A Consensus-Based Framework for Design, Validation, and Implementation of Simulation-Based Training Curricula in Surgery

Boris Zevin, MD, Jeffrey S Levy, MD, FACS, Richard M Satava, MD, FACS, Teodor P Grantcharov, MD, PhD, FACS

BACKGROUND: Simulation-based training can improve technical and nontechnical skills in surgery. To date, there is no consensus on the principles for design, validation, and implementation of a simulation-based surgical training curriculum. The aim of this study was to define such principles and formulate them into an interoperable framework using international expert consensus based on the Delphi method.

METHODS: Literature was reviewed, 4 international experts were queried, and consensus conference of national and international members of surgical societies was held to identify the items for the Delphi survey. Forty-five international experts in surgical education were invited to complete the online survey by ranking each item on a Likert scale from 1 to 5. Consensus was predefined as Cronbach's $\alpha \geq 0.80$. Items that 80% of experts ranked as ≥ 4 were included in the final framework.

RESULTS: Twenty-four international experts with training in general surgery (n = 11), orthopaedic surgery (n = 2), obstetrics and gynecology (n = 3), urology (n = 1), plastic surgery (n = 1), pediatric surgery (n = 1), otolaryngology (n = 1), vascular surgery (n = 1), military (n = 1), and doctorate-level educators (n = 2) completed the iterative online Delphi survey. Consensus among participants was achieved after one round of the survey (Cronbach's $\alpha = 0.91$). The final framework included predevelopment analysis; cognitive, psychomotor, and team-based training; curriculum validation evaluation and improvement; and maintenance of training.

CONCLUSIONS: The Delphi methodology allowed for determination of international expert consensus on the principles for design, validation, and implementation of a simulation-based surgical training curriculum. These principles were formulated into a framework that can be used internationally across surgical specialties as a step-by-step guide for the development and validation of future simulation-based training curricula. (J Am Coll Surg 2012;215:580–586. © 2012 by the American College of Surgeons)

The number of research studies addressing the role of simulation in surgical education has grown dramatically during the last decade. Multiple authors have reported on the beneficial effects of simulation-based training for technical[1-4] and nontechnical[5-7] skills. Two systematic reviews addressing this topic have also been published.[8,9] However, despite evidence from randomized controlled trials for the use of simulation in surgical education,[10-12] most surgical training programs continue to struggle with the integration of structured, simulation-based training into their curricula.[13,14] These difficulties are, in part, a result of the multitude of opinions on what should be included in such simulation-based curricula, how they should be validated, and how they should be integrated into the residency training programs. The result has been a multitude of different curricula (frequently on the same procedures) being developed with competing, contradictory, and duplicative approaches. The Association of Program Directors in Surgery and the American College of Surgeons have jointly tackled this issue by developing the *American College of Surgeons/*

Disclosure Information: Dr Satava receives a consulting fee from Karl Storz Endoscopy America, and a grant from Lockheed Martin. Boris Zevin is supported by the Canadian Institutes of Health Research Frederick Banting and Charles Best Canada Graduate Scholarship. All other authors have nothing to disclose.

Received April 10, 2012; Revised May 18, 2012; Accepted May 18, 2012.

From the Department of Surgery, University of Toronto, Toronto, Canada (Zevin, Grantcharov), Department of Obstetrics and Gynecology, Drexel University, Philadelphia, PA (Levy), and Department of Surgery, University of Washington, Seattle, WA (Satava).
Correspondence address: Boris Zevin, MD, St Michael's Hospital, 30 Bond St, 16CC-056 Toronto, Ontario, Canada M5B 1W8. email: boris.zevin@utoronto.ca

Published by Elsevier Inc.

ISSN 1072-7515/12/$36.00
http://dx.doi.org/10.1016/j.jamcollsurg.2012.05.035

Association of Program Directors in Surgery Surgical Skills Curriculum for Residents, which includes 20 modules that address basic surgical skills, 15 modules that address advanced surgical skills and procedures, and 10 modules that address team-based skills.[15]

Several authors[16-18] have emphasized the importance of cognitive teaching in a comprehensive surgical skills training curriculum. Such teaching should focus on the tasks required for correct execution of a particular procedure, including error detection, forward planning, and decision making. Tang and colleagues[17] demonstrated that the majority of errors performed by trainees during simulated training were not caused by a technical mistake, but rather by a knowledge gap in understanding of the correct sequence of steps in the particular task. A nonrandomized study of the impact of cognitive training on task execution emphasized that cognitive training not only improved understanding of a particular task, but also improved its execution.[18]

A comprehensive simulation-based training curriculum should incorporate evidence-based methodological principles, such as deliberate practice, distributed practice schedule, and proficiency-based training. Participation in such a curriculum should also be made mandatory.[19,20] Ericsson has studied the concept of deliberate practice and its impact on the acquisition of expert performance.[21,22] He suggested the following 5 conditions that are required for deliberate practice: a well-defined task to practice; provision of detailed and immediate feedback; motivation to improve; variable level of difficulty for the chosen task; and ample opportunity for repetition and gradual refinement of performance. Moulton and colleagues[23] studied massed vs distributed practice in a randomized controlled trial of 38 surgical residents learning a new skill of microvascular anastomosis. Residents who practiced in a distributed pattern exhibited superior retention of technical skills. Proficiency-based training has been shown to result in an improved operative performance[3,10] and reduced error rate in the operating room.[1]

Aggarwal and colleagues[24] and Stefanidis and Heniford[14] have tried to incorporate some of the methodological principles mentioned previously into frameworks for the design of simulation-based training curricula in surgery. Aggarwal and colleagues' 5-step framework includes a stepwise progression from knowledge-based learning, to deconstruction of the procedure into its component tasks, to training in a laboratory environment, to demonstrating transfer of skills to the real environment, and to eventually granting privileges for independent practice. Stefanidis and Heniford's[14] framework takes a trainee from baseline technical skill assessment to demonstration of procedure using video tutorials, to deliberate practice on a simulator in distributed sessions with feedback, to attainment of proficiency with overtraining, maintenance of training and post-training assessment. Both frameworks are similar in their recommendations for combining cognitive learning with training to proficiency on simulated models using deliberate practice and a distributed practice schedule.

The purpose of this study was to advance the field of surgical education and curricular development by developing a standardized, prescriptive, and internationally relevant framework for the design, validation, and implementation of a simulation-based surgical training curriculum. It used current evidence-based methodological principles for simulation-based training, international expert consensus, and a Delphi method.

METHODS

Study design

Classical and modified Delphi methods were used to achieve consensus among a panel of international experts in surgical simulation on the essential components of a standardized and comprehensive framework for the design, validation, and implementation of a simulation-based surgical training curriculum. The classical Delphi method, developed by the RAND Corporation in 1948, provides a means for obtaining opinions of experts in a systematic manner.[25-27] It is an anonymous process where ideas are presented to participants in the form of a questionnaire.[28] The responses are collated, analyzed, and presented back to the participants in an iterative fashion until consensus is achieved.[28] The modified Delphi method has the same basis; however, it is conducted in a planned consensus conference with participation of subject matter experts, using the same iterative methodology.

Participant selection

Participants were identified as opinion leaders in the field of surgical education as evidenced by their key roles within North American, South American, European, Asian, and Australasian surgical societies and organizations (Table 1). They were all members of the Alliance for Surgical Simulation Education and Training (ASSET) group—a group of key members of senior leadership from a diverse cross-section of 16 surgical societies in the United States and 9 internationally. Accrediting organizations, US Department of Defense and Veterans Health Administration were also represented. These individuals were purposefully invited for panel membership to represent opinions from different surgical specialties across a diverse geographic area (Table 2), although their opinions were not the official opinions or endorsement of their respective

Table 1. Surgical Societies and Organizations Surveyed for the Study

Location	Society and organization
North America	American Academy of Orthopaedic Surgeons
	American Academy of Otolaryngology-Head and Neck Surgery
	American Association for Thoracic Surgery
	American Association of Gynecologic Laparoscopists
	American Association of Plastic Surgeons
	American College of Obstetricians and Gynecologists
	American College of Surgeons
	American Hernia Society
	American Society of Colon and Rectal Surgeons
	American Society of Plastic Surgeons
	American Urogynecologic Society
	American Urological Association
	Arthroscopy Association of North America
	Association for Surgical Education
	Society of American Gastrointestinal and Endoscopic Surgeons
	Center for Advanced Medical Learning and Simulation
	Society of Laparoendoscopic Surgeons
	US States Department of Defense
	Veterans Health Administration
	Thoracic Surgery Directors Association
	National SimLEARN Center of the Veterans Health Administration
South America	Latin American Association of Laparoscopic Surgery
	Latin American Hernia Foundation
Europe	Royal College of Surgeons of England
	Royal College of Surgeons in Ireland
Asia	Asia Pacific Hernia Society
	China Hernia Society
Australasia	Royal Australasian College of Surgeons
Africa	South African Society for Obstetricians and Gynaecologists

societies. To ensure adequate representation of various surgical specialties across a wide geographic area, 45 individuals from 8 different surgical and 3 nonsurgical specialties from 9 countries were invited. Participation was voluntary and informed consent was implied if a panel member chose to participate.

Derivation of survey items

Survey items were based on evidence-based educational principles, published peer-reviewed literature, results of a modified Delphi consensus conference of ASSET group in June of 2011 and opinions of 4 international experts in surgical education. The complete list of survey items is available from the authors on request.

Survey administration

A complete list of survey items was presented to each participant in the form of an online survey (SurveyMonkey.com). Each participant was asked to rate each item on a Likert-type scale from 1 ("strongly disagree") to 5 ("strongly agree") with respect to the degree that s/he believed that item should be included in the final framework. Participants were given the opportunity to comment on each item and to clarify their ratings. One email reminder to complete the survey was sent 2 weeks after the initial invitation email. Data collection stopped 31 days after the original invitation email.

Table 2. Composition of the Delphi International Expert Participant Panel

Location	Surgical and nonsurgical specialty	No. of participants contacted	No. of participants responding
Canada	General surgery	3	2
USA	General surgery	11	5
	Obstetrics and gynecology	5	2
	Orthopaedic surgery	3	2
	Urology	2	1
	Plastic surgery	1	1
	Military	6	1
	PhD, MPH	3	2
	Otolaryngology	1	1
	Internal medicine	1	0
UK	General surgery	1	1
Brazil	General surgery	1	0
Australia	General surgery	1	1
Ireland	General surgery	1	1
	Vascular surgery	1	1
China	General surgery	1	0
South Africa	Obstetrics and gynecology	1	1
New Zealand	General surgery	1	1
	Pediatric surgery	1	1

Determination of consensus

A classical Delphi methodology analysis[28,29] was used to establish consensus among the participants. On the basis of the work of Graham and colleagues[28] and Palter and colleagues,[29] Cronbach's α was chosen as the statistical index to quantify a measure of consensus.[30] Cronbach's α of 0.80 was chosen to represent an acceptable measure of consensus for this study.

Creation of final framework

Once consensus among study participants was achieved (Cronbach's $\alpha \geq 0.80$), items that were rated as 4 ("agree") or 5 ("strongly agree") by >80% of participants were selected for inclusion into the final framework. Comments for each survey item were reviewed and modifications to the items were made based on the comments. A list of comments and modifications for each item is available from authors on request. The items in the final framework were then discussed and revised during a second modified Delphi consensus conference of ASSET in December 2011.

Data analysis

Descriptive statistics were calculated for all survey items. Cronbach's α was calculated as a measure of consensus among study participants. Missing data points, which resulted from submissions of incomplete surveys, were handled in the following ways: a mean participant score for that item was calculated and the missing data point was replaced with the mean score; missing data point was replaced with a 3 ("neutral") based on the assumption that because the question was omitted, the participant did not feel strongly about including or excluding that item from the final framework; and the missing data point was replaced with the mode for that item. Cronbach's α was calculated for each of these methods. All statistical analyses were performed using STATA software version 12.0 (Stata Corp).

RESULTS

Forty-five international experts in surgical education from 9 specialties and 9 countries were invited to participate in this study (Table 2). Twenty-four (53%) experts from 8 specialties and 7 countries completed the online survey (Table 2). The calculated Cronbach's α was 0.91 (missing data points replaced by mean), 0.89 (missing data points replaced by 3 "neutral"), and 0.91 (missing data points replaced by mode). Second round of the Delphi survey was not necessary, as consensus was achieved during the first round. The main components of the final framework included predevelopment analysis; cognitive, psychomotor, and team-based training; curriculum validation, evaluation, and improvement; and maintenance of training (Supplementary Appendix; available at: http://www.journalacs.org). Cronbach's α for each individual component of the framework ranged from 0.83 to 0.99 (Table 3). Percentage of missing data for each individual component ranged from 0 to 14% (Table 3).

Table 3. Agreement among Participants for Each Component of the Final Framework

Framework component	Cronbach's α	Missing data, %
Predevelopment analysis	0.89	0
Cognitive component		
Pretest of procedure-specific knowledge	0.97	0
Preprocedure patient assessment and preparation	0.89	0
Procedure-specific operative knowledge	0.83	0
Post-test of learned procedure-specific knowledge	0.95	0
Delivery of the cognitive component	0.92	0
Psychomotor skills		
Deconstruction of the procedure of interest into its component tasks/steps	0.89	12.0
Create a procedure-specific assessment tool using the identified tasks/steps/errors	0.84	8.0
Identify existing simulation models for training in the key tasks/steps of the procedure	0.91	12.0
Establish expert-level proficiency benchmarks for each simulation model	0.99	8.0
Define a specific practice schedule for trainees	0.90	12.0
Provide extrinsic feedback (during practice sessions) to trainees on their performance	0.86	12.0
Team-based component		
Develop a module for teaching nontechnical skills	0.86	8.0
Curriculum validation, evaluation, and improvement		
Validation	0.91	12.0
Evaluation and improvement	0.99	12.0
Maintenance of training	0.88	16.0

DISCUSSION

This study used classical and modified Delphi methodology to establish consensus among a panel of international experts in surgical education on the principles for design, validation, and implementation of a simulation-based surgical training curriculum. These principles were then formulated into a generic, prescriptive, uniform, and evidence-based framework for international implementation in any surgical specialty and for any surgical procedure of interest. Solicitation of experts from every surgical specialty and doctorate-level educators, from civilian and military backgrounds, and from diverse geographical locations (eg, North America, South America, Europe, Australia and New Zealand) strengthened the content validity of the developed framework and included perspectives not often included in curriculum development.

The final framework (Supplementary Appendix; available at: http://www.journalacs.org) has some similarities with the curricula development frameworks proposed by Aggarwal and colleagues,[24] as well as Stefanidis and Heniford.[14] Aggarwal and colleagues' 5-step framework includes:

1. Knowledge-based learning;
2. Deconstruction of the procedure into its component tasks;
3. Training in a laboratory environment;
4. Demonstration of transfer of skills to the real environment; and
5. Granting of privileges for independent practice.

Stefanidis and Heniford's framework takes a trainee from baseline technical skill assessment, to demonstration of procedure using video tutorials, to deliberate practice on a simulator in distributed sessions with feedback, to attainment of proficiency with overtraining, maintenance of training, and post-training assessment. These individual frameworks represent excellent analysis and recommendations from a single perspective of experience and evidence, although they lack a broad international component with input from multiple surgical specialties and regulatory organizations. The proposed framework in this study benefits from these pioneering studies, adds a consensus of 24 international experts from all surgical specialties, and is grounded in evidence-based methodological principles of educational theory. In addition to these fundamental principles, it supports the Dreyfus and Dreyfus model of training from novice to expert.[31] The comprehensive incorporation of all of these components is a major strength of the proposed framework.

The broad principals of educational theory constitute the infrastructure of this proposed framework. These principals have been previously reviewed in detail elsewhere.[32-34] They include acquisition of cognitive knowledge,[16-18] psychomotor skills training and nontechnical skills training (eg, communication, collaboration, professionalism, and management). Acquisition of cognitive knowledge can be organized into phases of conceptualization, visualization, and verbalization.[34] During the conceptualization phase, the learner learns the relevant anatomy, indications, contraindications, and complications related to the procedure. During the visualization and verbalization phase, the learner is required to describe the procedure from start to finish. Psychomotor skills training requires appropriate targets for proficiency,[1] deliberate[21,22] and distributed practice,[23] motivation and timely feedback,[33] overtraining and maintenance training.[14] Communication, collaboration, professionalism, and management are the nontechnical components of surgical competency. These components should also be made part of a comprehensive surgical training curriculum.

The classical Delphi method used in this study has several strengths. It offers the opportunity to conduct the questionnaire by mail or online; thereby improving feasibility and lowering costs.[27] Participants can be recruited from various geographical locations and clinical backgrounds,[28] and the anonymous nature of the Delphi method makes it challenging for a single influential participant to have a disproportionate impact on the survey's outcomes.[28] The use of classical Delphi method allowed for conclusion of this study at a minimal cost. Participants were contacted through email and the survey was administered online. A Cronbach's α value of >0.90 in this study implies that there were enough participants to achieve consensus, as well as a lack of controversy with regard to the proposed framework items. This lack of controversy is a result of the effort on the part of the authors to include only those items that were based on valid and evidence-based principles in surgical education.

The present study has a few limitations. The response rate for this study (53%) was moderate although acceptable[5,9,10] and consistent with survey response rates in health professional literature.[29,35,36] It is possible that participants who chose to not to participate in the survey were not checking their email regularly, were not comfortable with online surveys, or missed the reminder in their inbox. Extending the duration of data collection, incorporating additional e-mail and telephone reminders, and sending out the survey by mail could have potentially increased the response rate. However, to maintain participants' interest in the survey and anticipating a possible need for a second round of the survey, data collection was stopped after 31 days. With the finding of consensus among panel members after the first round of the survey, a second round was not required. In addition, the response rate must be judged with the knowledge of the population that this study attempted to recruit—international experts with senior leadership positions in various national and international surgical societies.

The distribution of participants' geographical locations was not uniform. The greatest proportion of experts was from North America, whereas continental Europe, Asia, and South America were not well represented. Consequently, results likely reflect an Anglo-American point of view and should be interpreted in the context of an Anglo-American education format. Future studies can address this limitation by surveying each geographical region separately using a Delphi method and then pooling the results.

The classical Delphi method, although considered to be one of the ideal means to elicit expert opinion and determine consensus, has been criticized because the breadth of the question under consideration is in part controlled by the investigators.[28,29] This study mitigated this criticism by deliberately constructing the Delphi survey to contain not only the preliminary information from the initial ASSET consensus conference, but also items from published peer-reviewed literature and recommendation from key informant interviews. The reliability of the results obtained using a Delphi method has also been brought into question.[37] However, reports of Delphi results being accurate 16 years after the initial survey administration suggest that results are reliable.[38] The present study attempted to mitigate this potential limitation of reliability by recruiting participants from a diverse geographical area and from all surgical specialties. Missing data points, which resulted from partial completion of the survey represent another limitation; however, regardless of how the missing data points were analyzed, Cronbach's α remained well above the predetermined cutoff of 0.80—suggesting consensus among the participants.

Lastly, it is important to discuss and clarify the role for simulation in surgical residency-training programs. During the past 100 years, time-based apprentice-type surgical training programs have produced many superb surgeons; however, with restrictions in resident work hours, increasing emphasis on patient safety, and rising costs for training in the operating room, trainees now have fewer opportunities to train in the operating room. Simulation-based training is not intended to completely replace training in the operating room; rather, it is intended to provide an effective means of overcoming the initial learning curves in the safety and comfort of a laboratory. Such training has been shown to improve technical performance,[11] decrease technical error rate,[1] and shorten learning curves in the operating room.[39] The major driving force behind the call to incorporate simulation into surgical training comes from the need to improve patient safety and to shorten the learning curves in the operating room.

CONCLUSIONS

This study used a cohort of international experts and both the classical and modified Delphi methods to develop a consensus on an evidence-based framework for the design, validation, evaluation, and implementation of a simulation-based training curriculum for any surgical procedure in any surgical specialty. A step-by-step progression through the proposed framework will provide surgical educators, program directors, and decision makers with an evidence-based guide for design and incorporation of a simulation-based surgical training curriculum into their respective residency training programs. Wide implementation of this framework will require the support of national organizations, professional societies, and residency review committees.

This curricular framework is the first step to establishing a uniform, interoperable international approach to the full life cycle of curriculum design, development, and validation. Additional research studies should focus on defining the means for establishing benchmark values for assessment, defining the appropriate intervals for retesting of both cognitive and technical skills and for maintenance practice. Once this is achieved, it will be possible to begin a transition from the current time-based surgical training paradigm to a proficiency-based paradigm with objective means of certification and recertification. This has the potential to improve the overall quality of patient care by standardizing the competencies of surgeons in training and in practice.

Finally, this article and its proposed standardized framework are intended to bring uniformity to the curriculum development process. This framework is a living document, intended for improvement. There are numerous aspects that could have been addressed and added, and will likely be incorporated in the future. However, by providing a "straw man" based both on scientific rigor and pragmatic implementation values, the goal is to achieve a baseline for those intending to develop curricula that would result in interoperability of their efforts with those of the community at large. Although the architecture backbone of the framework should be adhered to when possible, the final details do have enough flexibility to be adaptable to many skills, procedures, and specialty needs. Careful attention was paid to insuring that this curriculum template can be implemented in most any type of surgery (ie, open, laparoscopic, robotic, endoscopic, image guided, etc), as well as across surgical and procedural specialties to be able to conduct comparative effectiveness analyses between training and procedures with equivalent parameters, comparing apples with apples rather than apples with oranges. As in all guidelines, it is implicit that this framework will change over time, adapt to new discoveries in science, technology, educational principals, etc.

Author Contributions

Study conception and design: Zevin, Grantcharov
Acquisition of data: Zevin, Levy

Analysis and interpretation of data: Zevin, Levy, Satava, Grantcharov
Drafting of manuscript: Zevin, Levy, Satava, Grantcharov
Critical revision: Zevin, Levy, Satava, Grantcharov

Acknowledgment: The authors would like to thank members of the Alliance for Surgical Simulation Education in Training group for participating in this study.

REFERENCES

1. Ahlberg G, Enochsson L, Gallagher AG, et al. Proficiency-based virtual reality training significantly reduces the error rate for residents during their first 10 laparoscopic cholecystectomies. Am J Surg 2007;193:797–804.
2. Korndorffer J Jr, Dunne JB, Sierra R, et al. Simulator training for laparoscopic suturing using performance goals translates to the operating room. J Am Coll Surg 2005;201:23–29.
3. Seymour NE. VR to OR: a review of the evidence that virtual reality simulation improves operating room performance. World J Surg 2008;32:182–188.
4. Fried GM. Lessons from the surgical experience with simulators: incorporation into training and utilization in determining competency. Gastrointest Endosc Clin N Am 2006;16:425–434.
5. Kozmenko V, Paige J, Chauvin S. Initial implementation of mixed reality simulation targeting teamwork and patient safety. Stud Health Technol Inform 2008;132:216–221.
6. Merry AF. Human factors and the cardiac surgical team: a role for simulation. J Extra Corpor Technol 2007;39:264–266.
7. Black SA, Nestel DF, Kneebone RL, et al. Assessment of surgical competence at carotid endarterectomy under local anaesthesia in a simulated operating theatre. Br J Surg 2010;97:511–516.
8. Issenberg SB, McGaghie WC, Petrusa ER, et al. Features and uses of high-fidelity medical simulations that lead to effective learning: a BEME systematic review. Med Teach 2005;27:10–28.
9. Sutherland LM, Middleton PF, Anthony A, et al. Surgical simulation: a systematic review. Ann Surg 2006;243:291–300.
10. Seymour NE, Gallagher AG, Roman SA, et al. Virtual reality training improves operating room performance: results of a randomized, double-blinded study. Ann Surg 2002;236:458–463; discussion 463−464.
11. Grantcharov TP, Kristiansen VB, Bendix J, et al. Randomized clinical trial of virtual reality simulation for laparoscopic skills training. Br J Surg 2004;91:146–150.
12. Larsen CR, Soerensen JL, Grantcharov TP, et al. Effect of virtual reality training on laparoscopic surgery: randomised controlled trial. BMJ 2009;338:b1802.
13. Korndorffer JR Jr, Stefanidis D, Scott DJ. Laparoscopic skills laboratories: current assessment and a call for resident training standards. Am J Surg 2006;191:17–22.
14. Stefanidis D, Heniford BT. The formula for a successful laparoscopic skills curriculum. Arch Surg 2009;144:77–82; discussion 82.
15. ACS/APDS Surgical Skills Curriculum for Residents 2010. Available at: http://www.facs.org/education/surgicalskills.html. Accessed December 15, 2011.
16. Satava RM. Identification and reduction of surgical error using simulation. Minim Invasive Ther Allied Technol 2005;14:257–261.
17. Tang B, Hanna GB, Cuschieri A. Analysis of errors enacted by surgical trainees during skills training courses. Surgery 2005;138:14–20.
18. Van Herzeele I, Aggarwal R, Neequaye S, et al. Cognitive training improves clinically relevant outcomes during simulated endovascular procedures. J Vasc Surg 2008;48:1223–1230,1230 e1.
19. Stefanidis D, Acker CE, Swiderski D, et al. Challenges during the implementation of a laparoscopic skills curriculum in a busy general surgery residency program. J Surg Educ 2008;65:4–7.
20. Chang L, Petros J, Hess DT, et al. Integrating simulation into a surgical residency program: is voluntary participation effective? Surg Endosc 2007;21:418–421.
21. Ericsson KA. Deliberate practice and the acquisition and maintenance of expert performance in medicine and related domains. Acad Med 2004;79:S70–S81.
22. Ericsson KA. Deliberate practice and acquisition of expert performance: a general overview. Acad Emerg Med 2008;15:988–994.
23. Moulton CA, Dubrowski A, Macrae H, et al. Teaching surgical skills: what kind of practice makes perfect?: a randomized, controlled trial. Ann Surg 2006;244:400–409.
24. Aggarwal R, Grantcharov TP, Darzi A. Framework for systematic training and assessment of technical skills. J Am Coll Surg 2007;204:697–705.
25. Nathens AB, Rivara FP, Jurkovich GJ, et al. Management of the injured patient: identification of research topics for systematic review using the delphi technique. J Trauma 2003;54:595–601.
26. Williams PL, Webb C. The Delphi technique: a methodological discussion. J Adv Nurs 1994;19:180–186.
27. Fink A, Kosecoff J, Chassin M, et al. Consensus methods: characteristics and guidelines for use. Am J Public Health 1984;74: 979–983.
28. Graham B, Regehr G, Wright JG. Delphi as a method to establish consensus for diagnostic criteria. J Clin Epidemiol 2003;56: 1150–1156.
29. Palter VN, MacRae HM, Grantcharov TP. Development of an objective evaluation tool to assess technical skill in laparoscopic colorectal surgery: a Delphi methodology. Am J Surg 2011;201: 251–259.
30. Bland JM, Altman DG. Cronbach's alpha. BMJ 1997;314:572.
31. Dreyfus HL, Dreyfus SE, Athanasiou T. Mind Over Machine: The Power of Human Intuition and Expertise in the Era of the Computer. New York: Free Press; 1986.
32. Drew PJ, Cule N, Gough M, et al. Optimal education techniques for basic surgical trainees: lessons from education theory. J R Coll Surg Edinb 1999;44:55–56.
33. Kneebone R. Simulation in surgical training: educational issues and practical implications. Med Educ 2003;37:267–277.
34. Kovacs G. Procedural skills in medicine: linking theory to practice. J Emerg Med 1997;15:387–391.
35. Burns KE, Duffett M, Kho ME, et al. A guide for the design and conduct of self-administered surveys of clinicians. CMAJ 2008; 179:245–252.
36. Sierles FS. How to do research with self-administered surveys. Acad Psychiatry 2003;27:104–113.
37. Keeney S, Hasson F, McKenna HP. A critical review of the Delphi technique as a research methodology for nursing. Int J Nurs Stud 2001;38:195–200.
38. Ono R, Wedemeyer DJ. Assessing the validity of the Delphi technique. Futures 1994;26:289–304.
39. Aggarwal R, Ward J, Balasundaram I, et al. Proving the effectiveness of virtual reality simulation for training in laparoscopic surgery. Ann Surg 2007;246:771–779.

SUPPLEMENTARY APPENDIX

An evidence-based framework for the design of a simulation-based surgical training curriculum

Predevelopment analysis

1. Perform a needs assessment/define the problem
   a. Determine the goals and objectives for the curriculum
   b. Determine the desired learning outcomes, measures and proficiencies to be achieved
      i. Include governing authorities—Boards, Societies, ACGME, etc, when possible
   c. Determine the common errors that need to be taught
   d. Define the population of interest (novice, intermediate, advanced trainee, expert surgeon)
   e. Include program director, faculty, educator, user (eg, resident, medical student, etc) and boards
2. Determine the available resources at your institution, society, etc (optional)
   a. Knowledge content
   b. Physical space
   c. Equipment (simulators? types? high or low fidelity?)
   d. Personnel including faculty with formal educator expertise (ie, fellowship, masters, or PhD)
   e. Financial resources
   f. Support of faculty members, chair, chief, and/or program-director

Curriculum development

Cognitive component (focused on specific technical issues)

1. Pretest of procedure-specific knowledge
   a. Written test format
      i. Multiple choice
   b. Interactive, graphic-based image software that delivers procedural training
2. Preprocedure patient assessment and preparation
   a. Review of indications/contraindications for surgery
   b. Teamwork and environment considerations (functioning equipment, instruments, suture/stapling devices, etc.)
3. Procedure-specific operative knowledge
   a. Key steps of the procedure
   b. Anatomy: descriptions of important structures and tissue planes
   c. Instruments: identify types, indications, safety and limitations of instruments
   d. Psychomotor skills: demonstration of fundamental technical skills needed to complete the procedure
   e. Errors (most common and most critical though infrequent):
      i. Anticipation/avoidance
      ii. Recognition
      iii. Management
   f. Postoperative complications (potential and actual):
      i. Knowledge of
      ii. Anticipation/avoidance
      iii. Recognition and management
   g. Inclusion of video examples of failures/disasters
4. Faculty interaction: time for questions to the expert; answers and reflections by trainees
5. Post-test of learned procedure-specific knowledge
   a. Written test format
      i. Multiple choice questions
      ii. Interactive, graphic-based image software that delivers procedural training
   b. Mock operative note dictation
6. Delivery of the cognitive component
   a. Procedure-specific video-based tutorials with pauses in the video at critical steps of the procedure
   b. Interactive software that delivers procedural training
      i. Interactive graphic-based image of procedure has to be identical to psychomotor skill simulator (this is critical)
   c. Progress to the psychomotor skills component of the training is not permitted until cognitive part is passed, preferably with 100%.

Psychomotor skills (technical component)

1. Deconstruction of the procedure of interest into its component tasks/steps
   a. Acquire videorecordings of different approaches of experts performing the same procedure of interest
      i. Include only the tasks in common from multiple different approaches
   b. Perform a hierarchical tasks analysis (HTA) of the procedure (ideally should include experts in surgical education, behavioral psychologists, expert clinicians)
      i. Two or more independent reviewers to watch the video-recordings and write down all the tasks/steps of the procedure

   ii. Establish consensus between reviewers on the tasks/steps
   iii. Each tasks/steps should have a clear and unambiguous definition of the task/subtask, start point and end point, and quantitative measure (if possible) or clearly defined Likert Scale
   iv. By reviewing the video-recordings of novices, identify which tasks/steps are the most difficult for a novice to perform and what are the most common errors
2. Create a procedure-specific assessment tool using the identified tasks/steps/errors
   a. Use Delphi methodology to identify the tasks/steps/errors from the HTA to be included in the technical skills assessment tool
   b. Establish reliability of the tool
   c. Establish validity of the tool
   d. Demonstrate construct validity by demonstrating a difference in scores of novice surgeons (<10 procedures), intermediates (10 to 100 procedures), and experts (>100 procedures). The number of procedures at each level can vary significantly depending on the type of procedure.
   e. Establish the benchmark criteria for competency (Experts' mean ± 1 SD)
      i. Analyze experts' learning curves (stop when 2 consecutive trials show no improvement, ie, they have reached a plateau)
      ii. Establish the mean of all the experts' plateaus
      iii. Establish the mean ±1 SD of the mean of all the experts
      iv. The benchmark of competence is experts' mean minus 1 SD
   f. Define a cutoff value to be used in formative assessment (from the errors list)
   g. Define a cutoff value to be used in summative assessment to define proficiency (from the benchmark criteria)
3. Conduct initial assessment of trainees technical skill using the procedure-specific assessment tool
4. Identify existing simulation models for training in the key tasks/steps of the procedure
   a. If no procedure-specific simulator exists, determine how to develop the desired simulator
   b. Basic skills training
      i. Synthetic models
      ii. Low fidelity computer-based models
      iii. Aim to select models with embedded objective metrics
   c. Advanced (procedure-specific) skills training
      i. Synthetic models
      ii. Computer-based models
      iii. Cadaveric/tissue models
      iv. Live animal models
   d. Confirm or define validity for the chosen models
      i. Face
      ii. Content
      iii. Concurrent
      iv. Construct
      v. Predictive
5. Define different levels of difficulty for each simulation model
6. Establish expert proficiency benchmarks for each simulation model (see item 2e)
   a. Using measurement parameters imbedded in the model
   b. Using the previously developed procedure-specific assessment tool
7. Set acceptable level of proficiency as the "pass" score for each model in the curriculum (see item 2e)
   a. Forced proficiency at each level before progression to the next level (learn basic skills before move on to advanced levels)
8. Define a specific practice schedule for trainees
   a. Distributed practice
   b. Deliberate practice
   c. Variability of practice (variable levels of difficulty)
   d. Practice until acceptable level of proficiency is achieved
   e. Define amount of overtraining (if needed)
9. Provide extrinsic feedback (during practice sessions) to trainees on their performance
   a. Formative
      i. Immediate computer generated feedback
      ii. Immediate, face-to-face feedback by expert instructors with alternative approaches demonstrated and practiced in the presence of the instructor
   b. Summative
      i. Review areas of weakness based on the procedure-specific assessment tool

ii. Review of performance on video

Team-based component (supports technical aspect of procedure)

1. Develop a module for teaching nontechnical skills (teamwork, communication, situation awareness, etc) as related to the procedure of interest
   a. Identify the roles and responsibilities of allied health care team (individual and team training)
   b. Develop crisis-based scenarios for learning nontechnical skills in a simulated environment (appropriate for level of learner and of incremental difficulty)
   c. Use available assessment tools for formative evaluation of nontechnical skills (evaluate the reliability, validity and applicability of instruments; if not available develop new tools)
   a. Use an existing team training model (eg, TeamSTEPPS) if available
   b. Incorporate train-the-trainer programs for teaching faculty
      i. Provide specific training for debriefing skills

Curriculum validation, evaluation, and improvement
Validation

1. Conduct a randomized controlled trial to demonstrate transfer of learned skills to the real environment
   a. Compare simulator trained group with traditionally trained group
   a. Demonstrate difference in cognitive knowledge and technical skill
   b. Evaluate the impact of simulation-based training on the learning curves in the operating room
   c. Calculate the transfer-effectiveness ratio for the training curriculum

Evaluation and improvement

1. Prospectively collect data on curriculum validity
   a. Pre-post test learner performance data comparison
   b. Learner retention data
   c. Learner feedback
   d. Impact of curriculum on clinical outcomes
2. Perform periodic evaluations of curriculum and adjust curriculum based on feedback/experience
   a. Identify strengths and weaknesses and areas for improvement
   b. Develop methods for continuous program improvement

Maintenance of training

1. Define post-training intervals at which cognitive knowledge and technical skills tests will be administered to assure ongoing proficiency of the learner
2. Develop a specific "retraining" curriculum, based on principles listed. Generally, this will be less extensive and more focused than the full curriculum.
3. Include retraining to proficiency during testing sessions (maintenance practice) with formative feedback.
4. Study skill degradation for particular tasks.