Journal of Eye Movement Research
6(4):4, 1-11

Eye movements in surgery : A literature review

Frouke Hermens
School of Psychology, University of Aberdeen, UK

Rhona Flin
School of Psychology, University of Aberdeen, UK

Irfan Ahmed
Aberdeen Royal Infirmary & University of Aberdeen, UK

With recent advances in eye tracking technology, it is now possible to track surgeons' eye movements while engaged in a surgical task or when surgical residents practice their surgical skills. Several studies have compared eye movements of surgical experts and novices, developed techniques to assess surgical skill on the basis of eye movements, and examined the role of eye movements in surgical training. We here provide an overview of these studies with a focus on the methodological aspects. We conclude that the different studies of eye movements in surgery suggest that the recording of eye movements may be beneficial both for skill assessment and training purposes, although more research will be needed in this field.

Keywords: Eye movements, laparoscopic surgery, skill assessment, training

Introduction

Eye tracking technology has gone through extensive advancements in recent years. Initially, eye tracking devices, often applying a contact lens worn by the participant (Yarbus, 1967), were highly intrusive, allowing only for short recording times with a limited number of participants (often the authors themselves). In contrast, new eye tracking systems (e.g., the Eyelink systems from SR Research, the Tobii eye trackers, ASL mobile eye trackers or SMI systems[1] applying video analysis technology allow for the non-intrusive measurement of an observer's point of gaze for a range of tasks (such as visual search, reading, and economic decision making) and a range of settings (such as in driving, motor tasks and social interactions). These developments have not gone unnoticed by researchers interested in skill acquisition in surgery. We will here provide an up-to-date overview of the literature on eye tracking in surgeons to determine whether eye tracking aids the assessment of surgical skill, in addition the existing assessment techniques on the basis of tool and hand motion parameters and surgery outcomes (Richstone et al., 2010). Furthermore, we will examine whether eye tracking can be used for training of surgical skill (Wilson et al., 2011).

In our review of the literature, we found that work on eye movements in surgery almost uniquely involves laparoscopic surgery, also known as minimally invasive surgery or key-hole surgery. Laparoscopic surgery has a number of advantages compared to traditional surgery in terms of patient outcome and demand on hospital resources. Because smaller incisions are required, participants tend to experience less pain, have fewer complications and have a shorter recovery time (Monson, 1993) . These advantages come at a cost in terms of the level of technical skill required from the surgeon, with longer training times and more use of resources during training. Reasons for the higher technical complexity of laparoscopic surgery are the 2D viewing of the laparoscopic images, the separation between the field of view and instrument manipulation, and the restricted degrees of freedom of instrument movements, although some of these issues are alleviated in the more costly robotic surgery systems (Reiley, Lin, Yuh, & Hager, 2011). Because of these technical demands, techniques to facilitate the training process are much sought-after, and researchers have therefore started to examine whether eye movement measurements may be of use in this process.

When analyzing the literature, we also found that studies of eye tracking in surgery broadly cluster in four categories, based on the main question they aim to answer. The four questions are: (1) Do eye movement patterns differ between experienced surgeons and novices? (2) Can these differences in eye movements be used to assess the skill of the surgeon? (3)

Dr. Frouke Hermens, School of Psychology, William Guild Building, University of Aberdeen, Aberdeen AB24 3FX, UK. Tel. +44 (0)1224 272235.

[1] SR research: http://www.sr-research.com/, Tobii systems: http://www.tobii.com/, ASL eye tracking: http://www.asleyetracking.com/Site/, SMI systems: http://www.smivision.com/en.html.

DOI 10.16910/jemr.6.4.4

ISSN 1995-8692

How do the differences depend on the task the surgeon is performing (i.e., observing someone performing a surgery, practicing surgical skills, performing an actual surgery)? and (4) Can eye movement technology be used for educational purposes, to facilitate training of new surgeons? In the following, we address these questions one by one and discuss the studies into these questions.

Comparison of experts and novices

Studies comparing novices and experts (Table 1) have predominantly used a laparoscopic simulator. The use of a simulator has the advantage that novices and experts can perform the same task without any risks to patients. Tasks can involve virtual environments, but also handling of basic objects, or simulated surgery. While several studies focused on the difference between experts and novices, there appears to be a large variation in the analysis of the eye movement patterns. Despite these variations, eye movements seem to consistently differentiate between experts and novices.

In the first study of this type that we identified, Law and colleagues (2004) used a virtual environment and asked participants (four practicing and one retired surgeon, the experts, and five students, the novices) to touch a virtual target cube using a laparoscopic instrument (for more details, see Table 1). As expected, experts completed the task more quickly than the novices, however, they did not make significantly fewer errors. Analysis of the eye movement data revealed three dominant gaze patterns: (1) target gaze behavior (where participants look at the target), (2) switching behavior (alternating gaze between target and instrument or in between these objects), and (3) tool following behavior (where eye gaze follows the instrument on its way to the target). Experts showed more target gaze behavior, whereas novices more often demonstrated switching and tool following behavior, thereby suggesting more pro-active gaze behavior in experts.

Similar results were obtained by Wilson and colleagues (2010) who asked eight expert (more than 70 procedures performed) and six novice (fewer than 10 procedures) surgeons to touch a series of virtual targets using one of two instruments while their eye movements were recorded using a mobile eye tracker. As before, experts completed the task more quickly, but no difference in accuracy was found between the two groups. Furthermore, only for the left hand instrument experts demonstrated more economical movements. With respect to their eye movements, experts were found to fixate the targets more often than the instruments (a larger proportion of the trial), while novices fixated both types of objects in equal proportions.

These results were extended by Wilson and colleagues (2011), who specifically looked at 'quiet eye' times (the duration of the last fixation before an action, discussed in more detail later in this review). They compared eye movements of ten experts (>60 procedures) and fifteen novices (<10 procedures) during a virtual laparoscopy, involving three subtasks. Besides better performance, experts demonstrated a more efficient gaze strategy, with fewer eye movements between the tool and the target, while almost exclusively fixating on the target. Experts also showed longer quiet eye durations and made fewer grasp attempts.

Further studies examined broader patterns of eye movements, not only focusing on eye movements during tool use, but also on eye movements towards other sections of the operating room, thereby examining situation awareness (Schulz, Endsley, Kochs, Gelb, & Wagner, 2013) in experts and novices. One of the settings examined, involved a simulated gallbladder operation on one of two dummy patients (Tien et al., 2010, 2011; Zheng et al., 2011). One of the dummy patients was stable, whereas the other had a simulated arrhythmia. Vitals were simulated on a second monitor providing visual and auditory feedback on heart and breathing rate. Experts more often looked at the vitals screen, suggesting greater situation awareness. Interestingly, Zheng et al. (2011) also found more errors and longer completion times in experts. Experts also reported higher reported higher levels of frustration when physical demands were low, which may suggest that a simulator setting may not be optimal to examine expert performance. Using a similar setup Atkins and colleagues (2013) found that experts (13 surgical residence or practicing surgeons) fixated the anesthesia monitor more often and longer than novices (10 research fellows). They also found that longer and more frequent blinking was associated with a lower experienced workload.

The above studies compared experts and novices, but did not examine surgeons of intermediate levels, which would provide evidence about how eye movement patterns develop with experience. One exception is the work by Kocak and colleagues (2005), who compared eye movements of novice (non-medical hospital staff and medical students with no operative experience), intermediate (<100 cases) and expert (fellowship training) participants (8 participants in each group) in three laparoscopic simulator tasks. Global eye movement parameters were compared, but no information was collected about where participants looked during the task. Saccadic rates were significantly lower and peak velocity was significantly higher in experts, with intermediate values for the intermediately experienced participants.

As summarized in Table 1, the studies comparing experts and novices strongly suggest that eye movement patterns differ between the two groups, across a range of possible eye movement measures. Experts are found to spend more time fixating the target for an instrument movement than the instrument itself, they spend more time fixating a vitals monitor during a simulated opera-

DOI 10.16910/jemr.6.4.4

ISSN 1995-8692

Table 1
Overview of the methods of each of the studies comparing eye movements in experts and novices surgeons, listing the number of experts and novices in each study, the environment in which the eye movements were recorded, the task being performed by the research participants, and the eye tracking equipment used. The studies are ordered chronologically.

Study	Number of Experts	Number of Novices	Environment	Task	Eye tracker
Law et al. (2004)	5	5	Immersion Corp. Laparoscopic Impulse Engine	Touch target cube inside larger cube	ASL 504 remote
Kocak et al. (2005)	8	8 nov. + 8 int.	Stryker Surgical trainer box	'Loops', 'Rope' and 'Beans' task	Cyclops Eye Track saccadometer
Tien et al. (2010)	4	4	SurgicalSim VR Laparoscopic training simulator	Perform cholecystecotomy while monitoring patient vitals	Locorna eye tracker
Sodergren et al. (2010)	7	14	Visual presentation of sequences of images	Compare images of the same structure in different orientations	Tobii ET 1750
Wilson et al. (2010)	8	6	Lap Mentor VR simulator	Touch virtual targets	ASL head mounted system
Tien, Zheng, & Atkins (2011)	8	8	SurgicalSim VR Laparoscopic training simulator	Perform cholecystecotomy while monitoring patient vitals	Locorna eye tracker
Wilson et al. (2011)	10	15	Lap Mentor VR simulator	Grasp jelly, grasp target ball, place target ball into endobag	ASL head mounted system
Zheng et al. (Zheng et al., 2011)	13	10	SurgicalSim VR Laparoscopic training simulator	Laparoscopic procedure	Locorna eye tracker
Sodergren et al. (2011)	7	14	Visual presentation of sequences of images	Determine configuration of endoscope, orientation of image, orifice used for access	Tobii ET 1750
Eivazi et al. (2012)	4	4	Watching still images of surgery	Answer questions about images	Tobii T120
Atkins et al. (2012)	13	10	SurgicalSim VR Laparoscopic training simulator	Perform laparoscopic surgery while monitoring patient vitals	Locorna eye tracker

tion, and they display differences in general eye movements statistics, such as lower saccadic rates and higher peak velocity, independent of where these eye movements are aimed at. These results create the possibility of using eye movements as method to assess surgical skill, discussed in the next section. The differences between experts and novices may only hold when surgeons are engaged in a surgical task. Eye movements while watching recordings or still images from a surgical procedure have often been found to be similar between experts and novices (Sodergren et al., 2010, 2011) (however, see Eivazi et al., 2012). We will return to the role of the task and setting later.

Assessing a surgeon's skill

Assessing skills of surgeons and trainees is an important task, as its outcome can be used to decide which surgeon may be assigned to what complexity of procedures. Traditional methods in which an expert surgeon witnesses and assesses the skill of fellow surgeons and trainees, for example using a checklist, place a high demand on human resources and suffer from a lack of objectivity. Assessing the surgeon's skill on the basis of patient outcome is also problematic, because more experienced surgeons are more likely to take up more complex cases. More objective methods, based on the analysis of the movement patterns of the hands of the surgeons and the tools, have not lead to a completely satisfactory method of assessment either (Reiley et al., 2011). It is for these reasons that eye movements

DOI 10.16910/jemr.6.4.4

ISSN 1995-8692

have now started to be considered as a tool to measure the surgical skill (Table 2).

The studies in Table 2 differ from those comparing experts and novices (Table 1) in that they use machine learning algorithms to distinguish between experts and novices. The basic idea is that a set of eye movement data is fed into a computer program after which the parameters of the program are adjusted to yield an optimal classification of the surgeons in terms of skill or experience (usually a binary distinction between experts and novices). While the different studies use different machine learning techniques and eye movement parameters, they all report good classification performance, suggesting that eye movement parameters could be a valuable tool for skill assessment.

We would first like to mention a few studies that used machine learning algorithms, but did not directly examine skill. The first study in this context applied Markov models to eye movement data of seven medical students with no experience in laparoscopic surgery during the transection of a simulated blood vessel (Nicolaou et al., 2004). Eye movements ('transitions' in Markov models) between the tooltip, the transection point and the area between the tooltip were analyzed and plotted as a function of time. Eye movements directed to the transection point showed distinct patterns across participants in this analysis, suggesting that eye movement patterns can distinguish between different people, possibly reflecting differences in how the task is approached. A second study examined the classification of surgerical stages rather than participants. Three senior specialist registrars were asked to perform a laparoscopic cholecystectomy on a porcine model while their eye movements were recorded (James et al., 2007). The eye movements were automatically grouped into stages on the basis of either the eye movements alone (frequency and duration of gaze deviating away from the monitor) or the recorded images (features indicating a change in instruments) in combination with the eye movements using a Parallel Layer Perception (PLP). A larger (75%) correct automatic classification of the critical stage was obtained on a combination of both the images, but fairly good classification performance is obtained on the basis of eye movements alone (66% correct). A similar ability to classify surgical stages was obtained by Ibbotson and colleagues (1999), who used a coarser method of measuring and analysing eye gaze. They video-recorded live surgery and analysed the individual video frames for gaze directed at the monitor, external operative space, or elsewhere in the room. These gaze classifications were then plotted and visually inspected for patterns, and transition frequencies were examined, revealing specific sub-tasks, such as knotting and suturing.

Four subsequent studies examined the use of automatic classification techniques to distinguish between experts and novices on the basis of their eye movement patterns. Sodergren et al. (2010) used Hidden Markov Modeling (HMM) to identify different gaze patterns when watching images taken from different angles, simulating the effect of moving a flexible endoscope inside a cavity. While the HMM was not able to distinguish between experts (2 endoscopists and 5 surgeons) and novices (8 engineers, medical students and postgraduate researchers) it did reveal that eye movement patterns were different in participants with high and low performance on the task to describe the movement of the camera between images. In particular, participants with higher performance fixated central regions of the image for longer than participants with lower performance. Accurate classification between experts and novices was achieved by Ahmidi and colleagues (2010), who asked 5 participants with knowledge of sinus anatomy (the experts) and 6 participants without prior endoscopic experience (the novices) to touch structures in a cadaver using a nasal pointer. Instrument movement measurements as well as eye-gaze measurements (horizontal and vertical coordinates) were divided into stages (corresponding to the different tasks performed) and then fed into a Hidden Markov Model and a k-means clustering analysis. A reported 82.5% accuracy was obtained for skill level classification, in addition to a 77.8% accuracy for task classification, with better classification performance when both eye and instrument movement was used. This result was confirmed by Ahmidi and colleagues (2012) in 7 experts and 13 novices using similar methods. Excellent classification performance was also found in the detailed analysis of Richstone and colleagues (2010) of eye movements of three expert and eight non-expert surgeons (the distinction based on the number of surgeries performed, the number of years of postgraduate training and ratings from judges) in 40 live transperitoneal laparoscopic renal surgeries and 46 simulator tasks (across all participants). Linear and non-linear classifiers showed good discrimination performance on the basis of a mixture of fixation rates, vergence, blink rates and a measure of pupil diameter variability ('Index of Cognitive Activity' or ICA) , possibly reflecting the more focused attention of the experts. Analysis of repeated measurement of a non-expert also demonstrated that as the surgeon improved in skill, the eye movement data became more difficult to distinguish from those of the experts.

These studies show that eye movement data can be used to distinguish between experts and novices, as well as surgery stages, across a range of tasks and eye movement parameters. The studies also suggest that little preprocessing of the eye movement data is required for accurate classification and that it may therefore not be necessary to determine where exactly the surgeons were looking to obtain an accurate method to measure skill. The only study that did not obtain accurate classification performance between experts and novices measured eye movements while the participants examined images (Sodergren et al., 2010) instead

DOI 10.16910/jemr.6.4.4

ISSN 1995-8692

Table 2
Studies applying classifiers for skill assessment, listing the number of experts and novices that went into the analysis, the environment under which the eye movements were collected, the task participants were performing and the type of classifier being used.

Study	Number of Experts	Number of Novices	Environment	Task	Classification method
Nicolaou et al. (2004)	0	7	Laparoscopic simulator	Transect simulator blood vessel	Markov modeling
James et al. (2007)	3	0	Porcine model	Cholecystectomy	Perceptron
Richstone et al. (2010)	3	8	Live surgery and simulator	Various tasks	Linear and non-linear classifier
Sodergren et al. (2010)	7	8	Presentation of sequence of still images	Compare scene with picture of scene in different orientation	HMM
Ahmidi et al. (2010)	5	6	Cadaver	Touch given anatomies	HMM, k-means
Ahmidi et al. (2012)	7	13	Unclear	Endoscopic sinus surgery tasks	Statistical classifier

of being engaged in an active surgery task, suggesting an important role of task and context, which we will discuss in the next section.

Examining the role of context

Most of the studies discussed so far involved eye movements recordings while being engaged in a simulator task. This raises the question whether eye movements during simulated surgery provide the same information as those observed in live surgery. Furthermore, we also saw that eye movement patterns may differ between being engaged in a procedure (live surgery or simulator) and watching a procedure. As tracking surgeons' eye movements during a procedure is technically more challenging than measuring eye movements while watching a recorded surgery, demonstrating that similar results (e.g., similar classification of skill) can be obtained from eye movements while watching a procedure would make eye tracking methods available to a broader range of users. The studies listed in Table 3 examined the role of task and setting.

Richstone and colleagues (2010), introduced above, only found small differences in classification performance between live and simulated surgery, suggesting that simulator tasks may be a good proxy for live surgery when classifying between experts and novices on the basis of eye movements. Khan and colleagues (2012) compared eye movements during live surgery (16 laparoscopic cholecystectomies of 2 expert surgeons) and watching surgery (of video recordings of the surgeries; watched by the same two experts and several junior residents). Eye movements while self-watching and other-watching were compared using a 3 degrees of visual angle maximum distance overlap criterion, resulting in a higher overlap for self-watching (55% overlap) than for other-watching (43.8%). While this result suggests there is relatively little overlap between performing and watching a surgeon (possibly related to instances in which the surgeon changes the instruments), it is unclear what the consequences are for classification of expert performance. A limited overlap between eye gaze while watching and performing (simuluated) surgery was also observed by Tien, Atkins and Zheng (2012) who asked fourteen computing science graduate students to perform a grasp-and-transport task using a laparoscopic simulator. The video recordings of this task were later watched by seven other observers and the person who performed the task. Using 2.5 and 5 degrees (of visual angle) overlap measures, a higher overlap was obtained between being engaged in the same watching task (either the person performing the task or different viewers) than between tasks (within the same participant). Similar results were also obtained by Atkins, Jiang, Tien and Zheng (2012), who found a systematic delay of around 600ms between doing and watching and an overlap in gaze around 68% to 82% between the two tasks.

While the above studies suggest that eye movements differ between performing and watching procedure, they do not exclude a role of eye movements during observation for the assessment of skill. We touched this issue earlier, indicating that the results regarding this aspect were inconclusive. Eivazi and colleagues (2012) found reliable differences between eye movements (time to first fixation and mean fixation duration

DOI 10.16910/jemr.6.4.4

ISSN 1995-8692

Table 3
Studies examining the role of context in which eye movements are observed, listing the numbers of experts and novices that participated, the settings and/or tasks that were compared, and the analysis method used for comparison.

Study	Number of Experts	Number of Novices	Settings	Comparison
Richstone et al. (2010)	3	8	Live surgery and simulator	Statistical classifiers
Atkins et al. (2012)	0	17	Performing and watching simulator task	Start of saccades, percentage overlap
Tien et al. (2012)	0	14	Simulator task and watching simulated surgery	Percentage gaze overlap
Khan et al. (2012)	2	Unclear	Live surgery and watching surgery	Percentage gaze overlap
Atkins et al. (2013)	13, 0, 1	10, 14, 0	Simulator, Laboratory and Live surgery	Eye tracking methods involved

on certain regions, distribution of saccade amplitudes) of experienced and novice surgeons while watching still shots (presented for 10 seconds each). Such differences were, however, not found by Sodergren et al. (2010, 2011) when comparing dwell times on the different regions of endoscopic images. Because both Eivazi et al.'s and Sodergren et al.'s studies involved the analysis of fixation durations, the eye movement measure per se cannot explain the difference in outcomes. It cannot, however, be excluded that the choice of regions of interests may have played a role.

As mentioned earlier, technical aspects of measuring eye movements in the different settings differ, with lowest technical demands for watching, intermediate technical demands for simulator studies and highest technical demand for live surgery. These technical aspects were further described by Atkins and colleagues (2013) comparing the use of a remote eye tracker in the operating room, the same remote eye tracker in a simulator task (see above), and a mobile eye tracker in a situation awareness task. Difficulty of eye tracking in theater was related to surgeons moving around, stepping out of view of the camera of the remote eye tracker, and patient safety risks involved in using a head mounted (mobile) eye tracker.

Strong effects of task requirements and the observer's goal on eye movements have been observed in various other domains outside a surgical setting. For example, when viewing a painting with the aim of answering a range of questions, Yarbus (1967) reported different patterns of eye movements for the different questions (see also DeAngelus & Pelz, 2009; Greene, Liu, & Wolfe, 2012; Tatler, Wade, Kwan, Findlay, & Velichkovsky, 2010). Task and context influences were also observed during reading various types of text (Hermens, Loos, & Wagemans, 2012; Kaakinen & Hyönä, 2010; Kaakinen & Hyona, 2005; Rayner, Miller, & Rotello, 2008; Rothkopf & Billington, 1979). For example, Kaakinen et al. (2010) found leftward shifts of the initial landing position in words, shorter saccade lengths, longer first fixation and gaze durations, and higher refixation probabilities during a proofreading task than when participants read for comprehension. The influences of task and context effects found for surgical tasks are therefore well in line with findings in other domains.

The role of eye movements in training surgical skill

The observed differences in eye movements between experts and novices may also open the possibility of using eye movements in experts for training purposes. The effectiveness of eye gaze training was demonstrated in sports (Wilson et al., 2011). Eye gaze training may be particularly effective, because eye movements are often made subconsciously, and, as a consequence, trained eye movement strategies may be less susceptible to influences of stress. An important concept in gaze training is that of the 'quiet eye' (Vine, Moore, & Wilson, 2012), defined as the final fixation prior to a critical movement involved in a skill. As this final fixation tends to be longer for experts than for novices, it has been proposed as an objective measure of visuomotor skill. Furthermore, quiet eye can be trained using video modeling and verbal feedback (Vine et al., 2012).

The concept of quiet eye was adopted by Wilson and colleagues (2011) to examine whether laparoscopic skill can benefit from gaze training (see Table 4). In their study, they asked thirty medical trainees without laparoscopic experience to practice an eye-hand coordination task in a surgical simulator, with one third of the trainees receiving gaze-training, one third receiving movement training, and one third receiving standard feedback. The group receiving gaze training made the greatest progress in learning the skill (faster completion times), particularly under dual-task conditions, simulating the effect of a demanding surgery.

DOI 10.16910/jemr.6.4.4

ISSN 1995-8692

Table 4
Studies examining the role of gaze training in surgery, listing the number of participants in the study, the task being examined, and the training methods being compared.

Study	Particpants	Task	Training
Wilson et at. (2011)	30 novice medics	Touch virtual targets in LAP Mentor	Gaze training, movement training, or discovery training
Sodergren et al. (2011)	30 final-year medical students	Determine the orientation of images	Gaze training versus no training
Chetwood et al. (2012)	7 surgeons and 21 non-clinicians	Touch targets in simulator	Gaze cursor, verbal instruction, combination of gaze and verbal
Vine et al. (2012)	27 novice participants	Ball pick-up-and-drop task in MITS	Gaze training or discovery training
Vine et al. (2013)	36 novice participants	Ball pick-up-and-drop task and grasping and cutting task	Gaze training or discovery training

Although this study involved a relatively easy task, it suggests that gaze training could lead to faster improvement of skill in trainees, which is immune to distraction, such as in a multitask environment. These findings were supported in Vine et al. (2012) for 27 naive participants with no experience in laparoscopic procedures randomly assigned to a discovery learning group and a gaze learning group using a laparoscopic simulator task. Normal and time-pressured and evaluation-pressured conditions (in the latter of which the scores were publicly compared) showed that while both groups improved with practice, the gaze training group improved more strongly, both in terms of accuracy and in terms of completion time. The gaze training group adopted a target-locking viewing strategy more consistently, mimicking the gaze strategy of experts. In a further study, adopting the same ball pick-up-and-drop-task for training, Vine et al. (2013) found that the advantage of the gaze training group (both in terms of performance and eye movement patterns) extended to a bimanual grasp-and-cut task and was maintained after a 1 month retention interval. The advantage of gaze training was confirmed by Sodergren and colleagues (2011), who randomly assigned 30 final year medical students to either the intervention group or a control group. The intervention group was shown a video clip explaining the characteristics of gaze patterns of participants who were previously successful in the task they were going to perform (determining the orientation of a series of images from a laparoscopic cholecystectomy procedure). Participants receiving gaze training performed significantly better on the task (fewer errors and shorter inspection times).

A different approach was taken by Chetwood and colleagues (2012), who used eye movements as a way to convey information between a surgical tutor and a trainee by recording eye movements of both the tutor and the trainee during a laparoscopic simulator task. Verbal instruction (e.g., "touch the red object in between the two green objects") was compared with an eye gaze instruction (a cursor indicating the point of regard of the tutor) and a combination these two techniques. Task completion time, number of errors and the time to first look at the target were all positively influenced by gaze instruction (with or without verbal instruction).

Discussion

In the above, we have provided an overview of studies examining the role of eye movements in surgical skill. While we conducted our literature search with a broad focus on any type of surgery, we uniquely found studies into laparoscopic rather than open surgery. This may have two reasons. First, eye movements can more easily be tracked when eye-gaze is directed towards a screen. Eye trackers often have difficulties tracking eye movements across a large field of view, for example when a surgeon looks down to change instruments (however, see, Brandt, Glasauer, & Schneider, 2006). Second, laparoscopic surgery is technically more demanding and involves complex eye-hand coordination and may therefore be of more interest as a field of study.

We chose not to include eye movement studies of other surgical team members, which tend to focus more on decision making and situational awareness and less on eye-hand coordination. These studies, however, report very similar results regarding differences between experts and novices, which is why we would like to discuss them briefly. In comparing eye movements expert and novice scrub nurses (10 nurses in each group) during a series of caesarean section surgeries, Koh et al. (2011) found that expert nurses distributed their attention more effectively, in particular during high stress stages of the surgery, by better focusing on the important areas of interest for each stage. Similarly, in a pi-

DOI 10.16910/jemr.6.4.4

ISSN 1995-8692

lot study examining eye movements of four perfusionists (operating the heart-bypass machine during cardiac surgery), perfusionists with less experience were found to visually inspected key information areas less often than a more experienced colleague (Tomizawa, Aoki, Suzuki, Matayoshi, & Yozu, 2012), who was also found to distribute his attention more widely than his less experienced colleagues. In anaesthetists, Schulz et al. (2011) found that during periods of high workload, more experienced anaesthetists were able to increase the amount of time dedicated to manual tasks, while still continuing to monitor the patients' vitals.

One aspect that we observed across the studies that we reviewed is the relatively low number of participants the different studies were based on, in particular studies involving expert surgeons. There are plausible reasons for this as it is difficult to recruit surgeons because of their commitments. The low numbers of participants, however, may be an issue, because outcomes may depend on individual differences. One could argue that apparently the effect sizes, distinguishing experts from novices, are large, as most studies (however, see Sodergren et al., 2010, 2011) report significant differences across levels of experience, but a possible publication bias should be taken into consideration (e.g., Olson et al., 2002). A few studies performed a power-analysis (e.g., Sodergren et al., 2011) to determine the required number of participants, but these studies mostly involved naive participants only.

We also noticed that definitions of experts and novices differ strongly across studies. Some studies defined experts as people with surgical experience (the experts) with participants without any experience (novices), whereas other studies compared surgeons with extensive (the experts) and less extensive experience (the novices) in performing surgical procedures. Furthermore, reports often lack details about other aspects of the research participants, such as age and gender. These details may be important, as for example, gender differences in eye movements have been observed (e.g., Shen & Itti, 2012). Future studies would therefore benefit from a clearer distinction between experts and novices, possibly by defining different levels of expertise, so that it can be defined which particular levels (e.g., '2' and '4') are compared, and by more details about other aspects of the research participants.

More details may also be required regarding the methods of analyzing the eye movements, and the consequences of the choice of eye movement parameters. While some studies rely on global eye movement statistics, such as fixation durations and saccade amplitudes, or the coordinates of estimated eye gaze (e.g., Ahmidi et al., 2010; Kocak et al., 2005; Richstone et al., 2010; Tien et al., 2012), other studies examined dwell times on and saccades between certain objects, such as the instruments, monitors, or tissue (e.g. Sodergren et al., 2011; Wilson et al., 2011; Zheng et al., 2011). Despite these differences in eye movement parameters, the outcomes of the studies are generally consistent, demonstrating that eye movements differentiate between experts and novices. The methods relying on more basic eye movement parameters have the advantage that little preprocessing of the data is required and that the outcomes are little influenced by possibly subjective judgments about the important regions of interest. This comes at a cost, however, because methods that rely on global eye movement statistics provide little insight in how eye movement patterns differ between experts and novices, which may hamper their application in training. The development and evaluation of eye movement methods for analyzing eye movements in surgery may benefit from adopting and testing methods introduced in other domains, such as chess playing (Charness, Reingold, Pomplun, & Stampe, 2001), simulated flight (Kasarskis, Stehwien, Hickox, Aretz, & Wickens, 2001; Valerie et al., 2005), the interpretation of dynamic images (Jarodzka, Scheiter, Gerjets, & Van Gog, 2010), and distinguishing between psychiatric disorders (Clementz & Sweeney, 1990; ODriscoll & Callahan, 2008; Pallanti, Quercioli, Zaccara, Ramacciotti, & Arnetoli, 1998; Sweeney, Strojwas, Mann, & Thase, 1998), in addition to what is known from sports (e.g., Vine, Moore, & Wilson, 2011; Williams, Singer, & Frehlich, 2002).

Such investigations may aid to obtain a clearer picture of what are exactly the differences in eye movements between surgical experts and novices. Some studies have reported that experts tend to spend more time looking at the target ('target-locking'), while novices more often look at the tools (Wilson et al., 2010, 2011). Experts were also found to distribute their attention (reflected in their eye gaze) more efficiently than novices, looking at the important aspects of a scene at the right time (e.g., Koh et al., 2011; Tien et al., 2010, 2012). On the other hand, studies have demonstrated that it is well possible to distinguish between experts and novices on the basis of their eye movements without knowing what the participants were actually looking at (e.g., Kocak et al., 2005; Richstone et al., 2010).

Studies examining differences in laparoscopic surgeons' skill level have mostly compared experts and novices without examining gaze patterns of participants of intermediate skill. Only one study examining three levels of experience (Kocak et al., 2005) found differences between experts and novices, with intermediate levels of gaze parameters for surgeons of intermediate skill, but was unable to detect reliable differences between the intermediate level and the other two levels. One further study examining three levels of experience (Sodergren et al., 2010, 2011) found no differences between the three groups, which may have been due to the task being used (watching images rather than being engaged in surgery). A longitudinal study of one participant suggested that eye movements of a trainee became more difficult to distinguish from those of experts as training progressed (Richstone et al., 2010).

DOI 10.16910/jemr.6.4.4

ISSN 1995-8692

The strategy of using two groups (experts and novices) is problematic if the aim of the study is to develop an assessment method to evaluate skill, as it does not provide any information about whether eye movements change in a continuous fashion when gaining more experience, or whether there is a switching point in which eye movements of a novice turn into expert eye movements. The few studies that examined more than two levels of experience suggest that eye movement parameters change in a gradual fashion, but this should be established in a larger study before eye movements can be used for skill assessment.

From the studies so far, it is also unclear what are the best eye movement measures and task to assess surgical skill. Suggested measures seem to range from fixation durations, time to first fixation, saccade amplitudes, to an index of pupil diameter. It has been suggested that a combination of various measures leads to optimal classification (Richstone et al., 2010), but it is unclear at this stage whether all the relevant eye movements measures have been considered. It is also unclear in what settings the eye movements need to be collected. Perhaps eye movements while watching still images of surgery suffice (Eivazi et al., 2012), but the usefulness of eye movements in such a setting could not always be confirmed (Sodergren et al., 2010, 2011). Most studies applying simulators suggest that eye movements in this context provide a sensitive measure of skill (experts versus novices), although this observation was questioned by Zheng et al. (2011), who found that experts performed worse in this setting, possibly due to frustration with the task.

Eye tracking during live surgery is still a challenge, but this may be less of an issue with new eye tracking systems becoming available to a wider public such as Tobii and SMI glasses[2], which allow the recording of eye movements by means of a pair of glasses and a small device kept in a pocket or attached to a belt. More of a challenge will be the analysis of eye movements under such settings, because due to movement of the observer, the image on which the eye movements are superimposed, varies across observers (Segall, Taekman, Mark, Hobbs, & Wright, 2007). Possible strategies are to rely on general eye movement characteristics, such as fixation durations and saccade amplitudes (e.g., Richstone et al., 2010), or the use of regions of interest, which tend to be more similar across observers. Regions of interest will have to be defined, however, and although some automatic processing may be possible, the definition of these regions will be labor intensive and to some extent subjective.

Several studies suggest that eye movements may be beneficial for training purposes (Table 4). While these studies show promise, the effects of gaze training should be demonstrated in a live surgery setting.

Conclusion

Studies of eye movements in (laparoscopic) surgery have suggested that the recording of eye movements may be beneficial both for skill assessment and training purposes. Future work, however, is needed to refine these methods, measuring the relation between eye movements and skill across a broader range of participants and to determine in more detail what aspects of the eye movements are essential for assessment and training.

References

Ahmidi, N., Hager, G. D., Ishii, L., Fichtinger, G., Gallia, G. L., & Ishii, M. (2010). Surgical task and skill classification from eye tracking and tool motion in minimally invasive surgery. In *Medical image computing and computer-assisted intervention–miccai 2010* (pp. 295–302). Springer.

Ahmidi, N., Ishii, M., Fichtinger, G., Gallia, G. L., & Hager, G. D. (2012). An objective and automated method for assessing surgical skill in endoscopic sinus surgery using eye-tracking and tool-motion data. In *International forum of allergy & rhinology* (Vol. 2, pp. 507–515).

Atkins, M. S., Jiang, X., Tien, G., & Zheng, B. (2012). Saccadic delays on targets while watching videos. In *Proceedings of the symposium on eye tracking research and applications* (pp. 405–408).

Atkins, M. S., Tien, G., Khan, R. S., Meneghetti, A., & Zheng, B. (2013). What do surgeons see: capturing and synchronizing eye gaze for surgery applications. *Surgical Innovation*, *20*(3), 241–248.

Brandt, T., Glasauer, S., & Schneider, E. (2006). A third eye for the surgeon. *Journal of Neurology, Neurosurgery & Psychiatry*, *77*(2), 278.

Charness, N., Reingold, E. M., Pomplun, M., & Stampe, D. M. (2001). The perceptual aspect of skilled performance in chess: Evidence from eye movements. *Memory & Cognition*, *29*(8), 1146–1152.

Chetwood, A. S., Kwok, K. W., Sun, L. W., Mylonas, G. P., Clark, J., Darzi, A., & Yang, G. Z. (2012). Collaborative eye tracking: a potential training tool in laparoscopic surgery. *Surgical Endoscopy*, *26*(7), 2003–2009.

Clementz, B. A., & Sweeney, J. A. (1990). Is eye movement dysfunction a biological marker for schizophrenia? a methodological review. *Psychological Bulletin*, *108*(1), 77.

DeAngelus, M., & Pelz, J. B. (2009). Top-down control of eye movements: Yarbus revisited. *Visual Cognition*, *17*(6-7), 790–811.

Eivazi, S., Bednarik, R., Tukiainen, M., Fraunberg, M. von und zu, Leinonen, V., & Jääskeläinen, J. E. (2012). Gaze behaviour of expert and novice microneurosurgeons differs during observations of tumor removal recordings. In

[2] Even wider availability may be eminent considering Google has obtained a patent for eye tracking with their Google Glass systems. http://www.techradar.com/news/portable-devices/google-patents-eye-tracking-for-google-glass-1091428, retrieved 27/06/2013.

DOI 10.16910/jemr.6.4.4

ISSN 1995-8692

Proceedings of the symposium on eye tracking research and applications (pp. 377–380).

Greene, M. R., Liu, T., & Wolfe, J. M. (2012). Reconsidering yarbus: A failure to predict observers task from eye movement patterns. *Vision research*, *62*, 1–8.

Hermens, F., Loos, R., & Wagemans, J. (2012). An exploration of eye movements when reading texts with atypical spatial layouts. *i-Perception*, *3*(6), 394.

Ibbotson, J. A., MacKenzie, C. L., Cao, C. G., & Lomax, A. J. (1999). Gaze patterns in laparoscopic surgery. *Studies in Health Technology and Informatics*, *62*, 154–160.

James, A., Vieira, D., Lo, B., Darzi, A., & Yang, G.-Z. (2007). Eye-gaze driven surgical workflow segmentation. In *Medical image computing and computer-assisted intervention–miccai 2007* (pp. 110–117). Springer.

Jarodzka, H., Scheiter, K., Gerjets, P., & Van Gog, T. (2010). In the eyes of the beholder: How experts and novices interpret dynamic stimuli. *Learning and Instruction*, *20*(2), 146–154.

Kaakinen, J. K., & Hyona, J. (2005). Perspective effects on expository text comprehension: Evidence from think-aloud protocols, eyetracking, and recall. *Discourse Processes*, *40*(3), 239–257.

Kaakinen, J. K., & Hyönä, J. (2010). Task effects on eye movements during reading. *Journal of Experimental Psychology: Learning, Memory, & Cognition*, *36*(6), 1561.

Kasarskis, P., Stehwien, J., Hickox, J., Aretz, A., & Wickens, C. (2001). Comparison of expert and novice scan behaviors during vfr flight. In *Proceedings of the 11th international symposium on aviation psychology.*

Khan, R. S., Tien, G., Atkins, M. S., Zheng, B., Panton, O. N., & Meneghetti, A. T. (2012). Analysis of eye gaze: do novice surgeons look at the same location as expert surgeons during a laparoscopic operation? *Surgical Endoscopy*, *26*(12), 3536–3540.

Kocak, E., Ober, J., Berme, N., & Melvin, W. S. (2005). Eye motion parameters correlate with level of experience in video-assisted surgery: objective testing of three tasks. *Journal of Laparoendoscopic & Advanced Surgical Techniques A*, *15*(6), 575–580.

Koh, R. Y., Park, T., Wickens, C. D., Ong, L. T., & Chia, S. N. (2011). Differences in attentional strategies by novice and experienced operating theatre scrub nurses. *Journal of Experimental Psychology: Applied*, *17*(3), 233–246.

Law, B., Atkins, M. S., Kirkpatrick, A. E., & Lomax, A. J. (2004). Eye gaze patterns differentiate novice and experts in a virtual laparoscopic surgery training environment. In *Proceedings of the 2004 symposium on eye tracking research & applications* (pp. 41–48).

Monson, J. R. (1993). Advanced techniques in abdominal surgery. *BMJ*, *307*(6915), 1346–1350.

Nicolaou, M., James, A., Darzi, A., & Yang, G.-Z. (2004). A study of saccade transition for attention segregation and task strategy in laparoscopic surgery. In *Medical image computing and computer-assisted intervention–miccai 2004* (pp. 97–104). Springer.

ODriscoll, G. A., & Callahan, B. L. (2008). Smooth pursuit in schizophrenia: a meta-analytic review of research since 1993. *Brain and Cognition*, *68*(3), 359–370.

Olson, C. M., Rennie, D., Cook, D., Dickersin, K., Flanagin, A., Hogan, J. W., Zhu, Q., Reiling, J., & Pace, B. (2002). Publication bias in editorial decision making. *Journal of the American Medical Association*, *287*(21), 2825–2828.

Pallanti, S., Quercioli, L., Zaccara, G., Ramacciotti, A., & Arnetoli, G. (1998). Eye movement abnormalities in anorexia nervosa. *Psychiatry Research*, *78*(1), 59–70.

Rayner, K., Miller, B., & Rotello, C. M. (2008). Eye movements when looking at print advertisements: The goal of the viewer matters. *Applied Cognitive Psychology*, *22*(5), 697–707.

Reiley, C. E., Lin, H. C., Yuh, D. D., & Hager, G. D. (2011). Review of methods for objective surgical skill evaluation. *Surgical Endoscopy*, *25*(2), 356–366.

Richstone, L., Schwartz, M. J., Seideman, C., Cadeddu, J., Marshall, S., & Kavoussi, L. R. (2010). Eye metrics as an objective assessment of surgical skill. *Annals of Surgery*, *252*(1), 177–182.

Rothkopf, E. Z., & Billington, M. (1979). Goal-guided learning from text: Inferring a descriptive processing model from inspection times and eye movements. *Journal of Educational Psychology*, *71*(3), 310.

Schulz, C. M., Endsley, M. R., Kochs, E. F., Gelb, A. W., & Wagner, K. J. (2013). Situation awareness in anesthesia: concept and research. *Anesthesiology*, *118*(3), 729–742.

Schulz, C. M., Schneider, E., Fritz, L., Vockeroth, J., Hapfelmeier, A., Brandt, T., Kochs, E. F., & Schneider, G. (2011). Visual attention of anaesthetists during simulated critical incidents. *British Journal of Anaesthesia*, *106*(6), 807–813.

Segall, N., Taekman, J. M., Mark, J. B., Hobbs, G., & Wright, M. C. (2007). Coding and visualizing eye tracking data in simulated anesthesia care. In *Proceedings of the human factors and ergonomics society annual meeting* (Vol. 51, pp. 765–770).

Shen, J., & Itti, L. (2012). Top-down influences on visual attention during listening are modulated by observer sex. *Vision Research*, *65*, 62–76.

Sodergren, M. H., Orihuela-Espina, F., Clark, J., Darzi, A., & Yang, G. Z. (2010). A hidden Markov model-based analysis framework using eye-tracking data to characterise re-orientation strategies in minimally invasive surgery. *Cognitive Processing*, *11*(3), 275–283.

Sodergren, M. H., Orihuela-Espina, F., Clark, J., Teare, J., Yang, G. Z., & Darzi, A. (2010). Evaluation of orientation strategies in laparoscopic cholecystectomy. *Annals of Surgery*, *252*(6), 1027–1036.

Sodergren, M. H., Orihuela-Espina, F., Froghi, F., Clark, J., Teare, J., Yang, G. Z., & Darzi, A. (2011). Value of orientation training in laparoscopic cholecystectomy. *British Journal of Surgery*, *98*(10), 1437–1445.

Sodergren, M. H., Orihuela-Espina, F., Mountney, P., Clark, J., Teare, J., Darzi, A., & Yang, G. Z. (2011). Orientation strategies in natural orifice translumenal endoscopic surgery. *Annals of Surgery*, *254*(2), 257–266.

Sweeney, J. A., Strojwas, M. H., Mann, J. J., & Thase, M. E. (1998). Prefrontal and cerebellar abnormalities in major depression: evidence from oculomotor studies. *Biological Psychiatry*, *43*(8), 584–594.

Tatler, B. W., Wade, N. J., Kwan, H., Findlay, J. M., & Velichkovsky, B. M. (2010). Yarbus, eye movements, and vision. *i-Perception*, *1*(1), 7-27.

DOI 10.16910/jemr.6.4.4

ISSN 1995-8692

Tien, G., Atkins, M. S., & Zheng, B. (2012). Measuring gaze overlap on videos between multiple observers. In *Proceedings of the symposium on eye tracking research and applications* (pp. 309–312).

Tien, G., Atkins, M. S., Zheng, B., & Swindells, C. (2010). Measuring situation awareness of surgeons in laparoscopic training. In *Proceedings of the 2010 symposium on eye-tracking research & applications* (pp. 149–152).

Tien, G., Zheng, B., & Atkins, M. S. (2011). Quantifying surgeons' vigilance during laparoscopic operations using eyegaze tracking. *Studies in Health Technology and Informatics*, *163*, 658–662.

Tomizawa, Y., Aoki, H., Suzuki, S., Matayoshi, T., & Yozu, R. (2012). Eye-tracking analysis of skilled performance in clinical extracorporeal circulation. *Journal of Artificial Organs*, *15*(2), 146–157.

Valerie, A., Huemer, M., Hayashi, M., Renema, F., Elkins, S., McCandless, J. W., & McCann, R. S. (2005). Characterizing scan patterns in a spacecraft cockpit simulator: expert vs. novice performance. In *Proceedings of the human factors and ergonomics society annual meeting* (Vol. 49, pp. 83–87).

Vine, S. J., Chaytor, R. J., McGrath, J. S., Masters, R. S., & Wilson, M. R. (2013). Gaze training improves the retention and transfer of laparoscopic technical skills in novices. *Surgical Endoscopy*, *27*(9), 3205–3213.

Vine, S. J., Masters, R. S., McGrath, J. S., Bright, E., & Wilson, M. R. (2012). Cheating experience: Guiding novices to adopt the gaze strategies of experts expedites the learning of technical laparoscopic skills. *Surgery*, *152*(1), 32–40.

Vine, S. J., Moore, L. J., & Wilson, M. R. (2011). Quiet eye training facilitates competitive putting performance in elite golfers. *Frontiers in psychology*, *2*.

Vine, S. J., Moore, L. J., & Wilson, M. R. (2012). Quiet eye training: The acquisition, refinement and resilient performance of targeting skills. *EJSS*(2012), 1–8.

Williams, A. M., Singer, R. N., & Frehlich, S. G. (2002). Quiet eye duration, expertise, and task complexity in near and far aiming tasks. *Journal of Motor Behavior*, *34*(2), 197–207.

Wilson, M., McGrath, J., Vine, S., Brewer, J., Defriend, D., & Masters, R. (2010). Psychomotor control in a virtual laparoscopic surgery training environment: gaze control parameters differentiate novices from experts. *Surgical Endoscopy*, *24*(10), 2458–2464.

Wilson, M. R., McGrath, J. S., Vine, S. J., Brewer, J., Defriend, D., & Masters, R. S. (2011). Perceptual impairment and psychomotor control in virtual laparoscopic surgery. *Surgical Endoscopy*, *25*(7), 2268–2274.

Wilson, M. R., Vine, S. J., Bright, E., Masters, R. S., Defriend, D., & McGrath, J. S. (2011). Gaze training enhances laparoscopic technical skill acquisition and multi-tasking performance: a randomized, controlled study. *Surgical Endoscopy*, *25*(12), 3731–3739.

Yarbus, A. L. (1967). *Eye movements and vision.* Plenum press New York.

Zheng, B., Tien, G., Atkins, S. M., Swindells, C., Tanin, H., Meneghetti, A., Qayumi, K. A., Neely, O., & Panton, M. (2011). Surgeon's vigilance in the operating room. *The American Journal of Surgery*, *201*(5), 673–677.

DOI 10.16910/jemr.6.4.4

ISSN 1995-8692