

# Perceptual-motor coordination in an endoscopic surgery simulation

J. G. Holden,<sup>1,\*</sup> J. M. Flach,<sup>1</sup> Y. Donchin<sup>2</sup>

<sup>1</sup> Psychology Department, 3640 Colonel Glenn Highway, Wright State University, Dayton, OH 45435, USA

<sup>2</sup> Hadassah Hebrew University Medical Center and the Technion, Haifa, Israel

Received: 27 February 1998/Accepted: 30 April 1998

## Abstract

**Background:** This study examined perceptual-motor coordination with an apparatus that simulated a situation representative of endoscopic surgery.

**Methods:** Participants were trained with one arrangement of the apparatus, then tested with an alternative arrangement in which either the positions of the camera, the surgeon, or the objects in the surgical field were altered.

**Results:** Results showed that changes of either the camera's position or the surgeon's position disrupted performance. However, when the camera and surgeon positions were changed together, skilled performance was maintained.

**Conclusions:** This suggests that skill depends on a consistent mapping between the virtual hands and eyes, but not on the particular visual or motor orientations. The results suggest that movements of the camera during surgery can disrupt coordinated action. Also, in the design of training simulators, the mapping between camera and instruments may be more important than the static appearance of the displays or the topology of the movements.

**Key words:** Endoscopy — Perceptual-motor skill — training — Simulation

*Not in the past 100 years has such an upheaval in medicine occurred: The "discipline of surgery" is joining the technological revolution and advancing the state of the art with laparoscopic surgery. This represents a radical shift in the concept of surgical practice. The "great leap of faith" has occurred; for the first time in history, surgeons are performing surgical procedures without physically seeing or touching the organs they are removing or repairing [6].*

For centuries, surgeons have opened the belly in order to reach diseased organs. This procedure allowed the knowledge acquired by anatomists from the dissection of the human body to be transferred to the operating room. The ability to directly view and feel the internal organs has been

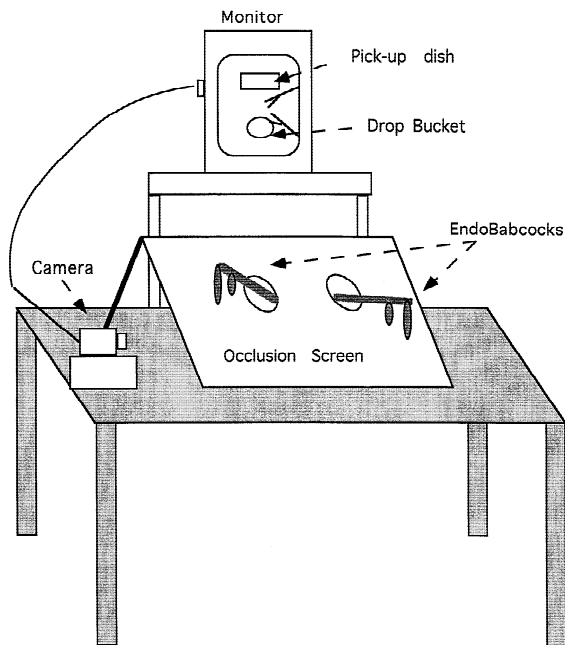
crucial to the perceptual-motor skill of expert surgeons. However, the introduction of the operating microscope and later the extensive use of the video camera have changed the way surgeons are coupled to the operating field. Direct vision of the surgical field has been replaced by a video image, and long, thin surgical instruments have replaced the fingers. The process of adaptation to this new set of eyes (i.e., the video camera) and new set of hands (i.e., the laparoscopic instruments) poses a challenge to both surgeons [3–6] and to psychologists [1, 2, 4, 7–9] interested in the development of perceptual-motor skill.

From the point of view of a perceptual researcher, one of the interesting features of endoscopic tasks, is that the surgical instruments and camera essentially become extensions of the surgeon's perceptual-motor system. The surgical tools serve as extended "hands" that allow for the manipulation of the patient's tissues. The video camera serves as an extended "eye" that is the source of information and feedback on the situation inside the patient. One consequence of this extension of the perceptual-motor system is an increase in the number of ways the components of the system can be coupled. Humans normally work using a relatively constrained coupling between their hands and eyes. Endoscopy, however, allows for many possible mappings or couplings between the hands and the eyes, because the video camera moves independently of the surgical instruments. For example, a movement of the right hand from left to right in the normal visual field may appear as a motion from right to left (or a motion in depth) on the video monitor, depending on the orientation of the camera.

The introduction of new hand-eye mappings can often disrupt coordination. New coordination patterns must be learned before one can perform efficiently. Furthermore, this particular artificial extension of the hands and eyes collapses certain sources of information about the spatial location of structures related to the task. The monocular view, for example, prevents the use of binocular stereopsis to determine the relative depth of the objects on the video monitor. Finally, control of the camera (eye) position and instruments (hands) is typically distributed over a team of surgeons. For example, in laparoscopic cholecystectomy,

\*Present address: Arizona State University, Tempe, AZ, USA

Correspondence to: J. M. Flach



**Fig. 1.** Perspective view of the experimental setting, the camera, the video monitor, and the occluding screen.

the camera is typically held by an assisting surgeon or nurse while a primary and an additional assisting surgeon manipulate instruments to retract, clip, and cut anatomical structures.

The difficulty of adapting to these various constraints was vividly described in a recent editorial:

No one could have predicted that here he would be, 56 years old and at the peak of his surgical prowess, grappling with a new technology that was changing the nature of the field right before his eyes . . . he was too young to retire, but he felt uneasy being a general surgeon who couldn't remove gallbladders. The question was simple. Was he fit to make it up the slope of his next learning curve? [3]

Over the past three years, we have been doing field and laboratory work to study human performance in the domain of minimally invasive surgery [1, 2, 4]. The purpose of the current study was to investigate the perceptual, motor, and cognitive factors that contribute to skilled performance—in other words, to characterize the learning curve associated with minimally invasive surgery.

A simple simulation of endoscopic surgery was developed, as shown in Fig. 1. The intent was not to create a virtual simulation of an endoscopic task, but rather one that reproduces many of the perceptual difficulties that surround endoscopic procedures (e.g., monocular visual display and perceptual-motor decoupling). The research participants were trained in a manual task that displaced their normal hand-eye coordination. An occluding screen blocked a direct view of the task space. A video camera was placed perpendicular to the direction the participant was facing while performing the task. The resulting side view of the task space was presented to a video monitor and provided visual performance feedback for the participants.

The experiment used a transfer paradigm in which participants were trained with one arrangement of the appara-

tus. They were then tested in an alternative arrangement in which either the position of the camera, the surgeon, or the objects in the surgical field was altered. The critical issue was the extent to which skills developed in the training phase transferred to the test phase. If the alterations changed dimensions that were critical to skill, then there should be little transfer from practice to test phases (i.e., performance should be worse, slower in the test phase). If the alterations preserved critical dimensions, then skills developed in training should carry over to good performance in the test phase (i.e., performance in the test phase should be at the same levels as at the end of training).

## Materials and methods

### Participants

Forty-eight students of psychology received either course credit or \$5 an hour for their participation in the study. Each participant had normal or corrected-to-normal vision.

### Apparatus

The low-fidelity surgical simulation is illustrated in Fig. 1. A small video camera was used to displace the participant's visual information about the task space. A standard 27-in (diagonal) color television monitor was used to present visual feedback. In addition, an occluding screen blocked the participant's direct view of the movements of their surgical instruments and the objects they were manipulating.

The occluding screen held a set of endoscopic grasping instruments. A  $2.5 \times 3$ -ft-wide piece of plywood was fitted, at its top edges, with a pair of hinged legs. This allowed the screen to sit on top of a table at an  $\sim 45^\circ$  angle. Two 4-in diameter holes were cut 4 in from the center of the screen. Nails were used to string heavy-duty rubber bands across the holes. The rubber bands were arranged so that they ran perpendicular to each other and intersected at the center of the holes cut in the occluding screen. Each of the surgical instruments was inserted inside the intersection of the rubber bands. The rubber bands served to hold the instruments so that they could move freely and center themselves automatically. An opaque piece of fabric was positioned over the top of the occluding board so the participants could not see through the instrument holes. Small plastic chips ( $\sim 50$  mm in diameter and 3 mm thick) of various shapes were placed in a pickup dish on one side of the space and a drop bucket was positioned on the other side of the space.

### Procedure

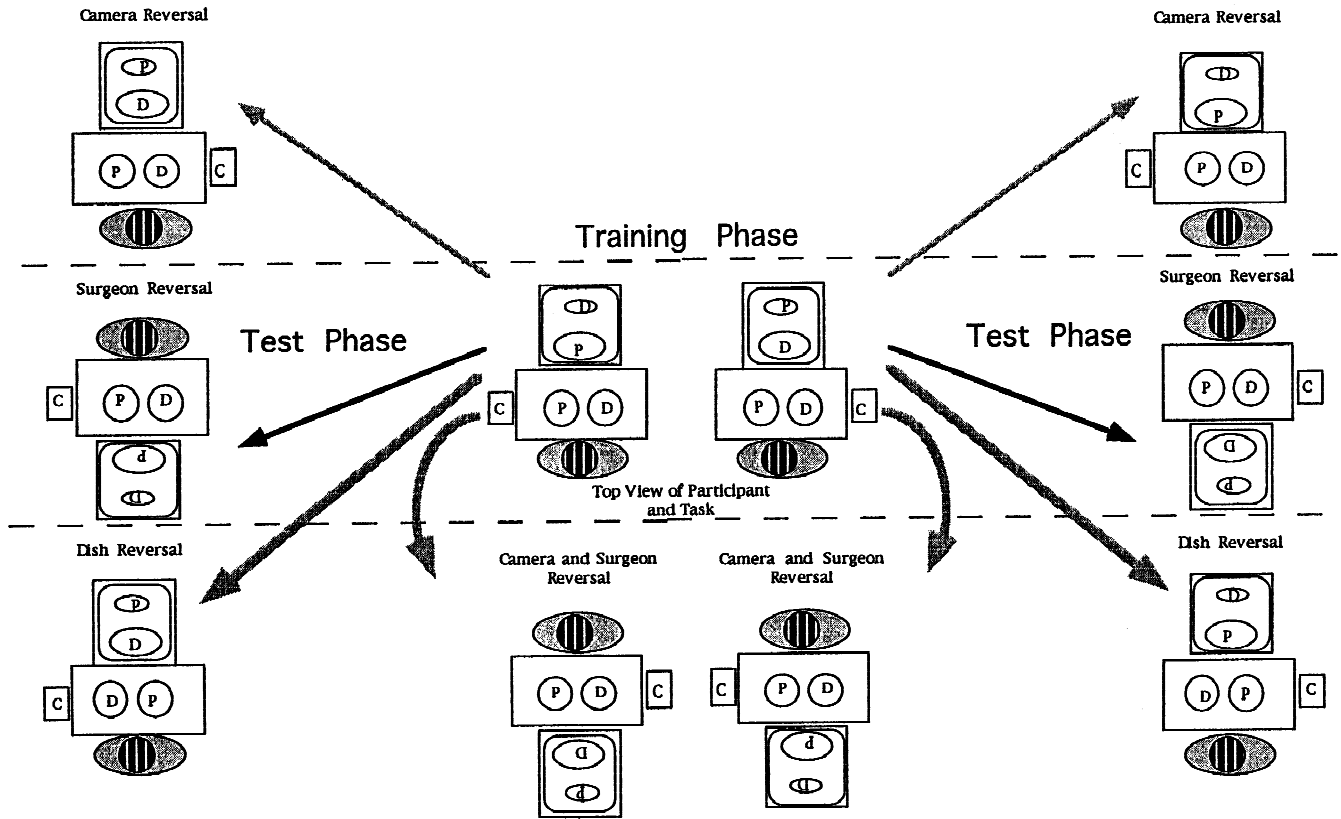
All participants were trained in a simple task using the surgical simulator. For the training task, the video camera was located on either the participant's left or right side and oriented at a right angle to the direction the participant was facing. In other words, the participant saw a side view of the work space.

During the training phase, the participants were asked to use their left hand to pick up plastic chips from a shallow pickup dish; to transfer them to the other instrument, controlled by their right hand; and finally to dump the chips in a small drop bucket on the other side. For a trial to be completed, the participant needed to transfer eight chips from the pickup dish to the drop bucket. A total of 10 training trials were conducted, seven on the first day and three on the second day. For half of the subjects, the camera was on the right for training trials; for the other half, the camera was on the left. The dependent measure was the amount of time, in seconds, to complete each trial.

After the 10 training trials were completed, each participant was assigned to one of four test conditions.

In the *camera reversal* condition, the video camera was simply moved to the opposite side of the work space, relative to its position during training. For example, if the participant was trained with the camera on the left, then a camera reversal placed the camera on the participant's right. This changed the visual perspective of the work space. Objects that were in the foreground during training would appear in the background in the test phase.

The second reversal condition, the *surgeon reversal*, was accomplished



**Fig. 2.** The layout of the eight transfer manipulations by view and reversal condition. Eight experimental manipulations created different combinations of changes in the position of the camera, the surgeon, or the objects in the surgical field.

by moving the participant and the video monitor to opposite sides of the work space. This changed the orientation of the hands to the work space. Actions that were previously performed with the right hand now had to be performed with the left hand.

The third reversal condition, the *camera-surgeon reversal*, combined the changes from the first two conditions. This was implemented by moving the participant and the video monitor to opposite sides of the work space. In addition, the camera was also moved opposite its location during the training trials. In this condition, both the visual perspective and the motor orientation changed from the training to test phases. However, the mapping from camera to hands was constant.

Finally, in the fourth reversal condition, the *dish reversal*, the camera and participant remained where they had been during training, but the positions of the pickup dish and drop bucket were reversed on the task table. This change also affected both the visual orientation (the position of the objects in the visual field) and the motor orientation (the role of the two hands), but a constant mapping was maintained from camera to hands. The eight conditions resulting from the factorial combination of the two camera positions for training and the four reversal conditions for the test phase are illustrated in Fig. 2.

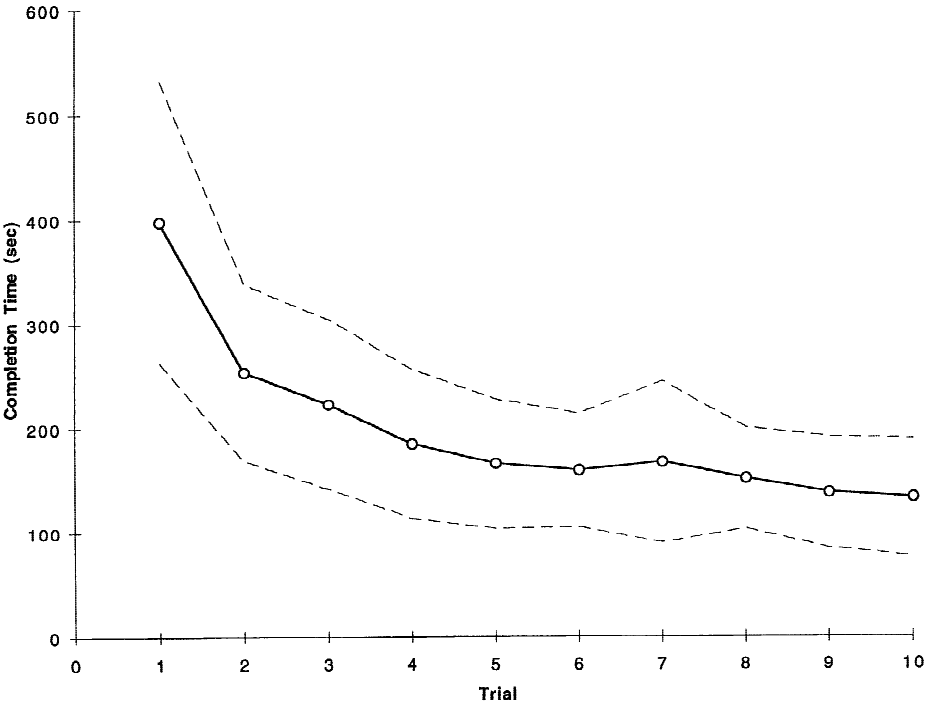
## Results

The data for each participant with the poorest initial performance in the training phase of the experiment were eliminated for each group. As a consequence, data for a total of forty participants were employed in the statistical analysis ( $n = 5$  in each view by condition cell). The completion times for the 10 training trials were evaluated using a 2 (view)  $\times$  4 (reversal condition)  $\times$  10 (trial) mixed design analysis of variance.

Figure 3 shows the significant effect for trials ( $F(9,288) = 74.09$ ;  $p < .001$ ). Each data point represents an average of over 40 participants. Performance improved from a mean time of 398 sec ( $SD = 135$ ) on trial 1 to 132 sec ( $SD = 57$ ) on trial 10. The dashed lines in Fig. 3 show plus and minus one standard deviation about the means. Note that there is a bulge in variability on trial 7. This was the last trial for day 1 of training, so the increased variability probably reflects fatigue. Note also that trial 8, which is the first trial on day 2 of the experiment, shows good retention of the skills learned on day 1.

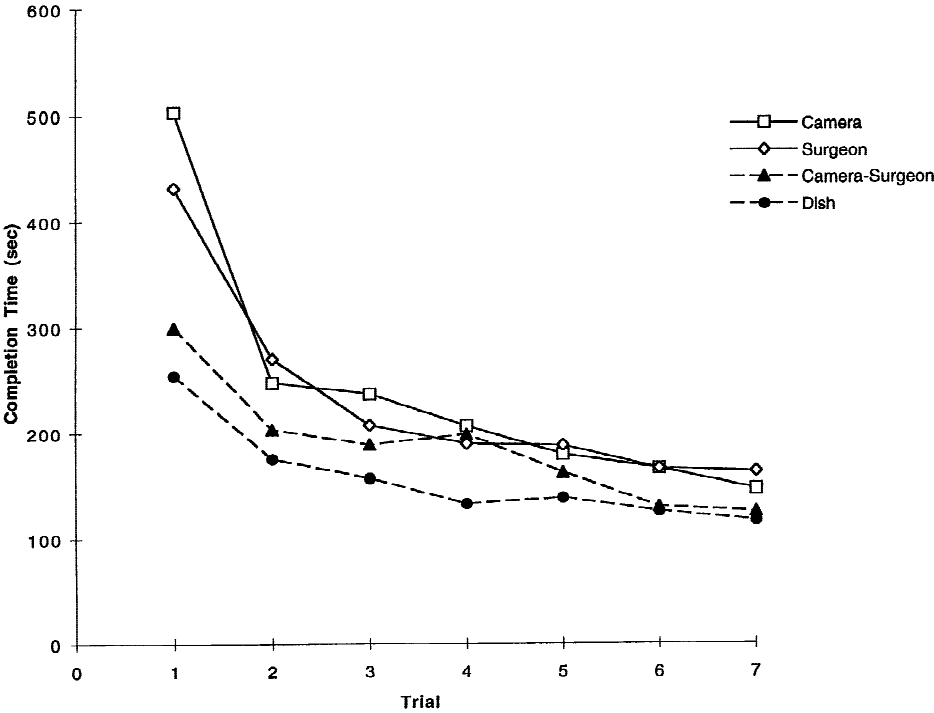
The data from the test phase of the experiment were analyzed using a 2 (view)  $\times$  4 (reversal condition)  $\times$  7 (trials) mixed design analysis of variance. As in the training phase, there was significant improvement over trials ( $F(6,192) = 42.55$ ;  $p < .001$ ). Mean completion time improved from 372 sec ( $SD = 222$ ) on the first transfer trial to 137 sec ( $SD = 47$ ) on the last trial. Also, there was an interaction between reversal condition and trials, as shown in Fig. 4 ( $F(18,192) = 2.25$ ;  $p = .036$ ). The *camera reversal* and *surgeon reversal* conditions showed a marked increase in completion time. Mean completion times for the first transfer trial in these two conditions were comparable to performance at the start of training, suggesting no transfer. The *camera-surgeon reversal* and the *dish reversal* conditions showed a much smaller increase in completion time for the first transfer trial, indicating positive transfer.

Three planned comparisons were evaluated using data



**Fig. 3.** Training trials. Mean completion time in seconds as a function of trial in the training phase of the experiment. Each point represents an average of >40 participants. The dashed lines represent plus and minus one SD about the mean.

**Transfer Trials**



**Fig. 4.** Transfer trials. Mean completion time in seconds as a function of trial and reversal condition in the test phase of the experiment. Each point represents an average of >10 participants.

from the first transfer trial. The first comparison evaluated the effects of changing only the visual orientation (*camera reversal*) relative to changing only the motor orientation (*surgeon reversal*). There was no significant difference between the mean for the *camera reversal* condition (502 sec) and the *surgeon reversal* condition (431 sec) ( $t(18) = 1.31$ ).

A second test, compared the *camera-surgeon reversal* condition with the *dish reversal* condition. These conditions change the visual appearance of the task and the topology of the actions, but they maintain a constant relation between the instruments and camera (i.e., the link between virtual hands and eyes is the same in transfer and training). There

was no significant difference between the mean for the *camera-surgeon reversal* (299 sec) and the *dish reversal* (253 sec).

A third comparison evaluated the effects of breaking the virtual hand-eye links created in the training phase (*camera reversal* and *surgeon reversal* conditions) versus conditions that changed the visual and motor orientation but maintained a constant linking between the virtual hands and eyes (*surgeon-camera reversal* and *dish reversal*). There was a significant difference between the combined mean over the *camera reversal* and *surgeon reversal* conditions (467 sec) and the combined mean of the two other reversal conditions (277 sec) ( $t(38) = 2.97$ ;  $p = .005$ ).

## Discussion

The learning curves (Figs. 3 and 4) show that participants adapted very quickly to the sensory rearrangements produced by the intervening camera and display. A trial involved picking up and transferring eight chips from one dish to another. This took, on average, >6 min to accomplish at the start of training. However, within 10 iterations, participants were accomplishing the same task in just over 2 min. When a second rearrangement occurred (due to the reversal manipulation), even when this rearrangement led to regression to levels of performance comparable to the start of training, participants quickly improved with practice. This is consistent with previous findings on adaptation to sensory rearrangement [9, 10]—humans are capable of adapting to almost any stable rearrangement.

The most important result, however, was the differential effects of the reversal conditions. The *camera reversal* and *surgeon reversal* conditions resulted in large increases in time to perform the tasks. With these reversals, participants regressed to performance levels that were comparable (in fact, slightly worse) to performance at the very beginning of training. This illustrates negative transfer from training to test phases. However, in the *camera-surgeon reversal* condition, in which both camera and surgeon positions were changed, there were significantly smaller increases in completion time.

This pattern of results suggests that skilled performance does not depend on a consistent visual orientation, as would result in the *surgeon reversal* condition, where the camera orientation did not change from training to test phases. Further, skill does not appear to depend on a consistent motor orientation, as would result in the *camera reversal* condition, where the orientation of the instruments did not change from training to test phases.

However, a consistent mapping from camera to motor orientation does appear to be important. In the *surgeon-camera reversal* condition, both the visual and motor orientations changed from training to test phases. However, they changed in a way that maintained a consistent mapping between the camera and instruments. This was also true for the *dish reversal* condition. For both of these conditions, skills that were developed in the training phase transferred to the test phase.

These results suggest that neither the retinotopic mapping (static video image) nor the movement topology (stable

motor orientation) were critical in and of themselves. Rather, the dynamic mapping over these two components was critical for transfer of skill. This finding suggests that adaptation cannot be understood as a recalibration of one component (e.g., visual or motor) that then directs the other components. Rather, adaptation reflects a dynamic reorganization to a new level of stability to which each component in the perception-action cycle contributes [7]. In sensory rearrangement experiments, it is not the changes on the retina that are critical but the changes over the coordination between vision and action.

There are several implications of this research for endoscopic surgery. First, the data show very clearly that moving the camera relative to the instruments can be temporarily very disruptive to the coordination. This conclusion is supported by observations in the operating room, where surgeons find that unexpected camera movements can be very disorienting. Good camera operators generally keep the camera in a consistent orientation, and devices are now available that will lock the camera into a fixed position for the surgeon. Also, surgeons can be disoriented by changing positions relative to the camera (e.g., an attending physician may have to move from the assisting position on one side of the patient to the primary surgeon position on the other side of the patient when difficulties arise that are beyond the skill of a resident surgeon). Surgeons talk about having to “turn their brain around” in order to adjust to the new perspective. Thus, it is good policy to keep the dynamic relation between camera and instruments as consistent as possible during surgery to maintain good coordination.

A second implication for surgery relates to simulator design. These results suggest that the dynamic coupling between instrument and camera are more important to transfer than the static appearance of the display or the topology of the movements. It seems that fidelity in terms of the appearance on the display (i.e., reproducing high-resolution images of the anatomy) or topology of the movements is not necessary for some degree of positive transfer. This is important, because it is currently very difficult and expensive to produce high-fidelity graphic displays of the internal anatomy. However, it is relatively easy to arrange an endoscope and endoscopic instruments into a dynamic configuration that preserves the relations found in surgery. Thus, it may be possible for surgeons to develop the coordination (or, in other words, experience or orient to the space) without having to reproduce all the details of the anatomy.

Finally, the results testify to the ability of surgeons (and people in general) to adapt and perform skillfully with perception-action couplings that are very different from the normal hand-eye couplings of everyday experience. Although any change to the relation between hand and eye will temporarily disrupt coordination, people can adapt to any stable (consistent) relation between the hand and eye. Thus, the rearrangements and constraints of micro- and telesurgery are not insurmountable obstacles to skilled surgical performance. As long as the dynamic mappings are consistent, we can expect surgeons to climb the learning curve to skilled coordination and control.

## References

1. Dominguez, C (1997) First, do no harm: expertise and metacognition in laparoscopic surgery [Dissertation]. Wright State Univ., Dayton, OH, 269 pp
2. Dominguez CO, Flach JM, Lake P, McKellar D, Dunn M (in press) The conversion decision: knowing your limits and limiting your risks. In: Smith K, Shanteau J, Johnson P (eds) Psychological explorations of competent decision making. Cambridge Univ. Press, Cambridge, England
3. Gaster B (1993) The learning curve. *JAMA* 270: 1280
4. Holden JG (1998) Hysteresis in hand-eye coordination. Proceedings of the 4th annual symposium on human interaction with complex systems. IEEE Computer Society, Los Alamitos, CA
5. Munsen JL, Sanders LE (1994) Cholecystectomy: open cholecystectomy revisited. *Surg Clin N Am* 74: 741–754
6. Satava RM (1993) Surgery 2001: a technologic framework for the future. *Surg Endosc* 7: 111–113
7. Satava RM, Ellis SR (1994) Human interface technology: an essential tool for the modern surgeon. *Surg Endosc* 8: 817–820
8. Thelen E, Smith LB (1994) A dynamical systems approach to the development of cognition and action. MIT Press, Cambridge, MA
9. Welch RB (1978) Perceptual modification: adaptation to altered sensory environments. Academic Press, New York
10. Welch RB (1986) Adaptation of space perception. In: Boff KR, Kaufman L, Thomas JP (eds) Handbook of perception and human performance; vol 1, sensory processes and perception. pp 24.1–24.45, Wiley, New York