Published in final edited form as:

J Surg Res. 2019 January ; 233: 351–359. doi:10.1016/j.jss.2018.07.040.

Evaluating Surgical Resident Needle Insertion Skill Gains in Central Venous Catheterization Training

Hong-En Chen, MS^a, Mary A. Yovanoff, MS^a, David F. Pepley, MS^b, Rohan S. Prabhu, BTech^b, Cheyenne C. Sonntag, MD, MS^c, David C. Han, MD, MS^{c,d}, Jason Z. Moore, PhD^b, and Scarlett R. Miller, PhD^{a,e,*}

^a Department of Industrial and Manufacturing Engineering, Pennsylvania State University, University Park, Pennsylvania

^b Department of Mechanical and Nuclear Engineering, Pennsylvania State University, University Park, Pennsylvania

^c Department of Surgery, Penn State Health Milton S. Hershey Medical Center, Hershey, Pennsylvania

^d Penn State Heart and Vascular Institute, Penn State Health Milton S. Hershey Medical Center, Hershey, Pennsylvania

^e School of Engineering Design, Technology and Professional Programs, Pennsylvania State University, University Park, Pennsylvania

Abstract

Background: Training for ultrasound-guided Central Venous Catheterization (CVC) is typically conducted on static manikin simulators with real-time feedback from a skilled observer. Dynamic haptic robotic trainers (DHRT) are an alternative method that simulates various patient anatomies and provides consistent feedback for each insertion. This study evaluates CVC needle insertion efficiency and skill gains of both methods.

Materials and Methods: Fifty-two first-year surgical residents were trained by placing Internal Jugular (IJ) CVC needles in manikins (N=26) or robots (N=26). Manikin-trained participants received verbal feedback from an experienced observer, while robotically-trained participants received quantitative feedback from the personalized DHRT learning interface. All participants were pre-tested on a Blue Phantom manikin; participants completed post-testing on a Blue-Phantom manikin (N=26) or a novel manikin (N=26) with different vessel depth and position. During pre- and post-tests residents were timed, motion-tracked, and scored on an IJ CVC checklist.

*Corresponding Author: Scarlett R. Miller. 213-P Hammond Building, University Park, PA 16802. Tel: 814-863-4143; fax: 814-863-7229. scarlettmiller@psu.edu.

Authors' Contributions: H.C. contributed with data analysis, statistical analysis, and manuscript preparation. M.A.Y. contributed with concept design, acquisition of data, and manuscript preparation. D.F.P. contributed with simulator development, acquisition of data, data trimming, and manuscript preparation. R.S.P. contributed with acquisition of data, statistical analysis, and manuscript editing. C.C.S. was responsible for participant recruitment, acquisition of data, and manuscript editing. D.C.H., J.Z.M. and S.R.M. contributed to the study design, data collection, robotic simulator design, and manuscript preparation.

Author Disclosure: The authors have no financial conflicts of interest in the research.

Results: (1) All skills on the IJ CVC checklist showed significant ($p < 0.014$) improvements from pre- to post-test; (2) Average angle of insertion, path length, and jerk improved significantly ($p < 0.005$); (3) Average procedural completion time for pre- and post-tests significantly decreased from pre-test ($M = 3.516$ min, $SE = 0.277$) to post-test ($M = 1.997$, $SE = 0.409$).

Conclusion: No significant group differences were observed in overall skill gains, but residents' average procedural completion time decreased significantly from pre- to post-test. Overall results support DHRT as an effective method for training IJ CVC skills.

Keywords

Needle insertion; surgical simulation; training and assessment; ultrasound manikins; robotic simulation; central venous catheterization

Introduction

Ultrasound guided Central Venous Catheterization (CVC) is a common procedure¹ with a high rate of complications². The likelihood of these complications, however, is significantly lower for surgeons who've inserted a CVC more than fifty times^{3,4}. There is no common standard for evaluating CVC placement competency outside of clinical applications. To determine the effectiveness of particular training methods, many studies use checklists filled out by supervisor⁵, such as the Global Rating Scale (GRS) Assessment tool (a 9-item Likert scale assessment intended for use generally across multiple skills)⁶, and a 24-item binary checklist specific to subclavian CVC skills⁷. The majority of these evaluation methods compare trainee skills to some baseline expectation and evaluate whether they 'pass' to be allowed to perform CVC in clinical situations^{8,9}. This approach, however, may not rule out incompetency¹⁰. While these tools may be effective at evaluating performance on a manikin, it is not clear that these evaluation methods can predict skill transfer from a static simulator to the variation seen in clinical practice¹¹. The number of *unsuccessful needle insertion attempts* significantly predicts cannulation failure¹², indicating the importance of training and properly evaluating needle insertion skill gains.

While checklists are useful for standardizing practices, they are imperfect and can encourage "teaching to the test"¹³. They are also prone to misuse where observers simply check off boxes rather than providing helpful and targeted feedback to the learner,¹⁴ undercutting a critical part of the learning process¹⁵. Continuous and qualitative feedback provided during training may help improve skill gains, supporting the need to distinguish between tools for evaluation and those for training¹⁶. This has led to efforts to develop tools specific for training¹⁷ that rely on motion analysis^{18,19}, such as the Imperial College Surgical Assessment device (ICSAD)¹⁸. These assessments use objective measures to evaluate skill performance through metrics such as path length, time to complete, steadiness of motion, velocity, total amount of movement, and overall efficiency of motion^{20–23}. The use of electromagnetic sensors to identify motion paths during CVC training show promise as a valid and objective assessment of skills²⁴ and these metrics can readily be captured using VR simulators²⁵.

The development of objective measures of performance can also enhance the quality of feedback that can be provided to novices during training. Research has shown that providing novices with detailed and appropriate feedback is a critical part of the learning process and that the opportunity to incorporate that feedback is crucial to the development of expertise in surgery¹⁵. A Dynamic Haptic Robotic Trainer (DHRT) system was designed to simulate various patient profiles and provide feedback on these metrics^{26,27}. The DHRT system includes a personalized learning interface that provides quantitative feedback after each needle insertion on metrics including the angle of insertion, number of attempts, and distance to the center of the vein²⁷. It has been suggested that the evaluation of competency of a trainee at a surgical procedure should include a combination of the trainee's self-perception of ability, an evaluator's perception of their ability, and an objective skills based evaluation of their ability²⁸. The purpose of this study was to compare needle insertion skill gains and procedural efficiency of surgical residents before and after training on a manikin with feedback provided by an experienced observer, or robot-based training system with personalized learning interface. Results from this study can inform future development of surgical training systems.

Materials and Methods

A study was designed to explore the effectiveness of the DHRT system and learning interface as compared to the standard manikin training for increasing CVC needle insertion skills. It first validated the use of a 10-item CVC checklist and qualitative motion tracking metrics to evaluate needle insertion skill gains, and then identified whether the different training modalities impacted resident procedural efficiency when encountering a novel patient scenario. Specifically, the study was developed to answer the following research questions:

RQ1. How do subjective observations of performance change after training and is this dependent on the training method? This question sought to address how a skilled observer's evaluation of first-year surgical residents' performance using a 10-item IJ CVC skills checklist changed from pre- to post- training and if these changes were different between training groups. It was hypothesized that there would be no differences in observed ratings on the IJ CVC skills checklist because prior research has indicated that checklists may not be granular enough to detect changes in performance²⁹.

RQ2. How do resident CVC skills quantified through motion analysis change after CVC training and is this dependent on the training method employed? This research question sought to address how the total path length of the needle tip, time to complete the CVC insertion, average angle of insertion of the needle, needle tip velocity, and jerk of the needle tip changed before and after CVC training and if there was a difference between training groups. It was hypothesized that these objective measures of skill gains would increase with training and the skill gains for the robotic training group would be greater than or equal to the manikin training group. This was hypothesized because our prior research found that the DHRT was able to increase performance on CVC surgical skills²⁹ as have other studies on the use of virtual reality (VR) simulators in other areas of surgical skills training³⁰.

RQ3. Does the training modality/environment impact the efficiency of CVC insertion? This research question sought to identify if there was any difference between robot-trained and manikin-trained participants in their time to complete the entire post-test procedure on a novel manikin. It was hypothesized that participants in the robotic training environment would complete the procedure more efficiently, because the variation in patient anatomies during training would better prepare them for novel situations.

Participants

First-year surgical residents at Hershey Medical Center (HMC) were invited to participate in the study during their “residency boot camp” in 2016 (26 participants) and 2017 (26 participants) which occurs during the first week of their surgical residency. Of the 52 participants, 12 were female and 40 were male. The specialties of the residents were as follows: General Surgery (18), Preliminary residents/Unspecified (12), Orthopedics (10), Urology (4), Plastic Surgery (4), Otolaryngology (3), and Vascular Surgery (1). For the purposes of this paper, we will highlight portions of the study that are relevant to this study below.

Research Design

The Institutional Review Board approved study consisted of a pre-test, CVC training, two full line practices, and a post-test, as shown in figure 1. The first session included an introduction to the study, obtaining consent, and a pre-test on a Blue-Phantom manikin. Participants from each year were randomly assigned to a training environment, manikin (N=26) or robotic (N=26), and completed their training on three separate days. Training included a demonstration of a CVC procedure and 22 needle insertions (2, 10, and 10 each day) with feedback given. After training was complete, residents completed two full central line practice insertions (starting with sterilizing the area and proceeding through suturing the line in place) on a manikin trainer. The final session included a Verification of Proficiency (VOP) post-test on a manikin for the full procedure, which also counted as a qualifying skills test for the participants’ residency requirements.

The pre-test and introduction to CVC placement was conducted during the first day of boot camp. A brief oral summary of the proposed research was provided, and participants viewed an 18 minute video³⁶ on central line placement without further instructions on how to perform the procedure. After viewing the video, each participant *individually* completed a pre-test where they inserted a needle for central line placement into a Blue-Phantom Gen II Ultrasound Central Line Training Model #BPH660 (Blue-Phantom manikin) while verbalizing what they were doing according to the think-aloud procedure³¹. During the pre-test, a trained senior surgical resident observed each participant and evaluated them using a modified Internal Jugular Catheterization (IJ CVC) evaluation form. The IJ CVC evaluation form used during the pre-test is a modified checklist (http://www.engr.psu.edu/britelab/resources/IJCVC_Checklist.pdf) focusing exclusively on the needle insertion portion of the procedure such as “selecting the appropriate site for venipuncture” and “locating the needle position on the ultrasound image.”³² Total time to complete the 10-items during the pre-test was timed by the observer. There was a 5-minute time limit for the pre-test and if the participant did not successfully stick a vessel during the 5-minute time period, the pre-test

was terminated. The needle pathway during the pre-test was tracked using an Ascension 3D Guidance trakSTAR (Shelburne, VT) electromagnetic position tracking system.

The Ascension 3D Guidance trakSTAR (Shelburne, VT) electromagnetic position tracking system works by using a transmitter to generate a magnetic field. This field is then measured by an electromagnetic sensor. The orientation of this sensor effects the field measurements and an output of the x-position, y-position, z-position, and probe angle can be calculated. In order to track the position of the syringe, the electromagnetic sensor was attached to the syringe at the 0mL mark and oriented with the tip pointing towards the needle. The needle tip was located 9cm in the axial direction from the end of the sensor, allowing its location to be calculated using the location and orientation of the sensor. A sample rate of 80 Hz was used for these tests. The set-up is shown in Figure 2.

After the pre-test, each participant was randomly assigned to one of two training conditions, manikin or robotic, for the remainder of the training sessions.

Manikin training (N=26). Participants completed training on a Blue-Phantom Gen II Ultrasound Central Line Training Model #BPH660 or a similar training manikin.

Robotic training (N=26). Participants completed training on the DHRT system.

Training was completed during three separate sessions. At the beginning of training, a senior surgical resident who is highly experienced in conducting CVC training sessions at HMC gave a demonstration of central-line placement (using a Blue-Phantom manikin) including how to use an ultrasound, how to identify and distinguish between the artery and vein, how to use anatomical landmarks to inform line placement, how to insert a needle, how to identify needle location based on ultrasound feedback and how to confirm needle placement using flash. After viewing the demonstration, participants individually completed 2 practice insertions in their assigned environment (manikin or robotic). Participants in the manikin training condition were provided with verbal feedback from the trained resident that corresponded to the full VOP CVC checklist, such as needle angle, ultrasound probe orientation, and identification of the artery and vein, see Appendix. Given the physical design of the manikin, the manikin training group received no variations in patient profile. Participants in the robotic training condition watched an instructional video on how to use the DHRT system, completed their 2 practice insertions, and had any questions about the robotic system answered. The robotic ultrasound image and needle forces for the following insertion were representative of the given profile. Performance feedback for the robotic environment was provided to participants through a personalized learning interface as described in Yovanoff, et al.²⁷. While a researcher was present in the DHRT rooms in case of glitches or problems operating the system, they provided no feedback to the participants on CVC performance. During each of the remaining two training sessions, participants completed 10 needle insertions and received feedback either orally from a resident in the manikin group or through feedback by the robotic learning interface. In total, participants completed 22 needle insertions in their respective training environments.

After the 22 training needle insertions were complete, participants completing two full practice sessions on the entire CVC insertion procedure (which extends beyond the scope of this study but includes the needle insertion procedure) on the Blue-Phantom Manikin and

their performance was evaluated, by a senior surgical resident, using the full 23-item IJ CVC checklist. Finally, participants individually performed a post-test needle insertion a manikin using the same procedures as the pre-test. However, the post-test manikin was different between the two cohorts (2016 and 2017). Specifically, the 2016 cohort completed VOP testing on the same Blue-Phantom manikin model used during training, while the 2017 cohort completed the testing on a newly manikin model that was created to mimic an obese patient³³. The novel manikin was designed to evaluate how trainees would perform when introduced to a new and unfamiliar patient profile. Regardless of the manikin used, residents were motion tracked and evaluated on the full procedural IJ CVC skills checklist by a senior surgical resident, see Appendix A. However, the 2017 VOP checklist included time to complete the entire procedure, as well as “Advancing the introducer needle slowly and steadily,” which was modified from the 2016 checklist.

For this study, we focus on participants’ subjective and objective performance during the pre-test (Blue-Phantom) and post-test (Blue-Phantom (2016) or custom obese manikin (2017)), including needle insertion performance and the IJ CVC skills checklist.

Metrics

For each participant trial, each data set was trimmed to include only the motion path while the needle was being inserted into the ‘skin’ of the manikin. To do this, two plots were created from the raw data including the needle depth versus time (s) and the needle angle versus time (s), see Figure 3. The combination of these two plots showed a distinct point in time when the needle slowed down and rested at a precise angle before entering the manikin. One researcher cut the data at the overlapping point in time when the data leveled out. Once the data were trimmed, five objective motion tracking metrics were calculated: time to complete, average angle, path length, average velocity, and average jerk. The following paragraphs outline how and why each of these metrics were calculated.

Time to complete the needle insertion was calculated by subtracting the start time from the end time. This was calculated as a metric because research has found that experts are faster than novices in completing basic surgical skills^{20–23}.

Average angle was calculated by taking the average of the angles between each position point. This was used as a metric because of how frequently verbal feedback on angle of insertion was shown to be provided on the manikin trainer²⁷, because of the specific feedback provided in the DHRT personalized learning system, and because of its presence on the IJ CVC checklist.

Needle tip *path length* was chosen as a metric because prior research has indicated that overall path length of hand motions can distinguish between expert and novice performance^{18,20}. However, path length of the needle tip in CVC insertion procedures has not been studied to date. Thus, path length was calculated as a summation of movement in the X, Y and Z coordinate system based on the points recorded by the DHRT system using the Euclidian distance formula.

Average velocity of the needle tip was chosen as a metric because prior research has indicated that the velocity of hand movements can distinguish between expert and novice performance^{23,34}. However, velocity of the needle tip has not been studied to date in CVC procedures. Average velocity was defined as the average velocity of the needle tip when it was below the skin surface. Velocity was calculated using the Savitzky-Golay method³⁵ which resulted in an x-, y-, and z-velocity. These were combined using the Pythagorean distance theorem to obtain a single measure of jerk with respect to time. The average velocity is the average value of the points contained in this vector

Average Jerk, the derivative of acceleration with respect to time, was chosen because it has been utilized as a method for calculating smoothness of motion in surgical procedures³⁶. Jerk was calculated for each vector using the Savitzky-Golay method³⁵ which resulted in an x-, y-, and z-jerk. These were combined using the Pythagorean distance theorem to obtain a single measure of jerk with respect to time. The average jerk is the average of the values contained in this vector.

Data Analysis

In order to answer our research questions, statistical analyses were performed on the IJ CVC checklist pass rates and each of the five needle motion path metrics for pre-test and post-test performance. The time to complete the procedure (pre-test versus post-test 2017), as well as the economy of motion and time (post-test 2016 and 2017) were also examined. Descriptive statistics are reported as mean (SE) unless otherwise reported. All analyses were conducted using SPSS (v. 25.0) at a significance level of 0.05.

To determine between-subject and within-subject effects on CVC checklist skills pass rates, a Generalized Estimating Equation was first run to determine if there were any differences between the two training groups for the pre- and post- IJ CVC skills checklist while controlling for year. Year was controlled for to account for different post-test manikins and differences in resident skills between the two cohorts. The between-subject variable was participant ID, the within-subject variable was the test order (pre-test or 1st attempt of post-test), the independent variable was type of training (manikin or robotic) with the year of training as a covariate, and the dependent variable was each question on the checklist. Each question was run as a separate binary logistic regression. Follow-up Wilcoxon signed rank tests were used in the post-hoc analyses.

To examine differences in motion tracked metrics, a repeated measures ANCOVA was run to determine the between-subject and within-subject training effects on each of the five motion tracked metrics. Specifically, the independent variables were the training group (manikin or robotic) and the skills test (pre- and post-); the dependent variables were the five motion tracked metrics; the year was conducted as a covariate.

Finally, a repeated measures ANOVA was conducted to examine within-subjects (pre- or post-test) and between-subjects (manikin or robotic) effects on time to complete the full procedure in 2017. Additionally, a Mann-Whitney U test was conducted to examine differences between training groups (manikin or robotic) for economy of time and motion evaluated on the post-test in both 2016 and 2017.

Results

A Generalized Estimating Equation was first conducted on each question as a binary logistic regression to determine if there was a significant impact of training group on pre- to post-test checklist pass rates when controlling for year. The analysis failed to reveal a statistically significant difference between training groups from pre- to post-test for all questions ($p > 0.137$). These results indicate that there were no differences between training groups for pre- to post-test CVC skills pass rates. Therefore, main effects of within-subject differences were examined. A Wilcoxon signed rank test was conducted for each question to understand whether there were median differences in pass rates from pre- to post-test. Results indicate that all checklist questions showed a significant improvement from pre- to post-test ($p < 0.014$), see Table 1. These results show that training in either environment significantly improved resident performance.

A repeated measures ANCOVA was conducted on all participants to determine the impact of manikin or robotic training on pre- and post-test motion tracked metrics while controlling for year of data collection. The full results of the repeated measures ANCOVA and the mean values for each metric are shown in Table 2. Assumptions were checked prior to conducting the test. During this assumption check, three outliers were identified due to having standardized values (Z-scores) greater than 4 standard deviations higher than the mean for any metric. Specifically, one of each of the following were removed: post-test path length, velocity, and jerk. Additionally, each pre-test and post-test motion tracked metric was checked for normality. The results showed that there was not a significant interaction effect for training condition on pre- and post-test motion tracked performance metrics ($p > 0.609$). Therefore, simple main effects were run. These results showed a statistically significant main effect of pre- and post-test on the average angle of insertion, $F(1, 42) = 16.874$, $p < 0.0005$, Wilks' $\Lambda = 0.713$, $\eta^2 = 0.287$, observed power of 0.980. Specifically, the average angle of insertion increased significantly from pre-, $24.3^\circ(3.1^\circ)$ to post-test, $40.2^\circ(1.4^\circ)$. In addition, there was a significant main effect of pre- and post-test on the average path length, $F(1, 41) = 8.728$, $p < 0.005$, Wilks' $\Lambda = 0.824$, $\eta^2 = 0.176$, observed power of 0.822; the average path length decreased from pre- to post-test, $(100.9(12.2)$ and $34.5(4.8)$, respectively). Finally, there was a significant main effect of pre- and post-test on the average jerk, $F(1, 41) = 31.886$, $p < 0.0005$, Wilks' $\Lambda = 0.563$, $\eta^2 = 0.437$, observed power of 1.000. The average jerk decreased significantly from pre-test, $5402(793)$ to post-test, $1822(236)$. All other dependent variables were not significant ($p > 0.177$).

Finally, a repeated measures ANOVA was conducted to determine whether training group (manikin and robotic) and time point (pre- and post-test) significantly impacted time to complete the procedure for the 2017 cohort ($N=26$). Results showed that training group did not have a significant interaction effect on time to complete, $F(1, 23) = 0.001$, $p = 0.971$, Wilks' $\Lambda = 1.000$, $\eta^2 = 0.000$, observed power of 0.050. Therefore, main effects were examined. The main effects showed that time to complete the procedure was significantly different from pre- to post-test, $F(1, 23) = 8.374$, $p < 0.008$, Wilks' $\Lambda = 0.733$, $\eta^2 = 0.267$, observed power of 0.791. Specifically, time to complete the procedure decreased significantly from pre-test ($M = 3.516$, $SE = 0.277$) to post-test ($M = 1.997$, $SE = 0.409$), see Figure 4. This metric is distinct from the time-to-complete calculated from motion tracking

in the previous research question, which focused solely on the time the needle entered the simulated skin to the time the needle was removed from the skin. Additionally, a Mann-Whitney U test was run to determine if there were differences in post-test economy of time and motion scores between manikin and robotically trained participants. Results showed that median scores for manikin-trained (3.0) and robotically-trained (3.0) participants were not statistically significantly different, $U = 239$, $z = -0.921$, $p = .357$, using an exact sampling distribution for U.

Discussion

Results from the first research question showed that with training, participants significantly improved pass rates in all checklist items from pre- to post-test. Contrary to our hypothesis, this study showed that when evaluated critically and properly, the checklist could distinguish proficiency of IJ CVC skills. It is important to note that no feedback to participants was given during pre- and post-test. Additionally, the results showed that training group did not have a significant effect on pre- to post-test performance. All participants improved checklist skills, regardless of training method.

Analysis of the motion tracked metrics showed that the angle of insertion, path length, and jerk changed significantly after training and provide more insight into specific needle insertion skill gains. These results support our hypothesis that objective measures of skill gains would improve with training. Additionally, there were no training group differences, which indicates that the robotic environment trained participants in needle insertion skills as effectively as the manikin environment.

Finally, time to complete the procedure in 2017 significantly decreased from pre- to post-test, regardless of training group. Additionally, there were no differences in the economy of time and motion between manikin-trained and robotically-trained participants. These results together show that the continuous quantitative feedback provided to the robotic group during training were equally effective to the verbal feedback provided to the manikin participants in improving ultrasound guided vascular access. Overall, these results show that the robotic training environment could provide more efficient training and prepare residents for novel situations as effectively as manikin training systems without the need of a trained preceptor.

There were a few limitations to this study. Specifically, the motion tracking data was trimmed to only track the final needle insertion attempt during the pre- and post-tests, and did not account for multiple previous attempts. Because the pre-test was the first time residents placed needles into a manikin, or patient simulation, several residents could not properly locate the vein or did not feel comfortable inserting the needle. Data was removed for individuals who could not complete the pre-test procedure within the allotted five minutes, while the post-test had no time limitations. The cost of robotic training materials in this iteration includes the haptic robotic arm (\$10,870), computer, and software packages that total to \$17,289. While this is more expensive than higher-end central line manikins, e.g. the Blue-Phantom model cost \$2,699, the haptic robotic arm has a longer lifespan and can be configured to represent more diverse anatomies and training experiences, which would reduce long-term costs. The research team has more recently converted to an open-

source software package and investigated different products to lower future costs of production.

Conclusion

The goal of this study was to examine the effectiveness of robotic training compared to manikin training. The main findings were as follows: (1) There were significant improvements from the pre- to post- test in all items on the IJ CVC checklist, but there were no training group differences; (2) There were significant increases in the average angle of insertion, path length, and jerk from pre- to post- test, but no training group differences for any of the 5-objective metrics; (3) Training significantly improved residents' ability to efficiently and effectively complete the procedure. These results support the use of the IJ CVC checklist and motion tracked metrics to evaluate needle insertion skills. Overall results support DHRT as an effective training method for improving needle insertion efficiency and for training ultrasound guided vascular access.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgements:

Research reported in this publication was supported by the National Heart, Lung, and Blood Institute of the National Institutes of Health under Award Number R01HL127316. The content is solely the responsibility of the authors and does not necessarily represent the official views of the National Institutes of Health.

References

1. Inhoff AW, Wang J. Encoding of text, manual movement planning, and eye-hand coordination during copytyping. *Journal of Experimental Psychology: Human Perception and Performance* 1992;18:437. [PubMed: 1593229]
2. McGee DC, Gould MK. Preventing complications of central venous catheterization. *New England Journal of Medicine* 2003;348:1123–33. [PubMed: 12646670]
3. Bernard RW, Stahl WM. Subclavian vein catheterizations: a prospective study. I. Non-infectious complications. *Annals of surgery* 1971;173:184. [PubMed: 5100094]
4. Sznajder JJ, Zveibil FR, Bitterman H, Weiner P, Bursztein S. Central vein catheterization: failure and complication rates by three percutaneous approaches. *Archives of Internal Medicine* 1986;146:259–61. [PubMed: 3947185]
5. Hartman N, Wittler M, Askew K, Hiestand B, Manthey D. Validation of a performance checklist for ultrasound-guided internal jugular central lines for use in procedural instruction and assessment. *Postgraduate medical journal* 2016;postgradmedj-2015–133632.
6. Doyle JD, Webber EM, Sidhu RS. A universal global rating scale for the evaluation of technical skills in the operating room. *The American journal of surgery* 2007;193:551–5. [PubMed: 17434353]
7. Huang GC, Newman LR, Schwartzstein RM, et al. Procedural competence in internal medicine residents: validity of a central venous catheter insertion assessment instrument. *Academic Medicine* 2009;84:1127–34. [PubMed: 19638784]
8. Laack TA, Dong Y, Goyal DG, Sadosty AT, Suri HS, Dunn WF. Short-term and long-term impact of the central line workshop on resident clinical performance during simulated central line placement. *Simulation in Healthcare* 2014;9:228–33. [PubMed: 24401920]

9. Scott DJ, Valentine RJ, Bergen PC, et al. Evaluating surgical competency with the American Board of Surgery In-Training Examination, skill testing, and intraoperative assessment. *Surgery* 2000;128:613–22. [PubMed: 11015095]
10. Walzak A, Bacchus M, Schaefer JP, et al. Diagnosing technical competence in six bedside procedures: comparing checklists and a global rating scale in the assessment of resident performance. *Academic Medicine* 2015;90:1100–8. [PubMed: 25881644]
11. Evans LV, Morse JL, Hamann CJ, Osborne M, Lin Z, D'Onofrio G. The development of an independent rater system to assess residents' competence in invasive procedures. *Academic Medicine* 2009;84:1135–43. [PubMed: 19638785]
12. Mansfield PF, Hohn DC, Fornage BD, Gregurich MA, Ota DM. Complications and failures of subclavian-vein catheterization. *New England Journal of Medicine* 1994;331:1735–8. [PubMed: 7984193]
13. Segall N, Bennett-Guerrero E. Should Checklists for Central Venous Catheter Insertion Be Used in Training, Credentialing, or Routine Clinical Care? *Critical Care Medicine* 2016;44:1943–4. [PubMed: 27635488]
14. Dong Y, Suri HS, Cook DA, et al. Simulation-based objective assessment discerns clinical proficiency in central line placement: a construct validation. *CHEST Journal* 2010;137:1050–6.
15. Ericsson KA. The influence of experience and deliberate practice on the development of superior expert performance. *The Cambridge handbook of expertise and expert performance* 2006:683–703.
16. Bjerrum F, Strandbygaard J, Rosthøj S, Grantcharov T, Ottesen B, Sorensen JL. Evaluation of Procedural Simulation as a Training and Assessment Tool in General Surgery—Simulating a Laparoscopic Appendectomy. *Journal of surgical education* 2017;74:243–50. [PubMed: 27717707]
17. Reznick RK, MacRae H. Teaching surgical skills—changes in the wind. *New England Journal of Medicine* 2006;355:2664–9. [PubMed: 17182991]
18. Datta V, Mackay S, Mandalia M, Darzi A. The use of electromagnetic motion tracking analysis to objectively measure open surgical skill in the laboratory-based model. *Journal of the American College of Surgeons* 2001;193:479–85. [PubMed: 11708503]
19. Moorthy K, Munz Y, Sarker SK, Darzi A. Objective assessment of technical skills in surgery. *BMJ: British Medical Journal* 2003;327:1032. [PubMed: 14593041]
20. Howells NR, Brinsden MD, Gill RS, Carr AJ, Rees JL. Motion analysis: a validated method for showing skill levels in arthroscopy. *Arthroscopy: The Journal of Arthroscopic & Related Surgery* 2008;24:335–42. [PubMed: 18308187]
21. Kim I, Miller SR, Frievalds A. Motion analysis as an evaluation framework for eye-hand coordination: A case study in ultrasound-guided catheter insertion. *ASME Design Engineering Technical Conference 26th Annual Conference on Design Theory and Methodology* Buffalo, NY 2014.
22. Varas J, Achurra P, León F, et al. Assessment of central venous catheterization in a simulated model using a motion-tracking device: an experimental validation study. *Annals of surgical innovation and research* 2016;10:1. [PubMed: 26807142]
23. Clinkard D, Holden M, Ungi T, et al. The development and validation of hand motion analysis to evaluate competency in central line catheterization. *Academic Emergency Medicine* 2015;22:212–8. [PubMed: 25676530]
24. Mohamadipanah H, Parthiban C, Nathwani J, Rutherford D, DiMarco S, Pugh C. Can a virtual reality assessment of fine motor skill predict successful central line insertion? *The American Journal of Surgery* 2016;212:573–8. e1. [PubMed: 27649977]
25. Stefanidis D, Sevdalis N, Paige J, et al. Simulation in surgery: what's needed next? *Annals of surgery* 2015;261:846–53. [PubMed: 25243562]
26. Pepley D, Yovanoff M, Mirkin K, Han D, Miller S, Moore J. Design of a Virtual Reality Haptic Robotic Central Venous Catheterization Training Simulator. *ASME 2016 International Design Engineering Technical Conferences and Computers and Information in Engineering Conference*; 2016; Charlotte, NC, USA: American Society of Mechanical Engineers p. V05AT7A033-V05AT07A.

27. Yovanoff M, Pepley D, Mirkin K, Moore J, Han D, Miller S. Personalized Learning in Medical Education: Designing a User Interface for a Dynamic Haptic Robotic Trainer for Central Venous Catheterization. Proceedings of the Human Factors and Ergonomics Society Annual Meeting Austin, TX: SAGE Publications Sage CA: Los Angeles, CA; 2017.
28. Dijksterhuis MG, Voorhuis M, Teunissen PW, et al. Assessment of competence and progressive independence in postgraduate clinical training. Medical education 2009;43:1156–65. [PubMed: 19930506]
29. Yovanoff M, Pepley D, Mirkin K, Moore J, Han D, Miller S. Improving Medical Education: Simulating Changes in Patient Anatomy Using Dynamic Haptic Feedback. Proceedings of the Human Factors and Ergonomics Society Annual Meeting Washington D.C.: SAGE Publications Sage CA: Los Angeles, CA; 2016:603–7.
30. Ruthenbeck GS, Reynolds KJ. Virtual reality for medical training: the state-of-the-art. Journal of Simulation 2015;9:16–26.
31. Ericsson KA, Simon HA. Verbal reports as data. Psychological review 1980;87:215.
32. Yovanoff MA, Chen H-E, Pepley DF, et al. Investigating the Impact of Simulator Functional Fidelity and Personalized Feedback on Central Venous Catheterization Training. Journal of Surgical Education 2018;75.
33. Pepley DF, Sonntag CC, Prahbu RS, et al. Building ultrasound phantoms with modified polyvinyl chloride: a comparison of needle insertion forces and sonographic appearance with commercial and traditional simulation materials. Simulation in Healthcare Accepted December 2017.
34. Emam TA, Hanna GB, Kimber C, Cuschieri A. Differences between experts and trainees in the motion pattern of the dominant upper limb during intracorporeal endoscopic knotting. Digestive surgery 2000;17:120–5. [PubMed: 10781973]
35. Savitzky A, Golay MJ. Smoothing and differentiation of data by simplified least squares procedures. Analytical chemistry 1964;36:1627–39.
36. Ghasemloonia A, Maddahi Y, Zareinia K, Lama S, Dort JC, Sutherland GR. Surgical Skill Assessment Using Motion Quality and Smoothness. Journal of surgical education 2017;74:295–305. [PubMed: 27789192]

Figure 1:

Flowchart of CVC training in 2016 and 2017. Pre-tests were completed on a Blue-Phantom (BP) manikin; post-tests were conducted on a BP manikin and a novel manikin.

Figure 2:
Attachment of the motion tracker (indicated by the red circle) to the syringe for the pre- and post- tests which was attached to the Ascension 3D Guidance trakSTAR electromagnetic position tracking system.

Figure 3: Plots used to identify point of needle insertion by plotting the needle vertical depth and needle angle against time (s). Dotted lines represent the start point and end point of the data trimming.

Figure 4:

Average time-to-complete the pre-test and post-test procedure in 2017 decreased significantly, $F(1, 23) = 8.374$, $p < 0.008$, from pre-test ($M = 3.516$, $SE = 0.277$) to post-test ($M = 1.997$, $SE = 0.409$).

Table 1:

Total number of pass rates and Wilcoxon signed rank results for all participants on the first test attempt during pre-test (N=52) and post-test (N=49) on the matching CVC checklist items. Data for three participants were not available for the post-test.

	Test Item	Pre-Test	Post-Test	N	Z	p-value
1	Selecting the appropriate site for venipuncture	46	48	52	-2.449	0.014
2	Selecting the correct ultrasound probe and using it in the appropriate orientation	35	49	52	-3.873	0.0005
3	Obtaining a clear image of the target vessels using the ultrasound machine	44	49	52	-2.828	0.005
4	Correctly distinguishing artery and vein: demonstrates compressibility of vein	42	49	52	-3.162	0.002
5	Inserting the needle at a 35–45 degree angle from the skin	30	49	52	-4.583	0.0005
6	Locating the needle's position on the ultrasound image	40	48	52	-3.317	0.001
7	Advancing the introducer needle slowly and steadily (2017)	16	26	26	-3.000	0.003
8	Venipuncture Successful	30	46	52	-3.719	0.0005
	Number of Insertion Attempts	101	67	-	-	-
9	Confirming vessel entry by aspiration of blood	24	47	52	-4.600	0.0005
10	Conducting the entire procedure without any mistakes	2	45	52	-6.481	0.0005
	Average Pass Rate (does not include number of attempts)	59.4	98.3	-	-	-

Table 2:

Mean, Standard Error, and ANCOVA results for participants from pre- to post-test.

		Pre-Test			Post-Test			ANCOVA	
		N	Means	SE	N	Means	SE	F	P-value
Needle Time (s)	51	200.7	67.9	45	66.1	24.9	1.882	0.177	
Angle (°)	51	24.3	3.1	45	40.2	1.4	16.874	0.0005*	
Path Length (in)	51	100.9	12.2	44	34.5	4.8	8.728	0.005*	
Velocity (in/s)	51	0.464	0.050	44	0.366	0.055	1.495	0.228	
Jerk (in/s^3)	51	5402	793	44	1822	236	31.886	0.0005*	

* Indicates significance at $p < .05$.