The computerized objective assessment of surgical skills: Considerations for counting the number of movements

Lawrence Grierson[1]

[1]McMaster University Faculty of Health Sciences

October 6, 2020

Abstract

Motion capture and analysis techniques are emerging in the surgical education and surgical education research literature as viable ways to augment the assessment of technical skills. In particular, these methods provide an opportunity to reveal objective information about the efficiency of surgical procedures, above and beyond the accuracy of procedural outcomes. One assessment that is very prevalent in the literature are counts of the number of movements a surgeon makes in completing a technical performance. In this commentary, the number of movements metric is explored from kinesiology and engineering perspectives; two disciplines that have contributed heavily to the development of rigorous motion analysis methods. Furthermore, the assumption that skill efficiency improves linearly as a learner progresses along the continuum of expertise is challenged. While movement efficiency does certainly improve, this assumption does not necessarily capture the way that learners flexibly prioritize particular aspects of performance in the intermediate stages of skill learning. By way of this commentary, important a priori decisions that should proceed effective motion capture and analysis are highlighted, a call for the standardization of procedures is made, and an opportunity to better understand the way that computerized movement analysis techniques may contribute (or be detrimental) to competency constructs in surgical education and assessment is realized.

Introduction

The medical education community has come to agree that direct measures of physician ability need to replace time-in-training as the main indicators of physician competence.[1, 2] Under the banner of competency-based education (CBE), the fundamental idea is that trainees will develop knowledge, skills, and behaviours at different rates, and that only through measurement of trainee abilities with reference to standards of achievement (i.e., milestones), can it be determined when a learner is competent enough to progress into independent practice.[3] In the surgical specialties, where many specific competencies are concerned with precision technical skills, this shift has been served by the development of a number of effective assessment approaches; including, the Objective Structured Assessment of Technical Skills (OSATS),[4] the Global Operative Assessment of Laparoscopic Skills (GOALS),[5]and the McGill Inanimate System for Training and Evaluation in Laparoscopic Surgery (MISTELS).[6] While each of these tools is psychometrically robust, they each still rely on subjective appraisals from qualified experts, which purports to be problematic as the time required to complete assessments encroaches deeper and deeper into the schedules of clinician-educators.[3]

As a consequence, the medical education community has explored the use of measurements from computerized systems, which have the ability to provide objective information to evaluators about a procedure, as a potential avenue to improving the process of assessing technical skills. This approach to assessment encompasses a wide variety of technologies, which are capable of measuring a number of surrogates of performance quality. In general, these types of measurements provide rich digitized metrics about the outcomes of a clinical performance. For instance, technologies that measure forces have been used to reveal the tensile strength of surgical knots,[7, 8] the consistency and accuracy of acupuncture needling,[9] and one's surgical expertise in bone-drilling tasks.[10, 11] However, in addition to outcomes, computerized measures can also provide objective assessments about the efficiency with which a skill is performed. Efficiency is often an important perspective on skill performance, as its optimization can have important impacts on patient safety and hospital operations; including, reducing patient exposures to radiation[12, 13] and the potential for infection,[14] improving the patient-to-patient flow of the operating theatre,[15, 16] and protecting the physician from fatigue.[17, 18]

Historically, the efficiency of clinical performance has usually been inferred through measurements of the time it takes to complete a procedure,[19-21] but more recently, medical educators and researchers have turned to kinematic measurements derived from motion capture analyses to assess technical skill efficiency.[22-25] Motion capture is used widely in a number of industries (i.e., filmmaking, video game development, military and sports) and for an array of different purposes (i.e., gait analysis, facial recognition, and computer animation). The process involves affixing markers to a performer's body, hands, or tools during a performance. These markers emit signals (i.e., electromagnetic, optoelectric, inertial, acoustic, etc.) that allow for their position to be recorded several times a second, permitting the determination of many things; including, the total distance traveled by the performer's limbs, the kinematic characteristics (i.e., displacement, velocity, acceleration) of the clinical movements, and the trial-by-trial spatial variability with which procedures are performed.

One particular measure of procedural efficiency derived from motion capture techniques that has become exceedingly popular in the surgical education literature is a count of the "*number of movements* " made by the practitioner during a technical skill performance.[22, 26, 27] The conceptual idea is that the technique performed with fewer movements is smoother, better planned, and more efficient, and thus more indicative of an expert clinician. Indeed, construct and concurrent validation studies have revealed the *"number of movements"* metric to differentiate expert and novice performances in a way that aligns with the ratings provided by subjective assessment scales.[26-28] To enact this measure as part of a competency-based technical skill education program, one may envision requiring learners to reduce their performances to below certain "*number of movements* " milestones in a simulation-based context before moving on to new entrusted activities in the criterion clinical environment. Although this approach to assessment has some appeal for its ability to ensure a certain degree of skill efficiency, above and beyond skill accuracy, before a learner progresses, it is not without its potential pitfalls. Specifically, the outcomes of motion capture and analysis are highly dependent on system-level decisions that are inputted by the assessor prior to the data collection period. Across cohorts of trainees, inconsistency in these decisions may have a major impact on the ability to distinguish learners' capabilities with respect to standards of competence, or even on our ability to set standards at all. Moreover, the relationship between number of movements and expertise is not necessary indirectly linear. As such, it is important that educators remember that strategic approaches to learning and the contexts of performance can have significant influence on the number of movements a trainee performs while practicing.

In this commentary, I consider the processes of counting and interpreting "*number of movements* " data from the perspectives of kinesiology and engineering science, which have contributed heavily to the rigorous application of motion analysis methods. In doing so, my goal is to advocate for consensus agreement on the setting of standards for motion capture in clinical assessment, which will allow for the determination of accurate and appropriate metrics that will ensure consistency in the way that measures of efficiency are utilized in competency assessment across surgical education programs.

Defining Movements

The first consideration when counting number of movements as a measure of performance efficiency is to understand the nature of the skills that are being evaluated. In this regard, the field of motor control kinesiology has typically classified actions as composed of combinations of movements that can be defined as either discrete or continuous. Discrete movements are those that have a recognizable beginning and end, like throwing a ball, turning a doorknob, or flipping a light switch; they are usually essential to skills that rely of the precise production of a distinct outcome. Continuous movements, on the other hand, have no recognizable beginning or end, and will continue until they are stopped arbitrarily by the performer. Skills composed of continuous movements can have precision constraints but are often concerned with the maintenance of an ongoing action. These skills include activities such as walking, swimming, and cycling.[29] In the surgical-medical domain, the clinical technical performances of interest can more often than not be characterized as serial actions. This refers to skills that are made up of a number of discrete movements that must occur in a very particular order. These actions can appear continuous, but usually have distinct components with very definitive beginnings and ends.[29] As such, any one clinical technical skill can be conceptualized in terms of the minimum number of discrete movements that would be needed for its successful execution. However, as the number of movements assessment construct suggests, this minimum is not always achieved.

That performances can contain movements in excess of the minimum required by the task is fundamental to the use of the number of movements metric as a measure of efficiency. Simply put, performances often contain errors or imperfect actions, which require corrections; and each erred movement and subsequent correction constitutes the production of additional movements. In this way, the hallmark error volume associated with novice performances has led to the natural assumption that new trainees will perform procedures with more movements, and therefore, more inefficiently. Given this position, the challenge for assessors of surgical technical skills is to determine where one movement within a procedure ends and the next one begins. In a motion capture and analysis protocol, the way that serial movements are usually disentangled from one another involves plotting the position function as a displacement profile and then differentiating and double-differentiating it to generate velocity and acceleration profiles respectively (Note: if the motion data is captured by accelerometer technology, then integrations are performed on the resulting profile to reveal velocity and displacement). From these profiles, assessors look for determinant characteristics within the action trajectories that indicate a new movement. Defining these characteristics becomes one of the most important decisions underpinning the effective use of the "*number of movements* " metric.

Reflecting on the way that errors and corrections emerge in a motor performance can be helpful in setting the appropriate motion analysis parameters for determining the onset and offset of a movement. Consider, for instance, the types of errors that require a correction. For one, an action can require a correction because the performer selects and executes the wrong movement. This type of error occurs, for instance, when the laparoscopic surgery trainee forgets that the display screen is incongruently rotated with respect to the work space and ends up moving a grasper to the right instead of the left. To correct these types of errors, the ongoing movement must be terminated and replaced with an entirely new movement. Sometimes, this involves reversing direction to return to where the action started, or stopping to reassess the situation before initiating a new movement in search of corrective solution. With this type of error and correction in mind, skill assessors may set a zero crossing in the velocity-over-time profile as the end and start points for successive movements. However, this type of new movement determinant can be insufficient when one considers that a series of movements can be executed without the limb coming to a complete stop between each.

It is necessary to understand that noise in the neuromuscular system means that the production of movements is inherently variable,[30] such that discrete precision actions usually require a subtle or not-so-subtle correction (or corrections) towards their conclusion in order to be successful even when the appropriate movement is selected and executed correctly by the performer.[31] In this regard, new movement determinants based on zero crossings in the acceleration-over-time profile are also problematic as they ignore the refined and controlled nature of human motion.[32, 33] Indeed, the findings from over a century's worth of experiments on the accuracy of voluntary actions reveal that any one precision movement includes complimentary impulses that move the limb toward its goal and then integrate response-produced sensory feedback to correct the overall movement for accuracy.[34] In this regard, the typical acceleration profile of a single discrete movement derived via motion capture techniques has 3 zero crossings, which characterize a large sinusoid (i.e., the initial impulse) that is followed by a second, smaller sinusoid (i.e., the corrective impulse).[31, 35, 36] The idea is that because variability in movement execution is so inherent to motor performance that its management becomes a fundamental challenge for learners as they move along the continuum of expertise. That is, novice performers struggle to determine whether their approach to performance introduces too much variability to correct, while skill performers understand the inherent variability and develop strategies that allow them to anticipate the type of corrections required for their movements. In this way, a single expert movement often includes periods of deceleration and re-acceleration.[31]

Setting Parameters

With consideration for this hard-wired type of error and the associated corrections, the most common method of disentangling separate movements within a serial action from a position function in kinesiology research is to set a threshold for a change in velocity. In this way, a movement is defined as an acceleration followed by a deceleration, but only when the resultant velocity exceeds the predefined threshold. The idea is that if the velocity of an action alters by more than a certain amount, in any direction, then a new movement can be inferred.[23] This method accounts for single movements that include graded accelerations or decelerations, permits new movements to be registered without a zero crossing in acceleration, and allows small re-accelerations to occur without necessarily registering a new movement.

The setting of the velocity threshold for a new movement can be one of the most important decisions to the calculation of the number of movements performed during an assessment. Consider, for instance, a small, controlled movement experiment in which a series of a known number of movements were counted under two different velocity thresholds for determining a new movement. In this experiment, a confederate performer was enlisted to slide a handle along a straight 25cm track with the goal of creating a single motion wherein the handle stopped gently against the stopper at the other end of the track. The confederate completed this action 10 times in each direction for a total of 20 known movements. These twenty movements were repeated 20 times while a motion capture device recorded the action. The device was the Imperial College Surgical Assessment Device (ICSAD), a custom software-hardware package that works to time stamp, filter, and digitize movement data by way of a Polhemus ISOTRAK II electromagnetic system (Polhemus, Colchester, VT, USA) with a positional resolution of 3mm from 1.5m away. Reports on the optimal operation of the ICSAD in medical education literature indicate a velocity threshold 15mm/s as appropriate for determining new movements.[37, 38] As such, we analyzed our confederate's movements with this velocity threshold, and for the purposes of the demonstration also at a velocity threshold of 7.4 mm/s. The results of our test revealed the ICSAD counted quite accurately at the 15 mm/s velocity threshold (21.9 $\pm$2.01 movements), but that reduction of the velocity threshold to 7.4 mm/s had a profound impact on the accuracy of the count (28.8 $\pm$ 4.16 movements).

Although the setting of the velocity threshold for a new movement is one of the most important decisions to the calculation of the number of movements performed during an assessment, it is not one that exists in isolation. The identification of new movements must also be considered with respect to the choices that are made regarding data filtering. This is because the technologies of motion tracking systems are unable to differentiate signals from meaningful movements of the sensor from those that result from hand tremor or other sources in the environment. This idea is similar to the way an electrocardiogram signal that is generated by the heartbeat of a baby *in utero* will be interfered with by the heartbeat of the mother. As a consequence, meaningful signals must be extracted from a context of considerable noise before they can be analysed.

In most signal processing applications, including motion analysis, this is accomplished via a Fourier transform, which works to decompose a signal over time into its constituent frequencies. The history and logic that underpin these mathematics fall outside of the scope of this commentary; it sufficient to understand that the result is a frequency distribution.[39] Given that we know that human movements occur a relatively low frequency,[40, 41]motion analysis techniques demand that a *low-pass* filter, which omits overly high frequencies, is applied, such that the total signal analyzed can be restricted as closely as possible with that that reflects the movement. The distribution is then transformed back so that the cleaned version is once again expressed as a signal over time. If the applied filter is not low enough or too low, then the frequency distribution will respectively preserve excessive noise or remove meaningful data from the final analyzed profiles. As such, the ability of the filter to accurately isolate the movement signals can interact significantly with the velocity threshold for new movements to have a major impact on the number of movements counted.

Consider again our small controlled movement experiment; however, note that our confederate's sliding track movements were also measured with a second device: the VICON optoelectric system (Vicon Motion Systems, Lake Forest, CA). The VICON is an integrated 13-camera system that is capable of providing 6 degree-offreedom digital position data for markers with an accuracy of 0.5mm from up to 16m away. Importantly, the VICON operates on the bases of custom MatLab scripts (MathWorks Inc., Natick, Massachusetts, USA) that allow assessors to pre-determine the parameters for data filtering. In this case, a conservative stance was taken and a low-pass Gaussian-Butterworth filter[22] with 5 Hz cut-off frequency was applied. The same sliding actions, recorded with this device, under this filtering protocol revealed accurate recordings at both the 15 mm/s (20.6 $\pm$ 1.4 movements) and the 7.4 mm/s (20.9 $\pm$ 1.6 movements) thresholds.

Although the VICON provides greater spatiotemporal resolution than the ICSAD, the differences in these two devices to accurately count movements at the lower velocity threshold is attributable to differences in the data filter processes. Specifically, the algorithms that underscore the ICSAD operations also use a Fourier transform method to filter data; however, they make the filter cut according to a standard deviation metric for the frequency distribution rather than at an absolute frequency measure (for e.g., 5Hz). That is, the ICSAD determines the standard deviation associated with the resulting frequency distribution and then sets the filter cut point based on a pre-set magnitude of that value. The ICSAD used in the small experiment was set to its default filter setting of 2, which means that all signals associated with frequencies above two standard deviations below the distribution mean were removed from the function prior to analyses. In this regard, the ICSAD allowed more noise to be incorporated into the analyzed function. While this noise was insufficient to alter movement determinations at the more conservative 15 mm/s velocity threshold, it was enough to register an increased number of movements at the lower 7.4 mm/s threshold. This is because the filter methodology determines the cut-off more so by the noise inherent in the measurement context rather than the frequencies of the target signals.

The intention of this demonstration is not to highlight the ICSAD as an inappropriate motion capture device for technical skill assessment. Indeed, the ICSAD has been lauded throughout the surgical education and assessment literature for the validity of its metrics, its ease of use, portability, and ability to capture data without constraining operative performance. Furthermore, the point is also not to advocate for the particular thresholds and filters tested here. Rather the goal is to emphasize that the effective practice of motion analysis for skill assessment involves a sophisticated understanding of the way that *a priori* decisions about data collection and analysis interact to influence outcomes. In this regard, it is essential that any efficiency standard of competency that is based on number of movements is established with appropriately and consistently applied analysis decisions; the determination of which will undoubtedly require numerous concurrent and criterion validation studies across the spate of skills of interest. Moreover, the assessment of trainees by way of motion capture techniques will require standards for measurement, as a differently adjusted motion capture system at one institution could artificially inflate or deflate a learner's performance relative to the performances of those at other institutions.

Interpreting Movements

As mentioned, it has largely been the case in the medical education literature that the number of movements an individual makes is indicative of their overall efficiency. That is, the fewer the movements, the more efficient the performance. While, construct validation studies show this to be the general case when expert and novice performances are compared,[27, 28] there is less clarity on the way that efficiency develops through the intermediate stages of learning. In particular, one perspective on the study of human motor control describes learning with respect to the way in which individuals vary and explore components of the action as they search for the optimal approach to performance.[42-45] As a consequence, different components of action can shift between phases of stability (i.e., low movement variability) and instability (i.e., high movement variability) as learners attempt to organize their movements. The idea is that these shifts occur in response to constraints placed on the performance by the task or environment, and even as a function of changes in the learner's ability and motivation. A good example of this is Guerin and Kunkle's (2004) study of individuals learning to kick a ball over a barrier onto a target. Measurement of the participants' kick height and accuracy over 12 extensive practice sessions demonstrated that they focused initially on ensuring that the ball crossed over the barrier, with little concern for accuracy.[46] However, as they became able to clear the height of the barrier consistently, their focus shifted to landing the ball accurately on the target. That is, as the learning experience progressed, the height constraint deteriorated in importance and the accuracy constraint emerged as increasingly more pertinent.

With respect to metrics of efficiency, such as number of movements, this type of shifting means that skill assessors need to understand the performance constraints to which learners are currently attending. A small study exploring the validity of an instrumented simulator for the assessment of surgical knot tying skills provides a nice example.[8] In this study, the simulator incorporated flexometer technology that measured the quality of the knots tied by pre-medical undergraduate students (i.e., novices), medical clerks (i.e., intermediates), and senior medical residents (i.e., experts). Interestingly, the technology also permitted the experimenters to measure the economy of action via the amount the walls of the simulated wound moved while the participants performed. Not surprisingly, the results showed that the completed knots were tighter and more sustainable as the performers increased in expertise. That is, the experts' sutures were better than those of the intermediates, and the intermediates' sutures were better than those of the novices. However, the movement economy metric revealed that the intermediates performed far more inefficiently than both the experts and the novices. Taken together, the two metrics reveal how the intermediate group had sacrificed attention to efficiency in order to achieve stronger final products. Given that learners will shift focus from outcome to efficiency at different stages of practice, it is essential that the assessor of clinical technical skills remembers that an objective computerized assessment of performance efficiency– defined in terms of number of movements or otherwise - exists independently of the outcome of the performance. As a consequence, the relationship between efficiency and expertise is not always direct.[47] Thus, as a construct for assessment, it may be necessary to avoid motion efficiency as a measure of competence until after the learner's proficiency at reliably producing quality outcomes is well established. In surgery, this may be particularly important, as learners will often value accuracy at the expense of efficiency.

Furthermore, when interpreting number of movements, one should also keep in mind the way in which contexts of performance and individual performer differences can impact how a metric such as number of movements is used as a competency standard. With respect to the former, differences in task rules or equipment can have significant impacts on movement performance. In this regard, for example, it would not be fair to compare the number of movements associated with open surgical performance to those associated with laparoscopic surgical procedure.[48] Indeed, even changes in the time a performer is allotted to complete a task has shown to disrupt the metrics of efficiency that so often characterize the most expert performers.[49] Moreover, skills that occur in closed, static environments are fundamentally different than those that occur in open, dynamic environments. The former afford approaches to skill performance that are largely anticipatory in nature, which allows for fuller action plans to be generated prior to initiation, while the latter requires more attention to be given to online control processes that determine the movements needed as a procedure unfolds.[31] Thus, it would also be unfair to compare performances that permit and don't permit prior planning, as the latter would most certainly be associated with higher movement counts.

Closing Thoughts

Taken together, the message for skill assessors considering the use of motion capture data for generating metrics of performance efficiency is that they must work to ensure a high degree of consistency between data processing techniques and assessment contexts; and given that acceptable ranges for movement efficiency can be established for particular clinical skills, progressions of efficiency should be evaluated specifically within the constraints posed on the task by the individual performer. In presenting this commentary, the hope is to convey that an incomplete knowledge of the tenets of objective movement capture can yield inaccurate results. It is my position, then, that a solid understanding of motion capture and human motor control is essential to the effective implementation of objective computerized competency-based assessment in medical education contexts. Beyond concerns about the way movements are defined and counted, this also includes consideration for the nature of learning progressions and challenges the assumption that number of movement measures will reflect greater efficiency as a learner progresses from novice to expert. In this regard, it is essential that educators and assessors are careful to reflect any efficiency data through the lens of task success.

Conflict of Interest : Lawrence Grierson has no conflict of interest to declare.

Ethical Approval : Ethical approval was not required for this work.

Acknowledgements : Lawrence wishes to acknowledge Dr. Daniel Garcia, Dr. Simran Ohson, Dr. Jim Lyons, and the Department of Kinesiology at McMaster University for making possible the data collection associated with the small controlled experiment described in this commentary. Lawrence also wishes to acknowledge Dr. David Rojas, who provided critical insight and input to the conversation regarding setting data processing parameters.

References

1. Frank JR, Snell LS, Cate O, Holmboe ES, Carraccio C, Swing SR, et al. Competency-based medical education: theory to practice. Med Teach 2010; 32: 638-645.
2. Whitehead C.R, Austin Z, Hodges BD. Flower power: the armoured expert in the CanMEDS competency framework?. Adv Health Sci Educ 2011; 16: 681-694.
3. Ferguson PC, Kraemer W, Nousiainen M, Safir O, Sonnadara R, Alman B, Reznick R. Three-Year Experience with an Innovative, Modular Competency-Based Curriculum for Orthopaedic Training. J Bone Joint Surg 2013; 95: e166-1.
4. Martin JA, Regehr G, Reznick R, MacRae H, Murnaghan J, Hutchison C, Brown M. Objective structured assessment of technical skill (OSATS) for surgical residents. Br J Surg 1997; 84: 273-278.
5. Beyer L, Troyer JD, Mancini J, Bladou F, Berdah SV, Karsenty G. Impact of laparoscopy simulator training on the technical skills of future surgeons in the operating room: a prospective study. Am J Surg 2011; 202: 265-272.
6. Derossis AM, Bothwell J, Sigman HH, Fried GM. The effect of practice on performance in a laparoscopic simulator. Surg Endosc 1998; 12: 1117-1120.
7. Leming JK, Dorman K, Brydges R, Carnahan H, Dubrowski A. Tensiometry as a measure of improvement in knot quality in undergraduate medical students. Adv Health Sci Educ 2007; 12: 331-344.
8. Rojas D, Cristancho S, Rueda C, Grierson L, Monclou A, Dubrowski A. The validation of an instrumented simulator for the assessment of performance and outcome of knot tying skill: a pilot study. St Heal T 2010; 163: 517-523.
9. Li J, Grierson LE, Wu MX, Breuer R, Carnahan H. (2013). Perceptual motor features of expert acupuncture lifting-thrusting skills. Acupunct Med, 2013; 31: 172-177.
10. Dubrowski A, Backstein D. The contributions of kinesiology to surgical education. J Bone Joint Surg 2004; 86: 2778-2781.
11. Praamsma M, Carnahan H, Backstein D, Veillette CJ, Gonzalez D, Dubrowski A. Drilling sounds are used by surgeons and intermediate residents, but not novice orthopedic trainees, to guide drilling motions. Can J Surg 2008; 51: 442.
12. Einstein AJ, Pascual TN, Mercuri M, Karthikeyan G, Vitola JV, Mahmarian JJ, et al. Current worldwide nuclear cardiology practices and radiation exposure: results from the 65 country IAEA Nuclear Cardiology Protocols Cross-Sectional Study (INCAPS). Eur Heart J 2015; 36: 1689-1696.
13. Mercuri M, Xie C, Levy M, Valettas N, Natarajan MK. Predictors of increased radiation dose during percutaneous coronary intervention. Am J Cardiol 2009; 104: 1241-1244.
14. Procter LD, Davenport DL, Bernard AC, Zwischenberger JB. General surgical operative duration is associated with increased risk-adjusted infectious complication rates and length of hospital stay. J Am Coll Surg 2010; 210: 60-65.
15. Gupta D, Natarajan MK, Gafni A, Wang L, Shilton D, Holder D, et al. Capacity planning for cardiac catheterization: a case study. Health Policy 2007; 82: 1–11.
16. Mercuri M, Natarajan MK, Holder DH, Xie C, Gafni A. Insight into capacity planning for cardiac catheterization services: policy lessons learned from "Looking in the Mirror" over a decade. Health Policy 2009; 91: 314-320.
17. Berguer R, Forkey DL, Smith WD. Ergonomic problems associated with laparoscopic surgery. Surg Endosc 1999; 13: 466-468.

18. Stock SR. Workplace ergonomic factors and the development of musculoskeletal disorders of the neck and upper limbs: A meta-analysis. Am J Ind Med 1991; 19: 87-107.
19. Cook DA, Hatala R, Brydges R, Zendejas B, Szostek JH, Wang AT, et al. Technology-enhanced simulation for health professions education: a systematic review and meta-analysis. JAMA 2011; 306: 978-988.
20. Pusic MV, Brydges R, Kessler D, Szyld D, Nachbar M, Kalet A. What's your best time? Chronometry in the learning of medical procedures. Med Educ 2014; 48: 479-488.
21. Ramsay CR, Grant AM, Wallace SA, Garthwaite PH, Monk AF, Russell IT. Statistical assessment of the learning curves of health technologies. Health Technol Assess 2001; 5:1–79.
22. Datta V, Bann S, Mandalia M, Darzi A. The surgical efficiency score: a feasible, reliable, and valid method of skills assessment. Am J Surg 2006; 192: 372-378.
23. Dosis A, Aggarwal R, Bello F, Moorthy K, Munz Y, Gillies D, Darzi A. Synchronized video and motion analysis for the assessment of procedures in the operating theater. Arch Surg 2005; 140: 293-299.
24. Lin Z, Uemura M, Zecca M, Sessa S, Ishii H, Tomikawa M, et al. Objective skill evaluation for laparoscopic training based on motion analysis. IEEE T Bio-Med Eng 2013; 60: 977-985.
25. Saleh GM, Gauba V, Sim D, Lindfield D, Borhani M, Ghoussayni S. Motion analysis as a tool for the evaluation of oculoplastic surgical skill: evaluation of oculoplastic surgical skill. Arch Ophthalmol 2008; 126: 213-216.
26. Hayter MA, Friedman Z, Bould MD, Hanlon JG, Katznelson R, Borges B, Naik VN. Validation of the Imperial College Surgical Assessment Device (ICSAD) for labour epidural placement. Can J Anesthes 2009; 56: 419-426.
27. Judkins TN, Oleynikov D, Stergiou N. Objective evaluation of expert and novice performance during robotic surgical training tasks. Surg Endosc 2009; 23: 590–597.
28. McNatt SS, Smith CD. A computer-based laparoscopic skills assessment device differentiates experienced from novice laparoscopic surgeons. Surg Endosc 2001; 15:1085–1089.
29. Schmidt RA, Lee TD. Motor control and learning: A behavioral emphasis. Human Kinetics Publishers, Champaign Illinois; 1999.
30. Schmidt RA, Zelaznik H, Hawkins B, Frank JS, Quinn JT. Motor-output variability: a theory for the accuracy of rapid motor acts. Psychol Rev 1979; 86: 415.
31. Elliott D, Hansen S, Grierson LE, Lyons J, Bennett SJ, Hayes SJ. Goal-directed aiming: two components but multiple processes. Psychol Bull 2010; 136: 1023.
32. Pagador JB, Sanchez-Margallo FM, Sanchez-Peralta LF, Sanchez-Margallo JA, Moyano-Cuevas JL, Enciso-Sanz S, et al. Decomposition and analysis of laparoscopic suturing task using tool-motion analysis (TMA): improving the objective assessment. Int J Comp Ass Rad 2012; 7: 305-313.
33. Smith SGT, Torkington J, Brown TJ, Taffinder NJ, Darzi A. Motion analysis: A tool for assessing laparoscopic dexterity in the performance of a laboratory-based laparoscopic cholecystectomy. Surg Endosc 2002; 16: 640-5.
34. Woodworth RS. Accuracy of voluntary movement. Psychol Rev 1899; 3: 77-86.
35. Grierson LE, Elliott D. Goal-directed aiming and the relative contribution of two online control processes. Am J Psychol 2009; 122: 309-324.
36. Grierson LE, Elliott D. Kinematic analysis of goal-directed aims made against early and late perturbations: an investigation of the relative influence of two online control processes. Hum Movement Sci 2008; 27: 839-856.
37. Moorthy K, Munz Y, Dosis A, Bello F, Darzi A. Motion analysis in the training and assessment of minimally invasive surgery. Minim Invas Ther 2003; 12: 137-142.
38. Munz Y, Moorthy K, Bann S, Shah J, Ivanova S, Darzi SA. Ceiling effect in technical skills of surgical residents. Am J Surg 2004; 188: 294-300.
39. Boashash, B., ed. Time-Frequency Signal Analysis and Processing: A Comprehensive Reference, Oxford: Elsevier Science; 2003
40. Franks IM, Sanderson DJ, Van Donkelaar P. A comparison of directly recorded and derived acceleration data in movement control research. Hum Movement Sci 1990; 9: 573-582.

41. Pezzack JC, Norman RW, Winter DA. An assessment of derivative determining techniques used for motion analysis. J Biomech 1977; 10: 377-382.
42. Berstein N. On dexterity and its development. In M.L. Latash & M.T. Turvey (Eds.), Dexterity and it development. Hillside, NJ: Lawrence Erlbaum Associates; 1996: pp. 3-244
43. Newell KM. Coordination, control and skill. Adv Psychol 1985 27: 295-317.
44. Renshaw I, Davids K, Shuttleworth R, Chow J. Insights from ecological psychology and dynamical systems theory can underpin a philosophy of coaching. Int J Sport Psychol 2009; 40: 540–602.
45. Williams AM, Davids K, Williams JG. Visual perception and action in sport, London: E and FN Spon; 1999
46. Guerin S, Kunkle, D. Emergence of constraint in self-organized systems. Nonlin Dynam Psychol 2004; 8: 131–46.
47. Adolph KE, Berger SA. Motor development. In Handbook of child psychology: Cognition, perception, and language. 6th ed, Vol. 2, ed. W. Damon and R. Lerner. New York: Wiley; 2006: pp. 161-213.
48. Starkes JL, Payk I, Jennen P, Leclair D. A stitch in time: cognitive issues in microsurgery. Adv Psychol 1993; 102: 225-40.
49. Mackay S, Datta V, Mandalia M, Bassett P, Darzi A. Electromagnetic motion analysis in the assessment of surgical skill: relationship between time and movement. ANZ J Surg 2002; 72: 632-634.