

**Validation of the Course for Operationally Relevant Patient Safety (CORPS)**  
An Avatar Guided Computer Game-Based TeamSTEPPS® Patient Safety Program

Andrea Parodi, RN, Ph.D.	Ross Gore, Ph.D.	Tom Frost, M.S.	Hector Garcia Ph.D. (candidate)
	Virginia Modeling, Analysis, and Simulation Center Old Dominion University Suffolk, VA		
<a href="mailto:aparodi@odu.edu">aparodi@odu.edu</a>	<a href="mailto:rgore@odu.edu">rgore@odu.edu</a>	<a href="mailto:tfrost@odu.edu">tfrost@odu.edu</a>	<a href="mailto:hgarcia@odu.edu">hgarcia@odu.edu</a>

### ABSTRACT

A VMASC research team and numerous active duty clinicians collaborated to devise an instructional assist program and methodology designed as a serious computer based game and training platform for military health care professionals. The learner is guided by an avatar-preceptor through course content and multi-media enrichment resources. She/he can present content through course completion using deliberative practice that progresses the learner to content mastery. In this case, the content is based on Team Strategies and Tools to Enhance Performance and Patient Safety (TeamSTEPPS®), a patient safety and communications program in response to the DOD and national imperative to reduce preventable patient harm due to errors in care. Currently, this content is taught using a didactic approach, necessitating students and instructors leaving their workspace. Few measures of student program success exist due to administrative workload. We automated capture of content learning measures using validated instruments. We call this program the Course for Operationally Relevant Patient Safety or CORPS. The main research objective was to determine the efficacy of CORPS in teaching concepts of teamwork and communications in comparison to a control group that also uses TeamSTEPPS v. 2 content but instructs using a traditional didactic presentation. This study supports military medical deployment team needs and was conducted at a large military healthcare facility. Participating in the study were 320 learners. Results indicate the experimental group performance was as good as the control. There was a slight trending toward superior performance, but the process generated significant administrative time and cost savings. In three months, the course director reported actual time-savings of over 320 man hours that could be used for team simulation practice. At present, the Department of Patient Safety at the test site continues to use the CORPS prototype because of its effectiveness and demonstrated savings of time and cost.

### ABOUT THE AUTHORS

V. Andrea Parodi, RN, Ph.D. is a Research Associate Professor and the Lead for medical/healthcare focus at Virginia Modeling, Analysis, and Simulation Center, a dedicated research laboratory within Old Dominion University (ODU) in Suffolk, VA. Her academic background includes a Ph.D. in Nursing from the University of Alabama at Birmingham with dual specializations of Health Policy Analysis and Higher Education Leadership. Her doctoral Health Policy residency involved working in the office of U.S. Senator Daniel Inouye (D-HA) in Washington, DC. Dr. Parodi also earned an MSN as a clinical nurse specialist in critical care from Vanderbilt University, a BA in Psychology from College of Mount St. Vincent on-the-Hudson and her BSN work was done at Pace University in New York. Completing 26 years as a U.S. Navy Nurse Corps officer, Dr. Parodi then joined VMASC at ODU. Current funded research includes the design and testing of TeamSTEPPS® and trauma knowledge assessment instruments and training systems for professional military clinicians. This current program was awarded “Best in Show” at the annual meeting of IMSH 2017 in Orlando, FL sponsored by the Society for Simulation in Healthcare (SSH).

Ross Gore, Ph.D. is a Research Assistant Professor and Lead for Big Data and Prediction at the Virginia Modeling, Analysis & Simulation Center at Old Dominion University. He also specializes in the testing, verification, and validation of computer models. He holds a Doctorate of Philosophy (Ph.D.) and a Master's degree in Computer Science from the University of Virginia and a Bachelor's degree in Computer Science from the University of Richmond. Dr. Gore has more than 10 years of research experience in problems that lie at the intersection of computer science and modeling and simulation. His work has yielded authorships on more than 50 conference and journal publications.

## BACKGROUND

### Patient Safety in an Operational Environment – Preparing to Jump into Clinical Practice

Military Medical Corps, Nurse Corps and clinical Medical Service Corps officers face many challenges that are unique to deployment clinical practice roles. The practice setting is likely more austere than stateside hospitals. The practice setting experiences frequent staff turnover that can de-stabilize the best of teams. Deployment has its own unique distractions, challenges and potentially threats of exposure to significant hazards, regardless of the assigned operational platform or Service. The greater the stress, the harder it is for clinicians to safely manage patient access, and care. When the environment has few back up options and high turnover, the need for a high performing team and turn-over process becomes especially critical and a team that practices the techniques of TeamSTEPPS® is better prepared to handle the challenges. Traditional didactic instruction, that forms the basis of most clinical training, has inherent variability, and inconsistency of quality, faculty and content emphasis for each member of the healthcare team. Training variability can only compound the challenge of getting the entire patient care team “on the same page” to promote quality care and reduce and prevent opportunities for medical error and patient injury or death. Joint Commission sentinel events data has identified the root cause of 70% of preventable medical errors cited involve communication issues. Of identified serious preventable medical errors, 80% involve mis-communication between caregivers during patient transfers or hand-offs. (Joint Commission Center for Transforming Healthcare, 2014) That is huge, unacceptable and preventable! Medical errors can be related to omission of appropriate care or commission of administering inappropriate care (America, 2001). In response to patient safety concerns and the desire for the medical enterprise to be composed of high performance, quality teams, the Department of Defense (DOD) mandates that all clinicians shall receive TeamSTEPPS® training. Scientifically rooted in more than 20 years of research and lessons learned from the application of teamwork principles, this DOD developed Patient Safety Program is a collaboration with the Agency for Healthcare Research and Quality. TeamSTEPPS® provides higher quality, safer patient care by producing highly effective medical teams that optimize the use of information, people and resources to achieve the best clinical outcomes for patients. Increasing team awareness and clarifying team roles and responsibilities helps prevent or resolve conflicts and improves information sharing by eliminating barriers to quality and safety. (“National Healthcare Quality & Disparities Report,” 2014) This study was designed to provide an operationally relevant TeamSTEPPS® course to military healthcare and medical personnel using a non-traditional pedagogical approach.

### Why Are We Testing a New Teaching Approach and What Is This New Approach?

This research team is using a “standardized” curriculum that is required by Army, Navy, and Air Force clinicians. As stated above, the goal of a TeamSTEPPS® trained team is to promote a high functioning team, catch or eliminate preventable patient injury, and maximize team efficiency. Testing an educational program and process on a widely used curriculum gave us the ability to study a highly representative group of learners working within a military healthcare facility. Anecdotally, the research team spoke to staff educators from different military medical centers. It was noted that more content and courses are required with less time allotted for intense training and there are limited resources to fund new programs. Staffing is often short, making it difficult to have a physician, nurse, or Physician’s Assistant leave their workspace, which makes student access more difficult. The amount of new knowledge coming from medical and nursing research alone is staggering. The challenge for all professionals is to stay current with the literature, practices and processes. This challenge necessitates, at least in part, that the professional needs to rely on technology. Our challenge is to identify ways to learn more content in a more time efficient way and ideally, retain this information for longer periods of time. Also, when reviewing the content of some “typical” pre-deployment medical training courses, there is too much content for the time allotted. An extensive study of cognitive absorption produced an in-depth analysis of education/learning concepts such as engagement, arousal, enjoyment, self-efficacy, control, challenge, computer skills and ease of use to name a few. This work addressed what we may call components of “learning readiness” (Agarwal & Karahanna, 2000). Learning readiness must be given greater consideration to improve the amount and duration of cognitive absorption. Consider the following elements when assessing what impacts readiness and can be influenced by the trainer/educator, and Command climate. These elements are: learning environment, state of arousal, fatigue, degree of interest, engagement, enjoyment, learner’s perceived value of the learning and the technology used or the lack of appropriate technology supports. Clearly, to actuate students with various learning styles and learning readiness levels, it is important to provide a course pedagogy and technology-supports that are complimentary to each other and the learner’s perceived needs and readiness.

“Just in time training” is often just too late. Ideally, a training process should be more flexible and mobile like our learners, able to reinforce important learning strategies like deliberative practice, mastery training standards, self-

spacing with content presentation and performance tracking. A hybrid educational approach could provide immediate feedback and guide the learner to a wealth of multi-media content resources, and training should try to maximize the cognitive absorption and avoid periods when the learner's time, attention, personal and professional needs all collide with the realities of pre-deployment anxiety and professional demands. Sitting in a classroom during the learner's pre-deployment phase for hours or days only adds to the learner's frustration. Innovation is needed to incorporate more of a data driven focus on individual training outcomes, learning duration and the incorporation of easy access to sustainment programs and apps. The end-game for the military clinician is to possess consistent and persistent knowledge and skills. This study is addressing only the knowledge component as skills maintenance is a greater challenge, but one that will require knowledge linkage as a foundation for the mastery of the skills. These are some reasons why new pedagogical processes need to be studied and developed. Additional insight into the history of learning will show the roots of some of the "new" education approaches come from the insights of much earlier times.

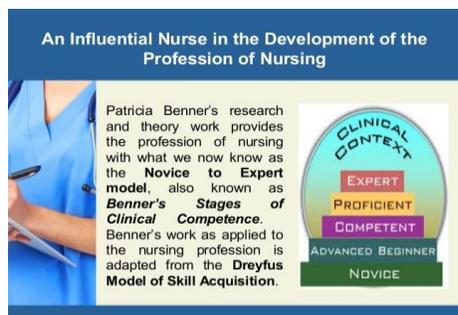
### **A Historical Look at the Training/Education Pedagogy Common to the Medical and Healthcare Disciplines**

In 1908, the American Medical Association asked the Carnegie Foundation to undertake an extensive study of medical education. The report authored by Abraham Flexner in 1910, called for greater university preparation by enacting higher admission and graduation standards, and to adhere strictly to the protocols of mainstream science in their teaching and research. Additionally, Flexner stated that a hospital for clinical training was as important to a medical school as a laboratory was for training in chemistry or pathology. He specified that adequate laboratories needed to be well equipped and supervised. And so, the hospital internship, acting as an apprenticeship, provided the learning laboratory for clinical practice in addition to providing service clinicians to the training hospital. (Markel, 2010) The impact of the Flexner Report was very significant. The medical community moved rapidly to adopt the recommendations made creating an enormous change in physician education, mentorship as well as striking a chord with other professional disciplines. For example, during this period of the early 1900s Nursing education was based on the new practices and science out of Europe, mainly France and England. Prior to the advent of science based training, nursing was in fact not a professional discipline but more like an indentured servitude of the untrained. Often these unprofessional nurses were working off debt or impressed "ladies of the night" who were assigned custodial duties or care of the sick. These untrained women were pariahs of society and tasked to work in lieu of jail time. With the advent of educated and "respectable" women and the use of nuns, nursing developed an extensive apprenticeship-based process within a diploma granting hospital training program, not a university. Seeing the impact the Flexner report had on the advancement of medicine and the stature of its people, Nursing commissioned the Rockefeller Foundation in 1920 to conduct a study of the education and training process for nurses. Published in 1923, the Goldmark Report was a research-based document that evaluated and identified inadequacies in the educational preparation of nurse educators, nurse administrators, public health nurses, and nursing students. (Norwood, 2000) The Committee recommended that schools of nursing educate nurses in a manner that was independent of hospitals and opposed utilizing nursing students as an exploited source of labor. Most importantly, the Goldmark Report recommended that the entry to practice should be the Bachelor of Science Nursing degree (BSN). Unfortunately, nursing has enacted the Goldmark recommendations at a rate far slower than the medical discipline. In fact, we still fail to recognize the BSN as the entry to practice degree. (Kalisch & Kalisch, 1995)

Fast-forward to the present era, and in 2010 the Carnegie Foundation created extensive reviews on the education and needs of physicians and the discipline of medicine. Additionally, another report was published on societal needs and the nursing education process today (Cooke, Molly & O'Brien, Bridget C., 2010) (Benner, 2010). Recommendations for medical education and training included the use of more simulators and simulations, and there was a greater emphasis on the integration of new technologies. Additionally, it was recommended that learners at all levels should not be obliged to spend time unproductively repeating clinical activities once they have mastered the competencies appropriate to their level. The report further stipulated that Medical education must make much more use of readiness assessments and design curricula that are sufficiently flexible to allow individual learners to engage at various levels of difficulty. This information resonated and is at the heart of looking at a viable alternative to didactic presentation and the typical, traditional classroom presentation and testing process. What seemed a natural fit to all these recommendations and observations was the use of various types of simulation, more spot check capability/knowledge assessments using artificial intelligence, and more targeted course content based on data acquisition from individual student competency measures. The use of computer based programs and testing could be designed to complement these recommendations. So, for this study, we looked at the integration of multiple types of

simulation with a training and testing platform that is web-based and populates a secure cloud with course grades or other content. This process reduces administrative work and eliminates the potential of grade transcription errors. Learning and testing locations could occur at almost any time or any place as long as Wi-Fi connectivity is available. Learners can and like to use their own personal mobile devices like a tablet, phone, or laptop.

### Theoretical Frameworks



**Figure 1. Benner's Novice to Expert Model**



**Figure 2. Kirkpatrick's Evaluation Model**

### Theoretical Frameworks Guiding this Study

This study is guided by two theoretical frameworks, Benner and Kirkpatrick. (Benner, P., 2001) (Kirkpatrick 14 Feb 2017). The Benner Model (See Figure 1.) guides our curriculum and scenario designs. The Kirkpatrick Model (See Figure 2.) will be used to assess the program's impact on the time to knowledge acquisition, confidence, satisfaction, and ultimately demonstrated performance of this research program. The theoretical framework developed by Patricia Benner is known as Benner's Stages of Clinical Competence. This framework is an adaptation from the Dreyfus Model of Skill Acquisition and relates directly to the work of clinical practice across the disciplines. The importance of using a guiding conceptual /theoretical framework for the curriculum and software design is that it assists the educator to target the learners needs and more effectively help to advance the learner to the next concept or skill level. This model has been widely and successfully used not only in the design and execution of "traditional" classroom learning experiences but has found its way to be most useful in high fidelity clinical and team simulations in nursing (Sanford, Pamela, 2010) (Turner & Parodi, 2012).

### Kirkpatrick Model of Program Effectiveness

Kirkpatrick's model of training evaluation is one of the better-known models of educational design and training effectiveness. Kirkpatrick identified four levels of training evaluation. Level I is called the reaction. This level strives to answer the following question: did the learners like the training? Did the learners feel the training was relevant to their work? Kirkpatrick called this level the "smile sheet" phase (Kirkpatrick & Kirkpatrick, 2006), (Kirkpatrick, 1998). He notes, that although a positive reaction does not guarantee learning, a negative reaction reduces the possibility for continued learning. Level II is the learning phase. This level shifts the emphasis beyond learner satisfaction to the assessment of the acquisition of new or advanced skills, knowledge or attitudes. Commonly, pre and post knowledge assessment tests are used to determine knowledge within a specific course or topic. This is not the only or the most accurate way to assess knowledge. The measurement and metrics of this level are more challenging to achieve. Level III is the behavior phase. At this juncture, it must be asked if the newly acquired skills, knowledge or attitudes stemming from the training are now being used in the everyday environment of the learner? For many educators, this level represents the truest assessment of a program's effectiveness. However, measuring at this level is very difficult as it is often impossible to predict when the change in behavior will occur and thus requires important decisions in terms of when to evaluate, how often to evaluate and how to evaluate. Level IV is the results phase. From an organizational perspective, this is the overall reason for a training program. Level IV data is difficult to show as direct measures; therefore, indirect measures of patient outcomes and organizational impact are more frequently linked to Level IV programmatic results. Kirkpatrick's Levels of Evaluation targeted measures and outcomes with their corresponding methodological approach, as well as providing a guide to the relevance and practicability. Our measurement goal is to collect and analyze data at the Level II for this study. Therefore, observed practice improvements of skill and/or concept linked behaviors will greatly influence the scenario approaches and metrics collected. Few training programs in the military conduct rigorous analysis describing curriculum correlates and educational pedagogy along with performance, knowledge, demographic,

satisfaction data and geographic data location. This situation exists for various reasons, however the data is still needed and must be easier to collect.

## METHODOLOGY

### Methods

This **Institutional Review Board (IRB) approved** study uses a convenience sample of military medical center orientees who are assigned to Command orientation, in this quasi-experimental study. Subject learners were representing all personnel required to take the TeamSTEPPS® training as part of their Command orientation. Focused subjects were active duty clinical learners who comprised the majority of students in the classes. All survey instruments used in this study were validated.

### Process Summary

TeamSTEPPS® is a patient safety and communications program based on the principals of crew resource management (CRM) emanating from military aviation safety systems. Because of the documented problems faced in healthcare today with large numbers of preventable errors, the military Surgeons General have mandated all military treatment facilities are to provide TeamSTEPPS® training to their staff. When practiced, TeamSTEPPS® has improved team work functions and has helped to prevent or catch medical errors. This study was designed to provide an operationally relevant TeamSTEPPS® course to military healthcare and medical personnel using a non-traditional pedagogical approach. This study takes place at a large Level One trauma center military treatment facility in Texas. We are using a quasi-experimental approach with a convenience sample of new employees that must take TeamSTEPPS® as part of their normal check-in process. The employees self-selected the date of choice to attend. **Inclusion criteria:** De-identified scores will only be collected from those persons who do not object to our using their data to assess the efficacy of the CORPS program used on their own personal mobile devices. **Exclusion criteria:** Those learners who could not read and speak English.

**Our power calculation plus drop-out rate required we have at least 150 subjects.** We requested a large increase in the number of subjects because as military members, they frequently can and do get re-tasked, pulled, deployed or go on additional temporary duties. Hence, the need to draw from a larger pool of subjects. We started with an N=320, and we used 310, as 10 students were unable to finish the course for extraneous reasons. The population sample of hospital workers included clinical and non-clinical people, English-as-a-second-language learners, those who had previously taken TeamSTEPPS® and those persons new to a healthcare profession. We proposed that with the presentation of course content introduced by a virtual mentor or avatar guide on a game based computer or mobile personal device platform, we could equal or better the performance of the learners in the traditional lecture based control group. The computer program is self-paced, and it will require learners to raise the knowledge bar to achieve mastery level content knowledge through deliberate practice. In other words, the learner will be required to repeat the module until a specified level of success has been achieved. This process continues until all modules are completed. The pre and post-testing materials include initial demographics, knowledge test, self-efficacy and post testing also includes satisfaction and ease of use with the design and function of the software and platform. Overall, learners found this program simple and highly intuitive. Upon the completion of the training and post-test, the learner brings their personal mobile device up to one of the program instructors and shows them the “thank you for completing this program” message. At that point, the instructor can see the scores as posted in near real time to a secure cloud. Now the learner is handed their course completion certificate. All grades are posted and able to be “seen” locally or as part of an academic dashboard for training tracking, QA, oversight and readiness determinations. There were significant savings of time, money and supplies generated by this process. And it is recommended that the learners demonstrate comprehension and application of the new knowledge through a teach-back to mentors during the execution of a scenario based simulation. This will elevate the level of content reinforcement, critical thinking at a Level III Kirkpatrick rather than just a Level II quiz. The simulation can be conducted in a simulation center or in situ “acting out” of their roles on the team and techniques for promoting safety, efficiency and team support. Using the simulation process immediately after computer course completion will add to the reinforcement of important information. What do we expect from this approach? First, the computer experimental group will be able to free up valuable time of instructors and students for repurposing into coaching the learners’ attempts to translate new knowledge into practice via post-testing team simulation practice. The results of our testing ranged from the predictable to the surprising. **We demonstrated that the avatar guided computer program taught TeamSTEPPS® as well as the control group in lecture.** There was a weak trend toward

improved performance over control by experimental group but that will require future study. There were some scores that formed a pattern that this PI would characterize as ‘going through the motions’ without any intent on improving their score. Up until now, this course was given without any pre or post-testing. The non-clinical team members did better than most of the clinical or experienced clinicians. Follow on study and analysis of this process that remains in place will be needed.

## KEYWORDS

CORPS, TeamSTEPPS®, trauma, simulation, training, military nursing, computer game, team structure, leadership, communications, teams, CRM, avatar, and game based computer training

## PRELIMINARY WORK – EARLY HYPOTHESIS TESTING

The purpose of this research study is to test the effectiveness of military learner knowledge and concept acquisition using an avatar-guided interactive computer game based teaching approach when compared to the control group that uses traditional didactic lecture presentations supplemented with video clips and PowerPoint presentations. If the computer based system can demonstrate evidence that it saves instructor teaching and administrative time as well as enabling the learner to acquire more than or equal to the amount of content acquired by the traditional didactic control group, then this research team can claim the adequacy and appropriateness of the use of this type of computer instruction.

### Preliminary Research Question

Do the post TeamSTEPPS® training scores of the experimental group statistically reflect superior post training scores when compared to the control (didactic) group?

To answer this question, we conducted a preliminary study using 114 subjects split into two groups: (1) a control group who was taught via traditional methods and (2) an experimental group that was taught TeamSTEPPS® concepts via our avatar-guided simulation-based training. Within our study, 56 subjects were assigned to the control group; and 58 subjects were assigned to the experimental group. This preliminary study was followed by the full study using our full complement of students. Before undergoing training, the subjects in each group took a pre-training validated concept knowledge test containing 23 questions to assess their knowledge of the TeamSTEPPS® protocol. After completing the pre-test, traditional didactic TeamSTEPPS® training was administered to this control group; and avatar-guided simulation-based training was administered to the experimental group. Once training was completed, the knowledge of the subjects was assessed again via a post-training test containing the same 23 questions as the pre-test. The change in the subjects’ scores from the pre-training test to the post-training test is shown in Figure 3. Figure 3 reflects a histogram where the y-axis denotes the number of subjects within a given group that answered the number of additional questions correctly that is shown on the x-axis.



**Figure 3. Change in Subjects’ Scores from Pre to Post Test – Preliminary Study**

### Study Research Questions

**Research Question 1.** Does avatar guided, interactive TeamSTEPPS® training improve post training TeamSTEPPS® knowledge as represented by the learner’s scores? Or, stated another way: Are the avatar-guided computer simulation-based interactive training program learners’ (experimental group) post TeamSTEPPS® training scores statistically as good as the post training scores of the control/ didactic group? This statement implies that both pedagogical approaches statistically approximate similar results related to concept and content knowledge. To answer this question, we initially started with a total of 320 participants. However, we are using 310 fully

completed sets of test materials for our total number of participants used in our calculations. This group of participants was divided into two groups: (1) an experimental group consisting of 201 participants and a control group consisting of 109 participants. Each participant in each group completed a pre-test before s/he completed TeamSTEPPS® training and a post-test after they completed the actual training. On average, the participants in the control group answered 1.37 more questions correctly on the post-test than the pre-test. Similarly, the participants in the experimental group answered 1.08 more questions correctly on the post-test than the pre-test. According to Cohen, the difference in these improvements reflects a non-effect ( $d = 0.1562$ ) because the effect size is below 0.20. (Cohen, J. 1992) Along with this non-effect, our sample size ( $n_{\text{experimental}} = 201$ ,  $n_{\text{control}} = 109$ ) and a significance level of  $\alpha = 0.001$ , the test only results in a test power of only 0.0369. This means that there is only a 3.69% chance that this simulation-based training is less effective than traditional TeamSTEPPS® training. Another way of performing this analysis is to conduct a one-sided t-test on the mean improvement for the control group and experimental group to test if the difference in the means is statistically significant. The result of this test is a p-value of 0.11. This value (0.11) is significantly greater than the largest p-value considered statistically significant (0.05). Since the p-value is larger than the largest p-value considered statistically significant, the difference in the mean improvement between the groups is not statistically significant. Based on this result and that the test power is only 0.0369, we conclude that this simulation-based training is not less effective (i.e. is as effective) as the traditional based training. This allows for **acceptance of the null hypothesis** (There is no statistically significant difference between the results of using interventions A or B). Frequently, researchers strive to reject the null hypothesis; however, in this case, the null still represents a successful intervention due to secondary program benefits.

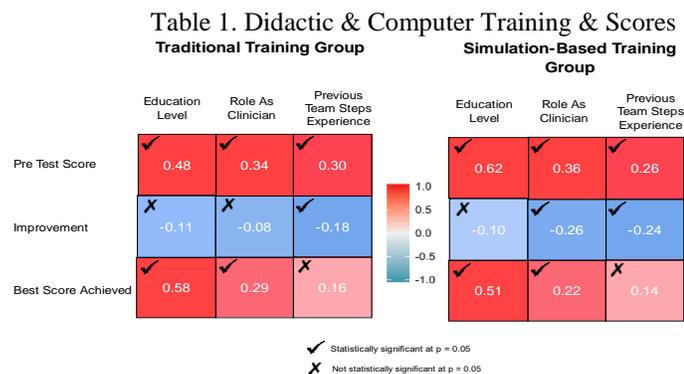
**Research Question #2 asks:** Is simulation based TeamSTEPPS® training at least as effective as conventional TeamSTEPPS® training (Cohen, S. R., Cragin, Wong, & Walker, 2012)? To answer this question, our 310 participants were divided into two groups: (1) a treatment group consisting of 201 participants and a control group consisting of 109 participants. Each participant in each group completed a pre-test before s/he completed training and a post-test after they completed TeamSTEPPS® training. The participants in the treatment group completed the pre-test, training and post-test on a simulation-based platform while the control group did through traditional means (PowerPoint and video clips). On average, the participants in the control group answered 1.37 more questions correctly on the post-test than the pre-test. Similarly, the participants in the experimental group answered 1.08 more questions correctly on the post-test than the pre-test. According to Cohen, the difference in these improvements reflects a non-effect ( $d = 0.1562$ ) because the effect size is below 0.20. (J. Cohen, 1992) Along with this non-effect, our sample size ( $n_{\text{experimental}} = 201$ ,  $n_{\text{control}} = 109$ ) and a significance level of  $\alpha = 0.001$ , the test only results in a test power of only 0.0369. This means that there is only a 3.69% chance that simulation-based training is less effective than traditional TeamSTEPPS® training.

### Self-Efficacy Scale Results

Subjects were asked to complete the general Self-Efficacy Scale. (Bandura, 1995) An option to create a specific scale was rejected as this research team wished to identify the subjects overall self-perspective. It has been postulated that those persons with a high degree of self-efficacy tend to be more successful at new activities and encounters. Our subjects completed this instrument. Some subjects were able to score a perfect 40 on the scale. They clearly and confidently consider themselves to be highly capable individuals. Subjects with a total self-efficacy rating of less than 40 across the 10 questions are considered to assess themselves as not so effective. For clinicians, we didn't see a difference in self-efficacy. In other words, their improvement on the pre-test and the post-test was not related to differences in perceived self-efficacy. Here are the comparisons: Effective vs. non-effective (pretest 19.37 - effective vs. 19.37 - non-effective). On the post-test (20.92 - effective vs. 21.31 - non effective). However, for non-clinicians, there was a difference. Those who assessed themselves as being effective had higher pre-test and post scores than those who did not. On the pretest (17.63 - effective vs. 16.72 - non effective). On the post-test (20.32 - effective vs. 18.18). From purely an educated observation, this researcher noted the clinicians were somewhat disengaged and disinterested. However, many of the non-clinicians asked questions about how to navigate the system and once we took them through the process, most were interested and some noted they thought the content useful and the process fun. A couple of people were excited to tell their grown children that they learned how to take a test on their phone. However, two people were found to be watching soap operas on their phones. When it was reinforced that they could not proceed or leave until they completed the program that required higher scores, their motivation and attitude improved significantly. It turns out the computer created a requirement for them to learn or fail their first encounter in their new hospital position, and this would be obvious to the training officer that they failed to meet a Command requirement. This was an unexpected administrative benefit.

## Factors of Demographic Groups

Within our study, we measured several factors of demographic groups of our participants in both the control and experimental group. For each factor and each group (control/experimental), we explored the extent to which it is correlated with: (1) performance on the pre-test, (2) improvement from the pre-test to the post-test and (3) best performance observed for a participant on either the pre-test or the post-test. The correlation among these demographics within the control and experimental group for the three observed outcomes is shown in Table 1. The control group is shown on the left hand side of the page, and the experimental group is shown on the right hand side. Within the figure, a red square indicates a positive correlation (0.0 to 1.0); and a blue square indicates a negative correlation (0.0 to -1.0). In addition, the shading of each square reflects the strength of the correlation. The darker the red block, the stronger the positive correlation. The darker the blue block, the stronger the negative correlation. Finally, a check denotes that the correlation is statistically significant at  $p = 0.05$  while an x denotes that the correlation is not statistically significant.



## SUMMARY OF FINDINGS AND SURPRISING RESULTS

Table 1 shows that both the pre-test and achieving the best score among the pre-test and the post-test participants in both groups performed better with more education, those who had a role as a clinician and those who had previous TeamSTEPPS® experience, were more likely to score higher when compared to those who did not fall into these demographic groups. Furthermore, the relative strength of these correlations in both groups is the about the same. For both the pre-test performance and best score among the pre-test and the post-test groups, in both the experimental and control group, the factor that was the most influential was the education level of the participant. Conversely, the factor that was the least influential was previous TeamSTEPPS® experience (this factor refers to exposure to a prior TS course. Without clinical performance there is no reinforcement of previously learned TS content). Finally, the correlation among participants with previous TeamSTEPPS® experience and the best score achieved on the pre-test or post-test is not statistically significant in either the control or the experimental group. This is an encouraging result for us and provides further evidence that in terms of: (1) assessing initial knowledge (pre-test score as compared to post test score) and (2) evaluating if mastery knowledge is achieved (best score achieved), the simulation-based training group was as effective as the traditional training group. However, what does it say about existing TeamSTEPPS® programs and the retention or likely use of the tenants of the program in the clinical environment? There is no evidence of TS concepts being reinforced by practice. It's not the course content, but maybe how we educate today's learner or the learner's attention to the course followed by lack of practice. Since no testing had been done previously, as it was labor intensive and costly, no feedback to Command or consequences to lack of success existed.

The factors driving improvement from the pre-test to the post-test in both groups demonstrates a difference between the two groups. In the simulation-based training group, there are stronger correlations among those without experience (were not clinicians, did not have previous TeamSTEPPS® experience) and improvement in the post-test score after training. This means that for those individuals who are less familiar with TeamSTEPPS® and the medical/healthcare domain the simulation-based training could offer an opportunity for more improvement. In contrast, for those who already possess background knowledge. A number of causal factors must be looked at and impartially analyzed. Knowledge must be aligned properly through active guided practice or simulation to support memory recall to long-term memory with rapid reach back capability. As deliberate practice is vital to long lasting content memory, so too is the importance of honing the content mastery via deliberative practice (real or virtual)

through concentrated effort and repetition. The best will include actual use of TS principals in the clinical environment. The patient's safety is dependent on this, and the integrity of the team's ability to be highly effective and in concert with one another is worth the investment in teaching with follow-on simulation and mandatory testing. This will help identify academic vs. behavioral deficiencies. This new training is now made more "affordable" due to an additional observation that the avatar-guided interactive computer game based process was estimated to have saved about 360 man hours of work time for the patient safety department at the military medical center. During the experimental testing cycle, this process significantly reduced the cost of testing, reporting and provided documentation and certificates of each person's course completion or lack there of. This would help free up SME mentors and coaches to support Simulation Center or in situ training for the course and reinforcement. It is also planned to have a sustainment set of games, much like the format we developed for trauma training games for mobile personal devices and the personal use of the learners. Recommendations include spot-checking TeamSTEPPS® performance behaviors. They should be observable during actual clinical practice and should be part of an area's unit Quality Assurance measures. Another mandate should be to incorporate TeamSTEPPS® behaviors into all mandated Continuing Education Unit and Graduate Medical Education clinical courses like Basic Cardiac Life Support or Advanced Cardiac Life Support as examples. After all, the use of TeamSTEPPS® can and will help prevent the preventable error; and our patients are worth the effort. It is how DOD will do business and acknowledged from the top of the Command to each employee that safety and team-work are mandated.

### Linear Regression Analysis

To elucidate the influence of the demographic group factors on the three outcomes (pre-test score, improvement and best score on either pre-test or post-test) in both the control and experimental group, we created linear regression models. For each regression model, we included the factors with statistically significant correlations shown in Table 1. Furthermore, for the improvement outcome, we included the pre-test score as a predictive factor. The pre-test score is included for improvement because it reflects the opportunity for enhancement that the individual participant has. For each model, we show the percentage of the variance in the outcome that the model accounts for. These results are shown in Table 2 and 3. Each model shown in Table 2 and 3 accounts for ~20-30% of the total variance in the outcome. Given that many other variables exist that are not and/or cannot be accounted for (i.e. participant focus) we are encouraged by these results.

Table 2: Linear Regression Models to Predict Each Study Outcome for Traditional Training (Control Group)

Outcome	# of Factors in Model	Names of Factors	% of Outcome Variance Model Accounts For
Pre Test Score	3	Education Level, Role as Clinician, Previous Experience	26.47 %
Improvement	2	Previous TeamSTEPPS Experience, Pre Test Score	27.86 %
Best Score Achieved	2 (3)	Education Level, Role as Clinician, (Pre Test Score)	29.63 % (69.53 % if Pre Test Score included)

Table 3. Linear Regression Models to Predict Each Study Outcome for Simulation-Based Training (ExperimentalGroup)

Outcome	# of Factors in Model	Names of Factors	% of Outcome Variance Model Accounts For
Pre Test Score	3	Education Level, Role as Clinician, Previous TeamSTEPPS Experience	15.72 %
Improvement	3	Role as Clinician, Previous TeamSTEPPS Experience, Pre /Test Score	31.86 %
Best Score Achieved	2 (3)	Education Level, Role as Clinician, (Pre Test Score)	19.30 % (78.98 % if Pre Test Score included)

This research demonstrates that the two training methodologies are equally effective and highlight situations where we are provided the opportunity to separate potential content or program presentation inadequacies from attitude and managerial issues or the loss of sustainable knowledge over time often from lack of use and or reinforcement.

## RESULTS SUMMARY

On average, the participants in the control group answered 1.37 more questions correctly on the post-test than the pre-test. Similarly, the participants in the experimental group answered 1.08 more questions correctly on the post-test than the pre-test. According to J. Cohen A power primer, Psychological bulletin, 112(1), 155, the difference in these improvements reflects a non-effect or no difference ( $d=0.1562$ ) because the effect size is below 0.20. (Cohen, J. 1992) Along with this non-effect, our sample size ( $n_{\text{experimental}} = 201$ ,  $n_{\text{control}} = 109$ ) and a significance level of  $\alpha=0.001$ , the test only results in a test power of only 0.0369. This means that there is only a 3.69% chance that this simulation-based training is less effective than traditional TeamSTEPPS® training. Another way of performing this analysis is to conduct a one-tailed t-test on the mean improvements for the control and experimental groups to test if the difference in the means is statistically significant. The result of this test is a p-value of 0.11. This value (0.11) is significantly greater than the largest p-value considered statistically significant (0.05). Since the p-value is larger than the largest p-value considered statistically significant, the difference in the mean improvement between the groups is not statistically significant. Based on this result and that the test power is only 0.0369, we concluded that this simulation-based training is not less effective (i.e. is as effective) as the traditional based training. Within our study, we measured several factors of demographic groups of our participants in both the control and experimental groups. For each factor and each group, we explored the extent to which it correlated with: (1) performance on the pre-test, (2) improvement from the pre-test to the post-test and (3) best performance observed for a participant on either the pre-test or post-test. On both the pre-test and in terms of achieving the best score among the pre and post-test participants in both groups with more education, who had a role as a clinician and had previous TeamSTEPPS® experience, were more likely to score higher than those who did not fall into these demographic groups. For both pre-test performance and best score among the pre and post-test in both the control and experimental groups, the factor that was the most influential was the education level of the participant. Conversely, the factor that was the least influential was previous TeamSTEPPS® experience. Finally, the correlation among participants with previous TeamSTEPPS® experience and the best score achieved on the pre or post-test is not statistically significant in either the control or experimental groups. This information provides further evidence that in terms of: (1) assessing initial knowledge (pre-test score) and (2) evaluating if mastery knowledge is achieved (best score achieved), the simulation-based training is as effective as the traditional training. We assert that there is sufficient evidence that teaching TeamSTEPPS® to military personnel via avatar-guided simulation-based training is at least as effective as teaching TeamSTEPPS® via traditional methods. Follow on study and analysis of this process that remains in place will be needed.

## ACKNOWLEDGEMENTS

Thomas F. Frost, MS holds a BS from the Georgia Institute of Technology and an MS from the Air Force Institute of Technology and is a retired USAF Colonel with 26 years' experience in C4I. After retirement, he supported the use of modeling and simulation in joint training at the Joint Forces Command (JFCOM). When JFCOM was disestablished, his expertise expanded by supporting programs in Healthcare and Medicine at the Virginia Modeling, Analysis and Simulation Center (VMASC). His responsibilities include Research Associate and Role 2 VMASC Project Operations Coordinator for projects, such as the avatar guided computer-based training Course for Operationally Relevant Patient Safety and Team Strategies & Tools to Enhance Performance and Patient Safety (TeamSTEPPS®), detailed Trauma Training Guidebook including TeamSTEPPS® & the ABCs of trauma, detailed state diagrams for patient scenarios on the ABCs of trauma and ten computer-based Trauma training games.

Hector M. Garcia M. Arch. Ph.D. (candidate) is a Senior Project Scientist at Old Dominion University's Virginia Modeling Analysis and Simulation Center, in the area of Visualization, Virtual Environments and Virtual Reality, integrating state of the art visualization systems with modeling and simulation applications, and the Scientist most closely involved with the CAVE (Cave Automatic Virtual Environment) on ODU's Norfolk Campus. Mr. Garcia received his Master's in Architecture from University of Houston in 1997. Mr. Garcia's expertise includes the use of large scale visual simulation display systems, the use of tracking devices, robotics, and haptic devices used in training and education. Mr. Garcia has more than 20 years experience developing highly interactive Virtual Environments for training. He has been involved in a variety of research projects funded by NASA, NSF, ONR,

AHRQ and private industry. Before joining Old Dominion University, Mr. Garcia spent 5 years as a Researcher at the University of Houston affiliated with the Virtual Environments Technology Laboratory working on several NASA projects for Astronaut training as well as NSF funded research for using Virtual Reality as a teaching tool.

## REFERENCES

- Agarwal, R., & Karahanna, E. (2000). Time Flies When You're Having Fun: Cognitive Absorption and Beliefs About Information Technology Usage. *MIS Quarterly*, 24(4), 665–694. Retrieved from <http://proxy.lib.odu.edu/login?url=http://search.ebscohost.com/login.aspx?direct=true&db=iih&AN=4386307&site=ehost-live&scope=site>
- America, I. of M. (US) C. on Q. of H. C. in. (2001). *Crossing the quality chasm: A new health system for the 21st century*. National Academies Press.
- Bandura, A. (Ed.). (1995). *Self-efficacy in changing societies*. Cambridge ; New York: Cambridge University Press.
- Benner, P. (2001). *From Novice to Expert: Excellence and Power in Clinical Practice*. Engelwood Cliffs, N.J.: Prentice Hall.
- Benner, P. E. (Ed.). (2010). *Educating nurses: a call for radical transformation* (1st Ed). San Francisco: Jossey-Bass.
- Cohen, J. (1992). A power primer. *Psychological Bulletin*, 112(1), 155–159. <https://doi.org/10.1037/0033-2909.112.1.155>
- Cohen, S. R., Cragin, L., Wong, B., & Walker, D. M. (2012). Self-Efficacy Change With Low-Tech, High-Fidelity Obstetric Simulation Training for Midwives and Nurses in Mexico. *Clinical Simulation in Nursing*, 8(1), e15–e24. <https://doi.org/10.1016/j.ecns.2010.05.004>
- Cooke, Molly, & O'Brien, Bridget C. (2010). *Educating Physicians: A Call for Reform of Medical School and Residency*. San Francisco, CA: Jossey Bass.
- Joint Commission Center for Transforming Healthcare. (2014). *Health Centers' Most Challenging Standards 2014* (No. 3/15) (p. 1). Retrieved from [https://www.jointcommission.org/assets/1/18/Challenging\\_NPSG\\_Standards-bphc\\_2014.pdf](https://www.jointcommission.org/assets/1/18/Challenging_NPSG_Standards-bphc_2014.pdf)
- Kalisch, P. A., & Kalisch, B. J. (1995). *The Advance of American Nursing*. Philadelphia PA: J.B. Lippincott Company.
- Kirkpatrick, D. L. (1998). *Evaluating training programs: the four levels* (2nd ed). San Francisco, Calif: Berrett-Koehler Publishers.
- Kirkpatrick, D. L., & Kirkpatrick, J. D. (2006). *Evaluating training programs: the four levels* (3rd ed). San Francisco, CA: Berrett-Koehler.
- Kirkpatrick14 Feb 2017. (2017, February 14). Kirkpatrick Model. Retrieved from <http://www.kirkpatrickpartners.com/OurPhilosophy/TheKirkpatrickModel/tabid/302/Default.aspx>
- Markel, H. (2010). Abraham Flexner and His Remarkable Report on Medical Education: A Century Later. *JAMA*, 303(9), 888. <https://doi.org/10.1001/jama.2010.225>
- National Healthcare Quality & Disparities Report. (2014). Retrieved from <https://www.ahrq.gov/sites/default/files/wysiwyg/research/findings/nhqdr/nhqdr14/2014nhqdr.pdf>
- Norwood, S. L. (2000). *Research Strategies for Advanced Practice Nurses*. Engelwood Cliffs, N.J.
- Sanford, Pamela. (2010). *Simulation in Nursing Education: A Review of the Research* (No. 4/2010). Retrieved from <http://nsuworks.nova.edu/tqr/vol15/iss4/17/>
- Turner, T. R. & Parodi, A. (2012). Theoretically-driven infrastructure for supporting health care teams training at a military treatment facility. *Military Medicine*, 177(2), 139–144.