SPECIAL SECTION

The Efficacy of a Medical Virtual Reality Simulator for Training Phlebotomy

Mark W. Scerbo, James P. Bliss, Elizabeth A. Schmidt, and **Sommer N. Thompson,**
Old Dominion University, Norfolk, Virginia

Objective: The present study compared the effectiveness of a virtual reality (VR) simulator for training phlebotomy with that of a more traditional approach using simulated limbs. **Background:** Phlebotomy, or drawing blood, is one of the most common medical procedures; yet, there are no universal standards for training and assessing performance. The absence of any standards can lead to injuries and inaccurate test results if the procedure is improperly performed. **Method:** Twenty 3rd-year medical students were trained under one of the two methods and had their performance assessed with a 28-item checklist. **Results:** The results showed that performance improvements were limited to those who trained with the simulated limbs, and a detailed comparison of the two systems revealed several functional and physical differences that may explain these findings. **Conclusion:** Participants trained with simulated limbs performed better than those trained with a VR simulator; however, the metrics recorded by the VR system may address some aspects of performance that could eventually prove beneficial. **Application:** The present study highlights the potential for medical simulators to improve patient safety by enabling trainees to practice procedures on devices instead of patients. Applications of this research include training, performance assessment, and design of simulator systems.

INTRODUCTION

Simulators have been a standard component of training for a variety of jobs, including military strategic and tactical command operations, managerial decision making, nuclear power plant operations, and space flight operations. The area in which simulators have had perhaps the most significant impact on training is aviation, which has used flight simulators for more than 60 years. The current class of high-fidelity flight training devices can reproduce almost all of the essential characteristics of piloting an aircraft, including cockpit displays and controls, out-the-window scenery, radio communications, various weather and traffic conditions, and in some cases even the appropriate pitch, roll, and yaw movements of the aircraft. In fact, it is not uncommon for flight schools to require that students have a minimum number of hours with a simulator to obtain a pilot's license. One could argue that the availability of simulators for training has played a critical role in the evolution of military and commercial aviation.

Medical Virtual Reality Simulators

Simulators have not played a significant role in the training and education of medical professionals, unlike other high-risk professions. Medical education has traditionally followed an apprenticeship model in which procedures are learned by the "see one, do one, teach one" approach. In fact, Dawson (2002) noted that this approach to medical education has not changed since ancient Egyptian times.

Medical simulators are relatively new. Although medical simulation *devices* have been around since the 1940s, most of them have been physical models with very limited functionality. Since their inception, medical simulators have evolved along two paths. One class of simulators is mannequin-based and provides a physical model of the patient as well as physiological responses to treatments

Address correspondence to Mark W. Scerbo, Department of Psychology, Old Dominion University, Norfolk, VA 23529-0267; mscerbo@odu.edu. *HUMAN FACTORS,* Vol. 48, No. 1, Spring 2006, pp. 72–84.

(Dawson, 2002). The second class features virtual reality (VR) technology. These simulators incorporate 3-D graphics of patients or their internal organs and provide computer-based training on a variety of procedures, including cholecystectomy (i.e., gall bladder removal), arthroscopy (e.g., knee surgery), upper and lower endoscopy (e.g., colonoscopy), and vascular access such as phlebotomy (i.e., drawing blood) and intravenous (IV) catheter insertion (Satava, 2001). It is this second class of simulators that is the focus of the present paper.

The first VR-based surgical simulator appeared in the early 1990s and provided a 3-D representation of internal organs in the upper abdomen that could be observed from different perspectives (Satava, 1993). Shortly thereafter, systems emerged that incorporated accurate representations of organs based on computed tomography scans. By the end of the 1990s, systems were being developed to address specific procedures such as knee, eye, and sinus surgery (Satava, 2001).

An important factor in the rapid evolution of this technology, and one that made medical VR systems more commercially viable, was the introduction of haptic interfaces. Medical professionals often rely on the sense of touch for performing procedures. For example, nurses use their fingers to take a patient's pulse. Surgeons use instruments to probe and cut organs and suture tissue. Haptic systems that provide force feedback allow users to touch virtual objects by holding facsimiles of medical instruments in their hands. The handles, grips, and controls are identical to the actual tools, but the operational ends of the tools are encased in a housing and modified with a force feedback system that monitors hand movements in 3-D space to allow the detection of collisions with virtual objects and surfaces (skin, organs, etc.) and to provide appropriate levels of force to impede further movement (see Basdogan & Srinivasan, 2002). Consequently, the instruments feel like they are contacting, probing, and penetrating tissue and organs.

There are several ways that medical simulators can help to improve patient safety. First, and most important, they allow medical personnel to train on devices instead of practicing on actual patients. This allows trainees to learn from their mistakes without injuring anyone and to meet minimal levels of acceptable performance before seeing a patient. Second, simulators also offer an environment to train specific skills in the absence of uncontrollable influences (e.g., variations in patient characteristics and pathology). Aside from these obvious advantages to patient safety, medical simulators offer other benefits. For instance, they can allow trainees an unlimited number of trials to acquire skills and provide immediate feedback on their performance. They can also expose medical students to rare or infrequent conditions. Last, they can help decrease the dependence on animals and cadavers to practice procedures.

A VR System for IV Procedures

A VR simulator was recently developed for IV procedures (see Ursino, Tasto, Nguyen, Cunningham, & Merril, 1999) and is available from Immersion Medical, Inc. (Gaithersburg, MD). The CathSim® system provides training on a subset of needle stick procedures, including catheter insertion (cannulation) and phlebotomy. Although both procedures require a needle to gain access to the underlying vascular system, the objective of each procedure is quite different. In cannulation, a needle is used to place a catheter into the patient's vein so that fluids may be introduced to the body through that portal. By contrast, the purpose of phlebotomy is to take blood samples from the body. Thus, a needle assembly is inserted into a vein and blood is withdrawn into one or more collection tubes.

CathSim® is a multimedia system that simulates IV procedures with visual, auditory, and haptic displays. The physical system consists of a PC accompanied by an AccuTouch® six-degree-of-freedom haptic feedback device that simulates the needle and a vinyl strip that simulates the skin. The AccuTouch® device allows trainees to experience the resistive forces associated with inserting a needle into the skin and vein.

The system includes a variety of unique case patient scenarios. For example, there is an adult male with no complications and pediatric and geriatric cases with varying complications. The system provides tutorial video clips to familiarize students about many procedural details; however, some background knowledge of procedures is required (e.g., how to select an appropriate needle gauge and knowledge of potential contraindications).

The user operates the system by making selections with the computer's mouse and manipulating

the AccuTouch® device to insert the needle. The system records many performance metrics that can provide trainees with immediate feedback to improve and refine their performance. These metrics include procedure success, angle of needle insertion, occurrence of hematoma, session time, tourniquet placement time, and a pain factor. The CathSim® system also incorporates two visual training aids that allow students to see either a representation of the needle entering the arm from a side view or a "transparent" view of the underlying vascular system.

To date, there have been few published studies examining the effectiveness of VR systems for teaching IV procedures. Prystowsky et al. (1999) studied 1st- and 3rd-year medical students and surgical residents with a VR system for cannulation that incorporated stereoscopic views and haptic feedback. They gave their participants background information about the procedure and had them attempt a cannulation on one another. The participants then practiced with the simulator for 12 min and subsequently attempted a second cannulation on one another. Performance ratings, however, revealed no differences between the first and second attempt for any group of participants. Reznek, Rawn, and Krummel (2002) conducted a similar study using the CathSim® VR system but found differences related only to years of medical training. Chang, Chung, and Wong (2002) compared the CathSim® system and plastic arms for teaching cannulation to student nurses. The participants were given 2 hr of training with either method and were assessed on their first attempt at cannulation with a genuine patient. Chang et al. (2002) observed that most participants were successful, but the success rate was slightly higher for those trained on the plastic arm. Interestingly, they suggested that the superior performance of participants in the plastic arm group might have been attributable to their greater experience with phlebotomy. More recently, Scerbo et al. (2004) also conducted a study comparing the CathSim® system and simulated limbs for training IV cannulation. They gave two groups of physician assistant students 2 hr of training on either method. Performance was measured before and after training with a standardized assessment form. The results showed that all students improved after training but that the degree of improvement was significantly greater for those trained with the simulated limbs. Engum, Jeffries, and Fisher (2003) also reported some advantages of the simulated limbs over the CathSim® system.

The Present Study

The results of those studies suggest that VR training for IV catheterization procedures is not as effective as other methods; however, existing studies have addressed only IV cannulation. None has examined phlebotomy. This is an unfortunate oversight because phlebotomy is performed more often than cannulation by medical personnel in hospitals, physician's offices, clinics, and even in mobile blood vehicles. Further, there are a number of concerns with respect to patient safety. According to Mishori (2004), there are no licensure requirements for phlebotomists and the field is largely unregulated. Thus, the level of competency of those performing the procedure can vary widely. In fact, in one study 35% of patients surveyed reported discomfort with the procedure (Howanitz, Cembrowski, & Bachner, 1991). Further, injuries can and do occur. Patients have suffered punctured arteries, chronic pain from needle sticks through a nerve, and concussions and bone fractures from fainting spells that occurred when they were left unattended (Mishori, 2004). There are also indirect effects on patient safety. An improperly performed procedure can generate erroneous lab results that either leave existing conditions undetected or indicate the presence of nonexisting conditions and subject the patient to additional unnecessary tests and associated risks (Howanitz et al., 1991).

Aside from the patient safety issues, there are other factors that affect the ability to perform the procedure. Many issues can complicate the process. Patients vary widely in terms of their age, size, obesity, general health, and personal habits such as smoking and IV drug use. All of these factors affect vein accessibility and elasticity and can make the procedure more difficult to perform. Further, Clover (2002) argued that the safety risks for both the phlebotomist and the patient are often underrated. There are a variety of safety issues unique to phlebotomy concerning the practice of collecting multiple blood samples during a single procedure and the use of sharps disposal containers. Moreover, it is estimated that 600,000 to 800,000 needle stick injuries occur in health care settings each year (National Institute for

Occupational Safety and Health, 1999) and that a large percentage of these carry the ever-present threat of blood-borne diseases. Last, there are no performance-based standards for assessing phlebotomy skills in health care workers. All of these factors make phlebotomy a particularly good candidate for simulation-based training. Thus, the primary objective of the present study was to examine the effectiveness of a VR system for training phlebotomy.

At present, a standard method for teaching phlebotomy uses plastic simulated limbs; therefore, the goal of the present study was to compare VR-based CathSim® training with the more traditional simulated limb training. It was expected that training under each method would improve performance; however, expectations regarding one method over another were unclear. On one hand, it is possible that the limited success observed in the previous studies with a VR system for cannulation would also be observed for phlebotomy. On the other hand, differences between the two procedures might make phlebotomy more amenable to VR simulation. In cannulation, once the needle and catheter are positioned, the catheter is slid off the needle into the vein and then the needle is withdrawn. By contrast, in the phlebotomy procedure, once the needle is positioned it must be held in place while the blood collection tubes are positioned, filled, and swapped. Thus, it is possible that needle placement and collection activities required for phlebotomy might be more easily achieved with the VR system.

The second objective was to examine the specific differences between the two training methods. Prystowsky et al. (1999) and Chang et al. (2002) did not offer detailed explanations for their findings. Scerbo et al. (2004), however, did describe differences between the two technologies they studied. The same procedure was followed in the present study. Toward that end, a physical and functional analysis of the two training systems was performed to address potential differences in training effectiveness.

METHOD

Participants

Participants were 20 3rd-year medical students from Eastern Virginia Medical School in Norfolk, Virginia. The students participated as part of their family medicine clerkship requirements. They ranged in age from 23 to 30 years (M = 25.45 years). All participants had some previous medical experience (e.g., working or volunteering in a nursing home, working as an emergency medical technician, as a research assistant, as a hospital volunteer, or as a certified nurses aid). Four students indicated some exposure to phlebotomy procedures, but none had ever attempted one. Nine students indicated some experience with subcutaneous or intramuscular injections. No participant reported experience with any other type of medical VR simulator, but all used computers on a regular basis (approximately 13 hr/week on average for word processing, E-mail, etc.) and 3 reported using other types of simulators (driving, games, etc.).

Materials

Simulated arm. A Nasco, Inc., (Fort Atkinson, WI) Life/Form® simulated arm was used for the pretest and posttest (see Figure 1). The simulated arm has a layer of vinyl skin covering a network of latex veins. A standard IV bag filled with a quart (0.946 L) of artificial blood is hung 2 feet (about 61 cm) above the arm and is connected to the tubing of the arm. Gravity draws the artificial blood into the system of veins and helps to fill the tubes when the vein is punctured with a needle. The simulated arm has several insertion sites, but only the antecubital fossa was used in this study. Other materials used in conjunction with the simulated arm were standard surgical gloves, tourniquets, alcohol swabs, gauze pads, 20-gauge needles, and various Vacutainer™ tubes and holders. Upon completion of the procedure, all needles were discarded in a biohazard sharps container.

CathSim® system. The VR simulator used in the present study was the CathSim® system from Immersion Medical, Inc., described earlier (see Figure 2; Immersion Medical, Inc., has released an upgraded version of the CathSim® System since this study was completed). The phlebotomy module includes six case patient scenarios. Students select a patient with the computer mouse and are then shown an image of the patient's arm. Next, they use the mouse to select a site for needle insertion, apply a tourniquet, palpate the vein, and cleanse the site by clicking and dragging objects on the screen. After the site has been prepared, the student manipulates the mouse and the

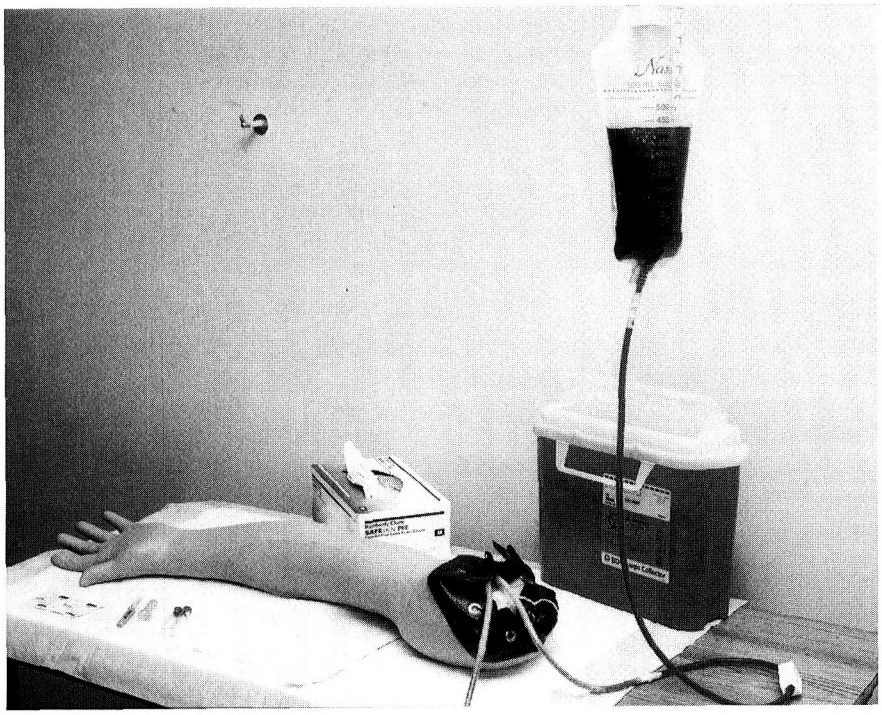

Figure 1. The Life/Form® simulated arm with standard IV bag brings artificial blood into the tubing of the arm.

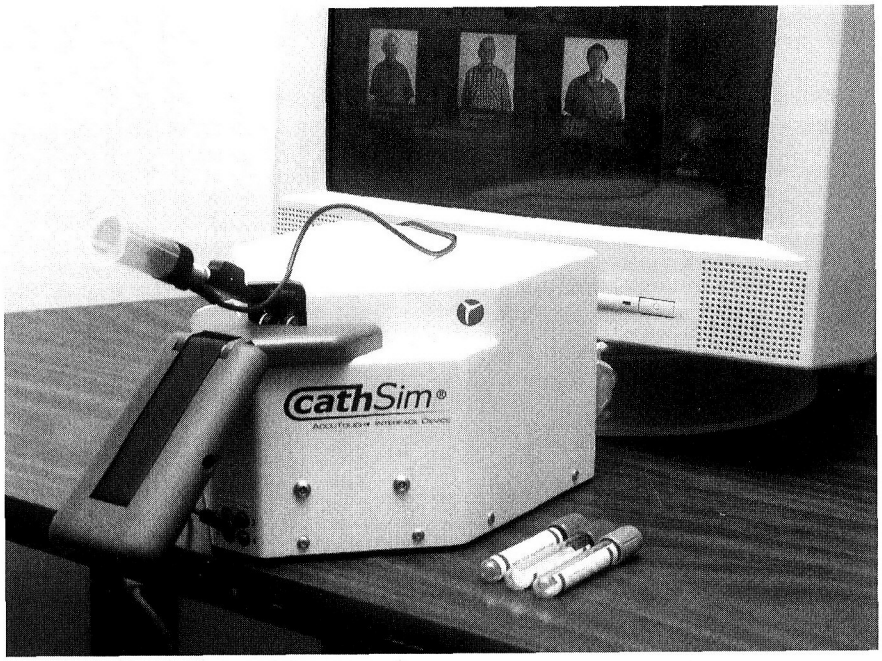

Figure 2. The CathSim® VR multimedia system with AccuTouch® six-degree-of-freedom haptic feedback device that simulates the needle and hub assembly and includes a vinyl strip to simulates the skin.

AccuTouch® device to position the needle. Skin traction is simulated by pulling down on the vinyl strip area of the device with the thumb. Next, the student inserts the needle by fully retracting the arm assembly of the AccuTouch® device and then pushing it forward to insert it back into the device while viewing the computer monitor to ensure vein access. Once the needle is positioned properly, the mouse is used to select the correct order of collection tubes, which are then inserted into and released from the Vacutainer™ holder component of the AccuTouch® device. The student completes the procedure by withdrawing the needle and depositing it in a simulated biohazard sharps container.

Four system-generated performance metrics were recorded. The first was a pain factor with a range of 1 through 10 (lower numbers reflect less pain). The pain score increases as the needle penetrates the flesh to deeper levels and if the needle is rotated after it is embedded in the skin. The second metric was the presence of a hematoma, which occurs if the needle passes entirely through the vein. Further injury increases the rate and intensity of the hematoma. The third metric was the duration that the tourniquet was in place, and the fourth metric was the success or failure of the procedure. None of the augmented view training aids (interior or side views of the underlying vascular system) was used.

Assessment

Performance was assessed using a checklist developed in the medical school to grade students on this procedure. The instrument addresses 28 specific steps of the procedure and provides scores for preparation, insertion, withdrawal/closure, and overall performance. The instrument was modified for the present study to include several rating scales so that performance on some steps could be assessed in a quantitative manner and weighted according to their importance. In addition, four steps were omitted because they could not be assessed with both simulators. The instrument had a range of 0 to 88 points, with higher scores reflecting better performance.

Procedure

All participants were asked to complete a background questionnaire. Afterward, they were shown an instructional video on phlebotomy. Half of the participants were assigned at random to the simulated arm group and the other half to the VR group. However, because of some scheduling changes, 9 students ended up in the VR group and 11 in the simulated arm group. Those participants who had indicated some exposure to phlebotomy were counterbalanced across the groups.

All participants then performed a phlebotomy pretest on the simulated arm. The next 2 weeks consisted of training. Participants were scheduled for two 1-hr labs held 1 week apart. They were trained in individual sessions and received instruction from the experimenter as needed. Those in the simulated arm group practiced until they were able to perform a successful procedure (as determined by the assessment form used for the pretest and posttest) with a tourniquet time of less than 90 s. Those in the VR group worked with the six different case patients. There were slight differences among the cases that made some more challenging than others (e.g., less visible veins or a steeper angle of needle insertion required to access the vein). Students began their training session with the easiest case and continued to practice with this case until they performed a successful procedure with no hematoma, a pain factor of 3 or less, and a tourniquet time of less than 90 s. Once they reached criterion, they moved on to another case patient. The students practiced on the remaining case patients in order of increasing difficulty.

Upon completion of training, all participants performed a posttest on the simulated arm. Their performance was assessed with the same instrument used for the pretest.

RESULTS

Full Procedure

Total scores on the assessment forms were computed for each student on the pretest and posttest. The means for each group are shown on the left side of Table 1. These scores were analyzed using a mixed-factor analysis of variance (ANOVA) with test as the within-subjects factor and group as the between-subjects factor. The results revealed a significant difference between the pretest and posttest, $F(1, 18) = 37.26$, $p < .001$, and a significant interaction, $F(1, 18) = 13.62$, $p < .002$. Differences among the means were examined with Tukey post hoc tests. The

TABLE 1: Mean Assessment Scores on the Pretest and Posttest for Both Groups on the Full Procedure and on Common Steps

	Full Procedure		Common Steps	
	Pretest	Posttest	Pretest	Posttest
Simulated arms	60.82 (8.6)	82.91 (3.5)	36.91 (6.5)	49.1 (2.3)
VR system	63.22 (13.8)	68.67 (7.2)	37.78 (8.6)	42.22 (5.3)

Note. Standard deviations appear in parentheses.

results of those analyses showed that the improvement between the pretest and posttest scores was significant only for those students who worked with the simulated arms ($p < .001$). Further, there was no significant difference between the two groups on the pretest, but posttest scores for the simulated arm group were statistically higher than those for the CathSim® group ($p < .01$).

Performance metrics were collected for all of the students in the CathSim® group during each training session. The means for both sessions are shown in Table 2. The results indicate that students were able to successfully complete about 95% of the procedures across sessions. The mean tourniquet time decreased from Session 1 to Session 2, $t(8) = 6.59, p < .001$. The overall instances of hematoma also decreased from the first to the second training session. A chi-square test on the frequencies indicated that this drop was statistically significant, $\chi^2(1) = 4.314, p < .05$. Last, the mean pain factor declined slightly from Session 1 to Session 2, but the difference failed to reach significance, $t(8) = 2.18, p = .061$.

Common Steps

During the study, it became clear that there were several important differences between the two training methods. Although both systems enable training on phlebotomy, neither device permits trainees to practice *all* of the steps required to successfully perform the procedure. Further, each device differs in the number of steps it addresses as well as the methods for performing those steps.

A post hoc analysis was performed to address these differences. The 24 steps for performing phlebotomy assessed in this study are listed in Table 3. Of those 24 steps, all can be practiced with the simulated arm but only 14 can be practiced with the CathSim® system. For example, both devices allow the user to practice inserting a needle (Step 15); however, only the simulated arm allows the user to practice placing gauze over the needle prior to needle withdrawal (Step 20). Further, those steps marked with a superscript *a* cannot be performed as they would with a genuine patient. For example, neither system allows the student to actually palpate the insertion site (Step 6). Similarly, although the CathSim® system requires the user to apply a tourniquet, the step is accomplished using the computer's mouse. The user cannot practice tying the tourniquet around a patient's arm.

TABLE 2: Performance Metrics for the VR System Participants on Each Training Session Collapsed Across Trials and Cases

Metric	Session 1	Session 2
Percent of successful procedures	94.1	95.4
Mean tourniquet time (s)	98.65 (12.33)	72.9 (14)
Frequency of tourniquet times exceeding 2 min	18	2
Frequency of hematoma		
Moderate	10	2
Serious	8	3
Mean pain score	3.05 (1)	2.18 (0.92)

Note. Standard deviations appear in parentheses.

TABLE 3: Steps of the Phlebotomy Procedure That Can be Practiced Using the Simulated Arm (S) and CathSim® System (C)

S		1. Gather and arrange all materials (tourniquet, Vacutainers™, gauze pads, etc.)
S		2. Wash hands
S		3. Position tourniquet
S	C[a]	4. Apply tourniquet
S[a]	C	5. Select phlebotomy site
S[a]	C[a]	6. Palpate venipuncture site
S		7. Put on gloves
S	C[a]	8. Prepare site with alcohol or Betadine stick
S		9. Sterilize collection tube tops with alcohol pad
S	C	10. Grasp needle holder behind the needle hub
S		11. Remove the needle's protective sleeve
S	C[a]	12. Stabilize vein
S	C	13. Hold needle at 15° to 30° angle
S	C[a]	14. Enter site bevel side up
S	C	15. Insert needle
S	C	16. Advance needle into vein
S	C	17. Depress tubes into Vacutainer until filling stops
S	C	18. Disengage collection tube
S	C[a]	19. Release tourniquet
S		20. Place gauze pad over needle
S	C	21. Withdraw needle
S		22. Press gauze pad onto puncture site and apply pressure
S		23. Properly label specimens
S[a]		24. Recheck puncture site for bleeding

[a]Indicates limited or partial ability to practice the step.

Given these differences, it is possible that the students who trained with the simulated arms achieved higher scores because they had the opportunity to practice more steps from the full procedure. Thus, a second analysis was performed on only those 14 steps common to both training methods. The results of that analysis are shown on the right side of Table 1. As can be seen in the table, the pattern of results is similar to that of the full procedure. Again, there was a significant difference between the pretest and posttest, $F(1, 18) = 28.19$, $p < .001$, and a significant interaction, $F(1, 18) = 6.1$, $p < .025$. Subsequent Tukey tests indicated a significant improvement for only those students who worked with the simulated arms ($p < .001$). Again, there was no significant difference between the two groups on the pretest, but scores for the simulated limb group were statistically higher than for the CathSim® group on the posttest ($p < .01$).

Specific differences between the two systems on the 14 common steps are shown in Table 4. Mean scores achieved on each step and the proportion of students who correctly performed and received full credit for each step are shown in the table. Because of the potential for inflated alpha levels, given the large number of comparisons, these data were not analyzed statistically; however, some trends are apparent. Performance was comparable on the two systems for 9 of the 14 steps. On the remaining 5 steps (marked with a superscript *a* in the table), however, there was an advantage for the simulated arm. For these steps, the overall mean scores were greater and a higher proportion of students were able to achieve the maximum score having trained with the simulated arm.

DISCUSSION

The primary goal of the present study was to compare the CathSim® VR system and simulated limbs for training phlebotomy. As noted earlier, there are no universal standards for training phlebotomists, and performance can vary widely from clinic to clinic and person to person. These systems offer the opportunity for students to acquire the skills needed to perform phlebotomy by practicing on devices instead of patients. Although the results showed that all participants benefited to some degree from their training, improvement

TABLE 4: Mean Scores and Proportions of Students Who Correctly Performed the 14 Common Steps of the Procedure Using the Simulated Arm (S) and CathSim® System (C)

Common Steps	Limited Fidelity	Sim. Arm Mean Score	CathSim Mean Score	Sim. Arm Proportion Correct	CathSim Proportion Correct
Apply tourniquet[a]	C	4.44	2.67	.82	.33
Select site	S	4.00	3.11	1.00	.78
Palpate site	S, C	2.00	1.77	1.00	.89
Prepare site[a]	C	4.00	2.67	1.00	.44
Grasp needle holder[a]		2.00	1.11	1.00	.56
Stabilize vein[a]	C	3.64	0.67	.82	.11
Angle needle properly		4.00	3.77	1.00	.89
Enter site, bevel up[a]	C	4.00	3.33	1.00	.67
Insert needle		3.82	3.77	.91	.89
Advance needle		6.00	6.00	1.00	1.00
Depress tubes		3.27	3.55	.73	.78
Disengage tube		4.00	4.00	1.00	1.00
Release tourniquet	C	1.82	2.00	.91	1.00
Withdraw needle		3.82	3.33	.91	.78

[a]Indicates an advantage for the simulated arm over the CathSim system.

from the pretest to posttest was statistically significant only for those students who practiced with the simulated arms. In this regard, these data are consistent with those of other studies showing either no advantage for a VR system over simulated limbs (Chang et al., 2002) or a disadvantage of VR compared with simulated limbs for training IV cannulation (Engum et al., 2003; Scerbo et al., 2004). The results from this study and the others suggest that there is no advantage for training with the CathSim® system over the more traditional simulated arms for either phlebotomy or cannulation procedures.

Comparison of the CathSim® and Simulated Limb Devices

Becoming adept at phlebotomy requires trainees to practice all steps of the procedure using equipment and activities that are as similar as possible to those used in the actual task. Doing so ensures that the learning experience is relevant and complete. It also allows for positive transfer from the training situation to the applied task. Classical approaches to training transfer emphasize similarity between the basic elements of the training task and the target task (identical elements theory) or between the principles conveyed by the training and the target tasks (transfer-through-principle theory; Holding, 1965).

Central to the success of any training method is the realism, or fidelity, of the training task. Hays and Singer (1989) explained that fidelity can be expressed both physically and functionally. These two types of fidelity may be mapped directly onto the conditions of "stimulus" and "response" similarity discussed by Holding (1965). Training systems that have high physical fidelity include many of the same sensory stimuli that are part of the actual equipment. Those that have high functional fidelity require many of the same responses or actions required by the actual equipment. Ideally, an optimal training system would have both high physical and functional fidelity.

Specifying physical and functional fidelity differences on the 14 common steps trained by the CathSim® system and the simulated limb may help clarify why the participants who trained with the simulated limb performed better on the posttest.

Regarding physical fidelity, the two devices vary most on the following steps: tourniquet application (an actual tourniquet is applied to the simulated limb; CathSim® displays an icon), preparing the site with alcohol or Betadine (an actual alcohol pad is used with the simulated limb for better monitoring of application technique; CathSim® displays an icon), stabilizing the vein (the simulated limb allows participants to experience feeling a vein with their fingers; CathSim®

has a flat rubber pad), and orienting the needle properly (the simulated limb uses a genuine needle so that one can monitor the orientation of the beveled end; CathSim® uses a facsimile of a needle without a beveled end).

With respect to functional fidelity, there are also differences between the two devices. This is particularly relevant because some researchers have argued that functional fidelity may be more important than physical fidelity for ensuring acquisition of skills (Gopher, Weil, Bareket, & Caspi, 1988). Because an actual tourniquet was used with the simulated limb, the trainees were able to practice tying the tourniquet. On the CathSim® system, the tourniquet is placed by clicking the tourniquet icon. This functional difference may have contributed to the lower posttest scores of CathSim® students on this step. Although both devices allow participants to cleanse the site with alcohol, only the simulated limb permits the opportunity to confirm that the site is dry before proceeding. This difference may have led to lower posttest scores for the CathSim® students on this step as well. In addition, the simulated limb provides some opportunity to practice stabilizing the vein (i.e., students can position their fingers on the vein, but the vein is largely fixed and does not move much). On the CathSim® system, however, this step must be performed on the simulated skin pad of the AccuTouch® device, which is flat and does not have any protruding areas to represent veins. This is another functional difference that may have led to lower posttest scores for the CathSim® students on this step. Finally, as noted earlier, ensuring that the needle enters the site with the bevel side up requires more proactive effort for participants using the simulated limb. This step cannot be performed properly with the CathSim® system and may also have contributed to the lower posttest scores for these students.

In addition to these differences, there are also some sensory deficiencies inherent in CathSim® system design. Visually, it is not possible to monitor all aspects of the phlebotomy procedure because of the system's limited field of vision. Auditorily, CathSim® provides utterances to signify the experience of pain by the patient; however, no other auditory signals are provided, such as heart rate, respiration, or patient commentary. There are two other design-oriented concerns regarding the CathSim® system. First, the CathSim® trainer is most easily used when seated, whereas phlebotomy is typically performed while standing. This difference influences other task performance steps, including the angle of needle insertion. Also, because CathSim® users must interact simultaneously with the computer mouse and the needle-hub assembly, there is potential for motor interference.

The physical and functional differences between the two systems may help to explain why the two groups attained different scores on the posttest and also provide guidance for designers of virtual phlebotomy systems and other medical training devices. The performance differences observed in the present study suggest that limitations in virtual simulators that compromise physical and functional fidelity can undermine the very training benefits espoused by proponents of this technology.

Despite these design limitations, the CathSim® system does provide other learning experiences. In particular, the system records a variety of performance metrics that are unavailable or not easily obtainable with simulated limbs. For example, students are typically instructed not to keep a tourniquet fastened around a patient's arm for longer than 2 min because it can affect the quality of the blood sample. As noted in Table 2, the mean tourniquet time decreased from about 99 to 73 s across training sessions. However, there were 18 instances in the first training session in which the tourniquet time exceeded the 2-min limit, and the mean of those 18 instances was 180 s. By the second session, there were only 2 instances that exceeded the limit, and the mean of those was 137 s. If these attempts had been conducted with genuine patients, the procedure would have needed to be terminated and started on the other arm. The CathSim® system also records an index of hematoma that is not observable with the simulated limbs. As noted previously, there was a significant decrease in occurrences of hematoma from the first to the second training session. Thus, one could argue that practicing on the simulator prevented 23 potential instances of hematoma from occurring in actual patients. This is an important point because performance in the present study was assessed on the simulated limbs and not genuine patients. Thus, the CathSim® system may provide a critical measure of performance that would be observable in actual patients but which

was not assessed with the simulated limbs used for the pretest and posttest in the present study.

In addition, the CathSim® system provides exposure to a range, albeit limited, of patient characteristics. Further, the system includes two auxiliary visual aids (a cutaway side view that shows the needle penetrating the arm and a transparency view of the underlying vascular system) that were not examined in the present study.

Although the functional and physical differences between the two training systems may explain the lower posttest scores for the CathSim® students, it is important to remember that neither system completely or faithfully represents the phlebotomy task as it is performed in practice. As noted in Table 3, neither device allows trainees to practice palpating the insertion site or to reinspect the site for bleeding, as one would do with a genuine patient. Also, site selection posed no real challenge. The veins in the simulated limb protrude from the arm, making site selection obvious. Further, the skin covering the simulated limb retains puncture marks that subsequent trainees can see and use for guidance. On CathSim®, the computer's mouse is used to select a site and will allow palpation only in specific areas of the arm. Further, neither system permits the user to "damage" nearby structures, and therefore students do not have an opportunity to learn about the risks associated with poorly executed procedures. Finally, neither the simulated limb nor the CathSim® system incorporates the stress, workload, or level of feedback present in an actual clinical setting.

In sum, there may be advantages and disadvantages to both the CathSim® system and simulated limbs for training phlebotomy. In fact, the optimal training program may be one that incorporates both systems and capitalizes on the unique advantages afforded by each. Moreover, the opportunity to practice on either system before attempting the procedure on genuine patients could potentially decrease risks to patients as well as the incidence of needle stick injuries across all health care facilities.

Future Research Issues

The results from the present study suggest that simulated limbs may be more effective than a VR system for training phlebotomy. However, it is important to understand that *both* groups were assessed on the simulated limbs. Thus, one could argue that the students who trained with the CathSim® system were at a disadvantage because they were assessed on a system that differed from the one on which they practiced. A decision was made to use simulated limbs for the pretest and posttest in the present study for several reasons. First, we wanted to obtain a baseline measure of performance on as many steps of the procedure as possible and, for obvious safety reasons, could not permit inexperienced students to draw blood from genuine patients without supervision or the possibility of having a supervisor intervene. Second, we wanted the baseline measure to address performance on the medical procedure and therefore be free from activities needed to familiarize oneself with the computer interface for the VR system. Last, the simulated limbs offered a standardized platform for assessment. Although this approach enables comparisons of performance before and after training, a more genuine and clearer picture of the efficacy of the CathSim® and simulated limb methods necessitates that performance be evaluated on a "neutral" target task. Ultimately, these training methods must be evaluated where it counts most – with genuine patients.

To date, most of the data supporting the benefits of VR medical simulators have not been gathered from actual patients. Instead, they have been generated from the systems themselves. There are numerous technical and practical issues that impact efforts to evaluate medical simulators in clinical settings (see Satava & Jones, 2002; Scerbo, in press); however, progress is being made. Recently, Seymour et al. (2002) reported some training benefits with a VR system for fundamental laparoscopic skills gathered from real patients.

Although evaluations based on system-generated metrics of the sort reported in this study are valuable, they are limited to demonstrating improvement over time and distinguishing between expert and novice performance. However, true measures of training transfer can be derived only from data collected in the target task environment (i.e., with real patients). Moreover, the ability to assess the predictive validity of system-generated metrics and to investigate the concurrent validity of simulator-based training with other procedures hinges upon gathering data with genuine patients. Accordingly, to address this limitation of the present study, a second experiment is

currently under way to examine the degree of training transfer afforded by the CathSim® and simulated limb methods for phlebotomy with genuine patients.

The results from studies such as ours are a necessary first step to help establish the effectiveness of VR medical simulators if the science and practice of medicine is to benefit from this technology. Ultimately, however, the degree to which time spent on simulators predicts performance with genuine patients needs to be understood so that the science and practice of medicine can be studied in a safe, controlled environment. Just as flight simulators allow researchers to examine pilot performance under a wide variety of conditions, medical simulators promise to allow the study of physicians, surgeons, and other medical personnel under an unlimited array of contexts. For example, simulators may be used to study the effects of fatigue, stress, workload, and work schedules on surgical skills, decision making, and problem solving. More important, research on potential moderating factors and countermeasures can also be studied. Medical VR simulators can help establish more efficient methods for training by capitalizing on principles of knowledge and skill acquisition that address task sequencing, part- and whole-task training, and integrative approaches (Proctor & Dutta, 1995); the development of automatic skills (Schneider, 1985); and adaptive training paradigms (Krahl & Scerbo, 1997). Medical VR simulators can also be targeted to individuals or teams of surgeons, residents, anesthesiologists, and nurses (Helmreich & Schaefer, 1994).

Conclusion

Recently, the Institute of Medicine issued a report, *To Err Is Human* (Kohn, Corrigan, & Donaldson, 1999), that suggested medical errors contribute to as many as 98,000 deaths annually in U.S. hospitals. A more recent report issued by Health Grades, Inc. (2004) suggested that the number of fatalities may be double the original estimate. In response to concerns about resident fatigue and patient safety, the American Medical Association's Accreditation Council for Graduate Medical Education, the organization that accredits teaching hospitals in the United States, recently set limits on the working hours for medical residents (American Medical Association, 2003). Although residents may welcome the reduction in working hours, the restrictions have created new pressures to train physicians and surgeons to higher levels of competency in shorter periods of time.

Medical VR simulators may provide a partial solution to these problems. They offer the opportunity for students, residents, and even practicing physicians to have access to alternative forms of training at any time of the day or night. Despite this obvious advantage, considerable work still needs to be done to validate the systems and establish rates of training transfer. The results from the present study indicate that a medical VR system for phlebotomy may not be as effective as a more traditional low-tech method, given the functional and physical differences between the systems. Clearly, there is a need for those in the human factors community to participate in the design and evaluation of medical VR simulators so that this technology can evolve with the greatest benefits for the end user.

ACKNOWLEDGMENTS

This study was a collaborative project between the Virginia Modeling, Analysis and Simulation Center (VMASC) at Old Dominion University and the Eastern Virginia Medical School. Partial funding for this study was provided by the Telemedicine and Advanced Technologies Research Center, U.S. Army Medical Research and Material Command, Contract DAMD17-03-2-0004, and the Naval Health Research Center through Naval Air Systems Command Training Systems Division under Contract N61339-03-C-0157, entitled "The National Center for Collaboration in Medical Modeling and Simulation." The ideas and opinions presented in this paper represent the views of the authors and do not necessarily represent the views of the Department of Defense.

REFERENCES

American Medical Association. (2003). *AMA delegates approve limits on resident working hours.* Retrieved August 12, 2003, from http://www.ama-assn.org/ama/pub/category/11106.html

Basdogan, C., & Srinivasan, M. A. (2002). Haptic rendering in virtual environments. In K. M. Stanney (Ed.), *Handbook of virtual environments: Design, implementation and applications* (pp. 117–134). Mahwah, NJ: Erlbaum.

Chang, K. K., Chung, J. W., & Wong, T. K. (2002). Learning intravenous cannulation: A comparison of the conventional method and the CathSim intravenous training system. *Journal of Clinical Nursing, 11,* 73–78.

Clover, S. (2002, July). *Phlebotomy safety: Sticking to good practices.* Retrieved January 11, 2006, from the Advance Newsmagazines for

Medical Laboratory Professionals Web site: http://www.advanceformlp.com/common/editorial/PrintFriendly.aspx?CC=6475

Dawson, S. L. (2002). A critical approach to medical simulation. *Bulletin of the American College of Surgeons, 87*(11), 12–18.

Engum, S. A., Jeffries, P., & Fisher, L. (2003). Intravenous catheter training system: Computer-based education versus traditional learning methods. *American Journal of Surgery, 186,* 67–74.

Gopher, D., Weil, M., Bareket, T., & Caspi, S. (1988). Fidelity of task structure as a guiding principle in the development of skill trainers based upon complex computer games. In *Proceedings of the Human Factors Society 32nd Annual Meeting* (pp. 1266–1270). Santa Monica, CA: Human Factors and Ergonomics Society.

Hays, R. T., & Singer, M. J. (1989). *Simulator fidelity in training system design.* New York: Springer-Verlag.

Health Grades, Inc. (2004). *Patient safety in American hospitals.* Lakewood, CO: Author.

Helmreich, R. L., & Schaefer, H. (1994). Team performance in the operating room. In M. S. Bogner (Ed.), *Human error in medicine* (pp. 225–253). Hillsdale, NJ: Erlbaum.

Holding, D. H. (1965). *Principles of training.* London: Pergamon Press.

Howanitz, P. J., Cembrowski, G. S., & Bachner, P. (1991). Laboratory phlebotomy: College of American Pathologists Q-probe study of patient satisfaction and complications in 23 783 patients. *Archives of Pathology and Laboratory Medicine, 115,* 867–872.

Kohn, L. T., Corrigan, J. M., & Donaldson, M. S. (Eds.). (1999). *To err is human: Building a safer health system.* Washington, DC: National Academy Press.

Krahl, K., & Scerbo, M. W. (1997). Performance on an adaptive automated team tracking task with human and computer teammates. In *Proceedings of the Human Factors and Ergonomics Society 41st Annual Meeting* (pp. 551–555). Santa Monica, CA: Human Factors and Ergonomics Society.

Mishori, R. (2004, June 1). Drawing without a license: Phlebotomists get little training, regulation. *Washington Post,* p. HE1.

National Institute for Occupational Safety and Health. (1999). *How to protect yourself from needlestick injuries.* Retrieved June 9, 2004, from http://www.cdc.gov/niosh/2000-135.html

Proctor, R. W., & Dutta, A. (1995). *Skill acquisition and human performance.* Thousand Oaks, CA: Sage.

Prystowsky, J. B., Regehr, G., Rogers, D. A., Loan, J. P., Hiemenz, L. L., & Smith, K. M. (1999). A virtual reality module for intravenous catheter placement. *American Journal of Surgery, 177,* 171–175.

Reznek, M. A., Rawn, C. L., & Krummel, T. M. (2002). Evaluation of the educational effectiveness of a virtual reality intravenous insertion simulator. *Academic Emergency Medicine, 9,* 1319–1325.

Satava, R. M. (1993). Virtual reality surgical simulator: The first steps. *Surgical Endoscopy, 7,* 203–205.

Satava, R. M. (2001). Accomplishments and challenges of surgical simulation: Dawning of the next-generation surgical education. *Surgical Endoscopy, 15,* 232–241.

Satava, R. M., & Jones, S. B. (2002). Medical applications of virtual environments. In K. M. Stanney (Ed.), *Handbook of virtual environments: Design, implementation and applications* (pp. 937–957). Mahwah, NJ: Erlbaum.

Scerbo, M. W. (in press). Medical virtual reality simulators. In W. Karwowski (Ed.), *International encyclopedia of ergonomics and human factors* (2nd ed.). Boca Raton, FL: CRC Press.

Scerbo, M. W., Bliss, J. P., Schmidt, E. A., Thompson, S. N., Cox, T. D., & Poland, H. J. (2004). A comparison of the CathSim system and simulated limbs for teaching intravenous cannulation. In J. D. Westwood, R. S. Haluck, H. M. Hoffman, G. T. Mogel, R. Phillips, & R. A. Robb (Eds.), *Medicine meets virtual reality 12* (pp. 196–198). Amsterdam: IOS Press.

Schneider, W. (1985). Training high-performance skills: Fallacies and guidelines. *Human Factors, 27,* 285–300.

Seymour, N. E., Gallagher, A. G., Roman, A. A., O'Brien, M. K., Bansal, V. K., Andersen, D. K., et al. (2002). Virtual reality training improves operating room performance: Results of a randomized, double-blinded study. *Annals of Surgery, 236,* 458–464.

Ursino, M., Tasto, J. L., Nguyen, B. H., Cunningham, R., & Merril, G. L. (1999). CathSim: An intravenous catheterization simulator on a PC. In J. D. Westwood, H. M. Hoffman, R. A. Robb, & D. Stredney (Eds.), *Medicine meets virtual reality VII* (pp. 360–366). Amsterdam: IOS Press.

Mark W. Scerbo is a professor in the Department of Psychology at Old Dominion University in Norfolk, Virginia. He received his Ph.D. in experimental psychology from the University of Cincinnati in 1987.

James P. Bliss is an associate professor in the Department of Psychology at Old Dominion University in Norfolk, Virginia. He received his Ph.D. in human factors psychology from the University of Central Florida in 1993.

Elizabeth A. Schmidt is a doctoral student of human factors psychology at Old Dominion University in Norfolk, Virginia. She received her M.S. in psychology from Old Dominion University in 2003.

Sommer N. Thompson is a customer experience specialist for a division of Landmark Communications in Norfolk, Virginia. She received her B.S. with honors in psychology from Old Dominion University in 2004.

Date received: January 23, 2004
Date accepted: December 27, 2004