly the same locations as in the model used for training (although previously rotated by 90°). Despite this clear advantage, dynamic model trainees found an average of 3.04 lumps on the third model, whereas static model trainees found only 2.54 lumps, although this suggestive trend would need to be followed up with further research.

The literature suggests that training with silicone breast models – in particular, the static model used in this experiment – can improve lump detection in real breast tissue (Clarke & Savage, 1999; Hall et al., 1980). The experiment reported here indicates a transfer of training from the dynamic training device to the static training device, but it does not test transfer of training from the dynamic device to human breast tissue. Although the results reported here are encouraging, this effect must also be formally tested in future research.

The training advantage afforded by dynamic model, resulting in improved performance on the static model, suggests that the presentation of unnatural stimuli aided the trainees in learning skills critical to the detection of realistic stimuli. However, several other differences between the breast models and the training protocol could be responsible for the performance difference. The CPM-F and CPM-S models have different geometries, use a nonhomogenous silicone interior, have different skin properties, and use lumps of a different shape and surface texture as compared with the dynamic model. These differences are a consequence of the design philosophy behind the breast models: The static model was created to provide realistic practice, and the dynamic model was created to train the skill of detecting small, deep tumors. The design philosophy difference that led to the most dominant distinction in breast model training is that the dynamic model provides the trainer with control over tumor presentation. By oscillating the water pressure when the trainee is confused, the trainer can help the trainee focus attention on the subtle tactile stimulation critical to the tumor detection task. By allowing the lumps to appear and disappear, the dynamic breast model also facilitates repeated practice by allowing the reconfiguration of lumps for multiple, unique test scenarios.

If the stimuli presented by the dynamic model were irrelevant to the lump detection task, then dynamic training would not be expected to benefit the realistic, clinically validated, static model lumps. The dynamic training benefit suggests that the presented cues are relevant to the lump detection task. If practice with the realistic

stimuli provided by the static device generally improved a trainee's ability to detect subtle, deep lumps, then static training would be expected to benefit the detection of small, deep lumps presented in the dynamic model. However, static model training did not improve the detection of these lumps. Consequently, although variations in the models' physical characteristics could have caused the observed performance differences, because the main physical features of the lumps were similar and because of the asymmetry in the training benefit, differences in feedback are more likely to account for the effect of training on performance.

CONCLUSIONS

The dynamic breast model presented in this work affords greater skill improvement than do conventional clinical breast examination training models. This improvement is most likely a function of pulsating tumor feedback and the dynamic model's capability to activate lumps for multiple, unique test scenarios. Specifically, the dynamic model training enables trainees to repeatedly practice with a variety of lumps at the same difficulty level, and augmented, pulsating feedback allows the trainee to focus on palpation and search technique independent from the formation of judgments. Training with the feedback afforded by the dynamic device is more effective than training with static models, at least for detecting simulated breast lumps.

Although the results suggest that pulsating haptic feedback can facilitate haptic training, it is important to note that the stimuli were selectively presented with oscillating water pressure. If all the dynamic tumors had always pulsated, the trainees might have overrelied on the unnatural cue and not have improved their performance with realistic lumps. A haptic training simulator can benefit from incorporating stimuli that augment the natural haptic task. However, the training protocol should later remove these perceptual crutches and allow trainees to practice the critical detection skill with stimuli that more closely approximate those in the natural task, in order to develop the interaction and perceptual skills demanded by the task in normal conditions.

The development of a reliable and valid tool to evaluate breast examination is critical to fu-

ture studies of CBE efficacy and utilization (Newcomb, Olsen, Roberts, Storer, & Love, 1995). The ultimate limit of CBE sensitivity and specificity is still undetermined. Improved, consistent training could improve screening effectiveness and have a significant impact on detection of early-stage breast cancer. The realization of this goal will, however, require further refinements in breast exam training devices, the validation of training with respect to clinical performance (especially with respect to transfer of training to real breast tissue), and the development of an objective skill assessment technique.

ACKNOWLEDGMENTS

REFERENCES

Adams, C. K., Hall, D., & Pennypacker, H. (1976). Lump detection in simulated human breasts. *Perception and Psychophysics, 20,* 163–167.

Bennett, S. E., Lawrence, R. S., Angiolillo, D. F., Bennett, S. D., Budman, S., Schneider, G. M., et al. (1990). Effectiveness of methods used to teach breast self-examination. *American Journal of Preventive Medicine, 6,* 208–217.

Berkley, J., Weghorst, S., Gladstone, H., Raugi, G., Berg, D., & Ganter, M. (1999). Fast finite element modeling for surgical simulation. *Studies in Health Technology and Informatics, 62,* 55–61.

Bloom, H. S., Criswell, E. L., Pennypacker, H. S., Catania, A. C., & Adams, C. K. (1982). Major stimulus dimensions determining detection of simulated breast lesions. *Perception and Psychophysics, 32,* 251–260.

Burdea, G., Patounakis, G., Popescu, V., & Weiss, R. E. (1999). Virtual reality-based training for the diagnosis of prostate cancer. *IEEE Transactions on Biomedical Engineering, 46,* 1253–1260.

Campbell, H. S., Fletcher, S. W., Pilgrim, C. A., Morgan, T. M., & Lin, S. (1991). Improving physicians' and nurses' clinical breast examination: A randomized controlled trial. *American Journal of Preventive Medicine, 7,* 1–8.

Clarke, V. A., & Savage, S. A. (1999). Breast self-examination training: A brief review. *Cancer Nursing, 22,* 320–326.

Coleman, E. A., Coon, S. K., & Fitzgerald, A. J. (2001). Breast cancer screening for primary care trainees: Comparison of two teaching methods. *Journal of Cancer Education, 16,* 72–74.

Coleman, E. A., & Heard, J. K. (2001). Clinical breast examination: An illustrated educational review and update. *Clinical Excellence for Nurse Practitioners, 5,* 197–204.

Fletcher, S. W., O'Malley, M. S., & Bunce, L. A. (1985). Physicians' abilities to detect lumps in silicone breast models. *Journal of the American Medical Association, 253,* 2224–2228.

Freund, K. M. (2001). *Clinical breast exam/BSE.* Retrieved July 26, 2005, from http://annieappleseedproject.org/rattecofclin.html

Gerling, G. J. (2001). *Dynamic simulator for clinical breast examination training.* Iowa City: University of Iowa.

Gerling, G. J., Thomas, G. W., & Weissman, A. M. (2003). *Dynamic simulator technical description* (Tech. Rep. R03-01). Iowa City: University of Iowa.

Hall, D. C., Adams, C. K., Stein, G. H., Stephenson, H. S., Goldstein, M. K., & Pennypacker, H. S. (1980). Improved detection of human breast lesions following experimental training. *Cancer, 46,* 408–414.

Hall, D. C., Goldstein, M. K., & Stein, G. H. (1977). Progress in manual breast examination. *Cancer, 40,* 364–370.

Jemal, A., Murray, T., Ward, E., Samuels, A., Tiwari, R. C., Ghafoor, A., et al. (2005). Cancer statistics, 2005. *Ca: A Cancer Journal for Clinicians, 55*(10), 10–30.

Korn, J. E. (1998). The clinical breast examination: Still important. *Medical Journal of Allina, 7*(2), 1–4.

Krouskop, T. A., Wheeler, T. M., Kallel, F., Garra, B. S., & Hall, T. (1998). Elastic moduli of breast and prostate tissues under compression. *Ultrasonic Imaging, 20,* 260–274.

Lee, K. C., Dunlop, D., & Dolan, N. C. (1998). Do clinical breast examination skills improve during medical school? *Academic Medicine, 73,* 1013–1019.

Newcomb, P. A., Olsen, S. J., Roberts, F. D., Storer, B. E., & Love, R. R. (1995). Assessing breast self-examination. *Preventive Medicine, 24,* 255–258.

Pennypacker, H. S., Bloom, H. S., Criswell, E. L., Neelakantan, P., Goldstein, M. K., & Stein, G. H. (1982). Toward an effective technology of instruction in breast self-examination. *International Journal of Mental Health, 11,* 98–116.

Picinbono, G., Lombardo, J., Delingette, H., & Ayache, N. (2000). *Improving realism of a surgery simulator: Linear anisotropic elasticity, complex interactions and force extrapolation* (Tech. Rep.). Rocquencourt, France: Institut National de Recherche en Informatique et en Automatique.

Pilgrim, C., Lannon, C., Harris, R. P., Cogburn, W., & Fletcher, S. W. (1993). Improving clinical breast examination training in a medical school: A randomized controlled trial. *Journal of General Internal Medicine, 8,* 685–688.

Ra, J. B., Kwon, S. M., Kim, J. K., Yi, J., Kim, K. H., Park, H. W., et al. (2001, February). *A visually guided spine biopsy simulator with force feedback.* Paper presented at the SPIE International Conference on Medical Imaging, San Diego, CA.

Salas, E., & Cannon-Bowers, J. A. (2001). The science of training: A decade of progress. *Annual Review of Psychology, 52,* 471–499.

Schmidt, R. A., Young, D. E., Swinnen, W., & Shapiro, D. C. (1989). Summary knowledge of results for skill acquisition: Support for the guidance hypothesis. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 15,* 352–359.

Shen, Y., & Zelen, M. (2001). Screening sensitivity and sojourn time from breast cancer early detection clinical trials: Mammograms and physical examinations. *Journal of Clinical Oncology, 19,* 3490–3499.

Stredney, D., Sessanna, D., McDonald, J. S., Hiemenz, L., & Rosenberg, L. B. (1996). A virtual simulation environment for learning epidural anesthesia. *Studies in Health Technology and Informatics, 29,* 164–175.

Swinnen, S. P. (1996). Information feedback for motor skill learning: A review. In H. N. Zelaznik (Ed.), *Advances in motor learning and control* (pp. 37–66). Champaign, IL: Human Kinetics.

Tendick, F., Downes, M., Goktekin, T., Cavusoglu, M. C., Feygin, D., Wu, R., et al. (2000). A virtual environment testbed for training laparoscopic surgical skills. *Presence, 9,* 236–255.

Wiecha, J. M., & Gann, P. (1993). Provider confidence in breast examination. *Family Practice Research Journal, 13*(1), 37–41.

Winstein, C. J., & Schmidt, R. A. (1989). Reduced knowledge of results enhances motor skill learning. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 16,* 677–691.

Wolfe, J. N. (1974). Analysis of 462 breast carcinomas. *American Journal of Roentgenology, Radium Therapy and Nuclear Medicine, 121,* 846–853.

Wulf, G., & Schmidt, R. A. (1989). The learning of generalized motor programs: Reducing the relative frequency of knowledge of results enhances memory. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 15,* 748–757.

Yagel, R., Stredney, D., Wiet, G. J., Schmalbrock, P., Rosenberg, L., Sessanna, D., et al. (1996). Building a virtual environment for endoscopic sinus surgery simulation. *Computers and Graphics, 20,* 813–823.

Gregory J. Gerling is an assistant professor at the University of Virginia in the Department of Systems and Information Engineering. He received his Ph.D. from the University of Iowa, Department of Mechanical and Industrial Engineering, in 2005.

Geb W. Thomas is an associate professor in the Department of Mechanical and Industrial Engineering and director of the GROK (Graphical Representation of Knowledge) Laboratory at the University of Iowa. He received his Ph.D. in industrial engineering in 1996 at Pennsylvania State University.