The American Journal of Surgery 217 (2019) 362–367

Contents lists available at ScienceDirect

The American Journal of Surgery

journal homepage: www.americanjournalofsurgery.com

Looks can be deceiving: Gaze pattern differences between novices and experts during placement of central lines

Hong-En Chen [a], Cheyenne C. Sonntag [b], David F. Pepley [c], Rohan S. Prabhu [c], David C. Han [b, d], Jason Z. Moore [c], Scarlett R. Miller [a, e, *]

[a] Department of Industrial and Manufacturing Engineering, Penn State, University Park, PA, 16802, USA
[b] Department of Surgery, Penn State Health Milton S. Hershey Medical Center, Hershey, PA, 17033, USA
[c] Department of Mechanical and Nuclear Engineering, Penn State, University Park, PA, 16802, USA
[d] Penn State Heart and Vascular Institute, Penn State Health Milton S. Hershey Medical Center, Hershey, PA, 17033, USA
[e] School of Engineering Design, Technology, and Professional Programs, Penn State, University Park, PA, 16802, USA

ARTICLE INFO

Article history:
Received 16 May 2018
Received in revised form
2 October 2018
Accepted 12 November 2018

Keywords:
Central venous catheterization
Medical training simulation
Eye tracking
Residency training

ABSTRACT

Background: The objective of this study was to determine whether gaze patterns could differentiate expertise during simulated ultrasound-guided Internal Jugular Central Venous Catheterization (US-IJCVC) and if expert gazes were different between simulators of varying functional and structural fidelity.
Methods: A 2017 study compared eye gaze patterns of expert surgeons ($n = 11$), senior residents ($n = 4$), and novices ($n = 7$) during CVC needle insertions using the dynamic haptic robotic trainer (DHRT), a system which simulates US-IJCVC. Expert gaze patterns were also compared between a manikin and the DHRT.
Results: Expert gaze patterns were consistent between the manikin and DHRT environments ($p = 0.401$). On the DHRT system, CVC experience significantly impacted the percent of time participants spent gazing at the *ultrasound screen* ($p < 0.0005$) and the *needle and ultrasound probe* ($p < 0.0005$).
Conclusion: Gaze patterns differentiate expertise during ultrasound-guided IJCVC placement and the fidelity of the simulator does not impact gaze patterns.

Introduction

Ultrasound-guided internal jugular central venous catheterization (IJCVC) is a procedural skill that has been taught on manikin simulators for over a decade.[1] While research has shown these simulators are effective for short term skill gains, these manikins have several known limitations, including a lack of long-term skill retention.[2] This has been attributed, in part, to the fact that these simulators only train surgeons on a single patient anatomy and thus do not represent the range of patient profiles surgeons face in clinical settings.[3] In addition, these simulators lack objective performance criterion and instead rely on a trained preceptor (e.g. faculty) to be present in order to subjectively evaluate trainee performance using a proficiency checklist[4–6] which includes evaluation on mechanical (e.g. motion and accuracy) and procedural-based skills (e.g. aseptic technique). Finally, current CVC simulators provide only basic summative feedback on performance (blue liquid is aspirated if the catheter hits a vein) and no concurrent or formative feedback. Because of these shortcomings and the over-reliance on faculty feedback, simulation-based surgical training has been criticized as being resource intensive.[7,8]

In light of these deficiencies, the Dynamic Haptic Robotic Trainer (DHRT) was developed to provide residents with IJCVC training on a *variety of patient anatomies* while providing *automated feedback* during the training process, see Fig. 1. While the DHRT system lacks some of the physical realism of manikin simulators, it has increased functional fidelity. Specifically, the DHRT provides training on ultrasound guided IJCVC needle placement by simulating variations in patient anatomy through changes in a virtual ultrasound image (e.g. size and depth of the vessels) and through haptic feedback provided through a robotic arm that simulates force changes of different types of tissues (e.g. skin, adipose tissue, vessel), see Pepley et al. for full details.[9,10] The DHRT also provides learners with automated feedback on their mechanical performance after each

* Corresponding author. 213P Hammond Building, University Park, PA, 16802, USA.
E-mail address: scarlettmiller@psu.edu (S.R. Miller).

https://doi.org/10.1016/j.amjsurg.2018.11.007
0002-9610/

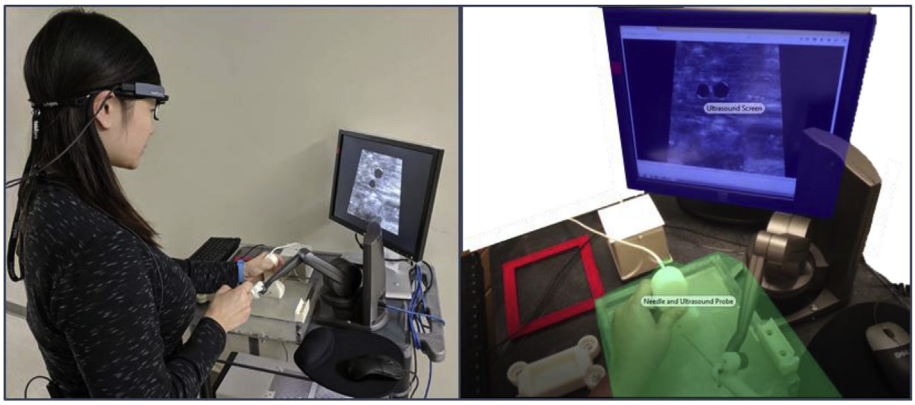

Fig. 1. (Left) Experimental setup with participant wearing Tobii Pro Glasses 2 at the DHRT system. (Right) Two main areas of interest (AOIs). The *ultrasound screen* and the *needle and ultrasound probe*.

needle insertion attempt without the need of a trained preceptor, including feedback on the number of insertion attempts, average angle of insertion, and the final distance of the needle tip from the center of the vein.[11] Prior research has shown that novice learners approach expert performance when training on the DHRT system[10] and indicated that there were no differences in pre-to post-test learning between manikin- and DHRT-based instruction.[12] While this prior work on the DHRT is promising, the feedback provided to learners is based purely on their mechanical skill acquisition (e.g. needle angle) and not on the cognitive skills necessary to complete this surgical procedure. This is important because prior studies have shown that cognitive skills training and mental imagery practice enhance the learning and acquisition of surgical skills.[13–15] One method of focusing surgeons' attention in the DHRT system is through the use of eye-tracking and gaze-training.

In medical domains, gaze behavior has been used for skills training[16] and expertise assessment.[17,18] For example, Tien and colleagues[19] differentiated expert and novice gaze behavior during open hernia surgery. Reviews of eye-tracking applications in medical and non-medical domains have also shown the reliability and validity for using eye-tracking as an objective assessment tool, as well as its potential in providing feedback and assessing surgical skills.[20,21] One study on virtual laparoscopic training found that experts spend more time gazing at target locations while novices spend more time tracking tools.[22] When using gaze training as a teaching tool in laparoscopic surgery, researchers have also found that gaze trained groups learned more quickly than self-guided groups,[23] and adopted expert-like gaze patterns by fixating more on target locations rather than on tool movement.[24] Other studies have developed objective methods to differentiate expert and novice surgeon skill levels based on tool movement and eye gaze patterns.[25,26] Finally, a pilot study in ultrasound-guided CVC found that experts displayed more efficient hand motions than novices and improved movement accuracy from gazing at the ultrasound monitor, indicating that experts used goal-oriented motion planning.[27]

While this prior work demonstrates the potential benefits of exploring gaze behavior during IJ CVC placement, this has not yet been empirically explored in CVC training. Thus, as a first step to improving training feedback, the goal of the current study was to assess gaze tracking as it relates to CVC placement and determine how surgeons with different levels of expertise allocate their visual attention. Knowledge from this study could contribute to the development of additional gaze training and assessment methods to aid a learner's path to competency in CVC training.

Research questions

The purpose of the current study was to determine whether gaze patterns could differentiate expertise during simulated ultrasound-guided Internal Jugular Central Venous Catheterization (US-IJCVC) and if expert gazes were different between simulators of varying functional and structural fidelity. Thus, the two research questions (RQ) explored in the current study were:

***RQ1:* Are expert eye gaze patterns consistent during central-line insertions between manikin and DHRT training systems?** This research question was developed to validate if eye gaze patterns were consistent between the manikin (higher structural fidelity or physical realism) and the DHRT (higher functional fidelity) systems. It was hypothesized that since the ultrasound-guided procedure for completing the central line was similar between the two systems, gaze patterns would not differ between the environments.

***RQ2:* How do expert, senior resident, and novice gaze patterns differ when placing central lines on the DHRT?** The second research question aimed to compare time that experts, senior residents, and novices spent fixating on the *ultrasound screen* versus the *needle and ultrasound probe* during placement of central lines on the DHRT system. This question was developed based on prior work that suggested that surgical expertise may be reflected in tool movement and eye gaze patterns[25–27] and on prior work in laparoscopic surgery that has shown that experts spend more time fixating on target locations, while novices split their time between the target location and tracking the tools.[22] In this study, it was hypothesized that novices would spend a higher percentage of time tracking the needle and ultrasound probe during these trials, while experts would spend more time focused on the ultrasound screen.[22] Senior residents, representative of a learner group that is often considered competent enough to perform the skill on patients with limited to no supervision,[28] were hypothesized to perform similarly to experts in this study.

Methods and materials

In order to answer these research questions, a series of experiments were conducted at Penn State Milton S. Hershey Medical

Center (HMC). The details of this experiment, conducted as part of a larger study, are discussed in this section.

Participants

Data for the current study was collected over two time periods. Specifically, the data from summer of 2017 included seven surgical interns (novice) ($N_n = 7$) while the data from fall of 2017 included eleven experts ($N_e = 11$) and four senior residents ($N_{sr} = 4$). To qualify expertise, the expert participants in this study included four fellows and seven attending surgeons who self-reported to have placed more than 50 central lines in their career, a procedural volume which has been stated to denote expertise in earlier research on central-lines.[29,30] On the other hand, the senior residents were all PGY4 or PGY5, and had placed a minimum of 10 central lines on patients in the course of their careers; only one of the novice residents had placed an Internal Jugular CVC prior to the study, reporting they had placed 3 prior lines. Senior residents were included in the current study to identify whether residents perform similarly to experts after several years of residency training. This was particularly interesting since a large portion of central-lines are placed by surgical residents at HMC while under the supervision of attending faculty.

Procedure

At the start of the study, the purposes and procedures were explained to all of the participants and informed consent was obtained according to an Institutional Review Board approved protocol. In the summer session, the novices were first shown a 15.5-min video on CVC placement followed by a live training demonstration of the procedure by a fourth-year surgical resident using a Blue Phantom Gen II Ultrasound Central Line Training Model (Model #BPH660). For the data collection, each participant was fitted and calibrated with the Tobii Pro Glasses 2,[31] see Fig. 1, and asked to complete one baseline needle insertion for central line placement on the Blue Phantom manikin. This was followed by an introduction to the DHRT through a 5-min video, after which, the participants individually completed two practice insertions on the DHRT in order to familiarize themselves with its functions. During these two practice insertions, the participants were guided through the functions of the DHRT and any questions were answered. Finally, the DHRT insertion data was collected, where each participant completed 10 insertions on the DHRT system. All of the participants received performance feedback from the DHRT learning interface, and no other external feedback was provided. Due to restrictions in the residency boot camp where the resident data was collected, novices had a day between their first two practice insertion and their 10 insertion trials. On the other hand, the 2 practice insertions and 10 insertion trials were conducted on the same day for the experts and senior residents, see Fig. 2 for timeline of study. It is important to note that the one-day break between the practice insertions and the 10 insertions did not result in a performance dip for novices. On average, novices in this cohort improved their overall performance score (a weighted score based on angle, insertion attempts, aspiration, etc.[10]) from practice stick 1 (74.6%) and practice stick 2 (85.9%) to the first two insertions on the following day, at 86.7% and 91.5%, respectively.

Eye tracking metrics

Eye gaze point data collected during the study was analyzed using the Tobii Pro Lab real-world automatic mapping functionality.[32] Once auto-mapped, an independent rater used the fixation filter to manually check all auto-mapped data. Specifically, raw data from trial 1 was analyzed by one rater and compared to the automatically mapped and checked data to validate the use of the auto-mapping feature. The inter-rater reliability was found to be very good, as classified by Altman (p.404),[33] Cohen's Kappa = 0.827, $p < 0.0005$. Thus, metrics exported using the default Tobii I-VT fixation filter[34] in the Tobii Pro Lab software were used to analyze participant eye gaze patterns.

Once the gaze points were properly mapped, data for each participant and each trial was reviewed for data quality and percentage of gaze samples captured during needle insertions. Participants with poor data quality as assessed by the Tobii Pro system were discarded. When using the Tobii Pro Glasses 2, gaze points are only captured when the participant is looking through the glasses range (e.g. not above or below the rim of the glasses). For this study, a lower threshold for percentage of gaze samples captured was set at 70% of each participant's eye gazes captured during needle insertions. During review of the collected data, one expert and one novice participants' gaze patterns were discarded due to having less than 70% (in many trials, less than 50%) of their gaze points captured during the needle insertions, leading to a total participant count of 6 novices, 10 experts, and 4 senior residents, see Fig. 2 for summary.

After participant gaze data for all 10 trials were reviewed for eye-tracking quality, the data were categorized into two main areas of interest (AOI): the ultrasound image on the screen, and the needle and ultrasound probe interface (Fig. 1). These two areas of interest were chosen because prior work in surgery has compared

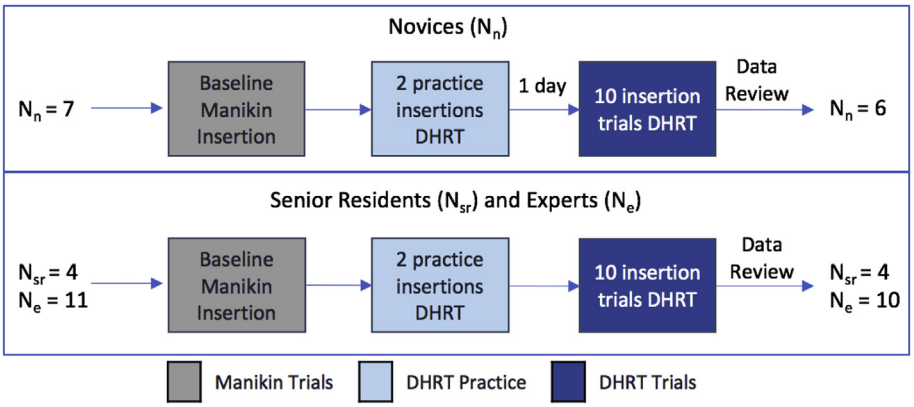

Fig. 2. Diagram summarizing the procedural flow.

the amount of time spent fixating on the target (e.g. ultrasound or image visualization), compared to the amount of time spent tracking the tools or patient (e.g. needle and ultrasound probe).[22] In order to compare participant gaze patterns, the total fixation duration for each needle insertion trial was used. Fixation data was analyzed based on the default Tobii Pro lab fixation filter settings. Additionally, because there was a large variability in the mean (SD) time to complete needle insertion trials among the participant groups – experts 22.42s (14.04), senior residents 24.79s (8.01), and novices 29.75s (12.60) – the percentage of time spent fixating on these two AOIs during each needle insertion was also calculated to normalize the data and account for these large discrepancies. Finally, to identify differences in scan patterns, the data was exported in 20 ms increments to analyze AOI hits in the *ultrasound screen* and the *needle and ultrasound probe*. An R script was executed to remove all lines where there were no AOI hits. The cleaned data was then analyzed for each trial to determine the number of times the participant switched between each AOI. For example, if a participant looked at AOI 1, AOI 2, and then AOI 1, this counted as 2 switches between AOIs.

Data analysis

All metrics were analyzed using SPSS (v. 25.0) with a significance level of 0.05. For the first research question, two paired-samples t-tests were computed to determine differences in the percent of time experts fixated on the *ultrasound screen* and the *needle and ultrasound probe* between the one insertion trial conducted on the manikin and the average of the ten trials completed on the DHRT system. The data had no significant outliers, as assessed by boxplots, and the assumption of normality was not violated, as assessed by boxplot and the Kolmogorov Smirnov test ($p > 0.200$).

For the second research question, three one-way multiple analysis of variances (MANOVA) were computed to determine if and how the total fixation duration and the percent time allocated (total fixation duration normalized to completion time) to the *ultrasound screen* or the *needle and ultrasound probe* were different among groups with different levels of central-line insertion experience throughout the 10 trials on the DHRT system. Groups were compared as a whole, with each needle insertion representing one data point. Preliminary assumption checking revealed that data were normally distributed, as assessed by boxplot and the Kolmogorov Smirnov test ($p > 0.200$); there were no univariate or multivariate outliers, as assessed by boxplot and Mahalanobis distance ($p > 0.001$), respectively; there were no linear relationships, as assessed by scatterplot; no multicollinearity ($r = -0.887$, $p < 0.001$). Follow-up post-hoc analyses were conducted to examine pairwise differences between groups. Finally, scan patterns for gaze switches between the two main AOIs were analyzed for group differences.

Results

Data are presented as mean ± standard error unless otherwise noted. Analysis of the areas of interests (AOI) showed that participants on average spent a total of 79.08 ± 0.67% of their time during the study trials fixating on the *ultrasound screen* and the *needle and ultrasound probe*. During the study, the participants spent an average of 63.88 ± 0.97% percent of the time fixating on the *ultrasound screen* and an average of 18.5 ± 0.79% of time tracking the *needle and ultrasound probe*.

The paired-samples *t*-test showed that there was no significant difference between the percent of time experts spent fixating on the *ultrasound screen* on the manikin (61.57 ± 3.22%) and the DHRT system (65.41 ± 3.94%), $t(9) = -1.082$, $p = 0.307$. Likewise, there was no significant difference between the percent of time experts spent fixating on the *needle and ultrasound probe* on the manikin (10.67 ± 2.42%) and the DHRT system (12.58 ± 2.39%), $t(9) = -0.946$, $p = 0.369$.

The first one-way MANOVA results showed that the differences among central line insertion experience (novice, senior resident, and expert) on the combined dependent variables (total time fixated on the *ultrasound screen* and *needle and ultrasound probe*) was statistically significant, $F(4, 394) = 9.072$, $p < 0.0005$, Pillai's Trace = 0.169, partial $\eta^2 = 0.084$, observed power of 0.999. See Fig. 3 for example gaze plots for the three groups. Follow-up univariate ANOVAs showed that the total time fixated on both the *ultrasound screen* ($F(2, 254.418) = 4.010$, $p < 0.020$, partial $\eta^2 = 0.039$, observed power of 0.712), and the *needle and ultrasound probe* ($F(2, 149.268) = 17.442$, $p < 0.0005$, partial $\eta^2 = 0.150$, observed power of 1.000) were significantly different between the participants from the different central-line experience groups, using a Bonferroni adjusted α level of 0.025.

The assumption of homogeneity of variances, as assessed by Levene's Test for Equality of Variance, was satisfied for the total time spent fixating on the *ultrasound screen* ($p = 0.149$) but not the *needle and ultrasound probe* ($p = 0.049$). Thus, the Games-Howell post-hoc pairwise comparisons were conducted for the total time fixated on the *needle and ultrasound probe*, while the Tukey-Kramer post-hoc pairwise comparison was conducted for the total time fixated on the *ultrasound screen*. The post-hoc results showed that novices (5.73 ± 0.36s) spent significantly more time fixating on the tracking the *needle and ultrasound probe* than experts (3.01 ± 0.33s) and senior residents (3.20 ± 0.31s). Novices (17.55 ± 1.20s) spent significantly less time fixating on the *ultrasound screen* than experts (14.17 ± 0.79s) but similar total times as senior residents (17.01 ± 0.87s). Specifically, differences between novices and experts (2.72s, 95% CI [1.56, 3.88], $p = 0.0005$) were as well as novices and residents (2.53s, 95% CI [1.39, 3.66], $p = 0.0005$) were statistically significant when gazing at the *needle and ultrasound probe*. On the other hand, novices spent significantly less time fixating on the *ultrasound* screen than experts (3.38s, 95% CI [0.31, 6.45], $p = 0.027$). All other comparisons were not statistically significant.

The second one-way MANOVA showed that the percentage of total fixation time allocated to the *ultrasound screen* and the *needle and ultrasound probe* was statistically significant, $F(4, 394) = 8.575$, $p < 0.0005$, Pillai's Trace = 0.160, partial $\eta^2 = 0.080$, observed power of 0.999. Follow-up univariate ANOVAs showed that both the percent time allocated to the *ultrasound screen* ($F(2, 1646.792) = 9.529$, $p < 0.0005$, partial $\eta^2 = 0.088$, observed power of 0.979), and the *needle and ultrasound probe* ($F(2, 1434.600) = 15.618$, $p < 0.0005$, partial $\eta^2 = 0.137$, observed power of 0.999) were significantly different between the participants from the different central-line experience groups, using a Bonferroni adjusted α level of 0.025. The percentage of time participants allocated to each AOI from different groups is shown in Fig. 4.

Because the assumption of homogeneity of variances was violated for both the percent time spent fixating on the *ultrasound screen* and the *needle and ultrasound probe*, as assessed by Levene's Test for Equality of Variance ($p = 0.0005$ and 0.025, respectively), the Games-Howell post-hoc test was used to compare pairwise differences between groups for time allocated to each AOI. The post-hoc results showed that novices (58.01 ± 1.85%) allocated significantly less time to the *ultrasound screen* than experts (65.40± 1.42%) and senior residents (68.86 ± 1.07%). In addition, novices (20.99 ± 1.44%) allocated significantly more time tracking the *needle and ultrasound probe* than experts (12.58 ± 0.92%) and senior residents (13.11 ± 1.22%). Specifically, differences between novices and experts (−7.40%, 95% CI [-12.93, −1.86], $p = 0.005$) as

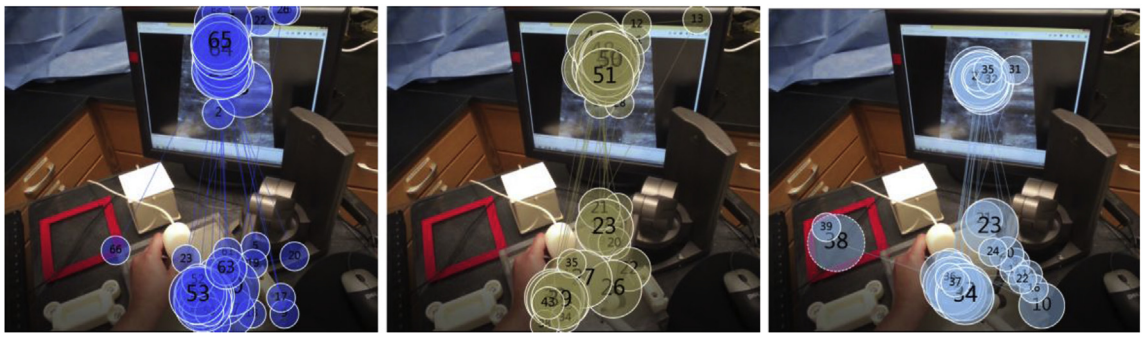

Fig. 3. (Left) Example novice gaze pattern on DHRT. (Middle) Example senior resident gaze pattern. (Right) Example expert gaze pattern. The size of each circle represents the fixation duration, while the number within each circle represents the chronological order of the fixation point.

well as novices and residents (−10.85%, 95% CI [-15.94, −5.76], $p = 0.0005$) were statistically significant when gazing at the *ultrasound screen*. On the *needle and ultrasound probe*, differences between novices and experts (8.40%, 95% CI [4.34, 12.46], $p = 0.0005$), as well as novices and residents (7.88%, 95% CI [3.39,12.37], $p = 0.0005$) were statistically significant. However, there were no significant differences between experts and senior residents in time allocated to the *ultrasound screen* ($p < 0.133$) or the *needle and ultrasound probe* ($p < 0.938$).

Finally, a one-way ANOVA showed that there was no significant difference between the scan path of novices (6.30 ± 0.504), senior residents (5.65 ± 0.424), and experts (5.11 ± 0.407), $F(2, 26.693) = 1.864$, $p < 0.158$, partial $\eta^2 = 0.019$, observed power of 0.385.

Discussion

The results for the first research question support our hypothesis that gaze patterns would not differ between the manikin and DHRT environments. This demonstrates that even in the absence of the physical realism, in the DHRT simulator, experts maintained similar gaze patterns. These results provide evidence that the DHRT system may require similar eye gaze patterns even in the absence of these physical features. Additionally, our results show that expertise in IJCVC can be differentiated using gaze behavior, supporting related work that differentiated expertise with gaze patterns during open hernia surgery and laparoscopic environments.[19,22] While the experienced surgeons in our study spent more time fixating on the ultrasound screen, the novices spent significantly more time tracking the needle and ultrasound probe. Additionally, while not significant, novices switched their attention between the AOIs more frequently than senior residents and experts. These results may point to the importance of skills needed to comfortably manipulate the needle and ultrasound probe, while simultaneously visualizing the target location on the ultrasound screen, during IJ CVCs.

Finally, it was hypothesized that experts would spend more time gazing at the ultrasound screen, while novices would spend more time tracking the needle and ultrasound probe, with senior residents falling along a continuum between the two groups. While not significant, an interesting discovery was that senior residents allocated the most attention on the *ultrasound screen,* and a similar amount of time as experts tracking the *needle and ultrasound probe*. When combined with prior work showing that senior resident performance was comparable with expert performance,[10] this additional similarity in gaze patterns indicate that senior residents attain expertise in CVC placement during their post graduate residency career. This may be due in part to the fact that IJCVCs are often performed by residents at this institution under the supervision of experts. This means that while our experts met the criteria of a minimum 50 central lines,[29] the senior residents in the study may have had more active line placement in the months leading up to the testing, which may have led to higher similarities in the performance among these two groups. Another explanation is that ultrasound guidance has shifted the learning curve and the number of placements necessary to perform similarly to experts.

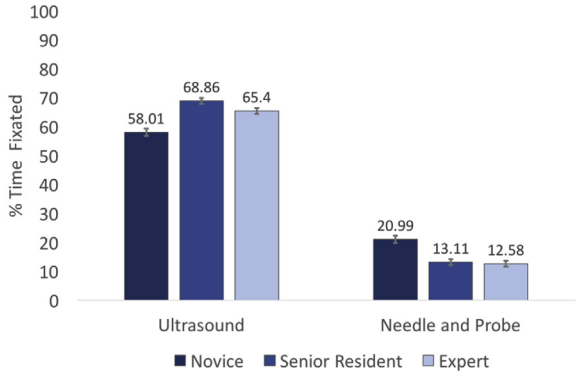

Fig. 4. Average percent of time each group allocated to the *ultrasound screen* and the *needle and ultrasound probe*. Novices spent significantly less time than experts and senior residents fixating on the *ultrasound screen,* and significantly more time tracking the *needle and ultrasound probe*. Standard error bars are represented for each group and AOI.

Limitations

This study was performed at a large academic institution. While the participant numbers were limited, the current study provides some of the first evidence for the use of gaze training in IJCVC education. All interns at our institution attend surgical bootcamp, thus protected time opportunity to participate in the study was present. Experts and senior surgical resident participation occurred during clinical work-days. Thus, time was not protected for study participation, potentially leading to a self-selection bias as well for individuals with higher interest or experience with CVC procedure. While this study did not collect external video data to supplement eye-tracking data, future work could explore in depth the relationship between hand movements and gaze patterns during IJ CVC placement. Finally, data was excluded for two participants (1 novice, 1 expert) due to poor tracking quality. Future work should examine more participants at various stages of expertise.

Conclusions

The current study differentiated expert and novice fixation and AOI scan patterns during IJCVC placement on the DHRT system and showed that experts performed similarly between DHRT and manikin-based simulation systems. These findings indicate that the DHRT training system effectively captures gaze behaviors of expert surgeons as well as manikin-based systems. Additionally, the results showed that differences in gaze patterns could be discerned between varying levels of expertise.

Funding

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.

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.amjsurg.2018.11.007.

References

1. Evans LV, Dodge KL, Shah TD, et al. Simulation training in central venous catheter insertion: improved performance in clinical practice. *Acad Med*. 2010;85:1462–1469.
2. Smith CC, Huang GC, Newman LR, et al. Simulation training and its effect on long-term resident performance in central venous catheterization. *Simulat Healthc J Soc Med Simulat*. 2010;5:146–151.
3. McGee DC, Gould MK. Preventing complications of central venous catheterization. *N Engl J Med*. 2003;348:1123–1133.
4. 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. *Postgrad Med*. 2016;93(1096):67–70.
5. Doyle JD, Webber EM, Sidhu RS. A universal global rating scale for the evaluation of technical skills in the operating room. *Am J Surg*. 2007;193:551–555.
6. Huang GC, Newman LR, Schwartzstein RM, et al. Procedural competence in internal medicine residents: validity of a central venous catheter insertion assessment instrument. *Acad Med*. 2009;84:1127–1134.
7. Ogden PE, Cobbs LS, Howell MR, Sibbitt SJ, DiPette DJ. Clinical simulation: importance to the internal medicine educational mission. *Am J Med*. 2007;120:820–824.
8. Sherertz RJ, Ely EW, Westbrook DM, et al. Education of physicians-in-training can decrease the risk for vascular catheter infection. *Ann Intern Med*. 2000;132:641–648.
9. Pepley D, Yovanoff M, Mirkin K, Han D, Miller S, Moore J. Design of a virtual reality haptic robotic central venous catheterization training simulator. In: *International Design Engineering Technical Conferences and Computers and Information in Engineering Conference*. Charlotte, NC, USA: American Society of Mechanical Engineers; 2016. https://doi.org/10.1115/DETC2016-59560.
10. Pepley DF, Gordon AB, Yovanoff MA, et al. Training surgical residents with a haptic robotic central venous catheterization simulator. *J Surg Educ*. 2017;74:1066–1073.
11. 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. In: *Proceedings of the Human Factors and Ergonomics Society Annual Meeting*. Austin, TX: SAGE Publications Sage CA; 2017. Los Angeles, CA.
12. Yovanoff M, Pepley D, Mirkin K, Moore J, Han D, Miller S. Improving medical education: simulating changes in patient Anatomy using dynamic haptic feedback. In: *Proceedings of the Human Factors and Ergonomics Society Annual Meeting*. Washington D.C: SAGE Publications Sage CA; 2016:603–607. Los Angeles, CA.
13. Hall JC. Imagery practice and the development of surgical skills. *Am J Surg*. 2002;184:465–470.
14. Bathalon S, Dorion D, Darveau S, Martin M. Cognitive skills analysis, kinesiology, and mental imagery in the acquisition of surgical skills. *J Otolaryngol*. 2005;34:328–332.
15. Kohls-Gatzoulis JA, Regehr G, Hutchison C. Teaching cognitive skills improves learning in surgical skills courses: a blinded, prospective, randomized study. *Can J Surg*. 2004;47:277–283.
16. Chetwood AS, Kwok K-W, Sun L-W, et al. Collaborative eye tracking: a potential training tool in laparoscopic surgery. *Surg Endosc*. 2012;26:2003–2009.
17. Wood G, Knapp KM, Rock B, Cousens C, Roobottom C, Wilson MR. Visual expertise in detecting and diagnosing skeletal fractures. *Skeletal Radiol*. 2013;42:165–172.
18. Khan RS, Tien G, Atkins SM, Zheng B, Panton ON, Meneghetti AT. Analysis of eye gaze: do novice surgeons look at the same location as expert surgeons during a laparoscopic operation? *Surg Endosc*. 2012;26:3536–3540.
19. Tien T, Pucher PH, Sodergren MH, Sriskandarajah K, Yang G-Z, Darzi A. Differences in gaze behaviour of expert and junior surgeons performing open inguinal hernia repair. *Surg Endosc*. 2015;29:405–413.
20. Ashraf H, Sodergren MH, Merali N, Mylonas G, Singh H, Darzi A. Eye-tracking technology in medical education: a systematic review. *Med Teach*. 2018;40:62–69.
21. Tien T, Pucher PH, Sodergren MH, Sriskandarajah K, Yang G-Z, Darzi A. Eye tracking for skills assessment and training: a systematic review. *J Surg Res*. 2014;191:169–178.
22. Wilson MR, McGrath JS, Vine SJ, Brewer J, Defriend D, Masters RS. Psychomotor control in a virtual laparoscopic surgery training environment: gaze control parameters differentiate novices from experts. *Surg Endosc*. 2010;24:2458–2464.
23. Wilson MR, Vine SJ, Bright E, Masters RS, Defriend D, McGrath JS. Gaze training enhances laparoscopic technical skill acquisition and multi-tasking performance: a randomized, controlled study. *Surg Endosc*. 2011;25:3731–3739.
24. Vine SJ, Masters RS, McGrath JS, Bright E, Wilson MR. Cheating experience: guiding novices to adopt the gaze strategies of experts expedites the learning of technical laparoscopic skills. *Surgery*. 2012;152:32–40.
25. Ahmidi N, Hager GD, Ishii L, Fichtinger G, Gallia GL, Ishii M. Surgical task and skill classification from eye tracking and tool motion in minimally invasive surgery. In: *International Conference on Medical Image Computing and Computer-assisted Intervention*. Springer; 2010:295–302.
26. Ahmidi N, Ishii M, Fichtinger G, Gallia GL, Hager GD. An objective and automated method for assessing surgical skill in endoscopic sinus surgery using eye-tracking and tool-motion data. In: *International Forum of Allergy & Rhinology*. Wiley Online Library; 2012:507–515.
27. Kim I, Miller SR, Freivalds A. Motion analysis as an evaluation framework for eye-hand coordination: a case study in ultrasound-guided catheter insertion. In: *International Design Engineering Technical Conferences and Computers and Information in Engineering Conference*. Buffalo, NY, USA: American Society of Mechanical Engineers; 2014. https://doi.org/10.1115/DETC2014-34575.
28. Pugh D, Cavalcanti RB, Halman S, et al. Using the entrustable professional activities framework in the assessment of procedural skills. *J Grad Med Educ*. 2017;9:209–214.
29. Sznajder JI, Zveibil FR, Bitterman H, Weiner P, Bursztein S. Central vein catheterization: failure and complication rates by three percutaneous approaches. *Arch Intern Med*. 1986;146:259–261.
30. Bernard RW, Stahl WM. Subclavian vein catheterizations: a prospective study. I. Non-infectious complications. *Ann Surg*. 1971;173:184–190.
31. Tobii AB. *Tobii Pro Glasses 2*; 2018. Retrieved from: http://www.tobiipro.com/product-listing/tobii-pro-glasses-2/ (Accessed April 15, 2018).
32. Tobii AB. *Real-World Mapping*; 2018. Retrieved from: https://www.tobiipro.com/learn-and-support/learn/steps-in-an-eye-tracking-study/data/real-world-mapping/ (Accessed April 15, 2018).
33. Altman DG. *Practical Statistics for Medical Research*. Chapman & Hall; 1991.
34. Olsen A. The Tobii I-VT fixation filter: algorithm description. Retrieved from Tobii Technology from: http://www/tobiipro com/siteassets/tobiipro/learn-and-support/analyze/how-do-we-classify-eye-movements/tobii-pro-i-vtfixation-filter.pdf.2012.