11 What Measures Represent Performance?

Tim Horeman, Kevin Sherman,
and Gabriëlle J.M. Tuijthof

Take Home Messages

- Motion- and time-based metrics can track instrument handling efficiency in endoscopy training.
- Force-based metrics can track tissue handling skills during training in endoscopy training.
- Motion, time and force information can be combined in specific metrics that indicate risks on hazards such as accidental tissue puncture or rupture.
- Metric-based post-task should be task dependent and easy to understand.
- Sufficient metrics are available to monitor training performance.

T. Horeman (✉)
Department of Biomechanical Engineering, Delft University of Technology, Delft, The Netherlands
e-mail: t.horeman@tudelft.nl

K. Sherman, MD
Department of Orthopedics, Castle Hill Hospital, Cottingham, East Yorkshire, UK
e-mail: kp_sherman@hotmail.com

G.J.M. Tuijthof
Department of Biomechanical Engineering, Delft University of Technology, Delft, The Netherlands

Department of Orthopedic Surgery, Academic Medical Centre, Amsterdam, The Netherlands
e-mail: g.j.m.tuijthof@tudelft.nl

11.1 Definitions

Assessing performance is one of the key elements that guide training and progress. To avoid discussions, the following definitions are made:

Metric, measure, parameter We define a metric as a quantity that in this context is supposed to reflect (part) of the performance of a trainee. Other terms that are considered as being synonym to a metric are *measure* and *parameter*.

Objective metrics are registered with sensors that are stand-alone or can be built-in a simulator. Optionally, the measured data from the sensors are post-processed to derive the metric.

Subjective metrics use expert judgments regarding performance behaviour. These can be partly objectified by scoring on rubrics using checklists.

Performance efficiency is an economic highly goal-oriented performance.

Performance safety is delicate tissue interaction and considerate instrument handling.

Proficiency in terms of instrument handling is defined as the optimal combination of performance efficiency and safety.

Direct feedback is given directly during the execution of a training task. Direct feedback is also named *real-time feedback*.

Post-task feedback is given after completion of a task. This type of feedback usually consists of several metrics and gives an overview of the entire task execution.

M. Karahan et al. (eds.), *Effective Training of Arthroscopic Skills*,
DOI 10.1007/978-3-662-44943-1_11, © ESSKA 2015

Subjective metrics are discussed in Chap. 13 as they are also suitable for application in the operating room to monitor complex tasks as they reflect a more holistic type of assessment.

11.2 Introduction

In previous chapters, the needs or training goals of arthroscopic skills were identified, as well as the possible means that is the various simulated training environments. This chapter focuses on the performance assessment or tracking of a trainee, which is the third key element required to provide a proper education environment. Establishing objective performance metrics is a challenging task, different approaches can be followed and many aspects influence the usability of a metric. One approach is to translate the training goals into measures; another approach is to translate psychomotor skills into measures. Both approaches require some form of decomposition into smaller elements that can be measured with a sensor. For example, the dexterity tools and tests described in Chap. 6 represent a decomposition into basic psychomotor skills that all can be measured with a single metric such as time. *However*, overall task performance of more complex tasks does not necessary equal the summation of the performance of several part-task goals or skills. Additionally, performance metrics should also be capable to discriminate between levels of expertise, easy to be interpreted, and to give directions to improve learning. One other aspect that should be taken into account is that objective metrics not only assess the performance of the trainees but also assess the performance of the simulator (Chap. 9). Bearing these considerations in mind, we do start by giving an overview of metrics that have been presented in literature to be useful in endoscopic training. They are categorized based on their suitability to represent performance efficiency or performance safety and suitability for direct or post-task feedback. Notice that not all presented metrics have been applied yet to training of arthroscopic skills. Based on this overview, the translation is made towards performance tracking and feedback by discussing several examples.

A standard setting is introduced based upon which most metrics are presented graphically. Figure 11.1 shows an arthroscope and a probe that are inserted in a phantom knee joint. The path each tip of the instruments has moved for a period of 2.2 s is represented by the two 3D curves. The data are actual data from an evaluation test where a *navigate and probe task* was performed. The instruments are drawn in the mean direction of the travelled path. This example will be used throughout the section to illustrate the metrics concerning motion.

11.3 Metrics Reflecting Performance Efficiency

11.3.1 Task Repetition

The first quantitative metric is the number of task repetitions required to achieve a certain level of completion. This metric gives insight in the capability to learn in a new training environment and basically reflects the learning curve in time. Task repetition is one of the few metrics that does not require sensors to be objectively documented as long as the definition of *satisfactory completion* is clear. Therefore, the metric is highly suitable in different kinds of training programs (Scott et al. 2000).

11.3.2 Task Error

Similarly, to the number of task repetitions, the number of task errors can be documented without sensors. Examples of task errors can be the number of missed abnormalities or landmarks during an inspection task in the joint (Bliss et al. 2005; Hodgins and Veillette 2013; Sherman et al. 2001), the number of dropped objects from an instrument during a pick and remove task (Pellen et al. 2009; Rosser et al. 2006) or the number of misplaced suture insertions when performing

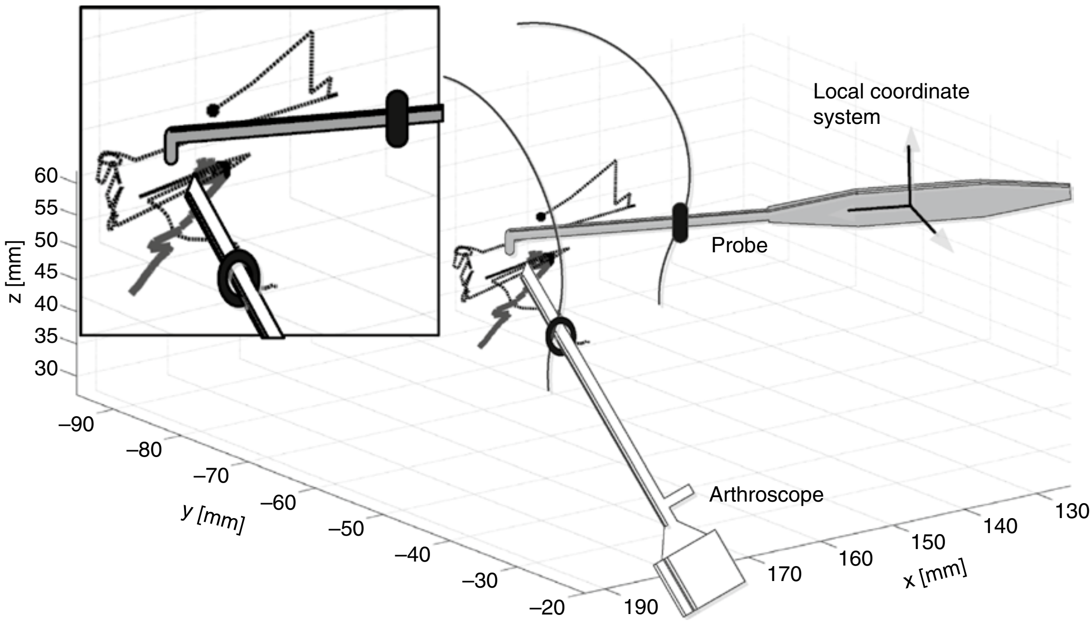

Fig. 11.1 Arthroscope and probe oriented in the lateral compartment of a phantom knee joint. The instruments are oriented in the mean direction of the travelled path. The two 3D curves represent the travelled path for a period of 2.2 s of the arthroscope (*grey*) and probe (*black*), respectively. Start of the paths are marked

meniscal suturing. As is shown by the examples, the errors can be knowledge based or skill based and reflect performance in such a manner that it is also applicable in real-life surgery. Notice that the task errors should be well defined to be able to document their frequency.

11.3.3 Task Time

Task time (t) is defined as the period of time elapsed between the start of a task and the first second after completion of the task (Fig. 11.2):

$$t = t_{end} - t_{start}$$

Task time is found to be most discriminating between levels of experience as it highly reflects economy of motion (Andersen et al. 2011; Gomoll et al. 2008; Howells et al. 2008; Martin et al. 2011; Martin et al. 2012; McCarthy et al. 2006; Oropesa et al. 2013; Pedowitz et al. 2002; Tuijthof et al. 2010; Tuijthof et al. 2011b; Verdaasdonk et al. 2007). Advantages are that task time is easy to understand and relatively easy to implement as it does not require high-end sensory equipment. A simple smartphone or timer is already sufficient.

11.3.4 Idle Time or State

Idle time (*it*) is defined as the percentage of the task time during which an instrument is held still (Chmarra et al. 2010; Oropesa et al. 2013). Idle state (*is*) is the number of instrument 'held stills' during task execution. For knee arthroscopy training, idle state was defined as the number of instances during which the subject looked away from the arthroscopy display unit to look at his or her hands while holding the arthroscope (Alvand et al. 2012).

Idle time or state reflects workflow interruptions due to a lack of knowledge (e.g. trainees hamper and do not know what kind of instrument to use) or task error (e.g. needle drops and need to be picked up). Therefore, this metric is in line with task time and path length in representing

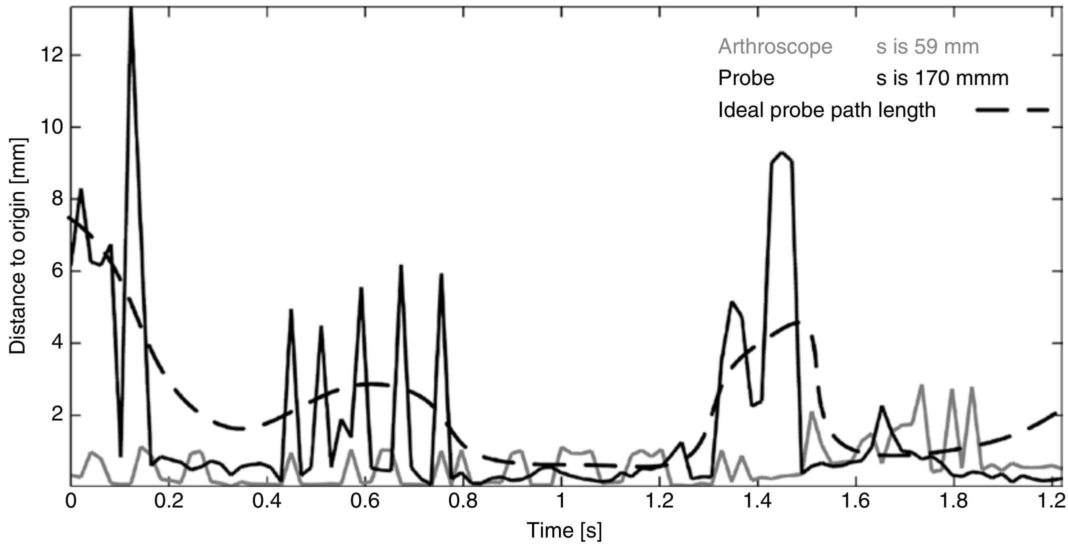

Fig. 11.2 Graphical presentation of task time, path length of arthroscope and probe and economy of movement, using the two 3D curves as presented in Fig. 11.1. The striped line shows an example of an ideal path length to demonstrate economy of movement

task efficiency. A study of Rosen and co-workers found that experts wasted less time between tissue manipulations in a box trainer compared with novices by looking at the idle states (Rosen et al. 1999).

11.3.5 Path Length

Path length (s) is defined as the total length of the path that the tip of an instrument or arthroscope has travelled in space (Fig. 11.2):

$$s = \sum_{i = start}^{end} \sqrt{\left(x_{i+1} - x_i\right)^2 + \left(y_{i+1} - y_i\right)^2 + \left(z_{i+1} - z_i\right)^2}$$

where the vectors x_i, y_i and z_i are the coordinates of the position in space of the tip for each time stamp. Physically, the path length can be measured by placing a rope at the start point and following the path with the rope until the end point. The total length of the rope needed to follow the trajectory is the path length of the instrument tip or arthroscope.

Path length is used as a measure to determine the efficiency of instrument and/or arthroscope motion (Andersen et al. 2011; Gomoll et al. 2007; Howells et al. 2008; McCarthy et al. 2006; Oropesa et al. 2013; Tashiro et al. 2009; Verdaasdonk et al. 2007). Path length has a role in goal orientation. Careful implementation of this parameter is required, as its value in reflecting performance depends on a well-defined trajectory. The reason is that experts do not necessarily take the shortest path to a target. This was nicely illustrated by Chmarra and co-workers (Chmarra et al. 2006). They found that a certain task had two clearly distinctive phases with the first being the 'seeking phase' when moving towards the target and the second being the 'retracting phase' when moving away from the target. Especially, in the 'seeking phase', differences were found between novices and experts with the experts demonstrating a significantly shorter path length, whereas the 'retracting phase' did not present differences.

Measuring the path length is easy when using virtual reality simulators as the simulated environment needs to know where the instrument is in the virtual space. So in order to let the simulation work, the orientation of the instruments is calculated anyhow and thus can be presented to the trainee. Tracking instrument motion in the physical world requires more effort, especially as a certain measurement accuracy is needed. Several tracking systems are indicated in Chap. 13, which can also be used in the operating room. A nice overview of tracking systems applied in endoscopic trainers is presented by Chmarra and co-workers (Chmarra et al. 2007). Once you have the 3D position data in time, other performance metrics can be deducted as well: economy of movement, motion volume, idle time, speed and smoothness.

11.3.6 Economy of Movement

Economy of movement (*em*) is defined as the percentage between the *ideal path length* (s_{ideal}) necessary to complete a task and the total path length (Bayona et al. 2008; Oropesa et al. 2013) (Fig. 11.2):

$$em = \frac{s_{ideal}}{s}$$

This implies that *em* gives a score between 0 and 100 %, with 100 % indicating *ideal economy of movement*. This metric is highly correlated to path length but compensates for the drawback of path length in defining the *ideal path length*, which is as indicated not necessarily the shortest length of the trajectory. Pedowitz and co-workers have used this metric and demonstrated differences between groups of different expertise (Pedowitz et al. 2002). A poor strategy to execute the task, task errors, steady hand, ambidexterity and manual precision are all technical skills that are reflected in the economy of motion metric.

11.3.7 Depth Perception

Depth perception (*dp*) is defined as the total distance travelled by an instrument along its longitudinal axis (Oropesa et al. 2013):

$$dp = \sum_{i=start}^{end} \sqrt{\left(l_{i+1} - l_i\right)^2}$$

With l_i being the distance for each time stamp. To determine *l*, the data first have to be oriented to the local coordinate system of the instrument, where one of the axes is defined in the direction of the longitudinal axis of the instrument (Fig. 11.1). Depth perception can be an indicator for poor instrument control when moving instruments perpendicular to the endoscopic image (Maithel et al. 2006; Rosen et al. 2006; Stylopoulos et al. 2004). This is caused by the fact that arthroscopic images are presented on a monitor showing a two-dimensional projection, whereas in reality instruments are navigated in a three-dimensional space inside the joint cavity. As this type of eye-hand coordination is completely different from everyday tasks, especially novices have difficulties in translating the three-dimensional environment to a two-dimensional representation, which results in poor positioning of the instrument tip. For application in arthroscopy, which uses a 30° angled arthroscope, *dp* should be determined of the instrument, and not of the arthroscope.

11.3.8 Volume of Motion

The volume of motion (V_{motion}) is defined as the volume of a 3D-dimensional ellipsoid spanned around the standard deviations of motion ($STD1_{motion}$, $STD2_{motion}$ and $STD3_{motion}$) along the three main directions of motion (Oropesa et al. 2013) (Fig. 11.3):

$$V_{motion} = \frac{4}{3}\pi\left(STD1_{motion} \cdot STD2_{motion} \cdot STD3_{motion}\right)$$

To determine the three main directions of motion, which do not necessarily coincides with the global coordinate system, a mathematical procedure is performed to convert the set of observations of possibly correlated variables into a set of values of uncorrelated variables called principal components (Chmarra et al. 2010; Horeman et al. 2012a). Subsequently, the standard deviations along those three main directions of motion are calculated and define the shape of the ellipsoid (Fig. 11.3).

Volume of motion is a measure for the space required by a trainee to complete the task. Similar as for path length, an *ideal volume of*

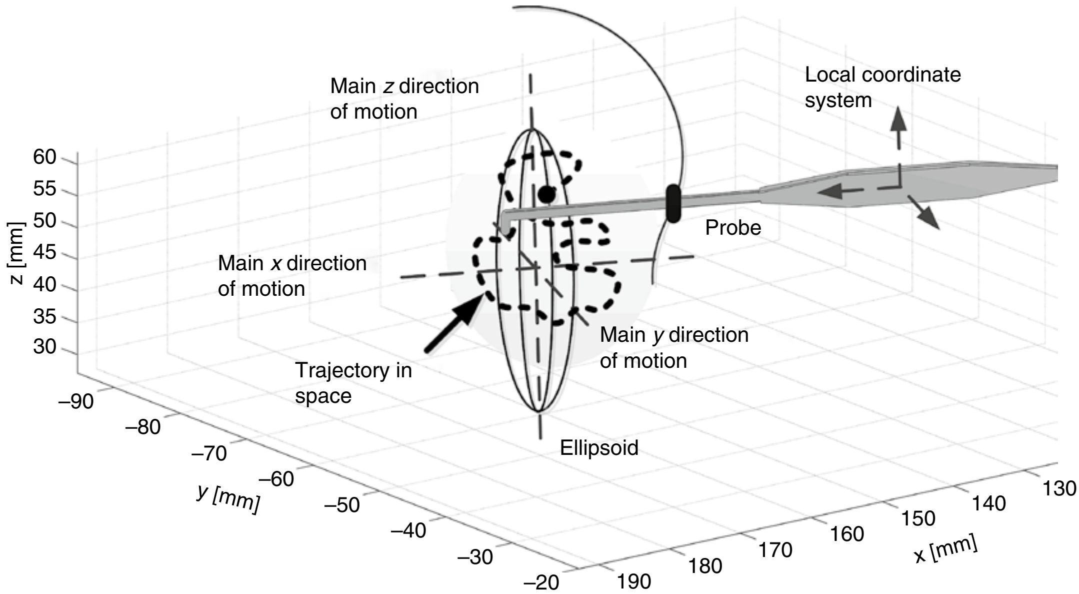

Fig. 11.3 Graphical presentation of the probe that is moved in 3D along a certain trajectory (*dotted line*). The solid lines represent a fitted ellipsoid which dimensions depend of motion (STD1 motion, STD2 motion and STD3 motion) along the three main directions of motion (indicated by straight dotted lines)

motion should be defined to compare trainee's performance with expert performance, as the smallest volume of motion does not necessarily reflect the optimal performance (Horeman et al. 2014b). Different from other metrics as path length, volume of motion is influenced by the direction of the instrument tip motion in three-dimensional space. For example, if the instrument tip is only moved along the instrument's longitudinal shaft, the path length increases, while the volume of motion remains zero until the instrument is also moved along all three axes of its local coordinate system.

11.4 Metrics Reflecting Performance Safety

11.4.1 Collision

A collision (*col*) is mostly defined as an instrument that unnecessarily touches surroundings. In this definition, a collision can be seen as a subset set of the metric task error, reflecting an error that potentially damages healthy tissue. So, the number of collisions reflects the number of times tissue might have been damaged during a task execution (Andersen et al. 2011; Gomoll et al. 2008; McCarthy et al. 2006; Pedowitz et al. 2002; Verdaasdonk et al. 2007). Similarly, to the path length, the number of collisions is a metric that is easily implemented in virtual reality simulators as this information needs to be determined to represent the virtual environment. Implementing collision detection in physical models is less straightforward. Some have applied electric wires that upon connection close an electric circuit which creates a buzz sound (Meyer et al. 1993; Tuijthof et al. 2003).

In other cases, collision detection can be determined based on the presence of certain force patterns in the recorded data (Horeman et al. 2014a). Collision detection can be based on the presence of frequencies in the recorded data for tasks such as pattern cutting, suturing and peg transfer in a box trainer (Smith et al. 1999). When applied for arthroscopy, such force frequency patterns might be detected when probing hard bony surfaces compared to probing softer fatty tissues, since the force build-up will be steeper in case of hard bone. It might be interesting to investigate if this theory can be applied, to distinguish between collision of instruments (hard surfaces) and collision with tissue (softer surfaces).

11.4.2 Out of View Time

The out-of-view time (*ovt*) is defined as the percentage of the task time that the instrument tip is not visible in the arthroscopic view (Horeman et al. 2014b), or as Alvand and co-workers defined "Prevalence of instrument loss" (Alvand et al. 2012). This metric reflects safety, as an instrument that is out of view can inflict unintended damage to the surrounding tissue when it is manoeuvred blindly. Out-of-view time can be quantified by analysis of recorded arthroscopic images. However, the out-of-view time can also be calculated if the 3D position and orientation of the arthroscope and instrument are measured as can be done with 3D tracking systems. Also, the arthroscope's view angle and diameter need to be known. The latter two parameters define the view cone, which can be considered a volume in space (Fig. 11.4). Using the orientation of the arthroscope in space, the cone is attached in a fixed 30° angle to the tip of the arthroscope. Subsequently, it is verified if the coordinates of the instrument tip coincide with the cone for every time stamp by calculating the distance between the instrument tip and the outer surface of the view cone.

11.4.3 Tip-to-Tip Distance

The tip-to-tip distance (*t2td*) is defined as the distance between the arthroscope and the instrument tip for the entire task trajectory (Fig. 11.5):

$$t2td = \sum_{i=start}^{end} \sqrt{\left(x_{iscope} - x_{iprobe}\right)^2 + \left(y_{iscope} - y_{iprobe}\right)^2 + \left(z_{iscope} - z_{iprobe}\right)^2}$$

With x_i, y_i and z_i being the coordinates of the instruments' tips for every time stamp. This metric reflects the zone in which the manipulation takes place and is correlated with out-of-view time, since a high maximum tip-to-tip distance suggests that the instrument might be out of the arthroscopic view. This finding suggests that safe tissue handling is compromised.

Although the discriminating power of *t2td* depends highly on the type of task, a high mean tip-to-tip distance can inform the trainee to improve his/her overall safety performance during a task (Horeman et al. 2014b).

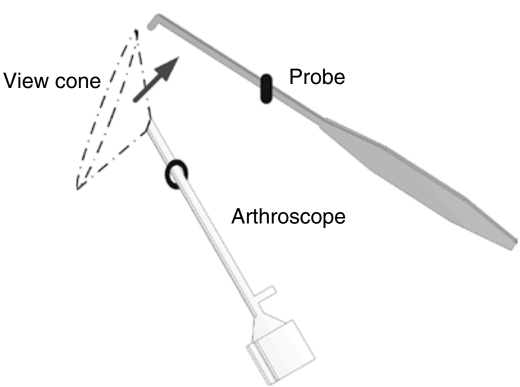

Fig. 11.4 Graphical presentation of the view cone of an arthroscope. The instrument has to be within this view cone to achieve full view time. The *arrow* indicates that the probe tip cannot be visualized by the arthroscope resulting in the start of registering the out-of-view time

11.4.4 Motion Speed

In general, motion speed (v) is defined as the distance travelled per time (Oropesa et al. 2013) (Fig. 11.6):

$$v = \frac{s}{t}$$

where s is the path length and t is the time (e.g. task time).

Motion speed links position information to time information. In arthroscopy, the average motion speed has been used to assess performance (Gomoll et al. 2007, 2008), which reflects

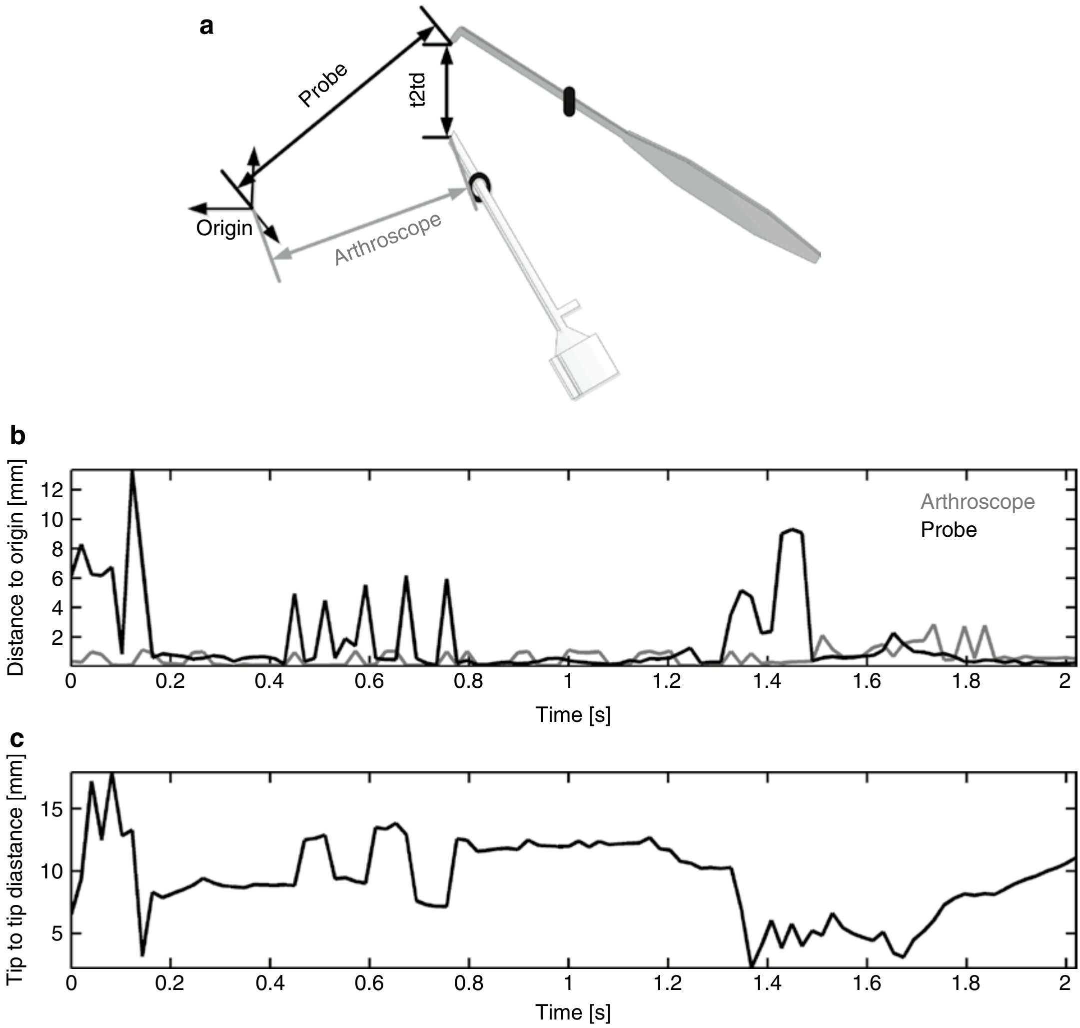

Fig. 11.5 (**a**) The tip-to-tip distance (*t2td*) is indicated in the sketch for one single time stamp. (**b**) The distance between point of origin and the distance to the instrument and arthroscope tips (Reproduced from Fig. 11.2). (**c**) From **b** the tip-to-tip distance curve is determined for this trajectory by subtraction of the distances between probe and arthroscope per time stamp

economy of movement. However, motion speed can also be used as follows. If the speed of the instruments is calculated per time stamp, the maximum speed can be determined for a training task. In a standard *position control* situation as is the case in robotic motion applications, a high instrument motion speed can be associated to overshoot and therefore poor position control. Returning to the clinical setting, instrument loading can build up due to contact between an instrument and the arthroscope or bony surface; when this loading is suddenly released, the instrument can overshoot and accidentally hit other tissues (Horeman et al. 2013). Thus, if uncontrolled instrument speeds occur during arthroscopy, it is possible that such events damage delicate anatomic structures around the operative zone. Due to force build-ups, instrument motion speed is linked to surgical safety (Horeman et al. 2013; Tarnay et al. 1999).

11.4.5 Motion Smoothness

Motion smoothness is defined as changes in instrument acceleration. Motion smoothness can

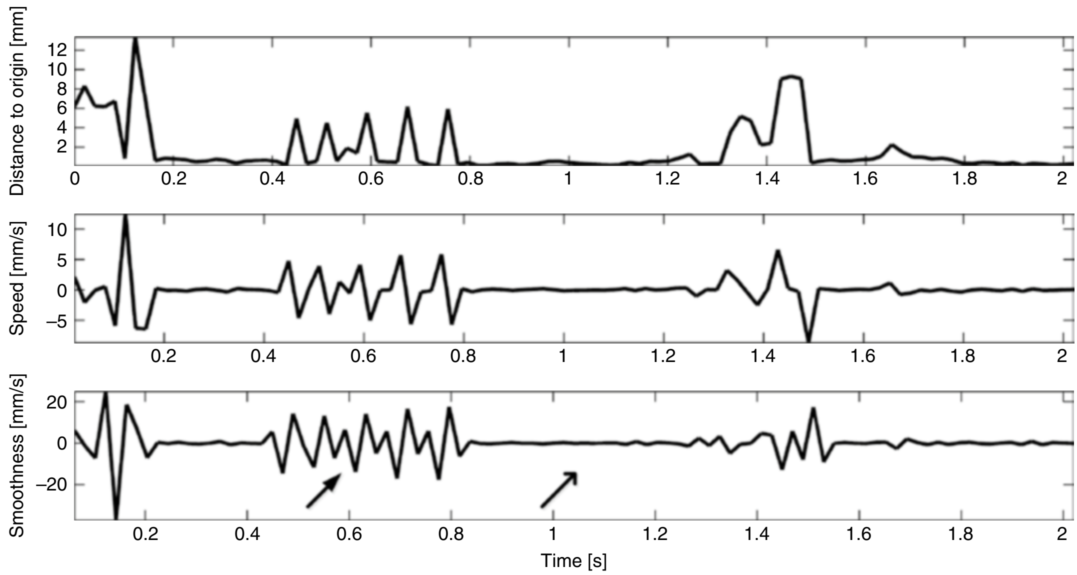

Fig. 11.6 Graphical presentation of the path length (s), the speed of motion (v) and motion smoothness (J) of the probe for the trajectory as presented in Fig. 11.1. In the motion smoothness curve, the closed arrow indicated a 'jerky' motion trajectory and the open arrow a *smooth* motion trajectory

be derived in various ways, we present one of the calculations suggested by Hogan and co-workers (Hogan and Sternad 2009), which is the root mean squared jerk (J) (Fig. 11.6):

$$J = \sqrt{\frac{1}{t}\sum_{i=start}^{end}\left(\frac{\Delta^3 x_i}{\Delta t^3}\right)^2 + \left(\frac{\Delta^3 y_i}{\Delta t^3}\right) + \left(\frac{\Delta^3 z_i}{\Delta t^3}\right)}$$

with $\frac{\Delta^3 xi}{\Delta t^3}$, $\frac{\Delta^3 yi}{\Delta t^3}$ and $\frac{\Delta^3 zi}{\Delta t^3}$ being the third derivative of the x-, y- and z-position in time. A high motion smoothness suggests jerky movements of the surgical instrument (Oropesa et al. 2013), which again can compromise safe tissue manipulation.

11.4.6 Force Magnitude

Analogous to the indication of a position in space having an x, y and z component, the magnitude of a force is composed from its F_x, F_y and F_z magnitudes:

$$F = \sqrt{\left(F_x + F_y + F_z\right)^2}$$

Besides quick and jerky motions, the force applied to surrounding tissue could lead to unintended damage of healthy tissue. This is especially true for healthy tissues in the intra-articular joint space that have limited healing potential (meniscus and cartilage) or vascular structures that heavily bleed. Tissue damage occurs if the tissue is loaded with magnitudes beyond the tissue's strength. Consequently, the force magnitude can qualify as a metric for monitoring safety. This suggests that calculation of the mean force exerted during a certain task might not sufficiently reflect safety performance. That is why the maximum peak force (Tashiro et al. 2009), as well as the standard deviation of forces (Chami et al. 2008; Horeman et al. 2010), and exceeding a certain threshold force (Obdeijn et al. 2014; Tuijthof et al. 2011a) have been suggested (Fig. 11.7). For all three options, significant differences were found between novices and experts.

The standard deviation of the (absolute) force indicates the ability of the trainee to apply a constant force on an object during a task. Especially in bimanual tasks as tissue stretching for dissection, anchor placement or needle driving a well-directed constant force on anchor, needle or tissue improves performance (Horeman et al. 2012a). Due to its nature, this component of the force is most informative in tasks that require

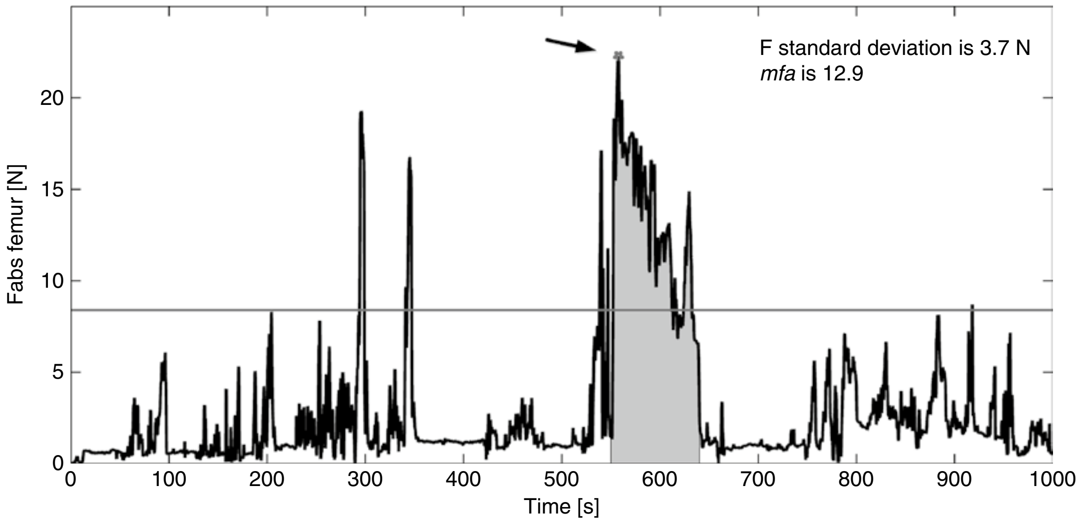

Fig. 11.7 Graphical presentation of the absolute force (*F*) when probing the femoral condyle in the PASSPORT simulator. Also the maximum peak force is indicated by the star (highlighted by the *arrow*), a threshold level is presented by the *grey line*, and the maximum force area (*mfa*) is indicated for this force pattern

continuous contact between instrument and task components.

Exceeding a certain threshold force is closely related to the collision metric but offers a more precise definition of the so-called collision as certain threshold needs to be exceeded for a certain period of time to qualify as collision (Fig. 11.7).

Measuring the actual forces during a task is not straightforward both in the virtual and the actual world. In virtual environments, haptic devices are used, which usually only give feedback on the tip forces. Also, their quality is not sufficient to mimic adequate haptic sensation, especially when machining of tissue is involved such as cutting, punching and drilling. In physical environments, instruments have been modified and equipped with sensors (Chami et al. 2006), 3D commercial force sensors that measure all force and moments (Tashiro et al. 2009) or a force platform that solely measure 3D reaction force of the tissue as result of instrument tissue manipulation (Horeman et al. 2010).

11.4.7 Force Direction

Analogous to the indication of a position in space having an x, y and z component, the direction of a force is composed from the ratios of the F_x, F_y and F_z magnitudes. An example is the expression of the force direction using two angles, with φ being defined as the direction of the force in the vertical plane:

$$tan\varphi = \frac{F_z}{F_y}$$

and γ being defined as the direction of the force in the horizontal plane.

$$tan\gamma = \frac{F_y}{F_x}$$

A study aimed to determine the force magnitude and directions exerted during arthroscopic navigation and inspection of a cadaver wrist showed that not so much the magnitude of the forces but the direction of the force differed significantly between experts and novices (Obdeijn et al. 2014). The experts executed forces containing a more perpendicular orientation on the cartilage tissue, whereas the novices executed forces containing a more shearing component. This might be an important difference, as navigation in the wrist is difficult due to the tide joint space and complex-shaped bones. In another study using the same setup of wrist arthroscopy training on a cadaveric specimen, it was found that novices did

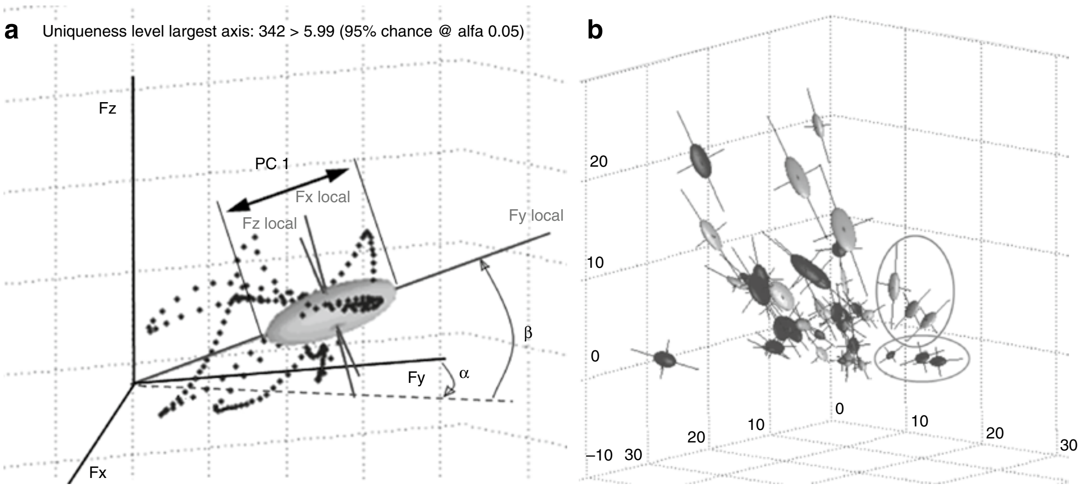

Fig 11.8 (**a**) 3D variability in forces. The *dots* represent the force in the global coordinate system (Fx, Fy, Fz). The ellipsoid is fitted on the force data and the orientation of PC1 force along Fx-local is defined by angles α and β. (**b**) Encircled ellipsoids of experts versus novices (not encircled) in a needle-driving task show that the force volume and direction can reveal consistency of an expert surgeon over three trials

not improve the loading of the tissues inside the wrist joint (Obdeijn et al. 2014). This suggests that force direction information can be indicative of a trainee's performance and the learning curve of novices.

11.4.8 Force Area

The maximum force area (m*fa*) is defined as the area of the absolute force peak in the force-time curve (Fig. 11.7):

$$mfa = \sum_{i=start}^{end} \sqrt{\left(F_{x,i} + F_{y,i} + F_{z,i}\right)^2}$$

With F_x, F_y and F_z being the x, y and z components of the force vector in space for every time stamp. In earlier work, *mfa* was referred to as force peak (Horeman et al. 2012a).

The starting time t_{start} and t_{end} can automatically be defined with different mathematical procedures, such as the point in time where the building of the absolute peak force is started or stopped, respectively, which can be deducted from the derivative of the force-time curve (Horeman et al. 2014b). The *mfa* indicates another aspect of the metric collision and can be considered as an elaboration of solely measuring the peak force by taking into account its duration as well. Peak forces that are only applied for a brief time period (e.g. less than 0.5 s) might not inflict any damage, whereas a relative high force that is applied for a prolonged time period could cause tissue damage. As indicated before, the aptness of this measure compared to other safety-related performance measures depends on the task to be trained.

11.4.9 Volume of Force

The volume of force (V_{force}) is defined as the volume of a 3-dimensional ellipsoid spanned around the standard deviations of force ($STD1_{force}$, $STD2_{force}$, and $STD3_{force}$) along the three main directions of motion (Horeman et al. 2012a) (Fig. 11.8):

$$V_{force} = \frac{4}{3}\pi\left(STD1_{force} \cdot STD2_{force} \cdot STD3_{force}\right)$$

This definition as setup analogous to the metric 'volume of motion' as the same mathematical procedure (principal component analysis) can be applied to the x, y and z components of the force (Chmarra et al. 2010; Horeman et al. 2012a).

Volume of force is a measure for the forces required by a trainee to complete the task. Similar as for the volume of motion, an 'ideal volume of force' should be defined to compare trainee's performance with expert performance, as the smallest volume of motion does not necessarily reflect the optimal performance. The discriminative power of volume of force was determined in a study by Horeman and co-workers for an endoscopic suture task and confirmed (Horeman et al. 2012a).

11.5 Metrics Summary

Figure 11.9 shows how time, motion and force metrics inform about efficiency and safety performance. In endoscopic surgery, efficiency is reflected most strongly by task time, followed by the metrics task repetition and idle time. These three added with economy of movement indicate how fast the trainee adapts to a new situation when training. Another time-based parameter, out-of-view time, reflects safety performance. Distance metrics can inform both on efficiency performance (e.g. large or short path length to complete a task) and on safety performance (e.g. tip-to-tip distance). Force metrics measuring interaction forces during tissue manipulation can exclusively be associated to tissue damage (e.g. peak force, force threshold, force direction), but cannot inform on safe instrument handling if there is no interaction with the tissue. Since the above mentioned metrics are composed of single parameter measurement and need no post-calculation with mathematical procedures, they qualify to be used for direct feedback on performance (Table 11.1).

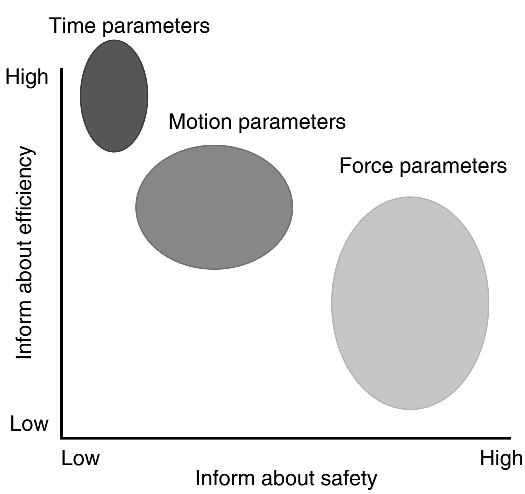

Fig. 11.9 The solid fields indicate the information that time and motion metrics can contain about efficiency and safety of a surgical action. The hatched field indicates the potential information that force metrics contain

Table 11.1 Categorization of metrics

	Performance efficiency	Performance safety
Direct feedback	Task time	Collision
	Idle time	Tip-to-tip distance
	Path length	Motion speed
		Force magnitude
		Force direction
Post-task feedback	Task repetitions	Collision
	Task errors	Out-of-view time
	Task time	Tip-to-tip distance
	Idle time	Motion speed
	Path length	Smoothness
	Economy of movement	Force magnitude
	Volume of motion	Force direction
	Task errors	Force area
	Task time	Volume of force

For other tasks, it can be helpful to use a metric that combines several parameters (Table 11.1). The combined information of time and motion is reflected by the metrics motion speed, motion smoothness and volume of motion. The combined information of time and force is reflected by the metrics force area and volume of force. These metrics of combined parameters tend to inform on safety performance (Table 11.1). The combined parameter metrics require post-processing or give a summary of the entire task performance, which makes them more suitable to use for post-task feedback.

11.6 Discussion

11.6.1 From Measuring Metrics to Training

Direct feedback on performance is probably most relevant to apply in training, when on aims to learn A) how to follow a protocol or a certain

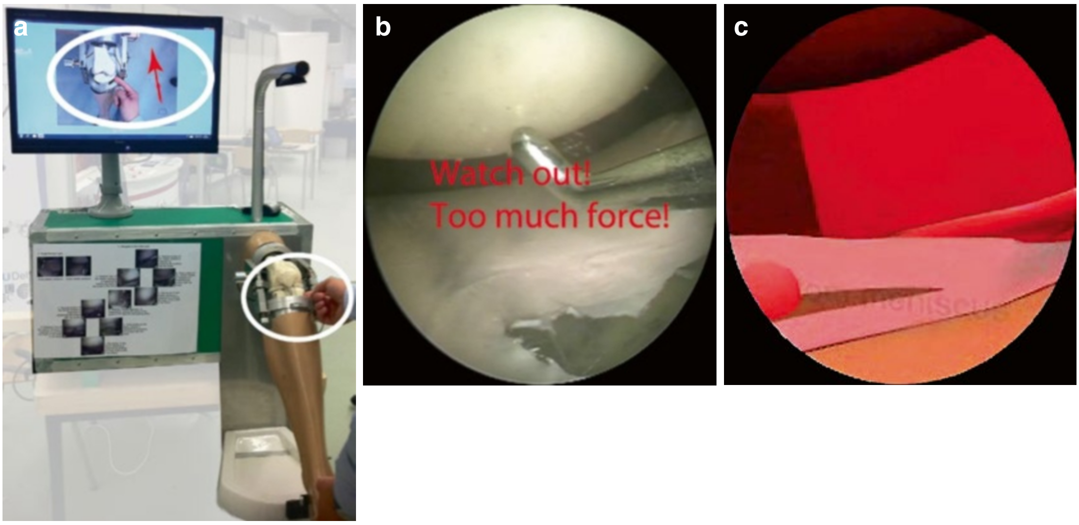

Fig. 11.10 (**a**) *Arrow* representation of the force magnitude and direction. The femoral condyle is pushed upward by the hand encircled in the right part of the picture; a resulting arrow is depicted on the interface screen to give visual feedback. (**b**) Textual guidance in PASSPORT user interface to guide a trainee to the next anatomic landmark. (**c**) Colouring red of the image in the SIMENDO arthroscopy, when tissue contact is too high. In all cases, the objects are displayed as an overlay on top of the arthroscopic image (© GJM Tuitjhof, 2014. Reprinted with permission)

sequence of manual actions and B) how to perform safe tissue manipulation (Table 11.1). For example, Horeman and co-workers have shown that novices can learn to endoscopically connect artificial tissue with less manipulation force by providing direct feedback of the force magnitude and direction (Horeman et al. 2012b).

When offering direct feedback, care has to be taken on how to inform the trainee and how to prevent mental overloading. The former depends on proper composition of the task and on defining the critical steps upon which should be reflected. Little to no literature on this aspect is available, other than general human factors and educational theories. For the latter, it should be noticed that people have three learning styles to receive and process information: oral, proprioceptive and visual (i.e. objects and text). Thus, all these type can be applied to give direct feedback. For oral feedback, an example was indicated for the collision metric where a buzz sound is given upon unallowed instrument tissue contact (Meyer et al. 1993; Tuijthof et al. 2003). Oral feedback in the form of alarms or buzzes is a common way to warn people that a dangerous situation is happening. This immediately poses a drawback to use in early learning stages of novices, as they are expected to make mistakes or errors at a frequent pace, which will set off the alarm signal many times during a task. If that happens, one tends to ignore such a signal and it no longer serves as adequate feedback signal. In multiple studies, researchers experimented to improve the sensation of the operating surgeons or to warn them if mistakes were made with haptic systems based on sensors, vibrating elements, motors and hydraulics (Westebring-van der Putten et al. 2008). For example, in the Daum-Hand (EndoHand) (Jackman et al. 1999; Melzer et al. 1997), the contact forces on the grasper were detected by membranes, amplified and transmitted hydraulically to membranes connected to the surgeon's fingertips to feel the average grasping force. A drawback of implementing such haptic way of giving feedback is that adding mass to the instrument handles or making modification to the instruments alters the instrument handling sensation which in itself is not ideal for training.

Visual or written feedback has also been implemented in arthroscopic simulators (Fig. 11.10). This is given in the form of object or text with or without colouring. In the study by Horeman and co-workers, the force direction and magnitude were simultaneously indicated by an arrow indicating the direction, its length, the size and its colour correct loading (Horeman et al. 2012b).

The real-time visual feedback in this study helps novices to choose the correct strategy on a skills- and knowledge-based level with information provided on a rule-based level (Dankelman et al. 2004; Wentink et al. 2003). This and other studies suggest that the human mind is capable of using additional visual information such as object colour and shape to improve tissue handling to some degree (Horeman et al. 2012b; Triano et al. 2006). Therefore, it seems that extra-visual information in the field of view of the trainee is easy to observe. On the other hand, if the visual feedback is given aside from the task area as presented on the screen, it can distract the trainee from the task or block other areas of interest. This aspect needs to be considered when simulating design, as during live surgery continuous focus is required on the surgical action. Thus, the manner in which feedback is given to the trainee is crucial to facilitate the training process.

So far, metrics have been presented that are more or less suitable for all or at least part of training tasks and can cover a wide range of different surgical actions. Concrete examples on how to use these for training arthroscopic skills are discussed.

The task time, tip-to-tip distance, motion smoothness and peak force can be applied in exercises where joint inspection is trained or tissue probing, and loose body removal. When punching or shaving meniscus tissue or executing meniscal suturing, the force area and volume of force can be applied, additionally. However, training of specific skills and tasks require additional performance assessment. For example, when performing a meniscectomy, it is highly relevant to measure the smoothness of the rim and the relative amount of removed tissue volume. This is possible but requires specific sensors (e.g. camera screenshots) and data processing. The last example is the training of drilling the holes in femur and tibia to prepare them for a cruciate ligament reconstruction. For this exercise, the direction of drilling is crucial. This could be measured by the direction of force, but in this case, the direction of the drill bit itself could indicate performance efficiency. This might be reflected by the sum of all angular rotation round the instrument's shaft length (called angular path). Conclusively, the learning goals per task or exercise need to be clearly defined and the task needs to be clearly described, before choosing the proper metrics that reflect the performance.

11.6.2 Learning from Feedback

It is evident that not all metrics used for performance monitoring can be used for educational purposes. Trainees that receive feedback about their 'volume of force' or 'economy of motion' during a training session are not likely to transform this kind of information into better performance. Therefore, it is recommended to translate those metrics into constructive (task dependent) oral or written feedback to be understandable for the trainee. Table 11.2 shows how post-task feedback can be provided in between

Table 11.2 Metrics-based comprehensible feedback

Metric	Informative instruction to trainee
Task time	Task time is high; more practice is needed
Part length	Try to minimize unnecessary instrument movements
Speed	Slow your pace; decrease your instrument motion
Motion volume	Try to minimize unnecessary instrument movements
Tip-to-tip distance	Keep the instrument in sight. This avoids unintended damage to surrounding tissue
Out-of-view time	Keep the instrument in sight. This avoids unintended damage to surrounding tissue and speeds up your efficiency
Max force	Forces are too high. Minimize pushing or pulling during manipulation
	Avoid high insertion forces; your instrument is probably directed in a wrong manner
Mean force	Too much contact between instruments and tissue. Watch out for unintentional contact with tissue
Force Volume	Too much jerks or collisions during manipulation. Lower your instrument speed during manipulation
Max force area	Lower your force during insertion and tissue manipulation
	Keep your instrument tip in sight to avoid unintended tissue damage

trials. This schedule can truly lead to autonomous learning without the need to have a supervising surgeon standing next to the trainee.

Bibliography

Alvand A, Khan T, Al-Ali S, Jackson WF, Price AJ, Rees JL (2012) Simple visual parameters for objective assessment of arthroscopic skill. J Bone Joint Surg Am 94(13):e97, available from: PM:22760398

Andersen C, Winding TN, Vesterby MS (2011) Development of simulated arthroscopic skills. Acta Orthop 82(1):90–95, available from: PM:21281257

Bayona S, Fernandez-Arroyo JM, Martin I, Bayona P (2008) Assessment study of insight ARTHRO VR arthroscopy virtual training simulator: face, content, and construct validities. J Robot Surg 2(3):151–158

Bliss JP, Hanner-Bailey HS, Scerbo MW (2005) Determining the efficacy of an immersive trainer for arthroscopy skills. Stud Health Technol Inform 111:54–56, available from: PM:15718698

Chami G, Ward J, Wills D, Phillips R, Sherman K (2006) Smart tool for force measurements during knee arthroscopy: in vivo human study. Stud Health Technol Inform 119:85–89, available from: PM:16404020

Chami G, Ward JW, Phillips R, Sherman KP (2008) Haptic feedback can provide an objective assessment of arthroscopic skills. Clin Orthop Relat Res 466(4):963–968, available from: PM:18213507

Chmarra MK, Bakker NH, Grimbergen CA, Dankelman J (2006) TrEndo, a device for tracking minimally invasive surgical instruments in training setups. Sensors Actuators A Phys 126(2):328–334, available from: ISI:000235431800007

Chmarra MK, Grimbergen CA, Dankelman J (2007) Systems for tracking minimally invasive surgical instruments. Minim Invasive Ther Allied Technol 16(6):328–340, available from: PM:17943607

Chmarra MK, Klein S, de Winter JC, Jansen FW, Dankelman J (2010) Objective classification of residents based on their psychomotor laparoscopic skills. Surg Endosc 24(5):1031–1039, available from: PM:19915915

Dankelman J, Wentink M, Grimbergen CA, Stassen HG, Reekers J (2004) Does virtual reality training make sense in interventional radiology? Training skill-, rule- and knowledge-based behavior. Cardiovasc Intervent Radiol 27(5):417–421, available from: PM:15383842

Gomoll AH, O'Toole RV, Czarnecki J, Warner JJ (2007) Surgical experience correlates with performance on a virtual reality simulator for shoulder arthroscopy. Am J Sports Med 35(6):883–888, available from: PM:17261572

Gomoll AH, Pappas G, Forsythe B, Warner JJ (2008) Individual skill progression on a virtual reality simulator for shoulder arthroscopy: a 3-year follow-up study. Am J Sports Med 36(6):1139–1142, available from: PM:18326032

Hodgins JL, Veillette C (2013) Arthroscopic proficiency: methods in evaluating competency. BMC Med Educ 13:61, available from: PM:23631421

Hogan N, Sternad D (2009) Sensitivity of smoothness measures to movement duration, amplitude, and arrests. J Mot Behav 41(6):529–534, available from: PM:19892658

Horeman T, Rodrigues SP, Jansen FW, Dankelman J, van den Dobbelsteen JJ (2010) Force measurement platform for training and assessment of laparoscopic skills. Surg Endosc 24:3102–3108, available from: PM:20464416

Horeman T, Rodrigues SP, Jansen FW, Dankelman J, van den Dobbelsteen JJ (2012a) Force parameters for skills assessment in laparoscopy. IEEE Trans Haptics 5(4):312–322, available from: ISI:000310384300003

Horeman T, Rodrigues SP, van den Dobbelsteen JJ, Jansen FW, Dankelman J (2012b) Visual force feedback in laparoscopic training. Surg Endosc 26(1):242–248, available from: PM:21858573

Horeman T, Kurteva DDK, Valdastri P, Jansen FW, van den Dobbelsteen JJ, Dankelman J (2013) The influence of instrument configuration on tissue handling force in laparoscopy. Surg Innov 20(3):260–267, available from: PM:22956398

Horeman T, Blikkendaal MD, Feng D, Van DA, Jansen F, Dankelman J, van den Dobbelsteen JJ (2014a) Visual force feedback improves knot-tying security. J Surg Educ 71(1):133–141, available from: PM:24411436

Horeman T, Dankelman J, Jansen FW, van den Dobbelsteen JJ (2014b) Assessment of laparoscopic skills based on force and motion parameters. IEEE Trans Biomed Eng 61(3):805–813, available from: PM:24216633

Howells NR, Brinsden MD, Gill RS, Carr AJ, Rees JL (2008) Motion analysis: a validated method for showing skill levels in arthroscopy. Arthroscopy 24(3):335–342, available from: PM:18308187

Jackman SV, Jarzemski PA, Listopadzki SM, Lee BR, Stoianovici D, Demaree R, Jarrett TW, Kavoussi LR (1999) The EndoHand: comparison with standard laparoscopic instrumentation. J Laparoendosc Adv Surg Tech A 9(3):253–258, available from: PM:10414542

Maithel S, Sierra R, Korndorffer J, Neumann P, Dawson S, Callery M, Jones D, Scott D (2006) Construct and face validity of MIST-VR, Endotower, and CELTS: are we ready for skills assessment using simulators? Surg Endosc 20(1):104–112, available from: PM:16333535

Martin KD, Belmont PJ, Schoenfeld AJ, Todd M, Cameron KL, Owens BD (2011) Arthroscopic basic task performance in shoulder simulator model correlates with similar task performance in cadavers. J Bone Joint Surg Am 93(21):e1271–e1275, available from: PM:22048106

Martin KD, Cameron K, Belmont PJ, Schoenfeld A, Owens BD (2012) Shoulder arthroscopy simulator performance correlates with resident and shoulder arthroscopy experience. J Bone Joint Surg Am 94(21):e160, available from: PM:23138247

McCarthy AD, Moody L, Waterworth AR, Bickerstaff DR (2006) Passive haptics in a knee arthroscopy

simulator: is it valid for core skills training? Clin Orthop Relat Res 442:13–20, available from: PM:16394733
Melzer A, Kipfmuller K, Halfar B (1997) Deflectable endoscopic instrument system DENIS. Surg Endosc 11(10):1045–1051, available from: PM:9381348
Meyer RD, Tamarapalli JR, Lemons JE (1993) Arthroscopy training using a "black box" technique. Arthroscopy 9(3):338–340, available from: PM:8323624
Obdeijn MC, Baalen SV, Horeman T, Liverneaux P, Tuijthof GJM (2014) The use of navigation forces for assessment of wrist arthroscopy skills level. J Wrist Surg 3:132–138
Oropesa I, Chmarra MK, Sanchez-Gonzalez P, Lamata P, Rodrigues SP, Enciso S, Sanchez-Margallo FM, Jansen FW, Dankelman J, Gomez EJ (2013) Relevance of motion-related assessment metrics in laparoscopic surgery. Surg Innov 20(3):299–312, available from: PM:22983805
Pedowitz RA, Esch J, Snyder S (2002) Evaluation of a virtual reality simulator for arthroscopy skills development. Arthroscopy 18(6):E29, available from: PM:12098111
Pellen MG, Horgan LF, Barton JR, Attwood SE (2009) Construct validity of the ProMIS laparoscopic simulator. Surg Endosc 23(1):130–139, available from: PM:18648875
Rosen J, MacFarlane M, Richards C, Hannaford B, Sinanan M (1999) Surgeon-tool force/torque signatures – evaluation of surgical skills in minimally invasive surgery. Stud Health Technol Inform 62:290–296, available from: PM:10538374
Rosen J, Brown JD, Chang L, Sinanan MN, Hannaford B (2006) Generalized approach for modeling minimally invasive surgery as a stochastic process using a discrete Markov model. IEEE Trans Biomed Eng 53(3):399–413, available from: PM:16532766
Rosser JC Jr, Colsant BJ, Lynch PJ, Herman B, Klonsky J, Young SM (2006) The use of a "hybrid" trainer in an established laparoscopic skills program. JSLS 10(1):4–10, available from: PM:16709348
Scott DJ, Bergen PC, Rege RV, Laycock R, Tesfay ST, Valentine RJ, Euhus DM, Jeyarajah DR, Thompson WM, Jones DB (2000) Laparoscopic training on bench models: better and more cost effective than operating room experience? J Am Coll Surg 191(3):272–283, available from: PM:10989902
Sherman KP, Ward JW, Wills DP, Sherman VJ, Mohsen AM (2001) Surgical trainee assessment using a VE knee arthroscopy training system (VE-KATS): experimental results. Stud Health Technol Inform 81:465–470, available from: PM:11317792
Smith S, Wan A, Taffinder N, Read S, Emery R, Darzi A (1999) Early experience and validation work with Procedicus VA – the Prosolvia virtual reality shoulder arthroscopy trainer. Stud Health Technol Inform 62:337–343, available from: PM:10538383
Stylopoulos N, Cotin S, Maithel SK, Ottensmeye M, Jackson PG, Bardsley RS, Neumann PF, Rattner DW, Dawson SL (2004) Computer-enhanced laparoscopic training system (CELTS): bridging the gap. Surg Endosc 18(5):782–789, available from: PM:15216861
Tarnay CM, Glass KB, Munro MG (1999) Entry force and intra-abdominal pressure associated with six laparoscopic trocar-cannula systems: a randomized comparison. Obstet Gynecol 94(1):83–88, available from: PM:10389723
Tashiro Y, Miura H, Nakanishi Y, Okazaki K, Iwamoto Y (2009) Evaluation of skills in arthroscopic training based on trajectory and force data. Clin Orthop Relat Res 467(2):546–552, available from: PM:18791774
Triano JJ, Scaringe J, Bougie J, Rogers C (2006) Effects of visual feedback on manipulation performance and patient ratings. J Manipulative Physiol Ther 29(5):378–385, available from: PM:16762666
Tuijthof GJM, van Engelen SJMP, Herder JL, Goossens RHM, Snijders CJ, van Dijk CN (2003) Ergonomic handle for an arthroscopic cutter. Minim Invasive Ther Allied Technol 12(1–2):82–90, available from: ISI:000184323300012
Tuijthof GJ, van Sterkenburg MN, Sierevelt IN, Van OJ, van Dijk CN, Kerkhoffs GM (2010) First validation of the PASSPORT training environment for arthroscopic skills. Knee Surg Sports Traumatol Arthrosc 18(2):218–224, available from: PM:19629441
Tuijthof GJ, Horeman T, Schafroth MU, Blankevoort L, Kerkhoffs GM (2011a) Probing forces of menisci: what levels are safe for arthroscopic surgery. Knee Surg Sports Traumatol Arthrosc 19(2):248–254, available from: PM:20814661
Tuijthof GJ, Visser P, Sierevelt IN, van Dijk CN, Kerkhoffs GM (2011b) Does perception of usefulness of arthroscopic simulators differ with levels of experience? Clin Orthop Relat Res, available from: PM:21290203
Verdaasdonk EG, Stassen LP, Schijven MP, Dankelman J (2007) Construct validity and assessment of the learning curve for the SIMENDO endoscopic simulator. Surg Endosc 21(8):1406–1412, available from: PM:17653815
Wentink M, Stassen LP, Alwayn I, Hosman RJ, Stassen HG (2003) Rasmussen's model of human behavior in laparoscopy training. Surg Endosc 17(8):1241–1246, available from: PM:12799883
Westebring-van der Putten EP, Goossens RH, Jakimowicz JJ, Dankelman J (2008) Haptics in minimally invasive surgery – a review. Minim Invasive Ther Allied Technol 17(1):3–16, available from: PM:18270873