SURGERY

Education

Isolating steps instead of learners: Use of deliberate practice and validity evidence in coronavirus disease (COVID)—era procedural assessment

Check for updates

Ingrid S. Schmiederer, MD[a,*], LaDonna E. Kearse, MD[a], Dana T. Lin, MD[a], Tiffany N. Anderson, MD, MHPE[b], James N. Lau, MD, MHPE, FACS[c], James R. Korndorffer Jr., MD, MHPE, FACS[a]

[a] Department of Surgery, Stanford University Medical Center, CA
[b] Department of Surgery, University of Florida, Gainesville, FL
[c] Department of Surgery, Loyola University Medical Center, Maywood, IL

ARTICLE INFO

ABSTRACT

Background: In surgical training, assessment tools based on strong validity evidence allow for standardized evaluation despite changing external circumstances. At a large academic institution, surgical interns undergo a multimodal curriculum for central line placement that uses a 31-item binary assessment at the start of each academic year. This study evaluated this practice within increased in-person learning restrictions. We hypothesized that external constraints would not affect resident performance nor assessment due to a robust curriculum and assessment checklist.

Methods: From 2018 to 2020, 81 residents completed central line training and assessment. In 2020, this curriculum was modified to conform to in-person restrictions and social distancing guidelines. Resident score reports were analyzed using multivariate analyses to compare performance, objective scoring parameters, and subjective assessments among "precoronavirus disease" years (2018 and 2019) and 2020.

Results: There were no significant differences in average scores or objective pass rates over 3 years. Significant differences between 2020 and precoronavirus disease years occurred in subjective pass rates and in first-time success for 4 checklist items: patient positioning, draping, sterile ultrasound probe cover placement, and needle positioning before venipuncture.

Conclusion: Modifications to procedural training within current restrictions did not adversely affect residents' overall performance. However, our data suggest that in 2020, expert trainers may not have ensured learner acquisition of automated procedural steps. Additionally, although 2020 raters could have been influenced by logistical barriers leading to more lenient grading, the assessment tool ensured training and assessment integrity.

© 2021 Elsevier Inc. All rights reserved.

Introduction

The need for procedural training to minimize complications associated with central venous catheter placement has been well-established.[1—3] Structured training programs focused on catheter

placement, and established bundles or algorithms have contributed to decreased incidence of central line-associated bloodstream infections and complications related to central venous catheter insertion.[4—6] To ensure patient safety, novice learners must be guided toward expertise in mastering the procedure itself, as well as in safe behaviors that include aseptic technique, ultrasound guidance, use of an assistant, and a clear checklist.[7—10] As such, successful training programs must design multimodal curricula that allow surgical residents the physical space, controlled environment, effective guidance, and the time to practice these procedural steps before performing procedures on patients.[11,12]

At Stanford University Medical Center, early first-year general surgery trainees undergo a multimodal, simulation-based training curriculum every year for central venous catheter placement. Verification of proficiency assessment is subsequently conducted with a 31-item binary checklist that has demonstrated strong validity evidence, and it occurs before the resident is allowed to perform the procedure in the clinical setting. At the start of the 2020 academic year, when the most recent cohort of surgical interns required central line training, institutional regulations regarding in-person training and safety measures due to the coronavirus disease (COVID) 2019 pandemic required alterations to training and assessment logistics. Learner group size was minimized to accommodate social distancing. Time for training and assessment was shortened to allow for all groups to undergo the curriculum within the same timeframe.

This study examines the stability of the central line training and its assessment tool results, despite changing external circumstances around resident procedural learning and performance. We hypothesize that despite increased logistical restrictions, our robust assessment tool allows for reliable, consistent assessment of technical skill, and it serves as an anchor for effective guidance and deliberate practice.

Methods

Participants

Trainees

Surgical residents in the Stanford University General Surgery Residency Program are required to undergo a simulation-based training curriculum and subsequent verification of proficiency assessment in central line placement during July and August of their internship years.

Trainers

Trainers include surgical faculty in the Department of Surgery, surgical critical care fellows, and surgical education fellows in the American College of Surgeons Accredited Education Institutes (ACS-AEI) program at Stanford University.

Raters

In 2018, surgical attendings and the ACS-AEI surgical education fellows (general surgery residents) served as raters for the verification of proficiency testing in its most current iteration of the program. In 2019, additional raters were trained to expand the assessment team. By 2020, raters included surgical attendings, surgical education fellows, midlevel residents, as well as nonmedical trained raters. Owing to constraints in 2020 regarding visitors and staff within the hospital setting, only 1 nonmedical rater was involved in rating, while the rest were physician raters (faculty and residents).

Rater training. All physician raters reviewed the checklist to confirm understanding of dichotomous steps to be evaluated, as all had extensive clinical experience in central line training and placement in order to be included in the assessment session.

In 2019, with the addition of nonmedical personnel within the surgical education office, nonmedical raters were trained using the proven method of Frame of Reference Training[13,14] in order to calibrate novice raters such that all raters were familiarized with standards and were able to conceptualize what an effective performance of the procedure looked like, using the checklist as a guide. Raters who were unfamiliar with performing central line catheterization themselves were guided through procedural steps by physician raters and offered repetitive practice and

demonstration. Questions were answered regarding the binary checklist. Finally, novice raters were given the opportunity to rate sample performances by the physician raters with feedback provided. This extensive rater training occurred during the 2019 assessment period.

Resident training and practice

Before 2020, first-year residents would attend a 2-hour session of training in central line placement focused on internal jugular vein catheterization. This training session included faculty demonstration and walk-through of aseptic technique, ultrasound use, and procedural steps. A Simulab CentraLineMan (Simulab, Seattle, WA) high-fidelity torso model was used during this training session to demonstrate anatomic landmarks of internal jugular and subclavian access with ultrasound guidance and catheterization. This previously occurred in a large group setting (~40 interns), after which interns were dispersed to different stations composed of about 10 learners for hands-on practice and guided instruction. A trainer supervised each station and provided individualized guidance and feedback to the interns. Each intern had the opportunity to complete the task in its entirety at least once.

In 2020, learner groups were restricted to 6 learners and 1 trainer in 1 room. Participants were spaced 6 feet apart from one another and the trainer. After demonstration and overview by the trainer, learners approached the simulation station 1 at a time for guided practice. Trainers generally stood 6 feet away, unless hands-on correction or instruction was required. After each use, the station was sanitized and reset for the next participant. Total time for each session was shortened to 1 hour, given the increased number of learner groups and added time required for sanitization. During 2020 preparations, trainers were prebriefed about time restrictions and were instructed to specifically focus on providing guided hands-on instruction for "the procedure itself," with emphasis on ultrasound guidance, Seldinger technique, and guidewire control. Other aspects and instruction of the procedure were at the discretion of the trainer in the room. Trainers demonstrated all items with varied emphasis, which ultimately led to more guided discovery learning of many checklist items.

Four weeks elapsed between the training sessions and the testing sessions. In that time, interns were traditionally given 24-hour access to the central line placement models for additional practice before testing. In 2020, access to these models was limited to appointment only, to ensure that accommodations were prepared with proper precautions and remote operations staff were available onsite. Appointments for guided deliberate practice with the surgical education fellows were available all 3 years.

Assessment

Validity evidence of checklist

This checklist assessment tool (Supplementary Table E1) was modified based on an Objective Structured Skills Assessment tool provided by the American College of Surgeons Resident Skills Curriculum and was evaluated in 2018 according to Messick's framework.[15,16] Our previous evaluation found this checklist to possess robust validity evidence.

- Content: Each task was reviewed, and 31 items were ultimately customized based on institutional expert consensus.
- Response process: Experts also determined that a binary checklist of "successful completion on first try" versus "unsuccessful on first try" was most feasible for scoring. Each of the dichotomous actions required by the checklist were based on expert performance, and all were deemed by experts to be

required and equally weighted for adequacy of placement of central lines during the initial curriculum development process.

- Internal structure: This tool has good interrater reliability (Cohen's kappa = 0.59) and good internal consistency (Cronbach's alpha = 0.84).
- Relationships to other variables: Raters in the last 3 years, with the most current assessment tool and scoring standards, have been required to follow the guidelines for scoring (objective score). To enhance frame of reference training and correlate scores with multiple measures, raters also circled "pass" (subjective score) at the end of the assessment.
- Consequences: Using the Angoff method, the assessment committee set a standard for the cut score (28/31, 90%) with tasks that were also designated as "critical fails" to emphasize those tasks for learner acquisition and patient safety. Tasks designated as critical fails were contamination of self or field and loss of guidewire control, as a means of emphasizing patient safety. Passing the assessment enabled the intern to perform central line placement under supervision in the clinical setting, while failure required remediation and retesting.

Objective scoring

Residents were required to complete all tasks for an internal jugular central venous catheterization on a simulated patient model, using ultrasound guidance and a central venous catheter kit. Upon arrival to the assessment area, the resident was familiarized with the simulated patient: paper chart, consent form and wristband, and a hybrid chicken tissue model[17] nested within a simulated torso base on a table. The resident executed the entire task, beginning with patient identification, chart review, and team introduction and then addressing skin prep, draping, equipment preparation, catheterization, suturing the line in place, sterile dressing, and finally, indication of a chest x-ray to confirm placement. The rater acted as a passive assistant for the resident if requested, while the rater also observed and rated the resident's performance using a 31-item checklist.

Subjective scoring

Raters deemed a task successful when the resident completed it on the first try in any order. If raters felt that the resident passed in a holistic, subjective assessment of performance (not based on the objective score), they circled "pass" at the end of the assessment checklist. Raters did not provide subjective scores during 2019, as this was the year in which raters were actively trained in objective assessment. Subjective scores from 2019 were not included in analysis.

Data analysis

After institutional review board approval, resident evaluations were retrospectively analyzed over the last 3 years (2018–2020) using the Kruskal-Wallis test, an independent t test, and χ^2 analysis to evaluate for changes in means, pass rates, or overall standards due to raters or to COVID-related regulations. Analyses were completed using IBM SPSS Statistics for Windows, version 27.0 (IBM Corp, Armonk, NY).

Results

Trainees

From 2018 to 2020, a total of 81 residents from general surgery and surgical subspecialties completed central line placement training and assessment: 27, 33, and 21 interns in 2018, 2019, and

Table I
Objective total assessment scores between years (Kruskal-Wallis test)

	N	Mean rank	
2018	27	41.96	
2019	33	42.44	P value = .72
2020	21	37.50	

2020, respectively. While there was opportunity for additional practice all 3 years, less than 10% of residents documented any additional simulated practice time. This extra practice session involved making an appointment with a surgical education fellow for 1-on-1 guided instruction. Given that reports were deidentified and that there were few appointments made, it was not possible to determine if these extra practice sessions were significantly advantageous for the trainee between the training session and the assessment period.

Raters

With the addition of multiple raters in 2019 and 2020, Fleiss' multirater kappa was calculated to account for the nonmedical trained raters and the physician raters at different training levels (faculty versus resident). Overall, this checklist and its use over 3 years with multiple raters demonstrated good strength of agreement (Fleiss' multirater kappa [(κ) = 0.64]).

Objective scoring

There were no significant differences in average raw total scores over the last 3 years (Table I) after analysis for right-skewed data nor were there statistically significant differences between 2020 trainees and pre-COVID trainees' scores (28.14 ± 4.07 vs 28.66 ± 3.25; P = .56). Objective pass rates using the established cut score and critical fail parameters also showed no significant difference between 2020 trainees and their pre-COVID counterparts (71.4% vs 73.3%; P = .87). Furthermore, objective pass rates for the specific years in which raters also gave subjective assessments (2018 and 2020) showed no significant difference (71.4% vs 74.1%, respectively; P = 0.84).

Upon examination of each task within the checklist, there was no significant statistical difference in critical failure items between 2020 and pre-COVID years (14.3% vs 20%; P = .562). However, 2020 trainees did have significantly lower first-time success rates than their 2018 and 2019 counterparts in certain skills: patient positioning, draping the patient, sterile ultrasound probe cover placement, and finder needle positioning before venous access (Table II). Most notably, only 71.4% of 2020 trainees correctly positioned the patient compared to 95.0% of pre-COVID trainees (P < .05), and only 71.4% of 2020 trainees successfully draped the patient on the first attempt compared to 90.0% of pre-COVID trainees (P < .05). A total of 76.2% placed the sterile probe cover correctly compared to 93.3% previously (P < .05). Finder needle positioning was successful on the first try for 85.7% of 2020 trainees compared to 98.3% of pre-COVID trainees (P < .05).

Subjective scoring

Significant differences were observed when evaluating subjective pass rates. The subjective pass rate of 2020 was 95.2% among physician raters and 1 trained rater, compared to 70.4% in 2018 among physician raters (P < .05). In fact, when comparing only expert physician raters who received the same rater preparation in 2018 and in 2020, the subjective pass rate of 2018 was 70.4%, and

Table II
Rate of successful first-time completion of 4 tasks

Successful first-time completion of task	Pre-COVID (2018, 2019)	2020	P value
Trendelenburg positioning	95.0%	71.4%	.003[*]
Draping patient	90.0%	71.4%	.039[*]
Sterile ultrasound probe cover placement	93.3%	76.2%	.031[*]
Finder needle positioning	98.3%	85.7%	.022[*]

COVID, coronavirus disease.
[*] $P < .05$.

the 2020 pass rate was 100% ($P \leq 0.05$). All physician raters of 2020 were more inclined than their 2018 counterparts to pass trainees based on subjective assessment.

Discussion

This study finds that modifications to simulation-based procedural training within current in-person restrictions did not adversely affect residents' overall performance and checklist scores. Simulation-based training methodologies in controlled, low-stakes environments have already been determined to be superior for learning.[18,19] In these controlled learning environments, trainers can identify specific steps of the procedure itself. A list of attainable and tangible steps, some of which experts may subconsciously perform in their own execution of the procedure, further ensures learner skill acquisition and deliberate practice for both the trainer and the trainee.[11,19]

The theories of deliberate practice suggest that it is more than the number of hours of practice but also the quality of practice that supports the development of expertise.[20] In a learning program that emphasizes a deliberate practice process, training is guided by faculty who are able to identify aspects of technical performance and provide specific guidance for improvement.[12] While this may not fully require Ericsson's 10,000 hours for expertise, there should be enough dedicated time and effective training for each task within these learning sessions. Training that is rushed, without full focus on improvement, will hinder potential improvement.[12] This established curriculum is meant to provide time and space within the first year of surgical residency for these trainees. Our study also evaluated the curriculum's effectiveness in facilitating this deliberate practice, as well as deliberate training of the procedure despite changing circumstances.

McGaghie et al expanded upon deliberate practice in medical education with a list of 9 requirements in 2009.[18] To further enhance and to evaluate our program's training, our curriculum development followed this list of requirements as a framework to ensure deliberate practice was implemented correctly. (1) "Highly motivated learners with good concentration"[18]: The caliber of these highly motivated learners has remained consistent since this curriculum was implemented. Within new constraints, one might argue that learner concentration further increased, because learning groups decreased in size, and there was potential for (2) increased "engagement among learners with a well-defined learning objective or task."[18] Learning objectives for both trainees and raters were described and explained, and tasks were clearly listed in the assessment tool for faculty, for raters, and for learners to remain engaged. Objective assessment scores of these residents in the last 3 years also remained consistent, which demonstrated effective motivation and clear objectives.

In the early development phase, the training program curriculum committee determined that this simulation-based central line training program was (3) at "an appropriate level of difficulty" and necessary for these incoming surgical residents. As such,

preplanned simulation-based training sessions were designed for (4) "focused, repetitive practice." Rather unexpectedly, this element of our curriculum design was altered due to COVID-related restrictions. Therefore, to safeguard focused and repetitive practice, trainers were encouraged to focus on the guided instruction of "high-yield" aspects of the procedure, such as ultrasound guidance, Seldinger technique, and guidewire control. This modification to the training was supported by the assessment tool, which allowed (5) "rigorous, precise measurements," as the procedural steps were broken down to a granular level that allowed for focused practice, standardized reproducibility, and precise measurement from trainers and raters. Using a checklist for both assessment and training ensured that all steps were detailed and addressed by both trainer and trainee. Counteracting automaticity with a clear checklist of procedural steps allowed for higher control of performance and guided instruction, as expert trainers and novice learners were assisted in the cognitive phase of deliberate practice.[21]

Simulation-based training sessions with guided instruction and the addition of subjective scoring encouraged (6) "informative feedback from educational sources (simulators or teachers)." In studies of assessment, it has been demonstrated that the juxtaposition of subjective and objective assessment together serves to optimize feedback and allows both the rater and the learner to manage the complex information necessary to improve and judge performance reliably.[22] In other words, a rater may deem an assessment subjectively failed for a resident who objectively met all objective measures for passing (or vice versa). This prompts a conversation regarding the breakdown of steps, the knowledge of the procedure, and the flow of motion, among other sources of assessment and validity. A subjective assessment uses expert detection of error[23] and also ensures that the learner is not simply relying on rote memorization or reading of a checklist. Comparing the physician raters of 2018 with the physician raters of 2020 also allows for more direct evaluation of subjective assessment by expert raters, as we found the trained raters contributed to a lower subjective pass rate in 2020. This finding warrants further evaluation into training experts, not only to enhance teaching proficiency, but for standardization of raters. Conversely, further training for nonmedical raters may also be needed to give more confidence in subjective assessment, rather than being influenced by the objective measurements—a likely byproduct of this subjective-objective juxtaposition for novice raters.

Scheduled individual assessment was designed (7) to "monitor, correct errors, and facilitate more deliberate practice." Further study might follow learners beyond the simulation environment into the wards for (8) more complete "evaluation to reach a mastery standard." However, passing assessment scores determined by our Angoff method process theoretically ensures that learners are ready to advance to more independent practice in the clinical setting with safeguards against changing subjective standards or borderline rater opinions. In its current iteration, this curriculum prepares trainees for (9) "advancement to another task or unit" in a safe, supervised method. After completion of training and assessment,

Table III
Rate of successful first-time completion of all tasks

Task	Pre-COVID (2018, 2019)	2020	P value (Pearson χ^2)
Chart reviewed: INR, platelets checked	96.7%	100.0%	.397
Consent verified	96.7%	100.0%	.397
Allergies confirmed	88.3%	90.5%	.788
ECG and SpO2 verified	96.7%	95.2%	.765
Team introduction	88.3%	100.0%	.102
Verify patient with 2 identifiers	98.3%	100.0%	.552
Hands washed	90.0%	95.2%	.462
Vein and artery identified with US (nonsterile)	95.0%	95.2%	.965
Mask, eye protection, hair cover	100.0%	100.0%	-
Adequate sterile prep with chlorhexidine (wait 3 min)	100.0%	100.0%	-
Patient placed in Trendelenburg[*]	**95.0%**	**71.4%**	**.003**
Gowned and gloved without breaking sterility	90.0%	95.2%	.462
Wide drape placed correctly[*]	**90.0%**	**71.4%**	**.039**
Ultrasound probe sheath applied correctly[*]	**93.3%**	**76.2%**	**.031**
All equipment verified, prepared, and placed (CVC kit)	93.3%	95.2%	.755
Catheter flushed and capped	88.3%	76.2%	.178
Location of IJ confirmed with US (sterile)	98.3%	95.2%	.431
Needle placed at proper location, aiming at the ipsilateral nipple[*]	**98.3%**	**85.7%**	**.022**
Vein accessed: no more than 3 needle passes	96.7%	90.5%	.260
Venous blood return verified	98.3%	95.2%	.431
Wire advanced 12–17 cm	86.7%	95.2%	.282
Wire visualized with US in longitudinal view	80.0%	61.9%	.098
Skin nick made with 11 blade	91.7%	95.2%	.591
Dilator deployed and removed properly (2–3 cm, no kinking of wire)	91.7%	81.0%	.179
Catheter advanced over wire, 12–17cm	93.3%	90.5%	.667
Control of wire maintained at all times	85.0%	85.7%	.937
Wire removed and stored without breaking sterility	93.3%	90.5%	.667
Catheter checked for blood return and flushed (all ports)	85.0%	90.5%	.528
Catheter secured in place with suture	85.0%	90.5%	.528
Sterile dressing applied	86.7%	85.7%	.913
Chest x-ray of correct placement confirmed	91.7%	95.2%	.591

Bolded values denote significant differences between groups in the given item.
COVID, coronavirus disease; *CVC,* central venous catheterization; *ECG,* electrocardiogram; *IJ,* internal jugular vein; *INR,* international normalized ratio; *US,* ultrasound.
  [*] Significant statistical difference ($P < .05$).

trainees receive a "passport" certificate that must be filled out by senior residents or faculty who monitor trainees placing central lines in the clinical environment before they are cleared to practice independently.

Within the curriculum, this study also analyzed each item of the assessment tool independently to measure if procedural steps were being monitored, corrected, and deliberately practiced, as they were designed to be. The decreased rate of successful completion of certain steps indicates that modifications may have negatively affected training by the trainers in the informative feedback and close monitoring of each task. Perhaps guidance was hindered by the modifications of 2020. The trainers' execution of each step for learner acquisition was more time restricted and physically distanced. Trainers may have felt constrained and deemphasized certain steps in order to focus on what they deemed to be more crucial procedural steps. One such crucial item in safe training and objective assessment was the critical fail element. Emphasizing guidewire control and sterile technical skill allowed learners to avoid the critical fails and may have compensated for the decreased performance in other steps. Overall pass rates remained consistent across 3 years, despite the decreased first-time success rate in multiple steps for 2020 trainees (Table III).

Outside of time constraints, the differences in successful completion of certain steps may also be explained by expert automaticity and modified guidance approaches. The expert trainers in 2020 may have demonstrated all steps of the checklist within the given timeframe like their previous counterparts, but their guided instruction focused more specifically on ultrasound-guided Seldinger technique steps, due to the abbreviated

sessions and social distancing. Perhaps due to constraints, the shift to discovery learning and observation was not effective for novice learners who did not know where to focus their attention during procedural demonstration or make sense of what was happening in their observations.[24] Persky et al explained that expertise might actually hurt learner acquisition when experts forget how difficult tasks were to learn. Efficient experts who are not proficient in teaching are not well-equipped to articulate their full thinking process or provide enough detailed description during demonstration, as the thinking and execution process has become too intuitive.[11,25] These limitations could have been mitigated by additional training for the expert trainers, within the curricular modifications of 2020.

Ultimately, checklists serve an important formative function during the development of trainee skills and provide an anchor for deliberate practice requirements. Composition of the list not only frames learning, but also conveys the program's priorities and objectives for assessment.[26] Using a checklist ensures that all steps are defined for both assessment and training with focused practice, close monitoring, and evaluation for mastery, in accordance McGaghie et al's requirements.[18] To inform feedback, the visible, tangible checklist of procedural steps counteracts automaticity and allows for higher control of performance and guided instruction. Expert trainers and novice learners are assisted in the cognitive phase of deliberate practice and correction of errors might be better facilitated.[21] Future studies might explore rearrangement of procedural training into multiple stations of fewer tasks for the deliberate practice repetition of each step rather than abbreviated focused training of all steps. This would condense time, prevent trainer cognitive overload, and allow

time for the learner to repeatedly execute all aspects of the procedure with close guidance and immediate feedback. This would also address gradual, monitored improvements on a granular level for each step[27] rather than the abbreviated, observational overview of all steps at once, arguably found within the 2020 training limitations.

A limitation to this study is that the utility of the checklist for these trainees or the consequence of their assessment in the real-world clinical setting (Messick's consequences validity evidence[15]: or McGaghie et al's mastery standard requirement[18]) were not fully taken into account. Correlating patient outcomes for this cohort of trainees who have undergone this training and assessment program would further strengthen validity evidence in consequences of this assessment and its relationships to other variables. In the context of consequences validity evidence and in evaluation to reach mastery for deliberate practice, one could argue that a ceiling effect occurs as many of these items are arguably easy to attain. However, if the goal is 100% successful completion of each item at first try, this data demonstrates that this has not been fully achieved by residents, and there is still room for improvement. Additionally, hard copies and electronic copies of the checklist are provided for testing purposes, but it is not tracked whether trainees use them when placing central lines on the wards. Previous studies agree that checklist algorithms and protocols enhance patient safety outcomes in the clinical and acute care setting[7,28] and that strong assessment tools enhance technical training. Still, the validity evidence is ongoing for many assessments and requires continued evaluation for optimization.[22,29]

In conclusion, the results of this study are 2-fold as trainee caliber, rater agreement, and total scores have remained consistent across multiple years and circumstances. The 2020 raters may have been influenced by constraints and had the inclination to pass learners more easily, knowing that (1) training was altered, and learner acquisition was potentially hindered or (2) that there was limited retesting time if remediation was required. In a brief comparison of pass rates based on the subjective assessment of physician raters, ~25% more trainees would have passed to the clinical supervision stage, potentially before they were ready. This assessment tool proved reliable, and the integrity of assessment was maintained when objective measures, such as cut score and critical fail guidelines, were followed.

This study suggests that despite potential external influences on training and assessment, sound education theory that includes the use of assessment tools with strong validity evidence, remains critical for training, deliberate practice, and assessment. Differences in successful completion of certain steps may have identified gaps in the execution of training curricula by the trainers themselves. Future studies might explore modification of procedural training approaches for the trainer, as well as the impact of these curricular interventions on procedural outcomes.

Conflict of interest/Disclosures

The authors have no related conflicts of interest to declare.

Supplementary Materials

Supplementary material associated with this article can be found, in the online version, at https://doi.org/10.1016/j.surg.2021.06.010.

References

1. Barsuk JH, Cohen ER, Potts S, et al. Dissemination of a simulation-based mastery learning intervention reduces central line-associated bloodstream infections. *BMJ Qual Saf.* 2014;23:749–756.
2. Criss CN, Gadepalli SK, Matusko N, Jarboe MD. Ultrasound guidance improves safety and efficiency of central line placements. *J Pediatr Surg.* 2019;54: 1675–1679.
3. Guembe M, Pérez-Parra A, Gómez E, et al, and the GEIDI Study Group. Impact on knowledge and practice of an intervention to control catheter infection in the ICU. *Eur J Clin Microbiol Infect Dis.* 2012;31:2799–2808.
4. Moureau N, Lamperti M, Kelly LJ, et al. Evidence-based consensus on the insertion of central venous access devices: definition of minimal requirements for training. *Br J Anaesth.* 2013;110:347–356.
5. Alexandrou E, Spencer TR, Frost SA, Mifflin N, Davidson PM, Hillman KM. Central venous catheter placement by advanced practice nurses demonstrates low procedural complication and infection rates—A report from 13 years of service.*. *Crit Care Med.* 2014;42:536–543.
6. Alanazi TNM, Alharbi KAS, Alrawaili ABR, Arishi AAM. Preventive strategies for the reduction of central line-associated bloodstream infections in adult intensive care units: A systematic review. *Collegian*; 2020. Available from: https://doi.org/10.1016/j.colegn.2020.12.001. Accessed March 1, 2021.
7. Chua C, Wisniewski T, Ramos A, Schlepp M, Fildes JJ, Kuhls DA. Multidisciplinary trauma intensive care unit checklist: impact on infection rates. *J Trauma Nurs.* 2010;17:163–166.
8. Timsit JF, Baleine J, Bernard L, et al. Expert consensus-based clinical practice guidelines management of intravascular catheters in the intensive care unit. *Ann Intensive Care.* 2020;10:118.
9. Practice guidelines for central venous access 2020: An updated report by the American Society of Anesthesiologists Task Force on Central Venous Access. *Anesthesiology.* 2020;132:8–43.
10. Schulman J, Stricof R, Stevens TP, et al, and the New York State Regional Perinatal Care Centers. Statewide NICU central-line-associated bloodstream infection rates decline after bundles and checklists; 2011. https://pediatrics.aappublications.org/content/127/3/436.long. Accessed March 3, 2021.
11. Persky AM, Robinson JD. Moving from novice to expertise and its implications for instruction. *Am J Pharm Educ.* 2017;81:6065.
12. Ericsson KA. Deliberate practice and acquisition of expert performance: A general overview. *Acad Emerg Med.* 2008;15:988–994.
13. Uggerslev KL, Sulsky LM. Using frame-of-reference training to understand the implications of rater idiosyncrasy for rating accuracy. *J Appl Psychol.* 2008;93: 711–719.
14. Newman LR, Brodsky D, Jones RN, Schwartzstein RM, Atkins KM, Roberts DH. Frame-of-reference training: Establishing reliable assessment of teaching effectiveness. *J Contin Educ Health Prof.* 2016;36:206–210.
15. American Educational Research Association, American Psychological Association, National Council on Measurement in Education, Joint Committee on Standards for Educational and Psychological Testing. *Standards for Educational and Psychological Testing.* Washington, DC: American Educational Research Association; 2014.
16. Ghaderi I, Manji F, Park YS, et al. Technical skills assessment toolbox: A review using the unitary framework of validity. *Ann Surg.* 2015;261: 251–262.
17. Nachshon A, Mitchell JD, Mueller A, Banner-Goodspeed VM, McSparron JI. Expert evaluation of a chicken tissue-based model for teaching ultrasound-guided central venous catheter insertion. *J Educ Perioper Med.* 2017;19: E503.
18. McGaghie WC, Siddall VJ, Mazmanian PE, Myers J, the American College of Chest Physicians Health and Science Policy Committee. Lessons for continuing medical education from simulation research in undergraduate and graduate medical education: effectiveness of continuing medical education: American College of Chest Physicians Evidence-Based Educational Guidelines. *Chest.* 2009;135(3Suppl):62S–68S.
19. McGaghie WC, Issenberg SB, Cohen ER, Barsuk JH, Wayne DB. Does simulation-based medical education with deliberate practice yield better results than traditional clinical education? A meta-analytic comparative review of the evidence. *Acad Med.* 2011;86:706–711.
20. Kuhlmann DO, Ardichvili A. Becoming an expert: developing expertise in an applied discipline. *Eur J Train Dev.* 2015;39:272–276.

21. Khalil MK, Elkhider IA. Applying learning theories and instructional design models for effective instruction. *Adv Physiol Educ*. 2016;40:147–156.
22. Pasquier M, Chéron C, Barbier G, Dugas C, Lardon A, Descarreaux M. Learning spinal manipulation: Objective and subjective assessment of performance. *J Manipulative Physiol Ther*. 2020;43:189–196.
23. Descarreaux M, Dugas C. Learning spinal manipulation skills: Assessment of biomechanical parameters in a 5-year longitudinal study. *J Manipulative Physiol Ther*. 2010;33:226–230.
24. Raîche I, Hamstra S, Gofton W, Balaa F, Dionne É. Cognitive challenges of junior residents attempting to learn surgical skills by observing procedures. *Am J Surg*. 2019;218:430–435.
25. Carraccio CL, Benson BJ, Nixon LJ, Derstine PL. From the educational bench to the clinical bedside: Translating the Dreyfus developmental model to the learning of clinical skills. *Acad Med*. 2008;83:761–767.
26. Yudkowsky R, Park YS, Riddle J, Palladino C, Bordage G. Clinically discriminating checklists versus thoroughness checklists: Improving the validity of performance test scores. *Acad Med*. 2014;89:1057–1062.
27. Ericsson KA. The differential influence of experience, practice, and deliberate practice on the development of superior individual performance of experts. In: Ericsson KA, Hoffman RR, Kozbelt A, Williams AM, eds. *Cambridge Handbook of Expertise and Expert Performance*. Second ed. Cambridge: Cambridge University Press; 2018:745–769.
28. Pageler NM, Longhurst CA, Wood M, et al. Use of electronic medical record-enhanced checklist and electronic dashboard to decrease CLABSIs. *Pediatrics*. 2014;133:e738–e746.
29. Vaidya A, Aydin A, Ridgley J, Raison N, Dasgupta P, Ahmed K. Current status of technical skills assessment tools in surgery: A systematic review. *J Surg Res*. 2020;246:342–378.