In Search of Characterizing Surgical Skill

David Azari, PhD,[*] Caprice Greenberg, MD, MPH,[*,†] Carla Pugh, PhD, MD,[‡] Douglas Wiegmann, PhD,[*] and Robert Radwin, PhD[*,§]

[*]Department of Industrial and Systems Engineering, University of Wisconsin-Madison, Madison, Wisconsin; [†]Department of Surgery, Wisconsin Surgical Outcomes Research (WiSOR) Program, University of Wisconsin-Madison, Madison, Wisconsin; [‡]Department of Surgery, Stanford University, Stanford, California; and [§]Department of Biomedical Engineering, University of Wisconsin-Madison, Madison, Wisconsin

OBJECTIVE: This paper provides a literature review and detailed discussion of surgical skill terminology. Culminating in a novel model that proposes a set of unique definitions, this review is designed to facilitate shared understanding to study and develop metrics quantifying surgical skill.

DESIGN: Objective surgical skill analysis depends on consistent definitions and shared understanding of terms like performance, expertise, experience, aptitude, ability, competency, and proficiency.

STRUCTURE: Each term is discussed in turn, drawing from existing literature and colloquial uses.

IMPLICATIONS: A new model of definitions is proposed to cement a common and consistent lexicon for future skills analysis, and to quantitatively describe a surgeon's performance throughout their career. (J Surg Ed 76:1348−1363. © 2019 Association of Program Directors in Surgery. Published by Elsevier Inc. All rights reserved.)

KEY WORDS: Objective performance assessment, Deliberate practice, Surgical skill

COMPETENCIES: Practice-Based Learning and Improvement

INTRODUCTION

Measures of surgical skill are increasingly shown to reflect clinical outcomes.[1,2] Common descriptors like superior performance (or elite performance), aptitude and ability, competency and proficiency, mastery, expertise and experience, however, often lack unique interpretation. Experience is sometimes used as a proxy for expertise, but also as a signal of professional status, without supporting evidence of best practices. Laufer and colleagues,[3] for instance, observed that clinicians of similar experience exhibit different approaches (sometimes radically) while completing simulated clinical breast exams. Should physicians with the greatest experience be considered experts even if − as Choudhry and colleagues[4] found − greater experience doesn't always produce greater quality of care? Or, are only some experts whose patients achieve better outcomes exhibiting what we would think of as "truly expert" behaviors?

Uncovering features to accurately describe surgical expertise is similarly an intricate challenge. Advanced performance in the operating room is known to depend on combining various skill sets,[5,6] for which any objective assessment must be tailor made[7] and rigorously tested.[8] Variable terminology hinders this effort, and makes it more difficult to validate assessments in line with robust evidentiary requirements of modern frameworks.[9,10] The rise of "surgical data science"[11] and engineering approaches to quantify surgery,[12] in part through simulation,[13,14] offers the opportunity to address boundary conditions of amorphous terms such as "expert" through quantifying behavior. Consistent terminology within and across studies will help provide appropriate context for these claims, and to promote standard and reproducible measures of surgical skill. This paper explores a series of common surgical skills terms and proposes a lexicon to encourage reproducibility among future studies of surgical performance.

What is Surgical Skill?

Surgical skills are commonly split into either technical or nontechnical categories,[6] despite known impacts of nontechnical skill on technical performance.[15] It is also widely accepted that "operative skills" are not just technical in nature.[16] Still, this artificial bifurcation has helped to frame studies examining hand motion,[17−22] errors and error management strategies,[23-26] cognitive readiness,[27] decision making,[28] and communication and teamwork.[29-31]

Funding: This research did not receive any specific grant from funding agencies in the public, commercial, or not-for-profit sectors.

Correspondence: Inquiries to Robert G. Radwin, Department of Industrial and Systems Engineering, 1550 Engineering Drive, Madison, WI 53706; e-mail: rradwin@wisc.edu

  1931-7204/$30.00
https://doi.org/10.1016/j.jsurg.2019.02.010

In a discussion of nontechnical skills, Yule et al.[6] proposed the interrelated category of cognitive skills to better describe features of surgical performance such as mental readiness, decision making and situational awareness, and interpersonal skills such as leadership and communication. This deconstruction is additionally supported by surveys of master surgeons (in this case, defined as those with high peer rankings and consistent involvement as trainers), who describe cognitive factors, innate dexterity, and personality as "important attributes" of surgical competence.[32] Greenberg et al.[33] advocate that nontechnical skills, including both cognitive and interpersonal skills should be integrated as part of the Wisconsin Surgical Coaching Framework (Fig. 1).

Madani and colleagues[5] have since developed a novel interoperative performance framework, composed of five inter-related performance domains: psychomotor skills (i.e., technical performance), declarative knowledge (i.e., recitable facts acquired outside the operating room), interpersonal skills (i.e., teamwork, leadership), personal resourcefulness (i.e., self-awareness and metacognition), and advanced cognitive skills (i.e., planning, error recovery).[34] This framework provides an excellent description of various domains or "competencies" in which surgeons perform. It facilitates continuing ontological understanding of a skill's "microstructure" – a necessary first step in designing "deliberate practice activities that allow performers to stretch their performance to a higher level,"[35] and helps to incorporate the "reasoning and motivation" behind successful operations in challenging professional settings.[36]

Madani and colleagues maintain that "competence has yet to be defined to a level that allows credentialing and licensing bodies to ascertain whether or not an individual has achieved the standards deemed to represent competent performance." In other words, it remains a challenge to establish "pass-fail standards," critical parts of competence-based education.[37] They are also careful not to refer to their framework as a "skills framework," presumably to avoid any confusion or contradiction between thinking of knowledge and performance as a skill or vice-versa. Rather, Madani's work presents a series of performance domains, the creative and fluid synthesis of which, through deliberate practice and rehearsal, may characterize surgical expertise. The authors adroitly navigate the difficulty in defining such terms and focus on the underlying surgical behaviors. They redirect emphasis from examining "skills" to interrelated "performance domains" that are widely applicable across procedures.

REVIEW OF SURGICAL SKILL LITERATURE

Using surgical skills terminology to describe how a surgeon's performance changes throughout their career is an intricate puzzle. A trainee's aptitude in one domain, for instance, hinges on a "working definition of superior surgical performance."[38] Yet what it means to be superior, generally considered a hallmark of medical expertise, presupposes measurement and attainment of competency.[39] Expertise, meanwhile, is interwoven with notions of both performance and skill. Murinson, Agarwal, and Haythornthwaite,[40] for example, frame expertise as the aggregation of "essential skills"; while Krawczyk and colleagues,[41] argue that expertise exists as a set of "exceptional skills" that can be measured and compared in laboratory tasks. In the same vein, Ericsson and Charness[42] describe experts as natural outliers: "performing at least two standard deviations above the mean level in the population."

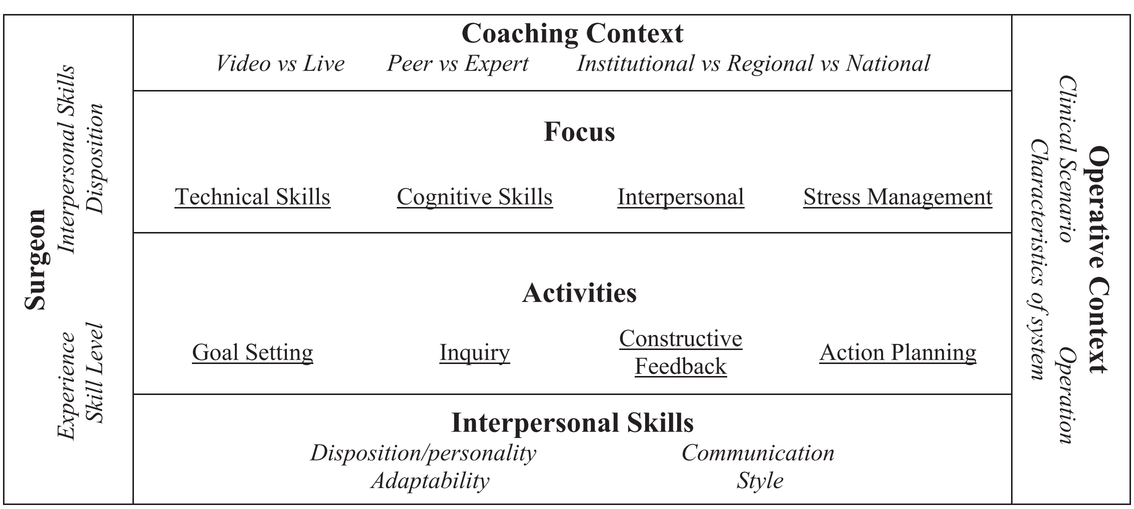


FIGURE 1. Wisconsin Surgical Coaching Framework (Adapted from Greenberg et al., 2017).

The implicit assumption that performance can be sufficiently (1) observed and (2) quantified to sort or rank performance in a meaningful way remains dubious: "[n]o single assessment method can provide all the data required for judgment of anything so complex as the delivery of professional services by a successful physician."[43] Instead, the Accreditation Council for Graduate Medical Education's (ACGME) Competency Based Medical Education (CBME) milestones approach[44] continues to promote operative autonomy through guidelines to document and show performance as a precursor. Any valid assessment to demonstrate these skills must be grounded within a robust basis of evidence.[9,45-47] Clear and consistent skill terminology will promote better evaluation and eventual application of assessment tools.

Performance

Like other performance domains (music, athletics, for example) surgical performance is repeatedly created anew at each opportunity, where contextual factors can change rapidly.[48] Subsequently, and stemming from notions that there exists a "maximal" level, performance is a dynamic, temporary, and alterable characteristic. Deliberate practice improves performance over time,[49] but performance is also subject to the context of the surgery[50] and various factors within the work-system.[51] Even those most practiced surgeons are not immune from committing errors, or not managing errors properly. Surgical performance thus represents the observable quality of a sequence of surgical actions at a specific point in time. It is possible (albeit unlikely) for a novice to outperform an expert, or an expert to underperform relative to their position. Such situations would be the exception, rather than the rule.

Defining maximum performance has a long history of debate. Francis Galton – contemporary of Charles Darwin, and who is generally credited with developing both "nature vs nurture" and "eugenics" terms – argued in *Hereditary Genius*[52] that maximal performance is a rigid and individually determined limit of one's genetic potential. Although he acknowledged that practice improves performance, Galton argued that "genius" depended foremost on your family tree. Later, Snoddy (as cited by Stratton et al.[53]) developed the now "ubiquitous" power law of learning, composed of the distinct cognitive, associative and autonomous classical stages of skill acquisition.[54,55] Galton's concept of an individual limit was re-envisioned as task-based performance limit – the asymptote of a "learning curve" governed by a power law. In an invited address to *Academic Medicine,* K. Anders Ericsson[49] expands this definition and argues that achieving an "expert" level of performance hinges on intentional and deliberate practice, implemented over long periods. Davids[48] reflects Ericsson's productive framework, in describing that "the power law of practice simply states that performance improves with practice, although there are eventual physical limits to this relationship."

Even though these definitions have evolved over time, surgical performance can be thought of as a temporary snapshot into the observable skills a surgeon brings to bear within a given situation. Performance can improve due to amount and style of deliberate practice but remains bounded to some asymptotic limit. Over longer periods, performance can improve or decline as surgeons age or switch to different types of operations. While transitioning to military service, for example, surgeons struggle to adapt their specialty based expertise to new challenges such as truncal hemorrhage or skeletal reconstruction from penetrating injuries,[56,57] while their other clinical skills, especially laparoscopy,[58] decay. Given the intense training required to achieve high performance in any surgical task, there is an ever-present interest in testing surgical residents for their abilities and aptitudes, to see who may be better equipped to gain operative skills with less training and coaching. Improving pedagogical techniques[59] may ease the difficulty for early learners, but the amount of training required for some individuals to achieve high surgical performance could still be prohibitive.

Techniques to measure performance limits continue to improve, prompted by increasing computational ability to quantify surgery in various contexts.[11] There are increasing improvements automatically measuring performance in open[18,60] and laparoscopic procedures,[61] as well as with eye tracking,[62] and automatic "stroke" recognition.[63] The vast majority of these efforts, however, are limited to benchtop simulations or robot-assisted devices.[14]

Aptitude and Ability

Surgeons may lament a lack of manual dexterity and psycho-motor coordination (i.e., coordination, balance, haptic force control) of incoming residents, behaviors largely unpracticed in medical school until clinical rotations and outside the scope of common premedical undergraduate programs. In contrast, elite musicians and professional athletes often begin deliberate practice in their field at an early age, engendering significant advantages later in life.[64] Decreased training time among residents,[65] little emphasis on mid-career training interventions,[16,32] and increasing complexity of the operating room,[66] have rekindled interest in "aptitude testing" to jumpstart selection and training of residents.[67-69] Even "intellectual prowess" and "emotional stability" have been proposed as potential avenues to test for surgical aptitude.[70]

Aptitude is commonly described as a "natural" advantage a trainee brings to the table.[71] This definition does not preclude new pedagogical techniques or better coaching from improving the performance of new trainees. Hislop et al.,[72] while examining aptitude for endovascular procedures, found that clinicians with extensive video game experience completed virtual reality tasks more quickly than those without prior video game experience. Willis et al.,[73] also found performance in virtual reality and video game to be related, suggesting that these pre-existing experiences may transfer well to some simulations. Both studies sought to connect pre-existing strengths to a relatively increased rate of performance gain during surgical training.

Existing literature, however, also tends to confound aptitude and ability. Ability is often qualified as "natural," "innate," or "fundamental," to describe an advantage someone brings to the selection process and training curriculum. Alfred Cuschieri[74] uses aptitude as an intermediary to distinguish between skills (i.e., trained) and abilities (i.e., untrained): "abilities are the innate aptitudes that people can bring to given tasks and determine the level of proficiency that individuals obtain with training." In that view, skills require training, while abilities (being innate aptitudes) are brought exclusively by the individual. Groenier et al.[75] similarly uses "ability" to describe incoming trainee cognitive and psychomotor performance scores. In that study, participants with higher psychomotor scores learned to complete laparoscopic tasks more quickly, and with greater efficiency of movement. Szasz et al.[76] also used "ability" to indicate a resident's likelihood of promotion, measuring how they could meet concurrent performance thresholds for both the Objective Structured Assessment of Technical Skills (OSATS) and Objective Structured Assessment of Non-Technical Skills (OSANTS). Moglia et al.,[67] refers to "innate aptitude" and "innate ability" synonymously while examining psychomotor performance scores utilizing the da Vinci Skills Simulator.

While individual strengths clearly impact outcomes of surgical training, merging aptitude and ability as a single concept creates problems defining performance. Using ability to express untrained strengths (i.e., "innate aptitudes") relegates to the role of training to second class status while describing what a surgeon can do, and undermines aptitude assessment in the training process. Yet, "ability" is also used to describe training outcomes: Mattar et al.[77] stipulates that many residents are "unable to operate for 30 unsupervised minutes of a major procedure" upon graduation. Residents would not be able to do this without adequate training, even if gifted with some pre-existing proclivity to learn the necessary skills (whether deemed fundamental, innate, or natural). Referring to individually different strengths in absence of training or outside practice as abilities confounds the distinct roles that individual strengths and training play in building surgical skill.

To prevent overlapping interpretation between ability and aptitude, and to promote unique definitions, an individual's pre-existing strength should be described as aptitude that would impact the rate of performance gain. This would free the term "aptitude" from needing to rely on "ability" as a stepping stone. It allows "ability" to represent formally trained techniques that need not be innate or fundamental, and integrates aptitude in the training process. Separating these definitions maintains the notion of pre-existing strengths (i.e., aptitudes) without sacrificing the importance of training to demonstrable surgical performance (i.e., abilities).

In other words, ability represents the maximal performance an individual can offer under ideal circumstances, based on exigent training and previous experience. Recognizing that someone is able to perform a task implies they consistently meet or exceed some arbitrary threshold of acceptable performance.[78] In this context, an ability would be demonstrable performance recognized as competent or higher. Ability grows over a career, commensurate with deliberate practice and exposure to difficult situations. Ability is also different than aptitude, as no trainee is considered "able" to perform a procedure because they score well on an aptitude test. However, both features are brought to bear in difficult operating room situations where surgeons use all advantages (trained or otherwise) to maximize performance.

Aptitude and ability are also domain-specific and nested within the taxonomy of Madani's framework discussed previously. For example, sewing aptitude and decision-making aptitude are quite different for early trainees, and ought to be described within the context of a relevant level of training and task, so as not to lose specificity or increase bias in selection. A software tool to provide quantitative feedback of performance without the need for coach intervention could test for aptitudes, and help to improve student abilities through targeted practice before starting clinical rotations.

Experience

As referenced by Weg Thomas,[79] Lord Smith, a past President of The Royal College of Surgeons of England said "it would take me one year to teach a trainee how to do an operation, five years to teach them when to do the operation, but a lifetime to teach them when not to do an operation." Experience describes the amount and breadth of familiarity a surgeon has in the operating room. Commonly expressed in the number of cases completed or amount of years in an attending role, experience also manifests through "war stories"[80] and efficient "case scripts" or "illness scripts"[81,82] that improve expectations in the operating room.

Robust mental scripts show how experienced clinicians compose a litany of creative, fail-safe approaches to help deal with new challenges. These are complex representations that include "kinaesthetic and visual imagery" that help in psychomotor planning and movement,[83] and may be improved by mental practice.[84] Previous experience helps form and represents critical patterns and cues which would otherwise go undetected. The set of expectations and scripts have also been described as a mental schema[81] to ease the burden of planning several steps ahead and recalling and integrating vast amounts of declarative knowledge on the fly.[85] By incorporating principles into patterns and schemas, surgeons reduce the demand on cognitive resources, enabling greater flexibility to direct attention where most needed.[27] Clinicians form crucial components of expertise by incorporating feature-based patterns and expectations of their own experience into these mental models.[82]

In the search for objective measures of performance, experienced surgeons offer a window into successful techniques that are tailored and honed over years of difficult practice. Carty and colleagues[86] found that variation in operative time steadily decreased as experience grew, plateauing after 10 years. Still, studying experienced surgeons – in the absence of more detailed objective measures – may fall short of guaranteeing positive outcomes. Geoff Norman summarized the insufficiency of experience as a measure of surgical expertise in his personal correspondence to David Cook, facetiously suggesting that "gray hair and baldness would be good measures of expertise when comparing senior surgeons and third-year medical students."[9] Clarified by Norman's wry attitude, experience is a necessary, albeit insufficient prerequisite for expertise. Some individuals who consider themselves more experienced may be more resistant to contradictory information.[87] Combating this trend, where it exists, and promoting so called "intellectual humility"[88] will be an ongoing effort.

Expertise

Early attempts by Simon and Chase[89] to define expertise emphasized direct testing of memory capacity. Although such measures provided a "convenient substitute for studies of actual performance,"[35] they did not provide sufficient explanations for the mechanisms supporting how expertise manifests across domains. Later approaches paralleled the proliferation of technology, describing the brain as a "computational device" which stored, responded and retrieved information in "motor programs" analogous to software operating on a computer.[48]

The early emphasis on memory recall has been reframed over time to describe how expert performance manifests as a set of "exceptional skills in a particular domain."[41] Growing efforts to observe, measure and compare expert behavior,[90] have been bolstered by procedural and technological advances in cognitive task analysis,[91] brain imaging,[41] protocol analysis and eye tracking,[39] simulation[92] and domain-specific factor analysis,[93] among others.

It is widely accepted today that expertise is achieved through deliberate practice over time.[27,42,64,94,95] Performers can target unique skills and reflect on their progress during planned periods where they "construct and seek out training situations in which the desired goal exceeds their current level of performance."[49] Exposure to difficult and variable situations is also critical,[96] enabling performers can become "adaptable for a range of varying performance characteristics" and "less vulnerable to transitory factors such as fatigue, audience effects, and anxiety."[48] Unfortunately, the chance to reflect on these lessons is limited in the rapid and stressful environment of resident training,[97] resulting in missed opportunities to promote metacognition and resident development.[98] Indeed, "practice without reflection and striving for continued improvement is a formula for mediocrity."[99]

Difficulty and expertise also have an intricate relationship. Experts "have adapted efficient ways to solve problems in their domains" and attack problems "by qualitatively different techniques" depending on the difficulty of the problem itself.[93] Sufficient surgical training develops cue recognition, planning and error recovery during cases complex enough to integrate both conscious (controlled) and unconscious (automated) actions.[100] An expert would thus be more responsive to the complexity of the task; able to conceptualize and plan a difficult operation at a high level of abstraction while integrating varying kinds of information.[101,102] Bond et al.[92] supports this idea in describing how surgeons exhibiting superior performance use "pattern recognition to be efficient at the mundane," and "recognize when the pieces do not fit," thereby adjusting their style of thinking to observe and adapt to evolving risks. Before the recent advances in developing surgical performance domains, recognizing these kinds of patterns, or "chunks" of knowledge, studied for years as part of successful performance in chess,[89,103] was criticized for yielding "little direction for improving education of medical students."[81]

Jerome Groopman, author of *How Doctors Think*[104] frames the advantages afforded by medical expertise as a break from traditional training: "studies show that while it usually takes twenty to thirty minutes in a didactic exercise for the senior doctor and students to arrive at a working diagnosis, an expert clinician typically forms a notion of what is wrong with the patient within twenty seconds." Groopman contends that practicing physicians pull cues in from all directions simultaneously — an adaptable and nonlinear mental process — but that the taught method is rigidly linear: "Medical students are taught that the evaluation of a patient should proceed in a discrete linear way; you first take the patient's history, then perform a physical examination, order tests and analyze the results. Only after all the data are compiled should you formulate hypotheses as to what might be wrong." In-context training for surgical residents is clearly crucial to build meaningful experience — despite the high cost to attending staff.[105]

Feltovich, Ford, and Hoffman,[50] creators of the TEMPEST model of expertise (Fig. 2), frame expert adaptability in terms of preplanned actions. Described as "predictive encoding," the authors emphasize the role of advanced cognitive skills and command of dynamic knowledge in demonstrating expert performance. Such strategies help the expert draw on relevant experience, select useful tools and information, and balance the numerous forces at play during an operation.

The TEMPEST model highlights the various experiences, goals, materials and strategies of familiar tasks that experts employ. Experts are driven and constrained by external forces like performance expectations, motives, and the rule of law. The "tail" — selection criteria, training, and professional standards — acts as a stabilizing force. The authors develop the model to represent completing a task.

The TEMPEST model does not differentiate between the relative quality of performance between experts, even though the various inputs and forces imply performance among each expert are different. No two experts will have the exact same background experience or motivations. A helpful sorting scheme for this purpose is presented by Ericsson and colleagues;[64] arranging performers into "Least Accomplished", "Good", "Best" and "Professional," but once again, these boundaries lack objective thresholds. Surgeons, similarly, who might not pass an "Olympic" or "excellent" bar of performance, may be recognized as experts nonetheless[16,106] as status may be conferred based on established case history, board certification, and leadership roles. Harald Mieg found that these kinds of signifiers of professional behavior serve as proxies for expertise in fields where "standards of best practice still need to be established." [107] Surgery fits the bill,[11,14] as clinicians "tend to conceptualize 'mastery' or 'expertise' as having conquered a specific set of skills, while other disciplines commonly associate these terms with a "continual learning state or perpetual devotion to improvement."[108] Achieving expertise, rather than practicing it, also reflects the strict social dichotomy between master (attending) and apprentice (resident) roles in surgical training. Such attributes are highly valued and enshrined in the professional model of medicine, but social signifiers are only a piece of the puzzle in pursuing

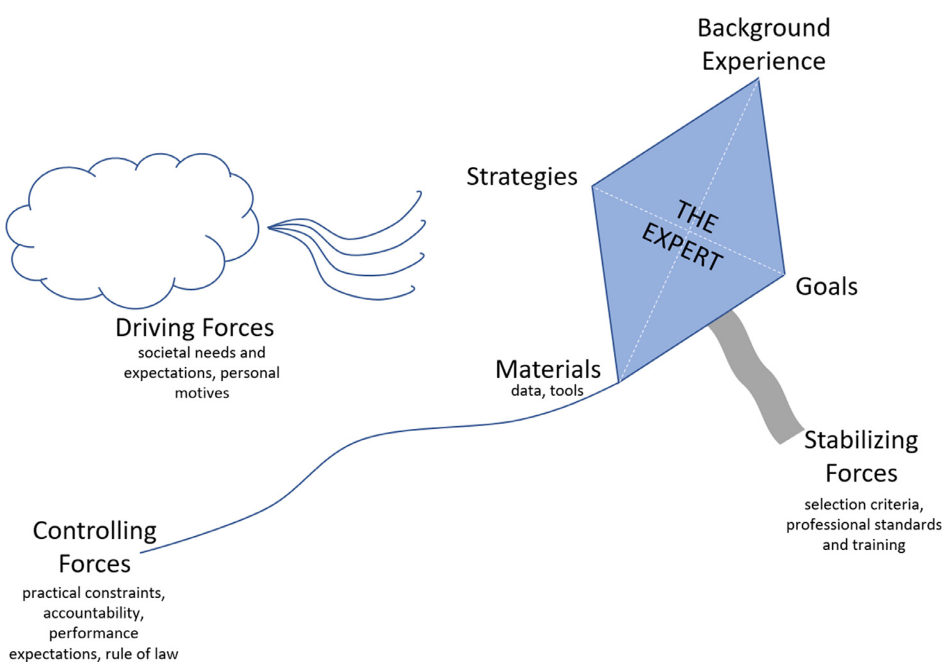


FIGURE 2. TEMPEST model describing the general framework of expertise (Adapted from Feltovich, Ford and Hoffman, 1997).

quantitative standards; they are notably absent in the TEMPEST model.

Expertise in surgery is thus challenging to pin down because it is used to refer both to the existing social hierarchy, and to the suite of practiced skills an individual can perform with consistency. In other words, surgical expertise continues to represent both professionalism and excellence; each of which is difficult, if not impossible, to define in their entirety. Professionalism is more easily understood; the Hippocratic Oath is recognized centuries over as part of the social contract of medicine. Measures of "excellence" as Mieg describes, however, continue to evolve and face scrutiny as technology-assisted measurements of skill and performance grow.[14] These latter advances may be more effective than social signifiers of professionalism at promoting guidelines of physician best practices.

Competency

Competency, like expertise, is dual-faceted. A competent surgeon, and a competent surgical performance, for example, may describe different ideas. George Miller[43] advocated that residents should achieve "competence," before performing and demonstrating skills on live patients. The US Accreditation Council for Graduate Medical Education (ACGME), meanwhile, is pushing for assessments across a series of job functions or "competencies," as part of the competency-based medical education (CBME) milestones project.[44] In light of these uses, a surgeon could be considered competent within a domain (i.e., consistently achieving a predefined rating in completing a procedure) or deemed competent overall (i.e., graduating from residency). For clarity, we define competency within a performance domain to mean meeting a quantifiable assessment threshold on a consistent and repeated basis. Referring to a clinician as a "competent surgeon," on the other hand, could also connote how that surgeon is perceived and trusted as a professional in the field, rather than how their skills have been quantifiably assessed. Firmly attaching competency to the underlying assessment content and context will help to promote consistent interpretation of a clinician's practiced skills as performance assessments become increasingly embedded in surgical training.

The colloquial understanding of competency as something an individual ought to conquer, rather than practice, much like expertise, is a driving force in defining competence as a testable threshold of performance. These kinds of performance-based competency assessments are currently under development for laparoscopy[109] and for various skills associated with professionalism.[110] Jelovsek et al.[8] provides a broad overview of reliability evidence for operative assessments. For a discussion on evolving educational approaches, consult Evans and Schenarts.[59]

Proficiency

Little emphasis is typically directed toward the difference between competency and proficiency. The Dreyfus Model of Skill Acquisition[111] (Table 1) offers a potential distinction, despite criticism for relying on intuition and omitting the utility of planning.[112]

In the Dreyfus Model, competency is exhibited by active decision making and categorizing information. Proficiency, on the other hand, is analogous to Miller's "shows how" stage and reflects gains in operative autonomy as residents become attending surgeons. This transition is earmarked by an increasing sense of responsibility commensurate with experience, and actively demonstrating abilities with decreasing oversight.

Competency in the Dreyfus Model can also be construed as the lowest suitable level of performance. Proficiency, in contrast, represents greater consistency and responsibility, albeit not yet at levels considered "expert." In analyzing surgical skill, competency would represent meeting a minimum required assessment and starting to "actively make decisions" with autonomy. George and colleagues,[113] in developing the Procedural Autonomy and Supervision System (PASS) on

TABLE 1. The "Dreyfus Five Stage Model of Mental Stages in Skill Acquisition"[111]

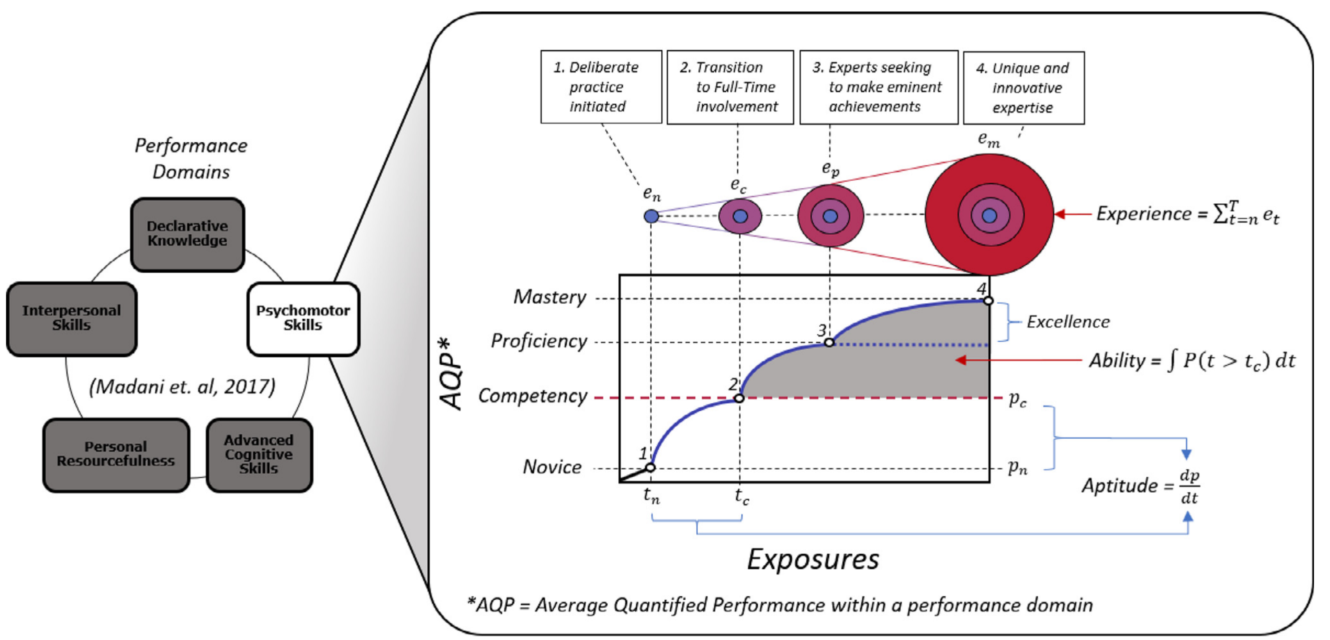
Stage	Autonomy	Mental Activities
1. Novice	Only feels responsible to follow the rules.	Follows specific rules for specific situations. Rules are not conditional.
2. Advanced Beginner	Still does not experience personal responsibility.	Begins to create and identify conditional rules. All decisions still follow rules.
3. Competent	Sense of responsibility arises from actively making decision.	Learns organizing principles. Information sorting by relevance begins.
4. Proficient	Sense of responsibility increases with experience.	Uses pattern recognition to assess what to do. Uses rules to determine how to do it.
5. Expert	Responsibility extends to others and the environment.	No analysis or planning. Pattern recognition extends to plan as well as action.

smartphones, reflect the first three stages of the Dreyfus model in describing progressions of autonomy as "show and tell," "active help", and "passive help," culminating in a later "supervision only" stage. Proficiency represents consistency in performance in excess of competent levels and would be characterized more by achieving repeated, stable, and efficient outcomes. Such thresholds would need to be drawn based on the procedure and task at hand and revisited as disruptive technologies like laparoscopy are introduced.

QUANTIFIED PERFORMANCE MODEL

The following model (Fig. 3) is newly proposed to represent each of these skills-based terms. The model represents performance as it develops over years of practice and exposure to challenging situations. Performance is represented as a quantifiable property over a surgeon's career and progresses through several stages, each with distinct exposure and responsibilities which change the rate of learning. Average performance is included to account for contextual and transitory factors. The model is adapted from the "three phases of development" model by Ericsson and Charness.[42] Competency is represented as an arbitrary, yet quantifiable threshold of performance and a gateway for increasing operative independence, with ability reserved to describe the sum of all actions for which a surgeon has already demonstrated competency – the integral under the continuous performance curve(s). Aptitude, on the other hand, is the rate (slope) of the curve from the time at which deliberate practice begins, to when a surgeon demonstrates competency on a regular basis. Experience, meanwhile, is represented as an expanding space and breadth of familiarity: the sum of unique exposures and instances of deliberate practice over the course of a career.

The Quantified Performance Model (QPM) is designed to strike a balance between literature describing expertise and colloquial uses for the various terms at issue in describing skill. Experience, for instance, while represented as a growing "bank" of lessons over the course of a career, is directly traceable to instances of deliberate practice and solving challenging cases as an autonomous professional. Excellent performance is superior to that of proficient and competent thresholds and associated with meaningful experiences. Because aptitude is typically used to describe the amount of effort or remedial training required at the beginning of surgical education, it is represented as the rate (slope) of the curve starting at the novice stage. The model assumes that the performance thresholds would be defined (and likely redefined) as surgical technique and approach evolves –



Note. Stages of performance curve from "Expert Performance: Its Structure and Acquisition" (Ericsson et. al, 1994). Average quantified performance would include valid measures of efficiency (e.g. fluidity, time to completion) and accuracy (e.g. successful completion).

FIGURE 3. Quantified performance model depicting relationship between common skill descriptors. Average Quantified Performance (AQP) is used to account for performance deviations due to transitory factors. The model combines both the Madani et al.[5] performance domains framework, as well as the Three-Stage model of expert performance development, put forth by Ericsson and Charness.[42]

much as the current Milestones project defining various surgical competencies evolves.[44] Although several performance thresholds may be drawn, the thresholds for each of the performance stages (competency, proficiency, mastery) are intended to represent developments over the course of a surgical career.

Consider, as an example, the University of Wisconsin-Madison Urology Department residency training approach. In PGY 2, resident operations are "completely supervised by an attending faculty. The attention is on learning proper surgical skills, instrument identification and handling, and the proper steps to simple surgical procedures. By the completion of that year, residents are expected to be able to perform all steps of simple surgical procedures with minimal guidance, but always under careful supervision."

Threshold performance at this stage is commensurate with the "novice" level. Operative autonomy is prohibited, and aptitude for general surgical practices is still being evaluated. Deliberate practice of in-vivo surgical skill is minimal, but grows over the course of the year; initializing a set of experiences from which to form basic "illness scripts." At this stage, it may be reasonable to use performance assessments to generate formative feedback, or to examine aptitude.

Between the third (PGY 3) and fourth (PGY 4) years, residents are expected to move up the performance curve. There are "increasing opportunities to conduct certain steps" as skills develop. Residents approach "conducting an entire procedure independently," albeit under direct supervision. Mentors provide "immediate feedback and remediation of any deficiencies" (PGY 4). Within the Quantified Performance Model, these improvements describe progress towards competent levels of performance in each domain. Instances of deliberate practice are increasing and building experience commensurate with the "active help" stage of PASS.

By the time a resident enters their 5th year, and as a chief resident, they are expected to "perform all steps of major urologic surgeries," and achieve "autonomy in performing basic surgical procedures." At this point, residents working with autonomy are not considered or consulted as experts, but their experience enables greater independence in the operating room and helps to inform those of lesser training. Objective assessments within the Quantified Performance Model may indicate a competent level in many of the performance domains for several procedures. Only after consistent passage, however, would that resident be considered "competent" for those procedures. Residents may also exhibit proficient levels of performance for a handful of simpler bedside or out-patient procedures. Completion of these five PGYs would be like completing the first three stages in the Dreyfus model of skill acquisition.

As residents become attending surgeons and transition to the new expectations of their full-time role, they would pursue a proficient level of performance. In contrast to residency, which places greater emphasis on focusing attention to enable periods of deliberate practice, time spent operating as an attending surgeon places greater emphasis on achieving efficient, positive outcomes, even in difficult situations. One study found that variation in operation time and complication rates during mammaplasty stabilized only after 12 years of active practice as an attending surgeon.[86] Still, lapses in planning, neglecting to complete or recall steps in an operation, or increased variability may indicate declines in performance at any stage in a surgeon's career. To enable consistent results in the operating room, these surgeons would need exhibit more efficient and robust error-management techniques. It is reasonable to assume that an attending performing at a master level have reached this point of stability. Within the proposed model, performance traits of an individual exhibiting mastery would serve as a template and resource to improve the rate of achieving proficiency for others in various difficult procedures.

As a surgeon progresses through these stages and strives for higher levels of performance, assessments would need to target more complex attributes of Madani's domains. Proficiency would need to be assessed through clinical simulations of increasingly difficult scenarios. At the same time, however, testing of previously surpassed performance thresholds would expose areas of needed practice to maintain skill with age, changes in life circumstance, or to demonstrate readiness to transition to another kind of surgery. Quantified performance testing throughout a surgical career may also serve to share expert strategies and mental models, while limiting patient exposure. Identifying features of performance commensurate with advanced tenure (i.e., expert and master surgeons) is an ongoing avenue of research.

Standard Lexicon

The standard lexicon proposed by the Quantified Performance Model (QPM) of surgical skills terminology (Fig. 3) focuses on defining performance as surgeons gain skills, age, and eventually retire. It applies to each of the five surgical performance domains (i.e., psychomotor skills, declarative knowledge, interpersonal skills, personal resourcefulness, and advanced cognitive skills),[5] and is particularly timely for increasing efforts to assess technical skill.[14] The model incorporates the role of deliberate practice in building expertise and paves the way to frame operative assessments as a consistent, repeated demonstration of performance rather than a one-time credential. The model does not, on the other hand, specify the content of these assessments. Described as an "instructional design problem," developing meaningful assessments is a continuing area of research subject to validity[47] and overall utility analysis.[114]

It is assumed that assessments will continue to adapt as technology and surgical techniques evolve. Procedure difficulty must also be considered. Planning, situational awareness, or other "advanced cognitive skills," for instance, may expose greater abilities in experienced clinicians than less complex assessments of salient psychomotor skills while suturing on simulated benchtop models.

The proposed model integrates potential assessment measures as an attempt to reach competency, and as a building block to proficiency (much like the Dreyfus model). Competency represents a transition to increasing responsibility and operative autonomy. The model reflects George Miller's focus on being able to show or perform skills. In addition, it uniquely frames performance as a repeated and consistent measure, to account for situational context and variation. It supports regular, repeated performance testing and reflects the ongoing push to demonstrate skills over time, even as they degrade due to advanced age or change in professional status. The model also supports the construction of various pass-fail thresholds, fitting well within the rhetoric of Madani, by encouraging active assessment for "a competent level of performance."[5]

Educational literature often uses similar terms, however, to detail a pedagogical approach. Consistently reaching a competent threshold, for example, is commonly referred to as "mastery" of the assessment topic.[78] In contrast, "master" surgeons are commonly described as those with substantial operative experience and involved in training efforts.[32] Performance, too, has held a unique educational meaning, connoting the final stage of one-on-one manual skill training.[115] Over time, the term "execution" has gained traction in these contexts,[116] with "performance" describing a more continuous scale of development[97] similar to the proposed model.

Over-simplifying any assessment framework poses a natural challenge to physician and patient autonomy – an evolving, yet fundamental tenet in the professional model of medicine.[117] If quantification of skill is implemented poorly and becomes anathema to the "secret glory" of medicine as a craft profession,[118] surgeons may opt instead to retreat to their respective corners; offering additional challenges to the already difficult prospect of competency-based medical education.[119] Worthwhile assessments could be overlooked before they have a chance to mature – undermining improvements to quality patient care and wasting valuable resources. Graham and Deary[38] argue that widespread adoption of such testing requires maturity of three things: robust understanding of skill, studies with subjective ratings as dependent variables, and an appropriate "working definition of superior surgical performance."

The proposed model in this paper offers a productive and traceable way to use surgical skills terminology in quantifying performance. The model integrates well with existing validity frameworks by promoting clear inferences and uses throughout a surgical career. To promote easier adoption among the medical community, the proposed definitions integrate existing literature and colloquial understanding.

Implications

This paper has focused on defining surgical skill terms that, despite their ubiquity, lack unique interpretations. Inconsistent terminology challenges development and comparison of valid metrics across studies by obscuring the relationship between measurements and underpinning constructs of skill. A novel model of terminology is proposed to assist in framing objective and feature-based surgical skills along a continuous scale of performance. Experience is represented as a growing "bank" of exposure to difficult situations; and includes the sum of instances of deliberate practice. Competency represents an arbitrary performance threshold, generally commensurate with graduation from residency and full-time involvement as an attending surgeon. Proficiency is characterized by decreasing variation and increasingly efficient outcomes. While attendings will pursue proficiency for the most difficult and complex operations, some residents may also reach proficient levels of performance for familiar operations and bedside procedures. Ability represents all performance a surgeon can offer in excess of a competent level, drawn as the integral under the performance curve. Aptitude is the rate at which one could achieve a competent performance level, given current pedagogical techniques. Mastery represents a performance threshold in excess of proficiency; characterized by excellent outcomes and novel techniques beyond those expected at proficient levels. Descriptors like elite and superior may be reserved for performances at the mastery level. Researchers can employ these representations to describe the scope of a nascent metric and map that measurement onto a unique construct of surgical skill.

Many of these definitions (consider competence, for example) depend on reaching a quantitative threshold of performance that has yet to reach maturity. Establishing validity evidence for such assessments in accordance with modern frameworks[45] is ongoing. The proposed model frames quantitative assessments within a continuous performance curve throughout stages of a surgical career. Each stage is associated with different training regiments and responsibilities, adapted from the "three phases of development" model by Ericsson and Charness.[42] To be considered competent or proficient to conduct an operation, a surgeon would need to consistently and repeatedly meet those relevant performance thresholds for relevant assessments in each surgical domain.[5]

As objective surgical skill analysis research continues to grow, consistent terminology will be critical in translating objective measures into formative feedback, and eventually, valid assessments. The Quantified Performance Model – accompanied by increasing abilities to measure performance – may aid in clarifying the duality of surgical expertise as an indication of professionalism and excellence. The model will also serve to frame the development of meaningful, targeted metrics. It may never be possible to quantify the artistry inherent in advanced surgery or define unique attributes of skill for complex operations. But, it may be possible to identify performance with enough specificity to discern surgeon progression from novice, to competent, proficient, and beyond. Measuring these thresholds with consistency could facilitate training, aptitude testing, placement, remediation, and timing of professional transition or retirement.

ACKNOWLEDGMENT

The authors acknowledge Sudha Pavuluri Quamme for her advice and assistance in completing this work.

REFERENCES

1. Birkmeyer JD, Finks JF, O'Reilly A, et al. Surgical skill and complication rates after bariatric surgery. *N Engl J Med.* 2013;369:1434–1442. https://doi.org/10.1056/NEJMsa1300625.

2. Hogg ME, Zenati M, Novak S, et al. Grading of surgeon technical performance predicts postoperative pancreatic fistula for pancreaticoduodenectomy independent of patient-related variables. *Ann Surg.* 2016;264:482–489. https://doi.org/10.1097/SLA.0000000000001862.

3. Laufer S, D'Angelo A-LD, Kwan C, et al. Rescuing the Clinical Breast Examination. *Ann Surg.* 2016;XX:1–6. https://doi.org/10.1097/SLA.0000000000002024.

4. Choudhry NK, Fletcher RH, Soumerai SB. Improving Patient Care Systematic Review: The Relationship between Clinical Experience and Quality of Health Care. 2005.

5. Madani A, Vassiliou MC, Watanabe Y, et al. What are the principles that guide behaviors in the operating room? *Ann Surg.* 2017;265:255–267. https://doi.org/10.1097/SLA.0000000000001962.

6. Yule S, Flin R, Paterson-Brown S, Maran N. Non-technical skills for surgeons in the operating room: a review of the literature. *Surgery*. 2006;139:140–149. https://doi.org/10.1016/j.surg.2005.06.017.

7. Yule S, Paterson-Brown S. Surgeons' non-technical skills. *Surg Clin N Am.* 2018;92:37–50. https://doi.org/10.1016/j.suc.2011.11.004.

8. Jelovsek JE, Kow N, Diwadkar GB. Tools for the direct observation and assessment of psychomotor skills in medical trainees: a systematic review. *Med Educ.* 2013;47:650–673. https://doi.org/10.1111/medu.12220.

9. Cook DA, Zendejas B, Hamstra SJ, Hatala R, Brydges R. What counts as validity evidence? Examples and prevalence in a systematic review of simulation-based assessment. *Adv Heal Sci Educ.* 2014;19:233–250. https://doi.org/10.1007/s10459-013-9458-4.

10. Cook DA, Reed DA. Appraising the quality of medical education research methods. *Acad Med.* 2015;90:1067–1076. https://doi.org/10.1097/ACM.0000000000000786.

11. Maier-Hein L, Vedula S, Speidel S, et al. Surgical data science: enabling next-generation surgery. 2017. https://arxiv.org/abs/1701.06482.

12. Rutherford DN, D'Angelo A-LD, Law KE, Pugh CM. Advanced engineering technology for measuring performance. *Surg Clin North Am.* 2015;95:813–826. https://doi.org/10.1016/j.suc.2015.04.005.

13. Scott DJ, Cendan JC, Pugh CM, Minter RM, Dunnington GL, Kozar RA. The changing face of surgical education: simulation as the new paradigm. *J Surg Res.* 2008;147:189–193. https://doi.org/10.1016/j.jss.2008.02.014.

14. Vedula SS, Ishii M, Hager GD. Objective assessment of surgical technical skill and competency in the operating room. *Annu Rev Biomed Eng.* 2017;19:301–325. https://doi.org/10.1146/annurev-bioeng-071516-044435.

15. Hull L, Arora S, Aggarwal R, Darzi A, Vincent C, Sevdalis N. The impact of nontechnical skills on technical performance in surgery: a systematic review. *J Am Coll Surg.* 2012;214:214–230. https://doi.org/10.1016/j.jamcollsurg.2011.10.016.

16. Bell RH. Why Johnny cannot operate. *Surgery*. 2009;146:533–542. https://doi.org/10.1016/j.surg.2009.06.044.

17. Azari DP, Pugh CM, Laufer S, et al. Evaluation of simulated clinical breast exam motion patterns

using marker-less video tracking. *Hum Factors*. 2015;58:427–440. https://doi.org/10.1177/0018720815613919.

18. Azari DP, Frasier LL, Quamme SRP, et al. Modeling surgical technical skill using expert assessment for automated computer rating. *Ann Surg*. 2017;XX:1. https://doi.org/10.1097/SLA.0000000000002478.

19. Glarner CE, Hu YY, Chen CH, et al. Quantifying technical skills during open operations using video-based motion analysis. *Surg (United States)*. 2014;156:729–734. https://doi.org/10.1016/j.surg.2014.04.054.

20. Frasier LL, Azari DP, Ma Y, et al. A marker-less technique for measuring kinematics in the operating room. *Surgery*. 2016. https://doi.org/10.1016/j.surg.2016.05.004.

21. Radwin RG, Azari DP, Frasier L, et al. A marker-less video tracking approach for quantifying open surgical skill. *Panel session presented at the Int Ann Meet Hum Factors Ergon Soc*. 2014.

22. Datta V, Chang A, Mackay S, Darzi A. The relationship between motion analysis and surgical technical assessments. *Am J Surg*. 2002;184:70–73. https://doi.org/10.1016/S0002-9610(02)00891-7.

23. Regenbogen SE, Greenberg CC, Studdert DM, Lipsitz SR, Zinner MJ, Gawande Aa. Patterns of technical error among surgical malpractice claims. *Ann Surg*. 2007;246:705–711. https://doi.org/10.1097/SLA.0b013e31815865f8.

24. Rogers SO, Gawande AA, Kwaan M, et al. Analysis of surgical errors in closed malpractice claims at 4 liability insurers. *Surgery*. 2006;140:25–33. https://doi.org/10.1016/j.surg.2006.01.008.

25. Law Forsyth K, DiMarco SM, Jenewein CG, et al. Do errors and critical events relate to hernia repair outcomes? *Am J Surg*. 2017;213:652–655. https://doi.org/10.1016/j.amjsurg.2016.11.020.

26. Nathwani JN, Glarner CE, Law KE, et al. Integrating postoperative feedback into workflow: perceived practices and barriers. *J Surg Educ*. 2017;74:406–414. https://doi.org/10.1016/j.jsurg.2016.11.001.

27. O'Neil HF, Perez RS, Baker EL. In: O'Neil HF, Perez RS, Baker EL, eds. Teaching and Measuring Cognitive Readiness, 9781461475, Boston, MA: Springer US; 2014. https://doi.org/10.1007/978-1-4614-7579-8.

28. Pugh CM, DaRosa Da. Use of cognitive task analysis to guide the development of performance-based assessments for intraoperative decision making. *Mil Med*. 2013;178(10 Suppl):22–27. https://doi.org/10.7205/MILMED-D-13-00207.

29. Dedy NJ, Fecso AB, Szasz P, Bonrath EM, Grantcharov TP. Implementation of an effective strategy for teaching nontechnical skills in the operating room. *Ann Surg*. 2016;263:937–941. https://doi.org/10.1097/SLA.0000000000001297.

30. Wiegmann DA, ElBardissi AW, Dearani JA, Daly RC, Sundt TM. Disruptions in surgical flow and their relationship to surgical errors: an exploratory investigation. *Surgery*. 2007;142:658–665. https://doi.org/10.1016/j.surg.2007.07.034.

31. Moorthy K, Munz Y, Adams S, Pandey V, Darzi A. A human factors analysis of technical and team skills among surgical trainees during procedural simulations in a simulated operating theatre. *Ann Surg*. 2005;242:631–639. https://doi.org/10.1097/01.sla.0000186298.79308.a8.

32. Cuschieri A, Francis N, Crosby J, Hanna GB. What do master surgeons think of surgical competence and revalidation? *Am J Surg*. 2001;182:110–116. https://doi.org/10.1016/S0002-9610(01)00667-5.

33. Greenberg CC, Ghousseini HN, Pavuluri Quamme SR, et al. A statewide surgical coaching program provides opportunity for continuous professional development. *Ann Surg*. 2018;267(5):868–873. https://doi.org/10.1097/SLA.0000000000002341.

34. Perdanasari AT, Hollier LH. Review of "What are the principles that guide behaviors in the operating room? Creating a framework to define and measure performance" by Madani A, Vassiliou MC, Watanabe Y, Al-Halabi B, Al-Rowais MS, Deckelbaum DL, Fried GM, Feldman LS in Ann Surg 265. *J Craniofac Surg*. 2017;28:842. https://doi.org/10.1097/SCS.0000000000003424.

35. Ericsson KA. Recent advances in expertise research: a commentary on the contributions to the special issue. *Appl Cogn Psychol*. 2005;19:233–241. https://doi.org/10.1002/acp.1111.

36. Ginsburg S, Regehr G, Lingard L. Basing the evaluation of professionalism on observable behaviors: a cautionary tale. *Acad Med*. 2004;79(10 Suppl):S1–S4. https://doi.org/10.1097/00001888-200410001-00001.

37. Reznick RK, MacRae H. Teaching Surgical Skills — Changes in the Wind. Cox M, Irby DM, eds. *N Engl J Med*. 2006;355:2664–2669. https://doi.org/10.1056/NEJMra054785.

38. Graham KS, Deary IJ. A role for aptitude testing in surgery? *J R Coll Surg Edinb.* 1991;36:70–74 http://www.ncbi.nlm.nih.gov/pubmed/2051420.

39. Charness N, Tuffiash M. The role of expertise research and human factors in capturing, explaining, and producing superior performance. *Human Factors*. 2008;50(3):427–432. https://doi.org/10.1518/001872008X312206.

40. Murinson BB, Agarwal AK, Haythornthwaite JA. Cognitive expertise, emotional development, and reflective capacity: clinical skills for improved pain care. *J Pain*. 2008;9:975–983. https://doi.org/10.1016/j.jpain.2008.07.010.

41. Krawczyk D, Bartlett J, Kantarcioglu M, Hamlen K, Thuraisingham B. Measuring expertise and bias in cyber security using cognitive and neuroscience approaches. In: *2013 IEEE International Conference on Intelligence and Security Informatics*. IEEE; 2013:364-367. https://doi.org/10.1109/ISI.2013.6578859.

42. Ericsson KA, Charness N. Expert performance: its structure and acquisition. *Am Psychol.* 1994;49:725–747. https://doi.org/10.1037/0003-066X.49.8.725.

43. Miller GE. The assessment of clinical skills/competence/performance. *Acad Med.* 1990;65:S63–S67. https://doi.org/10.1097/00001888-199009000-00045.

44. ACGME. Accrediation Council for Graduate Medical Eduction (ACGME). http://www.acgme.org/What-We-Do/Accreditation/Milestones/. Published 2013. Accessed 1 May 2018.

45. Kane M. The argument-based approach to validation. *School Psych Rev.* 2013;42:448–457.

46. Cook DA, Brydges R, Ginsburg S, Hatala R. A contemporary approach to validity arguments: a practical guide to Kane's framework. *Med Educ.* 2015;49:560–575. https://doi.org/10.1111/medu.12678.

47. Kane M. Validation. In: Brennan R, ed. Educational Measurement, 4th ed., Westport, CT: American Council on Education and Praeger; 2006:17–64.

48. Davids K, Button C, Bennett S. Dynamics of Skill Acquition - A Constraints-Led Approach. Champaign, IL, USA: Human Kinetics; 2008.

49. Ericsson KA. Deliberate practice and the acquisition and maintenance of expert performance in medicine and related domains. *Acad Med.* 2004; 79(10 Suppl):S70–S81. https://doi.org/10.1097/00001888-200410001-00022.

50. Feltovich PJ, Ford KM, Hoffman RR. Exertise in Context. Cambridge: The MIT Press; 1997.

51. Francis DMA. Surgical decision making. *ANZ J Surg*. 2009;79:886–891. https://doi.org/10.1111/j.1445-2197.2009.05139.x.

52. Galton F. Hereditary Genius. 2nd ed. Cleveland, OH: Meridian; 1892.

53. Stratton SM, Liu Y-T, Hong SL, Mayer-Kress G, Newell KM. Snoddy (1926) revisited: time scales of motor learning. *J Mot Behav*. 2007;39:503–515. https://doi.org/10.3200/jmbr.39.6.503-516.

54. Fitts PM, Posner MLI. Human Performance. Belmont, CA: Brooks/Cole Publishing Co; 1967. http://www.worldcat.org/title/human-performance/oclc/00480476.

55. Anderson JR. Acquisition of cognitive skill. *Psychol Rev*. 1982;89:369–406. https://doi.org/10.1037/0033-295X.89.4.369.

56. Tyler JA, Clive KS, White CE, Beekley AC, Blackbourne LH. Current US military operations and implications for military surgical training. *J Am Coll Surg*. 2010;211:658–662. https://doi.org/10.1016/j.jamcollsurg.2010.07.009.

57. Kelly JF, Ritenour AE, McLaughlin DF, et al. Injury severity and causes of death from operation Iraqi freedom and operation enduring freedom: 2003−2004 versus 2006. *J Trauma Acute Care Surg*. 2008;64:S21–S26. https://doi.org/10.1097/TA.0b013e318160b9fb. discussion S26-S27.

58. Perez RS, Skinner A, Weyhrauch P, et al. Prevention of surgical skill decay. *Mil Med*. 2013;178(10 Suppl):76–86. https://doi.org/10.7205/MILMED-D-13-00216.

59. Evans CH, Schenarts KD. Evolving educational techniques in surgical training. *Surg Clin North Am*. 2016;96:71–88. https://doi.org/10.1016/j.suc.2015.09.005.

60. Mackenzie CF, Watts D, Patel RR, et al. Sensor-free computer vision hand-motion entropy and video analysis of technical performance during open vascular surgery: proof of concept for methodology. *J Am Coll Surg*. 2016;223:e63. https://doi.org/10.1016/j.jamcollsurg.2016.08.166.

61. Aggarwal R, Grantcharov T, Moorthy K, et al. An evaluation of the feasibility, validity, and reliability of laparoscopic skills assessment in the operating room. *Ann Surg*. 2007;245:992–999. https://doi.org/10.1097/01.sla.0000262780.17950.e5.

62. Richstone L, Schwartz MJ, Seideman C, Cadeddu J, Marshall S, Kavoussi LR. Eye metrics as an objective assessment of surgical skill. *Ann Surg*. 2010;252:177-182. https://doi.org/10.1097/SLA.0b013e3181e464fb.

63. Ahmidi N, Poddar P, Jones JD, et al. Automated objective surgical skill assessment in the operating room from unstructured tool motion in septoplasty. *Int J Comput Assist Radiol Surg*. 2015;10:981-991. https://doi.org/10.1007/s11548-015-1194-1.

64. Ericsson KA, Krampe RT, Tesch-Römer C. The role of deliberate practice in the acquisition of expert performance. *Psychol Rev*. 1993;100:363-406. https://doi.org/10.1037/0033-295X.100.3.363.

65. Nasca TJ, Day SH, Amis ES. The new recommendations on duty hours from the ACGME task force. *N Engl J Med*. 2010;363:e3. https://doi.org/10.1056/NEJMsb1005800.

66. Bharathan R, Aggarwal R, Darzi A. Operating room of the future. *Best Pract Res Clin Obstet Gynaecol*. 2013;27:311-322. https://doi.org/10.1016/j.bpobgyn.2012.11.003.

67. Moglia A, Ferrari V, Morelli L, et al. Distribution of innate ability for surgery amongst medical students assessed by an advanced virtual reality surgical simulator. *Surg Endosc Other Interv Tech*. 2014;28:1830-1837. https://doi.org/10.1007/s00464-013-3393-6.

68. Roitberg B, Banerjee P, Luciano C, et al. Sensory and motor skill testing in neurosurgery applicants: a pilot study using a virtual reality haptic neurosurgical simulator. *Neurosurgery*. 2013;73(SUPPL. 4):116-121. https://doi.org/10.1227/NEU.0000000000000089.

69. Buckley CE, Kavanagh DO, Nugent E, Ryan D, Traynor OJ, Neary PC. The impact of aptitude on the learning curve for laparoscopic suturing. *Am J Surg*. 2014;207:263-270. https://doi.org/10.1016/j.amjsurg.2013.08.037.

70. McDonald P. Training for surgeons after the year 2000. *J R Soc Med*. 1998;91:401.

71. Schendel JD, Shields JL, Katz MG. Retention of motor skills: review. *Tech Pap*. 1974: 50.

72. Hislop SJ, Hsu JH, Narins CR, et al. Simulator assessment of innate endovascular aptitude versus empirically correct performance. *J Vasc Surg*. 2006;43:47-55. https://doi.org/10.1016/j.jvs.2005.09.035.

73. Willis RE, Gomez PP, Ivatury SJ, Mitra HS, Van Sickle KR. Virtual reality simulators: valuable surgical skills trainers or video games? *J Surg Educ*. 2014;71:426-433. https://doi.org/10.1016/j.jsurg.2013.11.003.

74. Cuschieri Sir Alfred. Lest we forget the surgeon. *Semin Laparosc Surg*. 2003;10:141-148. https://doi.org/10.1016/j.racsoc.2004.11.008.

75. Groenier M, Groenier KH, Miedema HAT, Broeders IAMJ. Perceptual speed and psychomotor ability predict laparoscopic skill acquisition on a simulator. *J Surg Educ*. 2015;72:1224-1232. https://doi.org/10.1016/j.jsurg.2015.07.006.

76. Szasz P, Bonrath EM, Louridas M, et al. Setting performance standards for technical and nontechnical competence in general surgery. *Ann Surg*. 2016;266:1. https://doi.org/10.1097/SLA.0000000000001931.

77. Mattar S, Alseidi A, Jones D, Jeyarajah D, et al. General surgery residency inadequately prepares trainees for fellowship. *Ann Surg*. 2013;258:440-449. https://doi.org/10.1097/SLA.0b013e3182a191ca.

78. McGaghie WC, Miller GE, Sajid AW, Telder TW. Competency-based curriculum development in medical education: an introduction. *Public Health Pap*. 1978: 1-90 http://eric.ed.gov/?id=ED168447%5Cnhttp://apps.who.int/iris/handle/10665/39703%5Cnhttp://whqlibdoc.who.int/php/WHO_PHP_68.pdf.

79. Thomas W. Teaching and assessing surgical competence. *Ann R Coll Surg Engl*. 2006;88:429-432. https://doi.org/10.1308/003588406X116927.

80. Hu YY, Peyre SE, Arriaga AF, Roth EM, Corso KA, Greenberg CC. War stories: a qualitative analysis of narrative teaching strategies in the operating room. *Am J Surg*. 2012;203:63-68. https://doi.org/10.1016/j.amjsurg.2011.08.005.

81. Norman G, Eva K, Brooks L, Hamstra S. Expertise in medicine and surgery. Ericsson KA, Charness N, Feltovich PJ, Hoffman RR, eds. Cambridge Handb Expert Expert Perform; 2006: 339-354.

82. Schmidt HG, Norman GR, Boshuizen HP. A cognitive perspective on medical expertise: theory and implication. *Acad Med*. 1990;65:611-621. https://doi.org/10.1097/00001888-199010000-00001.

83. Holmes PS, Collins DJ. The PETTLEP approach to motor imagery: a functional equivalence model for sport psychologists. *J Appl Sport Psychol*.

2001;13:60-83. https://doi.org/10.1080/10413200109339004.

84. Louridas M, Bonrath EM, Sinclair DA, Dedy NJ, Grantcharov TP. Randomized clinical trial to evaluate mental practice in enhancing advanced laparoscopic surgical performance. *Br J Surg*. 2015;102:37-44. https://doi.org/10.1002/bjs.9657.

85. Sweller J. Evolutionary bases of human cognitive architecture. In: Proceeding Fourth International Workshop on Computing Education Research - ICER '08; 2008:1-2. https://doi.org/10.1145/1404520.1404521.

86. Carty MJ, Chan R, Huckman R, Snow D, Orgill DP. A detailed analysis of the reduction mammaplasty learning curve: a statistical process model for approaching surgical performance improvement. *Plast Reconstr Surg*. 2009;124:706-714. https://doi.org/10.1097/PRS.0b013e3181b17a13.

87. Staats BR, KC DS, Gino F. Maintaining beliefs in the face of negative news: the moderating role of experience. *Manage Sci*. 2018;64:804-824. https://doi.org/10.1287/mnsc.2016.2640.

88. Gino F. Why too much experience can backfire. *Sci Am*. April 2018. https://www.scientificamerican.com/article/why-too-much-experience-can-backfire/.

89. Simon H, Chase W. Skill in chess. *Am Sci*. 1973;61:394-403.

90. Chi MTH. Theoretical perspectives, methodological appraoches, and trends in the study of expertise. *Expert Math Instr*. 2011:17-39. https://doi.org/10.1007/978-1-4419-7707-6.

91. Tofel-Grehl C, Feldon DF. Cognitive task analysis-based training: a meta-analysis of studies. *J Cogn Eng Decis Mak*. 2013;7:293-304. https://doi.org/10.1177/1555343412474821.

92. Bond W, Kuhn G, Binstadt E, et al. The use of simulation in the development of individual cognitive expertise in emergency medicine. *Acad Emerg Med*. 2008;15:1037-1045. https://doi.org/10.1111/j.1553-2712.2008.00229.x.

93. Prietula MJ, Feltovich PJ, Marchak F. Factors influencing analysis of complex cognitive tasks: a framework and example from industrial process control. *Hum Factors*. 2000;42:56-74. https://doi.org/10.1518/001872000779656589.

94. Palter VN, Grantcharov TP. Individualized deliberate practice on a virtual reality simulator improves technical performance of surgical novices in the operating room: a randomized controlled trial. *Ann Surg*. 2014;259:443-448. https://doi.org/10.1097/SLA.0000000000000254.

95. Hashimoto DA, Sirimanna P, Gomez ED, et al. Deliberate practice enhances quality of laparoscopic surgical performance in a randomized controlled trial: from arrested development to expert performance. *Surg Endosc*. 2015;29:3154-3162. https://doi.org/10.1007/s00464-014-4042-4.

96. Spruit EN, Band GPH, Hamming JF, Ridderinkhof KR. Optimal training design for procedural motor skills: a review and application to laparoscopic surgery. *Psychol Res*. 2014;78:878-891. https://doi.org/10.1007/s00426-013-0525-5.

97. Jeffree RL, Clarke RM. Ten tips for teaching in the theatre tearoom: shifting the focus from teaching to learning. *World J Surg*. 2010;34:2518-2523. https://doi.org/10.1007/s00268-010-0719-6.

98. Pugh CM. Getting a sense for the surgical touch [Video file]. https://www.tedmed.com/talks/show?id=2929972014.

99. Weinberggeer SE, Duffey FD, Cassel CK. "Practice makes perfect" . . . or does it? *Ann Intern Med*. 2005;142:302-303.

100. Pugh CM, Santacaterina S, DaRosa DA, Clark RE. Intra-operative decision making: more than meets the eye. *J Biomed Inform*. 2011;44:486-496. https://doi.org/10.1016/j.jbi.2010.01.001.

101. van Merriënboer JJG, Clark RE, de Croock MBM. Blueprints for complex learning: the 4C/ID-model. *Educ Technol Res Dev*. 2002;50:39-61. https://doi.org/10.1007/BF02504993.

102. Ruis AR, Rosser AA, Quandt-Walle C, Nathwani JN, Shaffer DW, Pugh CM. The hands and head of a surgeon: modeling operative competency with multimodal epistemic network analysis. *Am J Surg*. 2017. https://doi.org/10.1016/j.amjsurg.2017.11.027.

103. Burns BD. The effects of speed on skilled chess performance. *Psychol Sci*. 2004;15:442-447. https://doi.org/10.1111/j.0956-7976.2004.00699.x.

104. Groopman J. How Doctors Think. Mariner Books; 2008.

105. Babineau TJ, Becker J, Gibbons G, et al. The "cost" of operative training for surgical residents. *Arch Surg*. 2004;139:366-369. https://doi.org/10.1001/archsurg.139.4.366. discussion 369-70.

106. Alleman AM, Al-Assaf AF. Have you wondered about your colleague's surgical skills? *Am J Med*

Qual. 2005;20:78–82. https://doi.org/10.1177/1062860604273746.

107. Mieg HA. Two factors of expertise? Excellence and professionalism of environmental experts. *High Abil Stud*. 2009;20:91–115. https://doi.org/10.1080/13598130902860432.

108. Greenberg CC, Klingensmith ME. The continuum of coaching: opportunities for surgical improvement at all levels. *Ann Surg*. 2015;262:217–219. https://doi.org/10.1097/SLA.0000000000001290.

109. Miskovic D, Ni M, Wyles SM, et al. Is competency assessment at the specialist level achievable? A study for the national training programme in laparoscopic colorectal surgery in England. *Ann Surg*. 2013;257:476–482. https://doi.org/10.1097/SLA.0b013e318275b72a.

110. Hochberg MS, Kalet A, Zabar S, Kachur E, Gillespie C, Berman RS. Can professionalism be taught? Encouraging evidence. *Am J Surg*. 2010;199:86–93. https://doi.org/10.1016/j.amjsurg.2009.10.002.

111. Dreyfuss SE, Dreyfus HL. A five-stage model of the mental activities involved in directed skill acquisition. *Oper Res Cent*. 1980: 1–18. doi:ADA084551.

112. Peña A. The Dreyfus model of clinical problem-solving skills acquisition: a critical perspective. *Med Educ Online*. 2010;15:1–11. https://doi.org/10.3402/meo.v15i0.4846.

113. George BC, Teitelbaum EN, Meyerson SL, et al. Reliability, validity, and feasibility of the zwisch scale for the assessment of intraoperative performance. *J Surg Educ*. 2014;71:e90–e96. https://doi.org/10.1016/j.jsurg.2014.06.018.

114. van der Vleuten CPMM, Schuwirth LWTT. Assessing professional competence: from methods to programmes. *Med Educ*. 2005;39:309–317. https://doi.org/10.1111/j.1365-2929.2005.02094.x.

115. Peyton WM. Teaching in the theathre. *Teach Learn Med Pract*. 1998: 171–180.

116. Munster T, Stosch C, Hindrichs N, Franklin J, Matthes J. Peyton's 4-steps-approach in comparison: medium-term effects on learning external chest compression - a pilot study. *GMS J Med Educ*. 2016;33:Doc60. https://doi.org/10.3205/zma001059.

117. ABIM Foundation. Medical professionalism in the new millennium: a physician charter. *Ann Intern Med*. 2002;136:243–246. https://doi.org/10.7326/0003-4819-136-3-200202050-00012.

118. Donabedian A. The quality of care. How can it be assessed? *JAMA J Am Med Assoc*. 1988;260:1743–1748. https://doi.org/10.1001/jama.260.12.1743.

119. Touchie C, Ten Cate O. The promise, perils, problems and progress of competency-based medical education. *Med Educ*. 2016;50:93–100. https://doi.org/10.1111/medu.12839.

SUPPLEMENTARY INFORMATION

Supplementary material associated with this article can be found in the online version at doi:10.1016/j.jsurg.2019.02.010.