Verbal gaze instruction matches visual gaze guidance in laparoscopic skills training

Geoffrey Tien*
M. Stella Atkins†
Xianta Jiang‡
Simon Fraser University

Bin Zheng§
University of Alberta

Roman Bednarik¶
University of Eastern Finland

Abstract

Novices were trained to perform a unimanual peg transport task in a laparoscopic training box with an illuminated interior displayed on a monitor. Subjects were divided into two groups; one group was verbally instructed to direct their gaze at distant targets, while the other group had their gaze behaviour implicitly manipulated using distant target illumination. Both groups achieved similar task completion times post-training and developed peripheral vision strategies leading to delayed foveation on targets until the instrument was closer to its destination, although the ability to focus on targets earlier during manual movements as done by an expert surgeon was quickly regained by the verbal instruction group post-training. This suggests that care should be taken when employing visual attention cuing methods such as target highlighting for training eye-hand coordination skills, as simple verbal instruction may be sufficient to help trainees to adopt more expert-like gaze behaviours.

CR Categories: H.1.2 [Information Systems]: Models and Principles – User/Machine Systems—Human Factors;

Keywords: natural scene perception, gaze training, laparoscopic surgery

1 Introduction and background

Laparoscopic surgery is a minimally-invasive technique in which a surgeon uses long, slender instruments to operate inside a patient's abdominal cavity, guided by a video camera displaying the surgical site on an external monitor. Due to requiring only a few small incisions compared to traditional open surgery, patient morbidity and recovery times are greatly reduced, at the expense of requiring special training on the part of the surgeon to learn a new, non-intuitive mode of surgical manipulation. With indirect manipulation with reduced tactile feedback and indirect vision without depth perception, surgeons must learn the motor transformations to produce the necessary visual and surgical outcome [Ibbotson et al. 1999]. Since visual input becomes extremely important in such image-guided procedures, studies have been made to observe eye movement behaviours in laparoscopic operators as they progress in skill level.

Various studies consistently demonstrate gaze behaviour differences between expert and novice operators in various domains including laparoscopic surgery. In certain task-specific training scenarios, experts were found to make faster, larger saccades and longer fixations than novices [Kocak et al. 2005], as well as spending a higher proportion of task time gazing at the proper targets [Wilson et al. 2010b]. Another behavioural difference was established by Law et al. who observed that surgical novices perform tool-tracking rather than target-tracking as done by experts [Law et al. 2004]. Differences continue to hold under the extreme optical magnification conditions of microneurosurgery, where experts viewing static images captured from live operations were quicker to visually attend to critical areas of the scene and made longer fixations on these precise areas compared to novices viewing the same images [Eivazi et al. 2012]. In addition to performance and manual differences, expert surgeons were also found to experience lower cognitive loads than novices while operating, allowing the experts to be more aware of potential hazards in the working environment not directly occurring at the surgical site [Zheng et al. 2010].

It has been postulated that task learning in the surgical domain could be accelerated by training novices to adopt expert gaze behaviours [Wilson et al. 2010a]. Such studies were conducted, administering gaze training using explicit verbal instructions [Wilson et al. 2011] and implicitly with modifications to the visual scene [Vine et al. 2012; Vine et al. 2013].

Due to the tendency of novices to tool-track, which is considered harmful, this study used settings which allowed full vision of the scene, with the addition of illuminated targets to indicate the next destination for tool movement. Therefore, trainees were still able to make tool-tracking eye movements if they wished, but could have their attention drawn to a more salient target. This is unlike the setting used by Vine et al. to promote gaze training, which applied a software-based mask over the laparoscopic display, darkening non-target areas of the screen [Vine et al. 2012; Vine et al. 2013].

Under these new experimental conditions, we hoped to achieve similar results to Vine et al. [Vine et al. 2012; Vine et al. 2013] who demonstrated with a large sample of 27 to 36 subjects that trainees who underwent a gaze training protocol were able to complete a task roughly 25% faster than subjects who learned the task without any instruction. Furthermore we wished to compare the effect of an implicit visual gaze training using target illumination to an explicit verbal instruction of gaze targeting such as that administered by Wilson et al. [Wilson et al. 2011] but without any additional visual cues. Specifically, we hypothesize that after training, novices subjected to a visual highlight gaze training will more rapidly adopt gaze patterns similar to an expert surgeon, and that while both verbal and visual training groups will display increased task performance, the visual training group will produce shorter task completion times, in line with the behavioural and objective changes observed in the research of Vine et al.

On the other hand, if our findings can demonstrate that verbal instructions alone are sufficient in cuing attention of novices so that they adopt expert-like strategies, it would have significant implications into the design of eye-hand coordination training procedures.

---

*e-mail: gtien@cs.sfu.ca
†e-mail: stella@cs.sfu.ca
‡xiantaj@cs.sfu.ca
§e-mail: bzheng1@ualberta.ca
¶email: roman.bednarik@uef.fi

Permission to make digital or hard copies of part or all of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for commercial advantage and that copies bear this notice and the full citation on the first page. Copyrights for third-party components of this work must be honored. For all other uses, contact the Owner/Author.

ETRA 2014, March 26 – 28, 2014, Safety Harbor, Florida, USA.
2014 Copyright held by the Owner/Author.
ACM 978-1-4503-2751-0/14/03

2 Methods

We designed a laparoscopic peg transport environment with LED-illuminated targets linked to a series of electrical switches activated by placement or removal of the transported peg at the target locations, to visually cue learners to the next movement destination.

2.1 Apparatus

The training platform was based on existing laparoscopic education tools, with eye tracking performed by a Tobii X2-60-Wide remote eye tracker with a sampling rate of 60Hz [Tobii Technology]. A 3D Medical Services training box [Franklin, OH] was used, with its joystick camera removed and replaced with a fixed-position high framerate web camera, providing a laparoscopic view in 640×480 resolution at 60Hz. The data collection was managed using Tobii Studio 3.2.1 in 64-bit Windows 7. The training box contents were displayed on 4:3 aspect ratio LCD monitors with native resolution of 1600×1200. A photograph of the experimental setup at the University of Alberta is shown in Figure 1.

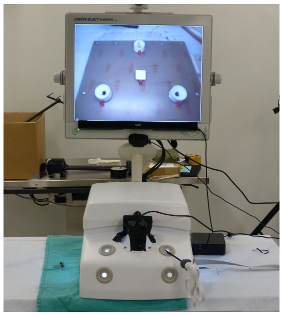

Figure 1: *Peg transport setup with Tobii X2-60 Wide eye tracker attached to Stryker Vision Elect HD monitor*

Inside the laparoscopic training box, three illuminated cups and a simple horizontal push-switch were used as training targets and arranged in a similar layout as earlier peg studies [Atkins et al. 2012]. Each target had one 5mm red LED affixed adjacent to it. The physical peg board is shown in Figure 2.

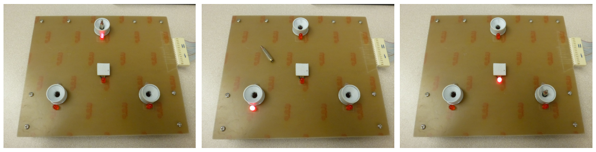

Figure 2: *Photographs of pegboard showing possible illumination states. From left to right: target dish at top illuminated; target dish at LHS, target dish at RHS, and target "Home" position.*

The cups had an outer diameter of 20mm and were arranged in an equilateral triangle with side lengths of 103mm. The central home button measured 15mm on each side and was given a tactile surface of textured hooks. The transported peg was a steel standoff with a 20mm hexagonal shaft with a 6mm threaded end which was wrapped with a thin layer of cloth tape to provide more compliance when held with the laparoscopic grasper.

The sequence of target illumination was controlled by a custom program on a microcontroller platform [arduino.cc]. Each target cup contained a mechanical switch which detected insertion and removal of the transported peg, and the home target was also constructed with a switch that could be depressed directly using the laparoscopic instrument. Based on activation of the various switches, targets were illuminated in the same order defined in an earlier study [Atkins et al. 2012]. With this electronic setup, illumination state changes were automatic and instantaneous, transitioning as soon as a subtask was completed without any reaction time lag from an experimenter-controlled manual interface. The subtask completion timestamps were electronically logged, eliminating the tedious and subjective step of manual subtask annotation from video as was done in earlier studies [Atkins et al. 2012].

2.2 Participants and procedure

Experimental participants included one qualified medical resident specializing in laparoscopic surgery, and eight novices who were naïve to laparoscopy. Performance data and gaze data recorded from the expert participant served as a reference point against which the novice data were to be compared. Subjects were provided a written description of the experimental goals and then asked to give signed consent to participate. Each completed a short questionnaire to collect demographic data; expert-level subjects were given an additional survey to determine their surgical experience score [Zheng et al. 2010]. Next, subjects were given a written/pictorial description of the experimental task and shown a demonstration by the experimenter. In a seated position, they were allowed to use the grasper to practice insertion and removal of the peg from each of the cups approximately five times. Subjects then underwent a 9-point calibration procedure in Tobii Studio.

After completing three untrained trials of the task as a baseline performance measure, novice subjects were randomly assigned to one of two training conditions: unlit targets with verbal gaze direction (similar to discovery learning as in Vine et al.'s study [Vine et al. 2012], DL, but with the important difference that these subjects were verbally instructed to gaze at the target rather than the tool), and illuminated targets without any explicit verbal direction (gaze-trained, GT). Subjects using the illuminated condition were only informed that the targets would be lit and were not explicitly asked to adjust their gaze in any way.

All novice subjects then completed six blocks of five trials each for a total of 30 trials under their assigned training conditions, with breaks as necessary in between blocks. After completing the sixth training block, subjects were asked to complete an electronic version of the weighted NASA Task Load Index survey as well as another brief survey to assess their perception of the training stimulus.

Finally, the training stimulus was removed and each novice subject completed three more trials of the experimental task to measure the retention of the skill. Full participation in the study typically lasted 30 to 45 minutes, with individual trials lasting from 25 seconds to 2 minutes.

2.3 Data analysis

Each trial was divided into 9 subtasks as detailed in [Atkins et al. 2012]: reaching and grasping tasks (RG) involved moving the grasper from the central home position to grasp the peg resting in one of the target cups; transport tasks (TR) were done by moving the peg from its original position to the next target cup; homing tasks (H) involve returning the empty grasper to the home position once the peg is inserted into its target cup – a single trial consists of three of each of these subtask types.

Subtask completion time was electronically recorded by the apparatus, as well as a measure of the spatial separation between the instrument and target location at the moment the target is visually acquired by the subject. This eye measure is reported as the Euclidean distance between the target and instrument tip at the moment of visual target acquisition, divided by the Euclidean distance

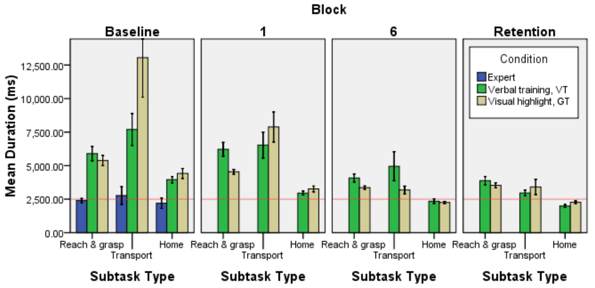

Figure 3: *Mean subtask (± 1 S.E.) completion time by experimental block (blocks 2 – 5 omitted). A horizontal red line roughly indicates the performance achieved by the expert surgeon. A shorter bar indicates better performance.*

	Subtask type		
VT vs GT	RG	TR	H
baseline	0.437	0.096	0.294
retention	0.324	0.461	0.060

Table 1: *Completion times compared using two-tailed t-test by subtask.* n *= 36 each for VT and GT (4 subjects × 3 of each subtask type × 3 trials).*

between the current target location and the previous target location. The step of division is simply done to normalize the measure over different inter-target distances - for example, it can be seen that TR subtasks necessarily involve movement of the grasper over larger distances than the RG and H subtasks.

The use of this tool-target separation measure contrasts with the "target-locking" eye measure score used by Vine et al. [Vine et al. 2013] which counted the proportion of time spent in fixations over the required target to time spent fixating on the instrument. In Vine's study, use of this target-locking measure made possible with a head-mounted eye tracker that could reliably maintain a steady view of the wearer's eye over the entire duration of the experimental trials. In our study, we were interested in more precisely observing the point of gaze within the surgical training scene, and this requirement for high spatial resolution (0.5° visual angle) was met with a remote Tobii X2-60-Wide eye tracker. However, with a remote eye tracker, subjects may move out of tracking range, resulting in intermittent data loss. Thus it was more appropriate to use instantaneous gaze measures for each subtask rather than relying on the availability of valid fixations over an entire trial to provide a measure such as a target-locking score.

3 Results

The collected data had a mean proportion of invalid/missing data at 0.09 (min. < 0.01, max. 0.26). Analysis was performed on filtered data with a moving average and gap-filling interpolation.

The mean duration of subtasks through progression of training is shown in Figure 3. The expert performed quickest at baseline without any practice. A table of independent-samples t-test differences in subtask performance between the two training groups is provided in Table 1, showing that the training groups did not show significantly different subtask completion times. A Bonferroni correction applied for a family-wise significance level of 0.05 with 12 tests yields individual test significance at $p \approx 0.004$.

The tool-target separation measure is illustrated in Figure 4. Training condition t-test differences for this measure are given in Table 2.

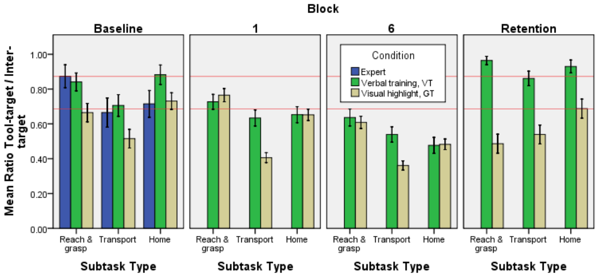

Figure 4: *Mean tool-target separation (± 1 S.E.), normalized by the distance between the movement origin and destination. The reference lines show expert behaviour at baseline for the RG task, and approximately for TR and H subtasks. A longer bar indicates more target-oriented gaze behaviour.*

	Subtask type		
VT vs GT	RG	TR	H
baseline	0.020	0.026	0.043
retention	<0.001*	<0.001*	<0.001*

Table 2: *Tool-target separation differences compared using two-tailed t-test by subtask. Values of* n *are the same as in Table 1.*

Significantly different table entries at the corrected level are marked with an asterisk.

Table 2 together with Figure 4 show that on average the novices begin training with gaze behaviour not significantly different from the expert. The GT novices remain the same as they started, whereas the VT group significantly increased their tool-target separation beyond expert-levels at retention.

4 Discussion

The results obtained were contrary to the expectation set by the study of Vine et al. [Vine et al. 2013]. While Vine was able to demonstrate a significant effect of his gaze training method to both task completion time and a target-locking measure, his findings were not supported in this particular experimental setting.

Subjects in both training conditions approached expert-level task performance by the end of the training period, with no statistically significant performance difference between the two training groups (Figure 3 & Table 1). Surprisingly, the GT group did not exhibit any change in performance when target illumination was removed during the retention phase. It may be concluded that the task was repetitive enough that the task's motor requirements were predictable and already learned well enough, so that the added visual cues were no longer needed.

The VT group at baseline exhibited more target-oriented behaviour on the tool-target separation measure than the GT group, although this may be largely attributed to population sampling - it can be seen in Figure 4 that while not significantly different, the VT group already demonstrated a relatively higher measure during the baseline phase.

During the training blocks, all the subjects had a progressively lower tool target separation measure, while their performance steadily improved. This may arise due to developing a strategy of directing the tool to the target by using their peripheral vision, made easier because of the predictable nature of the target locations.

Neither group was given any specific instruction about the task before the retention trials, which were performed after a time gap during which the post-training questionnaire was completed. During the retention trials, the gaze behaviour of the two groups became significantly different (Table 2 & Figure 4), when the VT group appeared to gaze at the target to direct the tool for the retention trials; recall that they had received the instruction to gaze at the target before their training blocks. However the GT group did not change their eye behaviour on the target between the training trials and the retention block, although the lights were switched off. Both groups performed similarly at the retention trials.

This implies that using verbal instruction to direct their gaze towards the next target might be as effective as more complex gaze training methods in a simple repetitive task where peripheral vision can be employed and the targets are totally predictable. This seems to contradict the findings of Masters et al. which indicate that manual skills learned under explicitly verbal instruction are not as robust as those learned implicitly by observation [Masters et al. 2008]. Although we did not find any significant difference in task performance between the two training groups, the high degree of tool-target separation at retention for the VT group suggests that a target-oriented gaze behaviour is still learned eventually, and the verbal instruction may allow trainees to bypass the step of discovering this target-oriented behaviour which, while not correlated with completion time for this simple, repetitive task, may be a useful ability when performing more complex and dynamic task sequences.

The use of target illumination for gaze manipulation allowed GT subjects to use peripheral vision; however, with a darkened mask as used by Vine, the more pronounced highlighting as well as partially obscuring non-target areas conceivably discouraged peripheral vision of targets. Thus the two different highlighting methods could each give rise to very different gaze strategies by trainees.

On the technical side, Vine et al. employed a head-mounted eye tracker, whereas a remote system was used here. One is stronger where the other is weak and vice-versa as discussed above, but as a result it is infeasible to collect the same gaze parameter measurements for direct comparison. Although it was necessary to use a different eye measure with our setup, the progress of the subjects' task performance could still be monitored, in spite of the task itself being quite different, requiring considerably high precision to insert the peg into the target cups at the correct angle.

Lastly, the participant pool in this study is smaller compared to the 36 subjects recruited in Vine et al.'s study. The results here were prone to be skewed by sampling error, for example with our VT group exhibiting noticeably different gaze behaviour from the GT group at baseline. The effects of such anomalies can be reduced if the study is in the future expanded to include more novice trainees.

5 Conclusion

Previous attempts to train novice surgical trainees to adopt expert-like gaze behaviours have shown success in both gaze behaviour manipulation and rapid skill acquisition under certain experimental conditions. This study presented an alternative difficult task and gaze training method using a fixed-camera laparoscopic simulation, which led to a peripheral-vision-oriented gaze strategy that yielded neither harm nor benefit over a simple verbal gaze instruction. In future work, a control group receiving neither verbal nor visual gaze training will be studied to assess the benefit of such gaze training effects and to investigate the transfer of gaze behaviors to more complex tasks.

References

ATKINS, M. S., JIANG, X., TIEN, G., AND ZHENG, B. 2012. Saccadic delays on targets while watching videos. In *Proceedings of the 2012 symposium on Eye tracking research & applications*, ETRA '12, 405–408.

EIVAZI, S., BEDNARIK, R., TUKIAINEN, M., VON UND ZU FRAUNBERG, M., LEINONEN, V., AND JÄÄSKELÄINEN, J. E. 2012. Gaze behaviour of expert and novice microneurosurgeons differs during observations of tumor removal recordings. In *Proceedings of the Symposium on Eye Tracking Research and Applications*, ACM, New York, NY, USA, ETRA '12, 377–380.

IBBOTSON, J. A., MACKENZIE, C. L., CAO, C. G. L., AND LOMAX, A. J. 1999. Gaze patterns in laparoscopic surgery. *Studies in Health Technologies and Informatics 62*, 154–160.

KOCAK, E., OBER, J., BERME, N., AND MELVIN, 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, Part A 15*, 6 (December), 575–580.

LAW, B., ATKINS, M. S., KIRKPATRICK, A. E., AND 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*, ACM, New York, NY, USA, ETRA '04, 41–48.

MASTERS, R., LO, C., MAXWELL, J., AND PATIL, N. 2008. Implicit motor learning in surgery: Implications for multi-tasking. *Surgery 143*, 1, 140–145.

VINE, S. J., MASTERS, R. S. W., MCGRATH, J. S., BRIGHT, E., AND WILSON, M. R. 2012. Cheating experience: Guiding novices to adopt the gaze strategies of experts expedites the learning of laparoscopic technical skills. *Surgery 152*, 1, 32–40.

VINE, S. J., CHAYTOR, R. J., MCGRATH, J. S., MASTERS, R. S. W., AND WILSON, M. R. 2013. Gaze training improves the retention and transfer of laparoscopic technical skills in novices. *Surgical Endoscopy 27*, 9 (September), 3205–3213.

WILSON, M. R., COLEMAN, M., AND MCGRATH, J. S. 2010. Developing basic hand-eye coordination skills for laparoscopic surgery using gaze training. *British Journal of Urology 105*, 10 (March), 1356–1358.

WILSON, M. R., MCGRATH, J. S., VINE, S. J., BREWER, J., DEFRIEND, D., AND MASTERS, R. S. W. 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., VINE, S. J., BRIGHT, E., MASTERS, R. S. W., DEFRIEND, D., AND MCGRATH, J. S. 2011. Gaze training enhances laparoscopic technical skill acquisition and multi-tasking performance: a randomized, controlled study. *Surgical Endoscopy 25*, 12 (December), 3731–3739.

ZHENG, B., CASSERA, M. A., MARTINEC, D. V., SPAUN, G. O., AND SWANSTRÖM, L. L. 2010. Measuring mental workload during the performance of advanced laparoscopic tasks. *Surgical endoscopy 24*, 1 (January), 45–50.