Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 17 October 2018 doi:10.20944/preprints201810.0379.v1

Peer-reviewed version available at *Information* **2018**, *9*, 316; doi:10.3390/info9120316

Article

Expert-based speed-precision control in early simulator training for novice surgeons

Birgitta Dresp-Langley[1]

[1] Centre National de la Recherche Scientifique, UMR 7357 ICube Lab, Université de Strasbourg Faculté de Médecine Pavillon Clovis Vincent; birgitta.dresp@icube.unistra.fr

* Correspondence: birgitta.dresp@unistra.fr; Tel.: +33-3-88119117

Abstract: Simulator training for image-guided surgical interventions may benefit from artificial intelligence systems that control the evolution of task skills in terms of time and precision of a trainee's performance on the basis of fully automatic feed-back systems. At the earliest stages of training, novice trainees frequently focus on getting faster at the task, and may thereby compromise the optimal evolution of the precision of their performance. For automatically guiding them towards attaining an optimal speed-accuracy trade-off, an effective control system for the reinforcement/correction of strategies must be able to exploit the right individual performance criteria in the right way, reliably detect individual performance trends at any given moment in time, and alert the trainee, as early as necessary, when to slow down and focus on precision, or when to focus on getting faster. This article addresses several aspects of this challenge for speed-accuracy controlled simulator training before any training on specific surgical tasks or clinical models should be envisaged. Analyses of individual learning curves from the simulator training sessions of novices and benchmark performance data of one expert surgeon, who had no specific training in the simulator task, validate the suggested approach.

Keywords: surgical simulator training; individual performance trend; speed-accuracy function; automatic detection; performance feed-back

1. Introduction

Technological development and pressure towards a reduction in time available for learning has radically changed the traditional apprenticeship model of surgical training. Simulation now offers the opportunity for repeated practice in safe and controlled environments and the most recent technological advances have led to the development of various simulators, which have already been introduced in surgical training. The complexity and reliability of available simulators vary considerably, and selecting an appropriate simulator for surgical skill training is in itself a challenge. Simulators that are used for specific surgical skills training are generally tested for the highest validity level [1], that of predictive validity, ensuring that assessments of performance in the specific simulator task are likely to predict future performance of the trainee in the same task in a clinical context (animal, patient). However, only a certain percentage of these surgical simulators give some kind of performance feedback to the trainee, and the feed-back systems as such are generally not validated. In other words, whether the feed-back given during training is actually truly useful to a novice is not known. Ideally, within a surgical curriculum, trainees should have dedicated time for simulation-based training with appropriate performance monitoring through truly effective feed-back systems, as the main advantage of computer simulators for surgical training is the opportunity they afford for independent learning. Yet, if the simulator does not provide relevant and truly useful instructional feedback to the user, then instructors need to be present to supervise and tutor the trainee.

Relevant performance metrics [2-8] are therefore essential to the development of surgical simulator systems for optimal independent training, and the presentation of such metrics to the user in a way that boosts independent learning by facilitating measurable skill improvement is clearly the most important aspect of developing effective simulator systems [9]. Metric-based simulation ensures that training sessions are more than just simulated clinical procedures and gets rid of subjectivity in evaluating skill evolution, as there is no ambiguity about the progress of training. Benchmarking individual levels of proficiency on the performance levels of experts on a validated metric-based simulation system has well-established intrinsic face validity [9] and appears a better approach than benchmarking on some abstract performance concept based on expert consensus, for example. Building expert performance as a basis for skill assessment into simulator training programs ensures that the "pass" level is defined on realistic criteria set directly by the proficiency levels of individuals who are actually very experienced at performing the actual clinical procedure [9-13]. Artificial intelligence provides well-suited concepts for reinforcement learning procedures and allows building prior benchmark knowledge on expert performance into simulator training most effectively. Specific control procedures [11] based on metric-based benchmark criteria for automatized performance comparisons, leading to decisions of the "if then" kind without supervision, not only enable the generation of learning curves in individual performance at the end of sessions but also, more importantly, enable trial-by-trial feed-back at any given moment in time during training to help individuals reach optimal performance as swiftly as possible [13-21]. This constitutes a far more effective approach compared with merely assessing end-of-session performance status, or differences between users after training. This article here focuses on a generic early simulator training model for automatic skill evolution, well before novices are trained on more specific surgical simulator systems, or in clinical trials on animal models. The approach discussed here is based on a simple and universal psychophysical human performance model [22-29] that allows telling apart individual strategies during motor learning on the basis of individual speed-accuracy trade-off functions. On the basis of this model, data obtained on a pick-and-place task simulator for image-guided eye-hand-tool coordination training are discussed. How automatic performance control and feed-back can be implemented at various steps and in the simplest possible way is then made clear. The training model proposed ensures that novice trainees, at the earliest stages of simulator training, reach optimal precision levels in any image-guided performance simulator system capable of generating reliable and discerning measures of skill evolution with respect to 1) the time of task execution from t_{zero} to t_n at any moment of the procedure and 2) the precision with which the task is performed at critical steps of the procedure.

2. Materials and Methods

The evolution of individual performance measures relative to task speed and precision was monitored using a specifically designed simulator platform for image-based analysis of performance data relative to the time and precision of hand-tool movements in a computer controlled simulator task (pick-and-place). The technical aspects of this platform, which was used in several experimental studies, are described in detail in some of our previous work [2-6].

For collecting the performance data shown in this paper here, a single camera view was generated through a 120° fisheye lens camera fully adjustable in 360°. The video input received from the camera was processed by a DELL Precision T5810 model computer equipped with an Intel Xeon CPU E5-1620 with 16 Giga bytes memory (RAM) capacity at 16 bits and an NVidia GForce GTX980 graphics card. Experiments were programmed in Python 2.7 for Windows using the Open CV computer vision software library. The computer was connected to a high resolution color monitor (EIZO LCD 'Color Edge CG275W'), which the Color Navigator 5.4.5 interface for Windows. The colors of objects visualized on the screen can be matched to LAB or RGB color space and the color coordinates for RGB triples can be retrieved from a look-up table at any moment in time. The task-action field consisted of a classic square shaped (45cm x 45cm) light grey LEGO board available worldwide in the toy sections of large department stores. Six square-shaped (4,5cm x 4,5cm) target areas were painted on the board at various locations in a medium grey tint (acrylic). In-between these target areas, small LEGO pieces of varying shapes and heights were placed to add a certain level of complexity to both the visual configuration and the task and to reduce the likelihood of getting performance ceiling effects. In the pick-and-place task, a small (3cm x 3cm x 3cm) cube made of very light plastic foam but resistant to deformation in all directions had to be placed on the target areas in a specific order. The video input received by the computer from the fisheye camera generated raw image data within a viewing frame of 640 pixels (width) x 480 pixels (height). These data were processed to generate show-image data in a viewing frame of 1280 pixels (width) x 960 pixels (height), the size of a single pixel on the screen being 0.32mm to ensure that the size of the task-action field viewed on the computer screen was identical to that in the real world. The training sessions were run under conditions of free viewing, with general illumination levels that can be assimilated to daylight conditions. The task-action field was illuminated by two lamps (40Watt, 6500 K) which were constantly lit during the sessions. Participants were comfortably seated at a distance of approximately 75cm from the RA from the screen. Trainees were generally right-handed, as those for whom data are shown here. They were instructed to position the cube, with their dominant hand and using a forceps-like tool, as precisely and swiftly as possible on the center of each target, in the right order, as explained to them. Data from fully completed trial sets only were recorded. A fully complete trial set consists of a set of pick-and-place operations, from target to target in the right order and without dropping the object accidentally. Ten fully completed trial sequences were recorded in each training sessions. For single trial, the computer program generated data relative to the performance measures 'time' and 'precision'. For ‘time’, the computer counted the CPU time (in seconds) from the moment the blue cube object was picked up by the participant to the time it was put on the next target. The rate for image-time data collection was between 25-30 Hz, with an error margin of less than 40 milliseconds for any of the time estimates. For ‘precision’, the computer program counted the cumulated number of blue object pixels at positions "off" the 3cm x 3cm central area of each of the five 4,5cm x 4,5cm target areas whenever the object was placed. The standard errors of these positional estimates, determined in a calibration procedure, were below 10 pixels. Individual time and precision data were written to an excel file by the computer program, with labeled data columns for the different conditions, and stored in a directory for subsequent analysis.

All experiments were conducted in conformity with the Helsinki Declaration relative to scientific experiments on human individuals with the full approval of the ethics board of the corresponding author's host institution (CNRS). All participants were volunteers and provided written informed consent. Their identity is not revealed. Some of the data shown here are from eight training sessions of two novices with no experience in image-guided or other surgical procedures (absolute beginners), as in [2-5]. The data relating to the expert performance measures, shown here for comparison, were recorded from a single training session of a highly skilled expert endoscopic surgeon with more than 30 years of experience in image-guided surgery but no training at all in this specific pick-and-place simulator task here.

3. Results

In several of our previous studies, simulator training data relative to time (in seconds) and precision (in pixels) of image-guided pick-and-place task performance were recorded from a total population of 30 individuals, including absolute beginners, novice surgeons without specific experience in image-guided simulator training, and expert surgeons with variable experience in image-guided surgery. Some of these data have been made available and discussed in our previous work [2, 3, 5].

Here, the training data of two novices with no experience in image-guided simulator training, and of one highly proficient expert surgeon with more than 20 years of experience in image-guided endoscopic surgery but no training in the specific simulator task are compared and discussed. Differences in the individual speed-precision trade-off functions and in the individual evolution of task execution times and task precision across sessions are brought to the fore to highlight spontaneously occurring individual learning strategies that lead to consider the AI enhanced training model proposed thereafter. The first analysis shows the individual trade-off functions between task execution time (in seconds) and precision, expressed here in terms of a score that takes the cumulated number of off-target pixels for a given session with ten repeated trial-sets of five pick and place operations. The higher the cumulated off-target score, the lesser the trainee's precision. To get these individual trade-off functions, the times taken for each of ten repeated trial sets are sorted in ascending order (x-axis) and plotted against their corresponding precision scores (y-axis) for each training session (Figure 1).

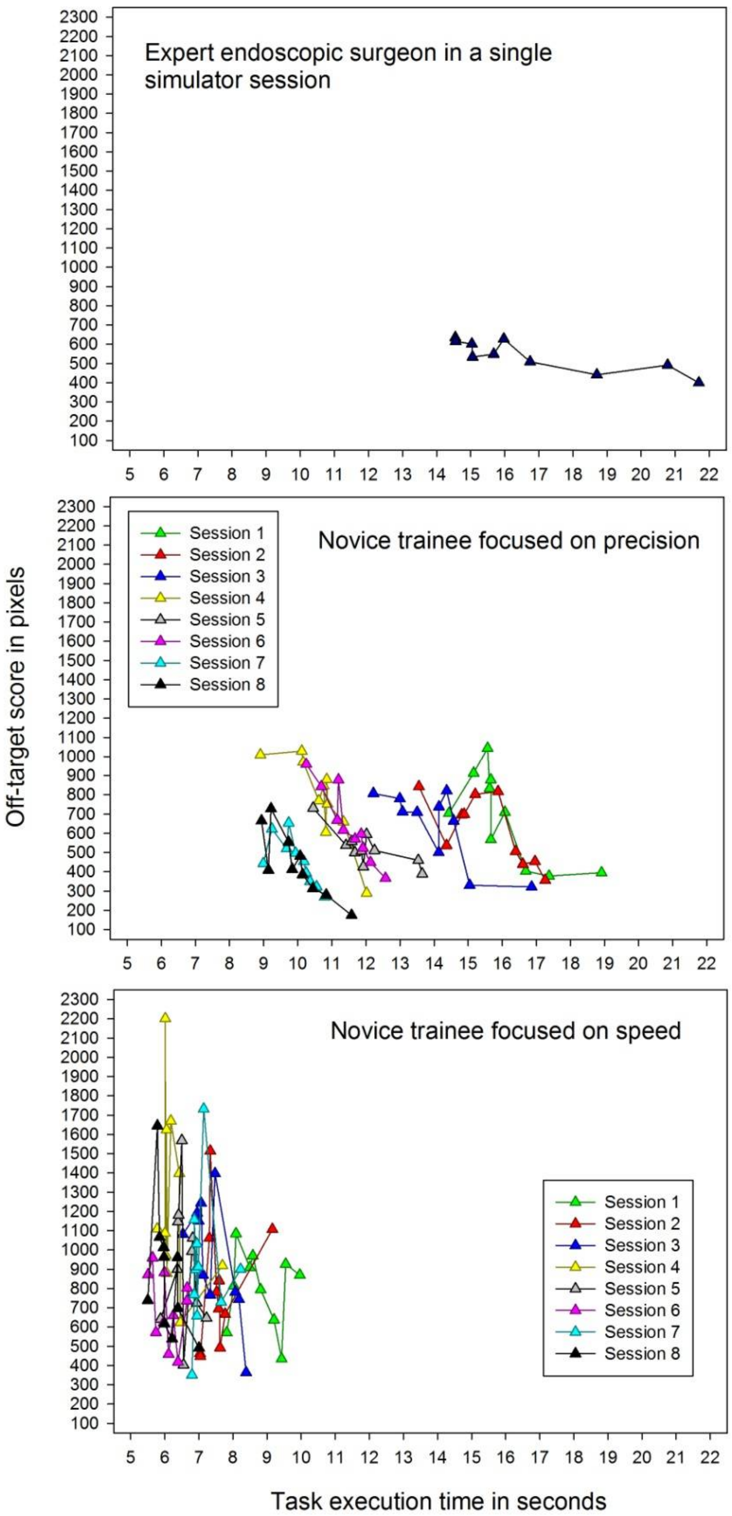


Figure 1. The individual trade-off functions between task execution time (in seconds) and precision, expressed in terms of the cumulated number of off-target pixels for a given session with ten repeated

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 17 October 2018 doi:10.20944/preprints201810.0379.v1

Peer-reviewed version available at Information 2018, 9, 316; doi:10.3390/info9120316

trial-sets of five pick and place operations. A higher cumulated off-target score reflects a lesser precision in a given session. The functions show data of an expert with more than 20 years of experience in image-guided precision surgery from a single simulator training session (graph on the left) for comparison with data of two novices from eight successive training sessions revealing two naturally occurring and radically different task strategies, one focused on precision (graph in the middle) and one focused on speed (graph on the right).

Comparison between the individual trade-off functions reveal two naturally occurring and radically different training strategies of the two novices. One of them clearly focused on speed (Figure 1, graph on right), in other words on doing the task as fast as he/she can, with task execution times less than half those of the expert with 20 years of experience. Across the eight training sessions, this trainee is getting progressively faster, but the precision score does not improve as training progresses and fails to evolve towards any stable performance pattern across the eight sessions, as shown by the extent of scatter in this trainee's individual speed-precision trade-off functions for each session. Even in the last session, the precision scores for the ten trial-sets are poorer and highly unstable compared with the consistent and rather stable precision scores of the expert's single session performance (Figure 1, graph on left), bearing in mind that this expert had no previous training in this specific simulator task here. The other novice's strategy is clearly far more focused on precision, as is made clear by the individual speed-precision trade-off functions of from the eight training sessions (Figure 1, graph in middle). This novice starts, in the first two sessions, with times in the range of those from the single session of the expert, then gets progressively faster and, ultimately, becomes faster than the expert who had no further training in the task. The precision scores of this novice also improve consistently as training progresses and, ultimately, in the last training session attain a level of stability that is comparable to that of the expert.

Plotting the data for time and precision of the same three trainees as a function of the sequence of the trials across the sessions (a single session for the expert and eight successive training sessions of the two novices) gives a clear overview of the individual evolution of each performance parameter and their relative variability/stability. The novice focused on precision (Figure 2, graph on left) displays a wider range of task execution times with a steeper learning curve compared with the data from the novice focused on speed (Figure 4, graph on left), whose task execution times are scattered around the lowest possible ceiling level, indicating that this trainee clearly started off way too fast. The expert gets naturally and consistently faster in a single session (Figure 3, graph on left). The most revealing learning curves are those for the parameter relative to precision, showing moderate scatter and a clear learning trend for the novice focused on precision (Figure 2, graph on right), and extensive scatter with no learning trend at all for the novice focused on speed (Figure 4, graph on right). The expert's precision data show almost no scatter at all, indicating a highly stable performance level and a slight trend towards better precision as trials progress in a single session (Figure 3, graph on right).

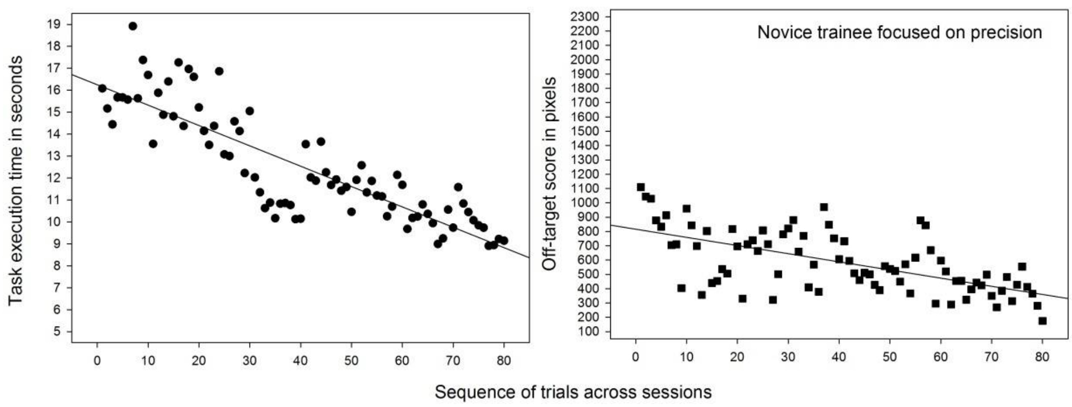


Figure 2. The evolution of task execution times (left) and precision (right) of the novice trainee focused on precision showing a clear learning trend for each performance parameter, with moderate extent of scatter (variability) in the data, shown here as a function of the sequence of the repeated trial-sets (10 per session) across eight successive training sessions.

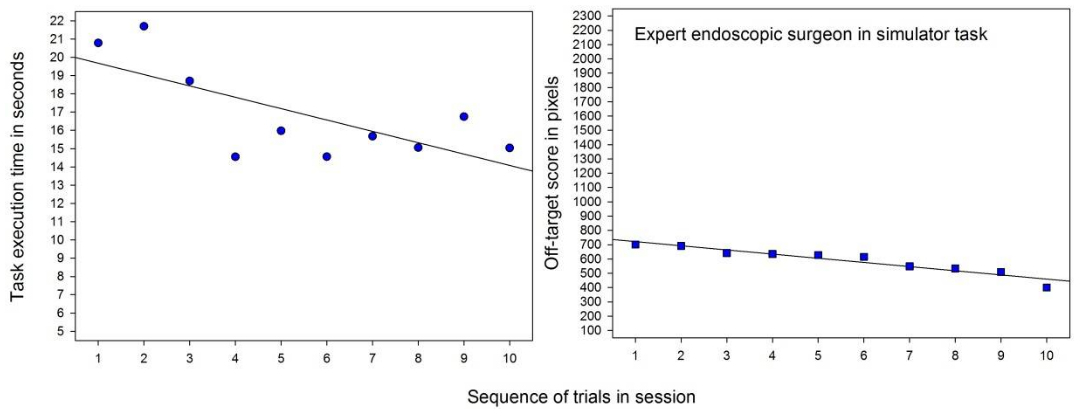


Figure 3. The evolution of task execution times (left) and precision (right) of the expert surgeon from a single simulator session showing a short-term training effect on task execution times and a close to perfectly stable performance level for precision.

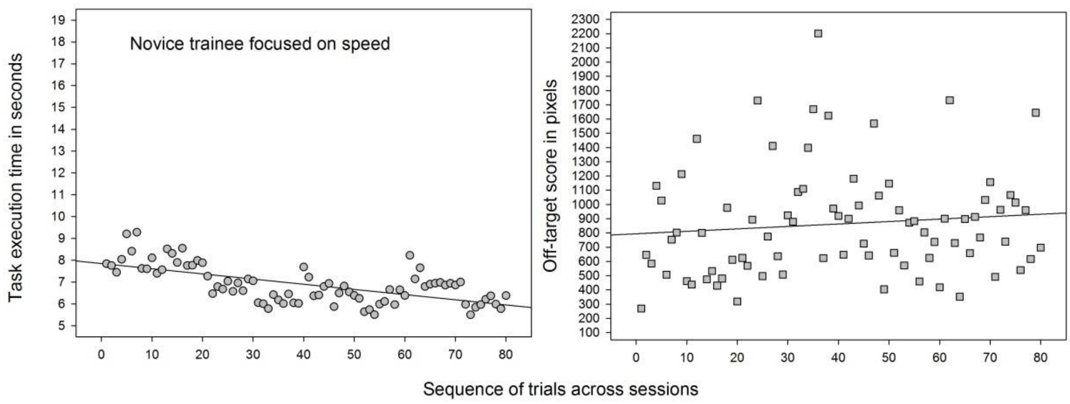


Figure 4. The evolution of task execution times (left) and precision (right) of the novice trainee focused on speed, with task execution times scattered around the lowest possible ceiling level, and extensive scatter of the precision parameter with no learning trend at all, indicating that this trainee started off much too fast.

The performance data relative to time and precision from the simulator training sessions of three individuals compared here and their differential evolution with training show three typical profiles providing clear proof of concept that the pixel-based precision measures of our simulator system are well-suited to give a reliable and discerning measure for task precision, i.e. a measure that allows 1) a clear performance distinction between a surgical expert and a novice trainee, and 2) a clear distinction between novices who adopt radically different task strategies during learning. The speed-accuracy trade-off functions provide direct insight into the nature of these strategies, as shown previously. They also provide a quantitative basis for what could be considered a general model of effective, and possibly AI-guided, simulator training for any image-guided task where precision matters critically, as in surgical training.

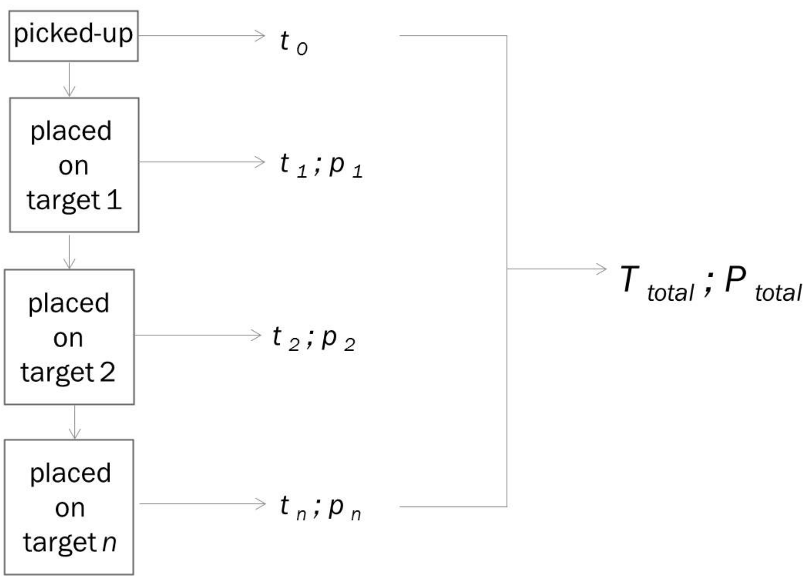
As shown and discussed in some of our previous work [2, 3, 5], the strategy differences between novices in simulator training for image-guided hand-tool movements vary between the two extreme cases shown here, which leads to general conclusions that need to be taken into account by any system which automatically monitors the evolution of individual performance in view of either a reinforcement or a correction of task strategies: The performance strategy of trainees who start off too fast needs to be corrected to enable effective precision learning, while the performance strategy of trainees who focus on being as precise as possible should be reinforced because they will naturally and without any further instructions get faster with training. When an individual precision performance can be considered optimal and stable, then and only then the trainee may be instructed to try to get even faster. The performance profile of an expert can serve as a benchmark profile to generate in-built system knowledge of what the desired performance profile of a novice should look like after successful training on a given simulator.

On these grounds, we propose a system that automatically corrects for individual strategy at any moment in time during training, and enables trainees to reach the optimal speed-precision strategy as swiftly as possible during training by receiving appropriate feed-back. Such a system must be able to:

- generate reliable and discerning measures (parameters) relative to time and precision of individual performance at any moment in time during training
- compare an individual parameter measure relative to time and precision at any moment in time during training with the desired parameter value based on an already known ("learnt") performance profile of an expert user
- provide feed-back to the user at any moment in time during training about what he/she needs to focus on to optimize his/her performance as swiftly as possible

How this may be achieved is illustrated on the example of a single five-step trial of the simulator task here (Figure 5). A single trial of the image guided pick-and-place task has several (here five, but it could be any *n* in any other system) successive steps. The system starts counting task execution time from the moment the object is picked up by the user with the surgical tool (t_0) and ongoing time can be communicated to the trainee at any moment (t_n) from then until the object is placed on the last of several (here five) successive targets.

Placing the object on a given target is a critical step of the procedure where precision matters as users are instructed to place the object with the surgical tool on the central area of each target as precisely as possible. This is a challenging task and involves specific visual attention to fine eye-hand-tool coordination for placing the target optimally as the precise borders of the target center are only known by the system in terms of pixel coordinates, but not visible to the user. The user only sees the borders of the targets as such in the image guiding his/her action, and the object that needs to be placed centrally is smaller than the target area.



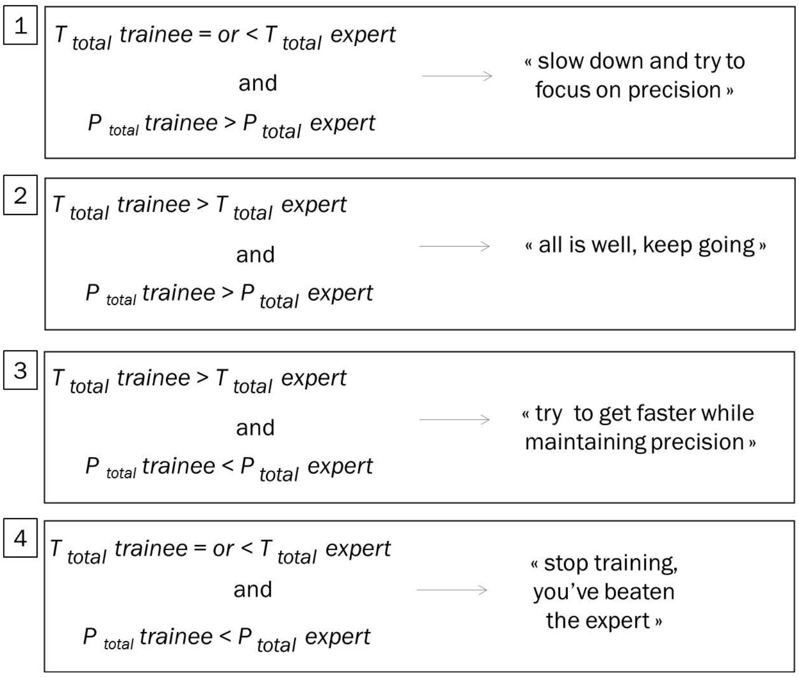
Peer-reviewed version available at *Information* **2018**, *9*, 316; doi:10.3390/info9120316

Figure 5. System flowchart of a single trial of the image guided pick-and-place task with *n* successive critical "place object" steps. The system starts counting task execution time from the moment the object is picked up by the user with the surgical tool (t_0). Ongoing time can be communicated to the trainee at any moment (t_n) from then together with the precision score (p_n) until the object is placed on the last targets. The data from a single training session shown in the graphs here above represent the cumulated values T_{total}x10 and P_{total}x10 from a sequence of ten successive trials per session.

At each such critical step of the procedure, the system counts the number of pixels corresponding to the object in the image that do not coincide with the pixels that define the central area of a target known by the system and are therefore "off-target" in terms of the task constraints as given ("place object as centrally on target as possible"). Hence, the smaller this measure, the greater the user's precision at a given critical step. This precision score (p_n) also can be communicated to the trainee an any critical moment in time (t_n) of the procedure.

On the basis of such a system, which automatically monitors the evolution of individual performance parameters at any given moment in time during training, it is possible to suggest a simple AI enhanced system for the control, in terms of either reinforcement or correction, of performance strategies of trainees to ensure effective precision learning. It goes without saying that priority needs to be placed on precision rather than speed, especially in surgical training, and trainees get faster naturally, as shown here above, once they have adopted the right strategy for working on their precision. As is shown here, a single dataset from a single expert can provide effective benchmark data for building prior knowledge into the system and these "learnt" data can be exploited for automatic performance feed-back to the user at any moment in time during training. The data from our expert here were from a single session. In an ideal world, expert data could be collected from multiple simulator sessions, as many as necessary, to allow for even more direct trial-by-trial comparisons where the observed data of a trainee at a given moment of the procedure for a given session S_n are compared to the "ideal" data of an expert for the corresponding moment of the procedure and session S_n of a training sequence.

Taken into account the general insight gained from the performance analyses here above (Figures 1-4), leads to suggest a generic AI enhanced control system that makes decisions based on in-built knowledge of expert data as shown here below (Figure 6).



Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 17 October 2018 doi:10.20944/preprints201810.0379.v1
Peer-reviewed version available at Information 2018, 9, 316; doi:10.3390/info9120316

Figure 6. Model for the automatic reinforcement/correction of individual performance strategies during training based on data from a single expert ("learnt" system knowledge) to generate appropriate feed-back. No more than four cases (1, 2, 3 and 4) need to be considered. Any simulator system which automatically and reliably collects individual performance data for time and precision (parameters) can be transformed into an AI enhanced system for the control, in terms of either reinforcement or correction, of performance strategies of trainees to ensure effective precision learning.

The four cases to be considered by the system may be summarized as follows. 1) At a given moment t_n in a training session S_n the trainee is as fast as or faster than the expert and less precise. In this case, the system needs to alert him/her to slow down and start focusing on precision. This is the classic case of a trainee focused on speed who tries to do the task as fast as he/she can and thereby compromises the swift evolution of his/her precision score. 2) At a given moment in a training session the trainee is slower than the expert and less precise. In this case, especially at early moments of training, the system needs to instruct the trainee to keep going, as he/she should get more precise and faster naturally. 3) At a given moment in a training session the trainee is slower than the expert and as precise or more precise. In this case, the system needs to instruct the trainee to try to go a little faster. 4) At a given moment the trainee is faster than or as fast as the expert and as precise or even more. In this case the trainee has beaten the expert. It this occurs, especially early in a training sequence, there is either a problem with the simulator task (i.e. the task does not produce adequate performance data that allow discriminating between levels of expertise, which is a problem that needs to be fixed), or the trainee is not a true novice.

4. Discussion

Surgical simulator training requires new forms of sensorimotor learning, an adaptive process that leads to improvement in performance through practice. This adaptive process consists of multiple distinct learning processes [28, 29]. Hitting a target, or even getting closer to it generates a form of implicit reward where the trainee increasingly feels in control. Successful error reduction, which is associated with specific commands relative to the specific motor task [24], can be optimized by gibing the trainee the right external feed-back. In this feed-back process, the integration of information from multiple senses (vision, touch, audition, proprioception) leads to improved adjustments in body, arm, or hand movements leading to perform the task with greater precision. Subjects are able to make use of error signals relative to the discrepancy between a desired and the actual movement or hand-tool-position, or a discrepancy between visual and proprioceptive estimates of body, arm, or hand positions [22, 28, 29]. The effective, if possible computer controlled, monitoring of strategies relative to speed-accuracy trade-offs in individual performance learning is a critical aspect of the skill assessment process given that cognitive theories of motor learning predict that strategy differences occur spontaneously when novices train to perform a motor task in a limited number of sessions [23-26], as is the case in laparoscopic simulator training. Conditional accuracy functions relate the duration of trial or task execution to a precision index reflecting the accuracy of the performance under conditions given. A variable relationship between speed and precision reflects hidden aspects of learning a beginner is usually not aware of [25]. In the fully trained expert, the trade-off between speed and precision does not vary. For the skill evaluator, the individual speed-accuracy trade-offs allow assessing whether a trainee is progressing, and this knowledge needs to be made available as early as possible in the training process. Simply comparing the skill levels of different trainees at the end of the process is not the right approach. What is needed are clear benchmark criteria for what the ideal performance of a successful trainee is to look like at the end of training. Such benchmark knowledge can be built into the computer controlling the simulator task on the basis of results from a certain number of training sessions of a surgical expert, as shown in our example here above.

Surgical simulators may be more or less task specific, or more or less "realistic" when compared with actual surgical task constraints they are supposed to train for. Many of them provide highly task specific feed-back, and the skills learnt on a given simulator may not transfer to other simulator

 doi:10.20944/preprints201810.0379.v1

Peer-reviewed version available at *Information* **2018**, *9*, 316; doi:10.3390/info9120316

tasks. One of the most important advantages of a simulation is to facilitate efficient training of skill outside the clinical context, which reduces the risk for patients. Different definitions of the notion of skill itself have produced different approaches to simulation-based surgical training. As pointed out by others previously [9, 10, 18], it is not always clear if more skilled individuals perform better on their assessments than less skilled or experienced individuals (construct validity), whether individuals who perform well on their evaluations will also perform well on other similar or vaguely related tasks (concurrent validity), or whether an assessment based on simulator training will predict future performance in the real-world context (predictive validity). Faced with this problem of providing reliable performance standards, it is essential that the system, the task, the metrics used to control performance learning during the task, and the mechanisms for providing feedback have somehow been validated by an expert to ensure that the training criteria and skill assessment provided by the system match those required real surgical tasks. Many different simulator tasks exist, which poses a problem of generalization of learning curves and skill transfer. The task model and control procedures suggested here should be implemented at the earliest stages of "dry lab" simulator training on eye-hand-tool coordination tasks that allow for computer controlled criteria for precision p at time t (Figure 7, for illustration). The early training task should successfully tell apart the performance levels of a surgical expert not trained on the simulator, as in the case discussed here above, and the performance levels of novice trainees. If an early training system satisfies this criterion, then it is indeed likely to measure critical aspects of surgical skill that will transfer to real surgical tasks and, ultimately, produce a valid selection of trainees that are likely to perform well in more specific tasks on physical models and in the clinical context, where direct supervision by experts will enable to promote individual expertise and excellence at the highest levels of surgical proficiency (Figure 8).

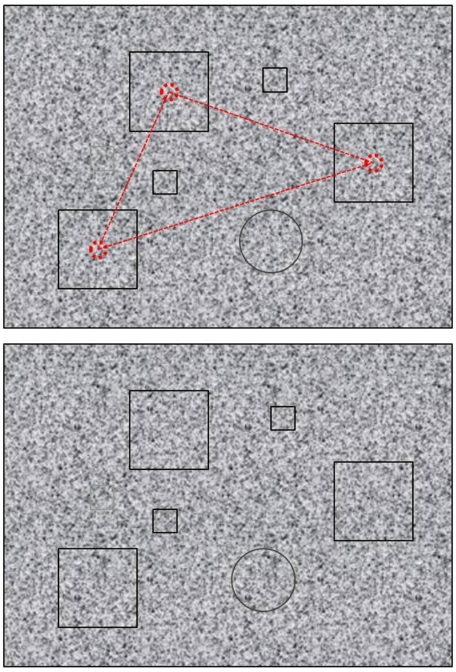


Figure 7. Illustration of computer controlled precision coding, based on pixel coordinates, for individual performance training in an image-guided simulator task of the pick and place type, as also described in our previous work [2, 5]. What the user sees on the screen in front of him/her is a camera view of an action space of the kind shown in the image at the bottom here. When the task instruction is, for example, to place a small object as swiftly and as precisely as possible on exactly the center of each of the three larger square shaped target areas, the computer may use the reference image coordinates, indicated here in the image on top by the red dotted lines, and compare them with the coordinates of the actual tool-object movements and/or positions of a given individual at a given trial momentum. Only the system "knows"

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 17 October 2018 doi:10.20944/preprints201810.0379.v1

Peer-reviewed version available at Information 2018, 9, 316; doi:10.3390/info9120316

the desired reference coordinates, and uses them to compute a task precision score (*p*) in terms of the number of pixels by which the tool-object position or tool-object trajectory produced by the user at a critical moment in time (*t*) during the task deviates from the supposedly ideal reference coordinates in any direction in space. Based on the known performance score of an expert built into the system, automatic feed-back may be provided to the user at any moment during training to control and/or correct his/her speed-accuracy trade-off.

Finally, whatever the simulator, a single performance metric inevitably gives a partial assessment of user performance [18]. Task completion time as a sole criterion has been explicitly demonstrated to be a poor or even misleading measure of surgical skill [1, 19]. Some metrics assume a simple global optimum value, such as a minimal tool path length, or a minimal completion time, and other quantities such as forces [4, 19] or velocities [21, 30-33], whose ideal values may vary in relation to changes in conditions, may have to be considered. Analysis of expert performance only can give insight into the nature of such dependencies and help develop better simulators.

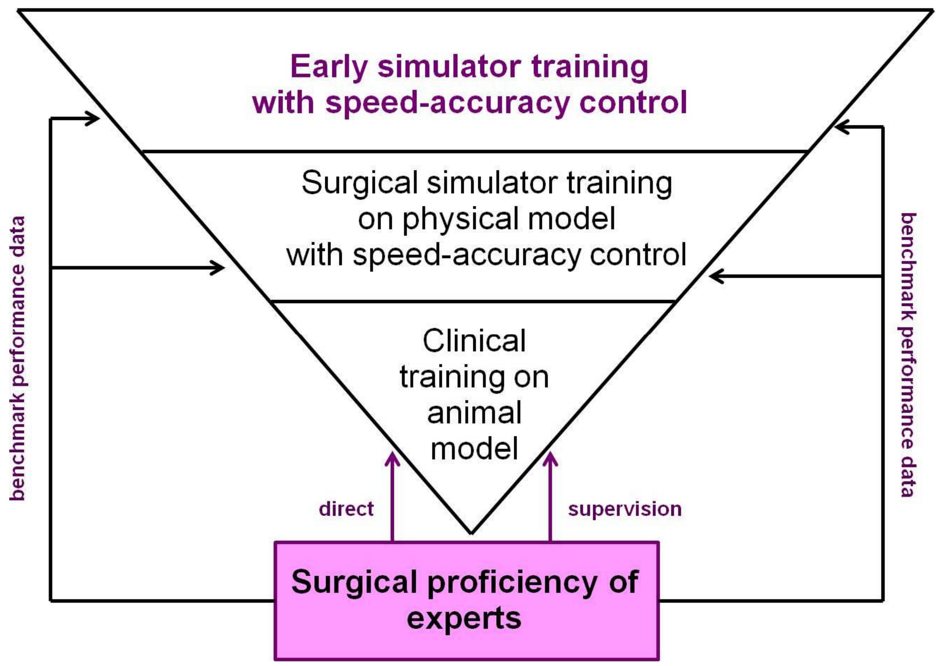


Figure 8. Speed-accuracy control on novice strategies in early training systems based on experts' benchmark performance data allows selecting for critical aspects of surgical skill that will transfer to further training on physical models in increasingly realistic surgical tasks. Ultimately, this will produce a valid selection of trainees that are likely to perform well at the highest levels of training under the direct supervision of an expert.

The fact that not all the important elements of surgical proficiency have been explored yet, does not change the heuristic validity of the general model for unsupervised training proposed here in this article. The model can, in principle, be adapted to any simulator system that exploits a criterion for task precision in task time. It is based on previously validated cognitive models of human performance learning showing that the individual speed-precision strategies of novices, which occur spontaneously and unconsciously [2, 22, 24, 25], can seriously compromise precision learning at all further stages of training if they are not controlled and corrected for as early as possible in the process. In robot-assisted surgical procedures, for example, where the camera moves along with the tool, for example [10-16], metrics such as camera movement frequency, camera movement duration, or camera movement interval are important indicators of technical skill, i.e. the proficiency/precision with which the trainee controls the tool, combined with other performance metrics such as task completion time, economy of tool motion, or master workspace range. A

Peer-reviewed version available at Information 2018, 9, 316; doi:10.3390/info9120316

decision model for unsupervised training procedures of the kind proposed here in this article could also be adapted to such performance criteria on the basis of device-specific expert performance benchmark data.

Funding: This research received no external funding.

Acknowledgments: Material support from the Centre National de la Recherche Scientifique (CNRS), Strasbourg University, and the Institut Hospitalo-Universitaire (IHU) de Strasbourg is gratefully acknowledged.

Conflicts of Interest: The author declares no conflict of interest.

References

1. Stunt JJ, Wulms PH, Kerkhoffs GM, Dankelman J, van Dijk CN, Tuijthof GJM. How valid are commercially available medical simulators? *Advances in Medical Education and Practice.* 2014; 5:385-395.
2. Batmaz AU, de Mathelin, M, Dresp-Langley, B. Getting nowhere fast: Trade-off between speed and precision in training to execute image-guided hand-tool movements. *BMC Psychology*. 2016a; 4: 55. doi:10.1186/s40359-016-0161-0.
3. Batmaz AU, de Mathelin, M, Dresp-Langley, B. Effects of indirect screen vision and tool-use on the time and precision of object positioning on real-world targets. *Perception.* 2016b; 45: Supplement, 286.
4. Batmaz AU, Falek M, Nageotte F, Zanne P, Zorn L, de Mathelin M, Dresp-Langley B. Novice and expert haptic behaviours while using a robot controlled surgery system. 2017 ; 13th IASTED International Conference on Biomedical Engineering (BioMed).
5. Batmaz, AU, de Mathelin, M, Dresp-Langley, B. Seeing virtual while acting real: Visual display and strategy effects on the time and precision of eye-hand coordination. *PLoS One*. 2017; 12: e0183789.
6. Batmaz, AU, de Mathelin, M, Dresp-Langley, B. Effects of 2D and 3D image views on hand movement trajectories in the surgeon's peri-personal space in a computer controlled simulator environment. *Cogent Medecine.* 2018; 5.
7. Gallagher, AG, O'Sullivan C. *Fundamentals In Surgical Simulation: Principles and Practice. Improving Medical Outcome - Zero Tolerance Series*. 2011; in P. Apell (Ed.). Springer Business Media, e-book.
8. Gallagher AG. Metric-based simulation training to proficiency in medical education: What it is and how to do it. *Ulster Med J* 2012; 81(3):107-113.
9. Gallagher AG, Ritter EM, Champion H, Higgins G, Fried MP, Moses G. Virtual reality simulation for the operating room: proficiency-based training as a paradigm shift in surgical skills training. *Ann Surg.* 2005 Feb; 241(2):364-72.
10. Dreyfus HL, Dreyfus SE, Athanasiou T. Mind over machine: the power of human intuition and expertise in the era of the computer. New York: Free Press; 1986.
11. Seymour NE, Gallagher AG, Roman SA, O'Brien MK, Andersen DK, Satava RM. Analysis of errors in laparoscopic surgical procedures. *Surg Endosc*. 2004; 18(4):592-5.
12. Van Sickle K, Smith B, McClusky DA, Baghai M, Smith CD, Gallagher AG. Evaluation of a tensiometer to provide objective feedback in knot-tying performance. *Am Surg*. 2005; 71(12):1018-23.
13. Van Sickle KR, Gallagher AG, Smith CD. The effect of escalating feedback on the acquisition of psychomotor skills for laparoscopy. *Surg Endosc.* 2007; 21(2):220-4.
14. A Reznick RK. Teaching and testing technical skills. *Am J Surg*. 1993; 165(3):358–361.
15. Chen C et al. Crowd-sourced assessment of technical skills: a novel method to evaluate surgical performance. *J Surg Res.* 2014; 187(1):65–71.
16. Moorthy K et al. Objective assessment of technical skills in surgery. BMJ. 2003; 327(7422):1032–1037.
17. Sewell C, Morris D, Blevins NH, Dutta S, Agrawal S, Barbagli F, Salisbury K. Providing metrics and performance feedback in a surgical simulator. *Computer Aided Surgery.* 2008; 13(2): 63-81.
18. Ritter EM, McClusky DA, Gallagher AG, Smith CK. Real-time objective assessment of knot quality with a portable tensiometer is superior to execution time for assessment of laparoscopic knot-tying performance. *Surgical Innovation*. 2005; 12(3):233–237.
19. Rosen J, Hannaford B, Richards CG, Sinanan MN. Markov modeling of minimally invasive surgery based on tool/tissue interaction and force/torque signatures for evaluating surgical skills. *IEEE Trans Biomed Eng.* 2001; 48(5):579–591.

20. Jarc AM, Curet MJ. Viewpoint matters: objective performance metrics for surgeon endoscope control during robot-assisted surgery. *Surg Endosc* . 2017 ; 31:1192–1202.
21. Batmaz AU, de Mathelin M, Dresp-Langley B. Effects of indirect screen vision and tool-use on the time and precision of object positioning on real-world targets. *Perception*, 2016; 45(ECVP Supplement): 196.
22. Fogassi L, Gallese V. Action as a binding key to multisensory integration. In Calvert G, Spence C, Stein BE, editors, *Handbook of multisensory processes.* Cambridge, MA: MIT Press. 2004; p. 425–441.
23. Bonnet C, Dresp B. A fast procedure for studying conditional accuracy functions. *Behavior Research, Instruments & Computers*. 1993; 25: 2-8.
24. Fitts P M. The information capacity of the human motor system in controlling the amplitude of movement. *Journal of Experimental Psychology.* 1954; 47: 381-391.
25. Meyer DE, Irwin A, Osman, AM, Kounios J. The dynamics of cognition and action: Mental processes inferred from speed-accuracy decomposition. *Psychological Review, 1988;* 95: 183-237.
26. Luce, RD. Response times: Their role in inferring elementary mental organization. 1986; New York: Oxford University Press.
27. Held R. Visual-haptic mapping and the origin of crossmodal identity. *Optometry and Vision Science*. 2009; 86, 595-598.
28. Henriques, DY, Cressman, E.K. Visuo-motor adaptation and proprioceptive recalibration. *Journal of Motor Behavior*. 2012; 44, 435-444.
29. Krakauer, JW, Mazzoni, P. Human sensorimotor learning: adaptation, skill, and beyond. *Current Opinion in Neurobiology*. 2011; 21, 636–644.
30. Goh AC et al. Global evaluative assessment of robotic skills: validation of a clinical assessment tool to measure robotic surgical skills. *J Urol.* 2012; 187(1):247–252.
31. Smith R, Patel V, Satava R. Fundamentals of robotic surgery: a course of basic robotic surgery skills based upon a society consensus template of outcomes measures and curriculum development. *Int J Med Robot Comput Assist Surg.* 2014; 10(3):379–384.
32. Aiono S et al. Controlled trial of the introduction of a robotic camera assistant (Endo Assist) for laparoscopic cholecystectomy. *Surg Endosc Other Interv Tech*. 2002; 16(9):1267–1270.
33. King BW et al. Towards an autonomous robot for camera control during laparoscopic surgery. *J Laparoendosc Adv Surg Tech.* 2013; 23(12):1027–1030.