APPLIED COGNITIVE PSYCHOLOGY
Appl. Cognit. Psychol. **20**: 487–503 (2006)

(www.interscience.wiley.com) DOI: 10.1002/acp.1198

Learning a Spatial Skill for Surgery: How the Contributions of Abilities Change With Practice

MADELEINE KEEHNER[1]*, YVONNE LIPPA[2],
DANIEL R. MONTELLO[1], FRANK TENDICK[3]
and MARY HEGARTY[1]

[1]*University of California, Santa Barbara, USA*
[2]*Miami University, Ohio, USA*
[3]*University of California, San Francisco, USA*

SUMMARY

We examined changes in performance as people learned to use an angled laparoscope, a challenging spatial skill that must be mastered by surgeons who perform minimally invasive techniques. In Experiment 1, novices took tests of spatial and general reasoning ability, and then learned to operate an angled laparoscope, simulated in a virtual environment, over 12 learning sessions. Initial performance showed considerable variability among learners, with performance related to general and spatial abilities. As learning progressed, interindividual variability diminished and all learners attained proficiency; the correlation with general ability diminished but the correlation with spatial ability remained significant. In Experiment 2, performance by highly experienced surgeons on the simulation was excellent from the first session, confirming its ecological validity. The findings contribute to theories of skill acquisition. They also have practical implications for the selection of surgeons and for the potential use of virtual environments in surgical training. Copyright © 2006 John Wiley & Sons, Ltd.

The question of how cognitive abilities predict task performance at different stages of learning has been a classic issue in psychology (Ackerman, 1987; Reed, 1931; Thorndike, 1908). Abilities such as spatial visualization are correlated with many real-world tasks, such as piloting aircraft and mechanical reasoning. But does this relationship change over time, as skill is acquired?

One domain in which this issue is currently relevant is surgery. Skill learning is an important aspect of surgical training, in which medical students must learn to perform complex procedures. Research has indicated that spatial ability is a key predictor of performance in surgical procedures. Scores on tests of spatial relations (Bennet, Seashore, & Wesman, 1981), which measure the ability to recognize a correspondence between two objects at differing spatial orientations, have been shown to predict the performance of surgical trainees on a microsurgery task (Murdoch, Bainbridge, Fisher, & Webster, 1994).

---

*Correspondence to: Madeleine Keehner, School of Psychology, Curtin University of Technology, GPO Box U 1987, Perth 6845, Western Australia. E-mail: m.keehner@curtin.edu.au

Contract/grant sponsor: National Science Foundation; contract/grant numbers: 9980122; 0313237.

Similarly, several studies have demonstrated a correlation between surgical skills and performance on the embedded figures test (Witkin, Oltman, & Karp, 1971), which measures the ability to distinguish 2D shapes within complex visual backgrounds (Gibbons, Baker, & Skinner, 1986; Gibbons, Gudas, & Gibbons, 1983; Steele, Walder, & Herbert, 1992). Using a comprehensive battery of psychometric tests, Schueneman and colleagues identified a cluster of cognitive functions underlying a factor that they labelled *complex visuo-spatial organization*. Surgical residents' scores on this factor proved to be a more reliable predictor of operative skill than either psychomotor ability or personality factors (Schueneman, Pickleman, Hesslein, & Freeark, 1984).

The relationship of spatial abilities to surgical performance has caused debate in medical education about whether spatial ability should be used as a factor in selecting surgery trainees (Gilligan, Welsh, Watts, & Treasure, 1999; Graham & Deary, 1991). In fact, pre-screening for spatial abilities is the current practice in another medical specialty, dentistry (Dailey, 1994; Graham, 1972). This debate hinges on the question of whether the relationship between spatial ability and surgical skill is transient (associated only with the early stages of learning) or enduring (continuing to affect performance even after much experience has been gained), and clearly has important implications for recruitment and training. If initial differences in performance endure or increase with practice, it may be important to select people for surgical training on the basis of their spatial abilities, whereas if individuals of all abilities can attain the same level of skill after practice, selection is not necessary.

Previous studies of the relationship between ability and skill acquisition in a variety of domains indicate that with practice and consistent feedback on a task, individual differences in performance diminish (Ackerman, 1986, 1987, 1988; Adams, 1987). Moreover, initial relationships between cognitive abilities and performance decrease with practice on tasks with consistent components. These results are interpreted in terms of theories of skill acquisition, which assume that skill learning passes from a cognitive or declarative stage through an associative stage to an autonomous or procedural stage (Anderson, 1982; Fitts, 1964; James, 1890; Shiffrin & Schneider, 1977). As skill becomes proceduralized, performance shifts from being controlled by attention-demanding cognitive processes to being automatic. According to this view, there should be a correlation of cognitive abilities with performance only at the early stages of skill acquisition, i.e. during the cognitive or declarative stage. This framework suggests that it is important to examine the correlation between abilities and performance in surgery at different stages of skill acquisition, in order to evaluate whether surgeons should be pre-screened for spatial ability. If the correlation is only present during the early stages of learning (as predicted by theories of skill acquisition), pre-screening is not indicated. However, if the relationship is enduring, then spatial ability may indeed be an important long-term predictor of surgical performance.

This study addresses this question in the context of minimally invasive surgery. This type of surgery is a recent and important innovation in which operative procedures are conducted from just outside the body, avoiding the need for a large incision. In *laparoscopy* (minimally invasive surgery of the abdomen), several small openings or ports are created in the patient's abdominal wall, each large enough to accommodate a narrow tube or cannula. Through one of these, a *laparoscope* (a camera mounted on a long tube with internal optics) is passed into the abdominal cavity. Long-handled surgical instruments are inserted through the other ports. The laparoscope transmits a video image from inside the abdomen to a monitor, which the surgeon uses to guide activities and manoeuvre instruments at the operative site. These techniques have significant advantages for patients, such as lower post-operative morbidity rates and faster recovery times.

Although clinically beneficial for the patient, these methods pose a substantial challenge to the surgeon. For one thing, the available perceptual information is highly restricted. The operative space cannot be viewed directly, nor can it be explored with the hands. Although the operative site is visible via the laparoscope, the 2D image provided on the monitor lacks binocular depth cues and thus differs substantively from that of open surgery. A particular challenge that we focus on in this paper comes from the laparoscope's angled lens. If surgeons used a laparoscope with a straight objective lens, the only areas visible would be those located directly in front of it. Insertion through the abdominal wall creates a fulcrum that limits the laparoscope's range of possible motion, so a straight lens would provide an overly restricted range of viewing perspectives. For many procedures, this limitation is overcome with an objective lens fitted at an angle (e.g. 45°) with respect to the laparoscope's longitudinal axis (see Figure 1). This expands the field of view considerably and allows the surgeon to look underneath, above, and partly around internal structures as the scope is rotated. However, it also alters the relationship between perceptual and motor events.

Rotating an angled laparoscope about its longitudinal axis produces highly unusual, non-intuitive changes to the available view. Imagine holding a laparoscope as in Figure 1, where the task is to look into the upper box. To bring the opening into view, a laparoscope with a *straight* lens would need to be *pitched up*. By contrast, an *angled* laparoscope has to

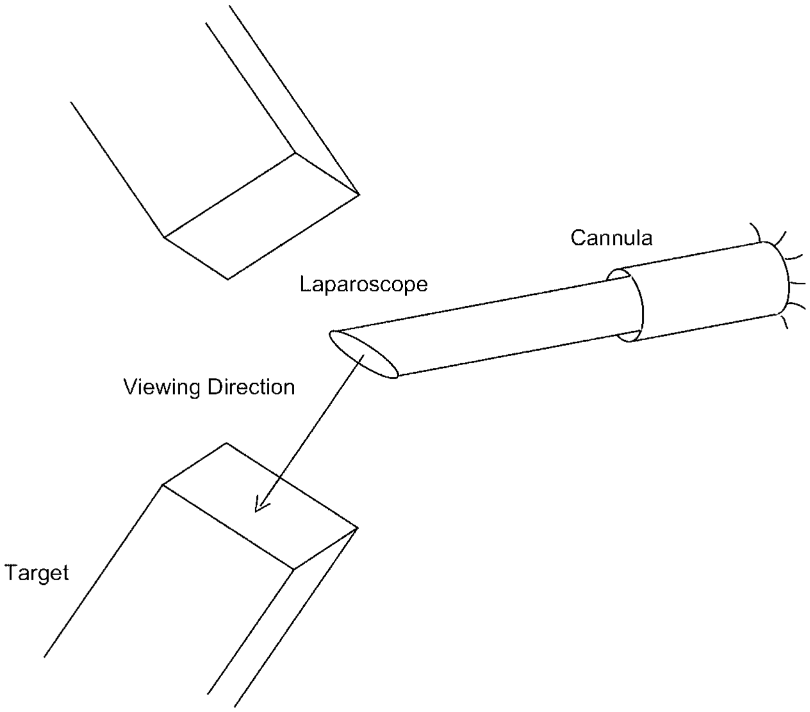

Figure 1. Illustration of an angled laparoscope. The laparoscope is inserted through a cannula, which constrains its motion to four degrees of freedom. The objective lens is angled with respect to the longitudinal axis of the laparoscope

be *rotated* through 180° to bring the lens into an upwards-facing orientation. Surgeons new to angled laparoscopy must master this novel relationship.

The constraints of using an angled laparoscope may be one reason why cognitive abilities, and spatial ability in particular, are important for performance under laparoscopic conditions. In a correlational study with novice laparoscopic surgeons, Keehner et al. (2004) found a significant relationship ($r = 0.4$) between surgical skills rated *in vivo* by experts and scores on the *Paper Folding Test* (Ekstrom, French, Harman, & Dermen, 1976), which measures complex spatial visualization abilities. Similarly, Wanzel et al. (2003) found correlations of 0.41 to 0.73 between surgical skill ratings and scores on spatial tests among surgical novices. Risucci, Geiss, Gellman, Pinard, & Rosser (2001) found a similar relationship among surgeons attending a laparoscopic skills training course. They reported significant correlations ranging from 0.2 to 0.5 between skill in simulator-based laparoscopic dexterity drills and suturing trials, and spatial subtests from the *Cognitive Laterality Battery* (Gordon, 1986) including mental rotation, form completion, and touching blocks. In a study with undergraduates, using a virtual reality simulation and abstract targets to assess the specific skill of controlling a laparoscope, Eyal & Tendick (2001) found significant correlations, ranging from 0.4 to 0.6, with tests of spatial ability.

Most of the research in this field is cross-sectional and much of the data comes from relatively inexperienced practitioners, such as surgical trainees and residents, so it is not known how experience affects this relationship. The primary purpose of this study was to examine the relationship between cognitive abilities and acquisition of a surgical skill in a longitudinal study. To fulfil this purpose, we had to focus on a specific surgical skill that would be learnable and measurable in the context of a laboratory study. We chose the skill of manipulating an angled laparoscope, as this has been identified as a particularly challenging skill in surgery (Tendick, et al., 2000).

The first question we addressed in this study was whether all individuals would be able to acquire the skill. This is related to the question of whether initial individual differences in performance increase, decrease or remain stable with practice. Based on previous studies of skill acquisition (Ackerman, 1987), we predicted that all individuals would acquire the skill and that individual differences would diminish with practice.

The second question concerned the relationship between cognitive abilities and performance at different points on the learning curve. Since manipulation of an angled laparoscope is a spatially demanding skill, we predicted that spatial ability would be correlated with initial performance of this skill. We also assessed non-verbal reasoning as a more general measure of cognitive ability. This should allow us to establish how much of the relationship between spatial ability and performance is due to content ability (i.e. facility in manipulating spatial information) versus more general cognitive ability.

Prevailing theories of skill acquisition (Ackerman, 1988; Shiffrin & Schneider, 1977) argue for a diminishing role for cognitive processes as skills become increasingly automatic. Thus, it has often been shown that the predictive power of cognitive abilities (including domain-general and content-based abilities) diminishes with practice on a task (Ackerman, 1988; Fleishman & Rich, 1963). However, there are exceptions to this rule. Using a longitudinal learning paradigm and a simulated air traffic control task, Ackerman (1992) demonstrated that the correlation of performance with reasoning ability remained stable over trials and the correlation with spatial ability actually increased. Ackerman interpreted this finding in terms of a distinction between tasks with consistent and inconsistent information-processing demands. In most early studies, the tasks had consistent components so that they became automatic with practice, whereas the air-traffic control task had inconsistent components, so that processing did not become automatic and cognitive abilities remained predictive of performance. In a later study, Ackerman and Cianciolo (2002) found that another important predictor of ability-performance relations was the content of a task, such that if a task is highly spatial, for example, the correlation with spatial ability will remain high after practice. Content was found to be even more predictive than consistency.

The laparoscope task used in the present research is an example of a consistent task according to Ackerman's framework, but its content is highly spatial. Because these factors are known to affect ability-performance correlations in opposite directions, it is difficult to make *a priori* predictions about the relationship between cognitive abilities and performance at different points in the learning curve. If consistency is the more important factor, we might expect the correlations of both general ability and spatial ability to decrease with practice. On the other hand, if the spatial content of the task is the more important factor, we might expect the correlation with spatial ability to remain high.

The third question was whether this surgical skill could be acquired using a virtual reality simulation. This is a critical issue because the dominant method for surgical training is currently an apprenticeship model, in which trainees practice their skills on real patients, supported by more experienced practitioners. If surgical skills can be trained using virtual reality this will avoid obvious risks to patients and provide a more flexible learning environment, without the need for an instructor to be present. An advantage of using a virtual reality simulation in the present study was that it allowed us to measure the acquisition of skills by individuals with no prior experience of surgery. We used a task that simulated the novel perceptual-motor relationships found in laparoscopic surgery, but required no surgical expertise or anatomical knowledge. In Experiment 1 (with undergraduates), we measured acquisition of this skill and individual differences in cognitive abilities, and examined their correlation over a series of learning trials. In Experiment 2, we administered our simulation task to a small group of experienced laparoscopic surgeons, to verify that manipulation of the angled laparoscope in our simulation depended on the same skills as it does in surgical practice.

EXPERIMENT 1

Method

Participants

Forty-four students from the University of California, Santa Barbara were recruited using advertisements on campus, and were paid for participation in the study. None was enrolled in medical education programmes or had any prior experience with surgery or with the experimental apparatus. All 44 students participated in the pre-learning session, which involved the administration of psychometric tests (18 males, 26 females; mean age 22.4 yrs, *SD* 3.9 yrs). Thirty-six of those participants embarked on the learning phase of the study (17 males, 19 females; mean age 22.4 yrs, *SD* 4.0 yrs), and 22 completed all 12 learning sessions plus one transfer session (11 males, 11 females; mean age 23.2 yrs, *SD* 3.6 yrs). The loss of participants was due to technical problems with the apparatus, or (in a few cases) scheduling problems or motion sickness. The ability profiles of the participants who completed the study did not differ substantially from those who began (see Table 1), indicating that the dropout did not reflect any systematic biasing or self-selection process.

Table 1. Experiment 1 participants' mean scores on psychometric tests

	All participants tested ($N = 44$)	Participants who completed 12 learning sessions ($n = 22$)
Abstract reasoning	27.0 (7.5)	26.8 (7.4)
Mental rotation	34.9 (18.5)	38.9 (19.4)
Visualization of views	11.5 (7.8)	12.8 (8.3)

Values in parentheses indicate standard deviations.

(Independent-samples *t*-tests confirmed that the finishers and non-finishers did not differ significantly on any of the tests).

Psychometric tests

Prior to the start of training, the participants' spatial and general cognitive abilities were assessed as part of a battery of psychometric tests (other measures administered during this phase were not relevant to our hypotheses in this paper and will not be discussed further here). Two paper-and-pencil tests assessed spatial visualization ability. In the *Mental Rotation Test* (MRT; Vandenberg & Kuse, 1978), participants viewed a drawing of a 3D test item and decided which 2 of 4 other items were rotated versions of the test item (2 × 10 items, 3 minutes per section). In an adapted version of Guay's (1976) *Visualization of Views Test* (Eliot & Macfarlane-Smith, 1983), participants saw a depiction of a 3D object in the centre of a cube. The same object from a different viewpoint was depicted below the cube. The task was to indicate the corner of the cube from which the new view was taken (24 items, 8 minutes).

General reasoning ability was assessed via the *abstract reasoning* task of the Differential Aptitude Test battery (Bennet et al., 1981). The test comprised sequences of geometric figures with elements changing systematically according to a logical rule. The task was to infer the rule and choose from five possibilities the next element in the series (40 items, 20 minutes). The test items are similar to those in Raven's Standard Progressive Matrices (Raven, Court, & Raven, 1977), which is often used as an index of general fluid intelligence (e.g. Engle, Kane, & Tuholski, 1999).

Materials and apparatus

The *angled laparoscope* task was presented via a virtual reality system for training laparoscopic surgical skills (Tendick et al., 2000). The simulation and data collection were controlled via custom software written in C and OpenGL on a Linux operating system. The virtual viewing environment was rendered on a Princeton Graphic System Ultra 20 monitor. The tip of a laparoscope (43 cm long) was attached to a modified PHANToM haptic interface (Sensable Technologies), which measured its motion in four degrees of freedom: translations in the vertical, horizontal, and depth axes, and rotations about the laparoscope's longitudinal axis. The kinematics were identical to a real laparoscope and served as input to the simulation (see Figure 2, or for a more detailed description see Tendick et al., 2000).

Participants watched a visual display of a virtual environment, within which they maneuvred the laparoscope. The image showed a large open virtual box (39.5 cm × 25 cm × 21 cm), inside which five smaller virtual boxes (2 cm × 2 cm × 3 cm) hovered in different locations and orientations. Using these general constraints, we constructed six sets of trials that fulfilled different purposes. One set of *practice* trials (used to familiarize

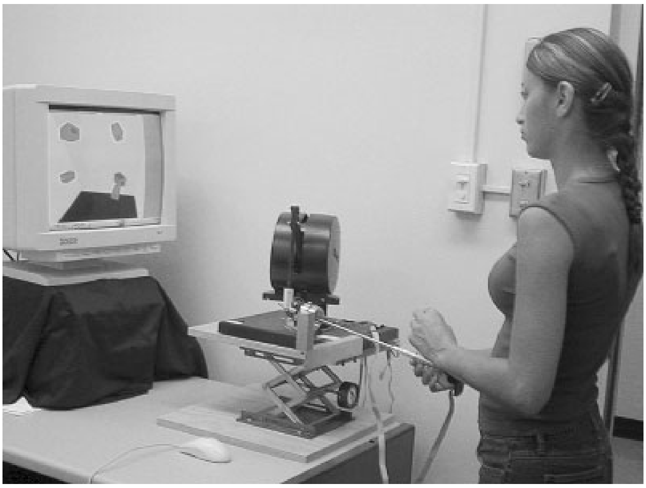

Figure 2. Virtual reality set-up in the simulated laparoscope task. See text for details

participants with the task) comprised five target boxes that were relatively easy to solve because they required laparoscope rotations that deviated no more than ± 59° from the starting position. Four sets of *learning* trials were more difficult; each set comprised five target boxes requiring rotations between ± 60° and ± 180°. A final *transfer* set of trials comprised five novel target boxes with the same difficulty parameters as the boxes in the learning trials. In all sets of trials, the five target boxes were distributed across space in a quasi-random configuration.

Procedure

The experiment was divided into one pre-learning session (in which the psychometric tests were administered) and 12 learning sessions (in which the laparoscope task was administered), scheduled over three weeks. In the first laparoscope learning session, participants were informed about the general purpose and procedure of the study. The experimenter described the angled lens without giving information that would suggest a specific strategy for its use, as this is the norm in real surgical training. Participants familiarized themselves with the laparoscopic workstation by completing the five simple practice trials. Thereafter, participants completed 12 learning sessions, each comprising 4 sets of 5 boxes each. In each session, the four sets of learning trials were presented in a random order, determined by a Latin square. The first learning session lasted up to 60 minutes, and subsequent learning sessions lasted between 30 minutes and 5 minutes (on average, time spent in the learning sessions decreased as learning progressed). Participants had no more than one learning session per day. Sessions were scheduled on consecutive days where possible, with a maximum of one day's break allowed per week (with the exception of the two intervening weekends, when no learning sessions occurred). In the final session, participants completed the transfer trials immediately following the last set of learning trials.

Each set began with the participant zooming out (by drawing the laparoscope back) to increase the field of view so that all five boxes were visible. When the participant was ready, the experimenter initiated the trial via a key press, and one target box turned from red to green. The participant's task was to manoeuvre the laparoscope to look directly into the target box, so that the inside of the box could be seen straight on. When the laparoscope had been positioned correctly, the colour of the box changed back to red, indicating successful completion of that target. A box in a different location then turned green, indicating that it was the next target. It was generally necessary to zoom out between targets to bring the next target into view. Participants proceeded until they had looked directly into all five boxes. The order of boxes was randomized in each set of trials. Participants were instructed to complete the task as quickly as possible. Response time was recorded automatically. The maximum time allowed per box was 3 minutes. If a box was not solved within this period, a time of 180 seconds was assigned to that trial.

Results

Participants reached the target within the time limit on 99.4% of trials. As the task forced accuracy on the vast majority of trials, our analyses were based on the dependent measure of solution time (time in seconds between onset and offset of the green colour for each box). For each session, we calculated mean latency over the twenty targets in the four learning sets. Unless otherwise stated, the analyses reported refer to the 22 participants who completed all 12 learning sessions, and the alpha level was 0.05.

First we examined how the acquisition of skill proceeded. Figure 3 shows individual latencies, group means and standard deviations. Learning advanced continuously, with the rate of improvement decreasing as practice increased. A repeated-measures ANOVA found a significant main effect of session, with time to solve a target decreasing with practice, $F(11, 231) = 9.86$, $p < 0.001$, partial $\eta^2 = 0.91$, and time to target decreased significantly between session 1 and session 12, $F(1, 21) = 59.31$, $p < 0.001$, partial $\eta^2 = 0.74$. Paired comparisons of group means from adjacent learning sessions (with alpha = 0.0045 to adjust for multiple contrasts) showed that participants' overall performance differed significantly

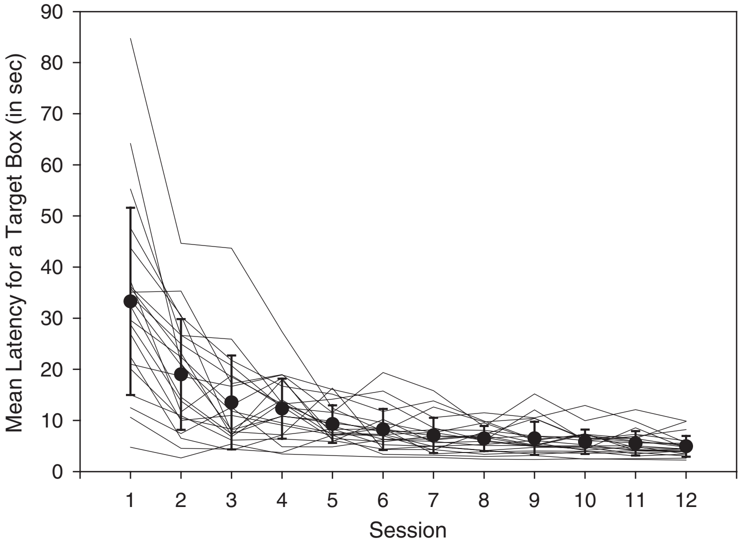

Figure 3. Performance on the simulated laparoscope task in Experiment 1. The lines depict individual learning curves obtained in the laparoscope task. The filled circles indicate mean latency of the group and the error bars show ±1 standard deviation

between session 1 and session 2 (t (21) = 5.87, $p < 0.001$; partial $\eta^2 = 0.62$), and between session 2 and session 3 (t (21) = 3.46, $p = 0.002$; partial $\eta^2 = 0.36$). After session 3, the difference between performance means from adjacent learning sessions ceased to reach significance at the adjusted alpha level. The standard deviation bars in Figure 3 show that practice attenuated the initial variability among participants. The group *SD* diminished 89% between session 1 and session 12. Mean time to target on session 1 was 33.29 seconds (*SD* = 18.32). By contrast, mean time to target on session 12 was 4.93 seconds (*SD* = 2.03).

To determine how well learning transferred to a novel set of targets, we compared performance in session 12 (*M* = 4.93 seconds, *SD* = 2.03) to performance in the transfer set (*M* = 5.70 seconds, *SD* = 2.61). There was no significant difference in performance (t (21) = −1.30, $p = 0.21$; partial $\eta^2 = 0.08$). This indicates that the participants had learned the skill of manipulating the laparoscope, and not simply the locations of the targets in the learning trials or how to move the laparoscope to these locations.

To examine the relationship between abstract reasoning ability and spatial ability, we created a correlation matrix based on all 44 initial participants (see Table 2). The correlation in parentheses is a partial correlation, which excludes variance arising from abstract reasoning ability (i.e. abstract reasoning test scores have been partialed out). Since both zero-order and partial correlations showed that mental rotation and visualization of views correlated strongly, we averaged *z*-scores from the MRT and visualization of views tests to produce an aggregate of the spatial ability measures, hereafter referred to as *spatial ability*. Consistent with previous research (e.g. Lohman, 1996; Marshalek, Lohman, & Snow, 1983), spatial ability shared some variance with abstract reasoning ability ($r = 0.55$). Using regression analysis, we calculated a new variable, referred to as *residual spatial*, equal to spatial ability with abstract reasoning ability partialed out (i.e. a measure of the variance in spatial visualization ability that is not shared with abstract reasoning ability).

Figure 4 shows the correlations of abstract reasoning ability, spatial ability and residual spatial with performance on the task, and plots the changing *r*-values over the course of the 12 learning sessions. Examination of the trends shown in Figure 4 indicates that initial performance was correlated with both the abstract reasoning and spatial ability measures. After the first few sessions, the contribution of abstract reasoning ability diminished. However, the correlation with spatial ability remained more stable. Figure 5 shows the performance means of higher and lower ability groups (determined by a median split) over the 12 learning sessions.[1] Whereas abstract reasoning ability differentiated performance only in the first few sessions, spatial ability continued to differentiate performance

Table 2. Correlations (*r* values) among psychometric tests for all participants in the pre-learning session (*N* = 44)

	Abstract reasoning	Mental rotation
Mental rotation	0.56*	
Visualization of views	0.45*	0.70*(0.60**)

*Correlation is significant at the $p < 0.01$ level (2-tailed).
**Correlation is significant at the $p < 0.001$ level (2-tailed).
Value in parentheses indicates partial correlation, controlling for abstract reasoning.

---

[1] Means (and *SD*s) for the higher-spatial and lower-spatial groups were 50.64 (12.72) and 27.09 (18.00), respectively on the Vandenberg Mental Rotation test and 20.17 (2.91) and 5.41 (3.91) on the Guay Visualization of Views test. Means (and *SD*s) for higher-and lower-abstract reasoning groups were 32.15 (4.28) and 21.45 (5.80) on the DAT Abstract Reasoning test.

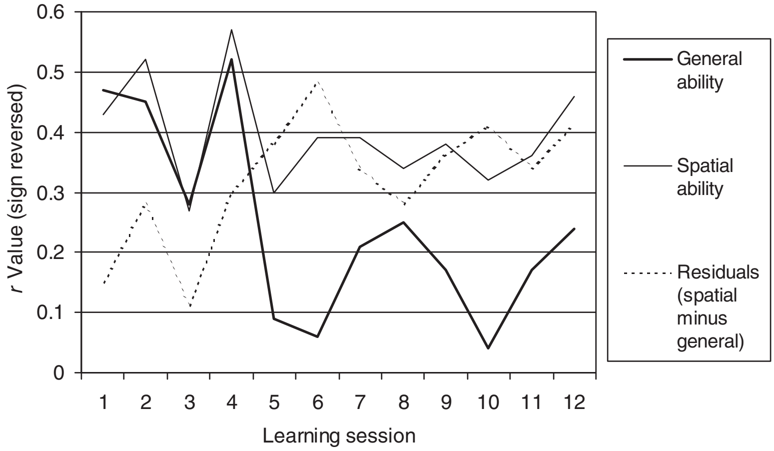

Figure 4. Changing correlations (*r*-values) in Experiment 1 between task performance and ability measures (general, spatial, residuals) over the course of 12 learning sessions. Correlations greater than +/−0.42 are significant at $p < 0.05$ for $N = 22$

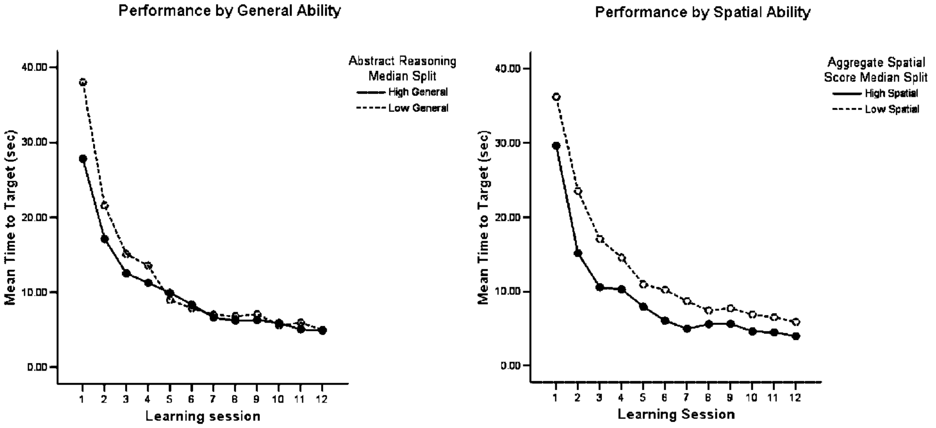

Figure 5. Performance means in Experiment 1 across learning trials of high and low general and spatial ability sub-groups, defined via median splits. Data points represent mean latencies

throughout learning. The residual spatial measure was comparatively unrelated to performance on the early trials, but became more strongly correlated with performance in later trials (see Figure 4).

In order to examine these relationships more closely, we averaged performance on the first three sessions (when participants were still learning) and on the last three sessions (by which time they had reached asymptote performance). We then correlated these *initial* and *final* performance averages with abstract reasoning ability, spatial ability and the residual spatial measure. We were testing two competing hypotheses: 1) that the correlations with both abstract reasoning and spatial abilities would diminish with practice; 2) that the correlation with abstract reasoning ability would diminish, but the correlation with spatial ability would remain significant, indicating that this factor continues to be important as skill is acquired. Table 3 shows that abstract reasoning ability correlated with initial

Table 3. Correlations of ability measures with initial and final performance ($N = 22$)

	Initial performance (mean of trials 1,2,3)	Final performance (mean of trials 10,11,12)
Abstract reasoning ability	-0.46 ($p = 0.03$)	-0.16 ($p = 0.48$)
Spatial ability (composite measure)	-0.46 ($p = 0.03$)	-0.42 ($p = 0.05$)
Spatial-residual (spatial ability controlled for abstract reasoning)	-0.19 ($p = 0.18$)	-0.43 ($p = 0.04$)

(Note that because the measure of performance is latency, smaller values indicate better performance, so a negative correlation is predicted.)

performance, but was not predictive of final performance. By contrast, spatial ability correlated with both initial and final performance. The residual spatial measure showed the opposite pattern to abstract reasoning ability, correlating with final performance but not with initial performance. The correlation with spatial ability was similar to abstract reasoning ability on the early sessions but similar to the residual spatial measure on later sessions. We compared initial and final correlations using a test for non-independent samples (Steiger, 1980; Williams, 1959), to establish whether the correlation changed significantly between early and late sessions. For abstract reasoning ability the correlation with performance was marginally stronger at the beginning of learning than at the end of learning ($t\,(19) = 1.67, p = 0.05$, one-tailed). For the residual spatial measure, although the correlation at the end of learning was somewhat greater than at the beginning, the difference was not significant ($t\ (19) = 1.29$, $p = 0.10$, one-tailed).[2]

Discussion

Over the course of twelve learning sessions, the performance of all participants reached asymptote, and indicated that all individuals learned to use the angled laparoscope to a similar level of proficiency. Even the participant who experienced the most difficulty at the beginning, performing well above the group's mean, had joined the group distribution by the fifth session. The steady reduction in the group *SD* over trials indicates that practice attenuated the variability within the group. The form of the data indicates a typical learning curve, with the rate of improvement decreasing as practice increased. Performance on the transfer set of targets indicated that participants had learned a general skill, and not merely to manipulate the laparoscope to the 20 targets on which they had practiced.

Consistent with theories of skill acquisition (Ackerman, 1988; Shiffrin & Schneider, 1977), the correlation of performance with general reasoning ability diminished after the first few sessions. By contrast, the correlation with spatial ability persisted even after the group variance had diminished. The correlations with the residual spatial measure showed an opposite trend to the general ability correlations, being weakest early in learning and becoming stronger with practice. These findings suggest that content-based cognitive abilities may remain important after skill has been acquired. They also argue against the

---

[2] The correlation of performance on the transfer trials was -0.39 with spatial ability, -0.34 with abstract reasoning ability, and -0.22 with the spatial residual. Because these correlations were based on only five target boxes, as compared to 60 target boxes for the initial and final performance correlations, it would be premature to draw conclusions on the basis of these data.

idea that the reduction in correlation between general ability and performance merely reflects a statistical artefact of variance attenuation.

EXPERIMENT 2

The purpose of the second experiment was to establish the construct validity of our virtual laparoscopic task. In order to do this, we assessed the performance of experienced laparoscopic surgeons on the simulation. If our simulation requires the same skill as manipulating an angled laparoscope in surgery, experienced surgeons' initial performance on the simulation should be similar to that of highly-practiced novices. On the other hand, if performance in the simulation depends on different skills to those used in real surgery, performance of surgeons on the first trial in our simulation should be more like the performance of novices at the beginning of training.

Method

Participants

Six surgery fellows specializing in laparoscopic techniques were recruited from the University of California, San Francisco (UCSF) Department of Surgery (5 males, 1 female). All were experienced laparoscopic surgeons, and were regularly employed as trainers in the UCSF Advanced Laparoscopic Skills course. Their mean years of surgery experience was 7.5 ($SD = 1.1$), including 4.6 years of laparoscopic surgery ($SD = 2.3$); the mean number of laparoscopic procedures they performed per week at time of testing was 4.2 ($SD = 3.3$). Four of the surgeons had no prior experience with the simulation; two had used the simulation briefly at least two years earlier. The fellows' mean score on Guay's (1976) *Visualization of Views Test* was 17.7 ($SD = 5.8$), which was somewhat higher than the novices in Experiment 1 (mean = 12.8, $SD = 8.3$).

Materials, apparatus, and procedure

The set-up of the laparoscopic simulation task was identical to that in Experiment 1. Due to time constraints, only Guay's (1976) *Visualization of Views Test* (Eliot & Macfarlane-Smith, 1983) was administered, using a procedure identical to Experiment 1. A single session of the laparoscopic simulation task was administered using instructions and procedure identical to Experiment 1. As in Experiment 1, the practice trials were followed by the 20 learning trials. Due to the fellows' time constraints, only four of the six (all males) took the spatial visualization test.

Results

The surgery fellows' mean response latency was 5.0 seconds ($SD = 1.5$). To provide a fair comparison, we selected a matched subset of novices from Experiment 1 ($N = 13$), who had comparable Guay spatial test scores (fellows' and matched novices' Guay means = 17.6 and 18.8; *SD*s = 5.8 and 4.2). Using an independent samples *t*-test we compared the fellows' response latencies to those of the matched novices on their first and last sessions. The surgery fellows were significantly faster ($M = 5.0$ seconds, $SD = 1.5$) than the matched novices on their first session ($M = 27.0$ seconds, $SD = 14.5$; $t\ (17) = 3.65$, $p = 0.002$; partial $\eta^2 = 0.44$), but did not differ significantly from the novices' final performance, on session 12 ($M = 4.1$ seconds, $SD = 1.1$; $t\ (17) = -1.37$, $p = 0.19$; partial $\eta^2 = 0.10$). The variance in performance among the fellows was substantially smaller than among the matched novices on session 1 ($SDs = 1.5$ versus 14.5), but by session 12 the variance among the matched novices had dropped to a level similar to that among the fellows ($SDs = 1.1$ and 1.5, respectively).

We did the same set of comparisons between the fellows' performance and that of the higher- and lower-spatial ability groups of undergraduates (determined by a median split in Experiment 1) whose mean response latencies are shown in Figure 5. The fellows' performance was better than that of both groups on the first session ($t\ (15) > 3.5$, $p < .01$ for both comparisons) but was not significantly different from either group on session 12 ($t\ (15) < 1.6$, $p > 0.1$ in both cases).

Discussion

In a single session, the experienced laparoscopic surgery fellows performed significantly better than novices on their first session but no better than the undergraduates after 12 learning sessions. These results held even with a subset of novices who were matched on spatial abilities, so the fellows' superior initial performance could not have been an artefact of higher spatial abilities. Rather, it appears that their surgical experience transferred to our task, suggesting that the simulation called for at least some skills similar to those acquired through laparoscopic practice. We acknowledge that the number of fellows was rather small for performing tests of significance, but even as a purely qualitative comparison their performance lends credibility to the construct validity of this task.

The variance in response latencies among the fellows was much smaller than among the novices on their first session (but comparable to the novices after 12 sessions). This finding supports the hypothesis that the fellows' experience with laparoscopic surgery had led to an attenuation of interindividual variability in performance on the simulation.

GENERAL DISCUSSION

The first question addressed in this study was whether individuals of all abilities would be able to acquire the skill of manipulating an angled laparoscope. The results indicate that, with practice, all learners acquired proficiency. Even individuals with relatively poor initial performance could learn the novel perceptual-motor relationships required for the task. That is, experience attenuated individual differences. Furthermore, participants showed no drop in performance when administered the transfer trials indicating that they had learned the skill of manipulating the laparoscope, rather than how to move it to the specific targets used in the learning sessions.

Our second purpose was to assess the relationship between cognitive abilities and performance at different points on the learning curve using a longitudinal study. We contrasted two possible accounts of this relationship. The first possibility is that the effects of all cognitive abilities on performance diminish as skill is acquired. The second hypothesis is that while the effects of general ability are reduced, content-based abilities (in this case spatial ability) remain correlated with performance even after skill is acquired

Consistent with prevailing theories of skill acquisition, Experiment 1 showed that the correlation between performance and general reasoning ability decreased as skill was acquired. Our task was a consistent task, so that participants should have developed automaticity with practice, and this would account for the diminishing correlation with general reasoning ability (Ackerman, 1988). By contrast, the correlation between performance and spatial ability did not diminish over time, but remained significant even after participants had reached asymptote, and the spatial residual component became more predictive of performance with practice. These results can be interpreted in terms of the content of the task, which has been shown to be more important than consistency (Ackerman & Cianciolo, 2002). Intuitively, our task is more demanding of spatial processing than previous tasks studied in this literature, including air traffic control tasks, which are substantially rule-based.

Why should spatial ability continue to be important with training in this task? One possible hypothesis is that performance on the task depends on learning the locations of the targets in space and spatial ability confers an advantage in this process. However, if this were true, we would not have observed transfer to the novel targets in the transfer set. Furthermore, some of our participants showed asymptote performance at the beginning and the surgery fellows showed this level of performance, despite having never seen the targets before.

An alternative explanation is that the changing correlations reflect a shift from a reliance on strategic or executive processes to spatial transformation processes. Recent theories have characterized differences in general fluid intelligence as reflecting the central executive component of working memory, which is involved in the control of attentional resources to maintain task goals, inhibit distracting information and schedule different sub-processes required to accomplish a complex task (Engle et al., 1999; Miyake & Shah, 1999). In our study, the early correlation between spatial ability and performance arose from variance shared by spatial and general ability, whereas the later correlation with spatial ability was due to the variance unique to spatial ability and *not* shared by general fluid intelligence. The shared variance between spatial and general fluid intelligence (which is also correlated with tests of executive working memory) has been interpreted as reflecting the *strategic* component of spatial tests, whereas the variance unique to spatial abilities has been interpreted as reflecting the ability to maintain and accurately transform spatial information (Miyake, Rettinger, Friedman, Shah, & Hegarty, 2001). Viewed within this framework, our data suggest that with practice, the strategic component of performance on the laparoscope task decreases, whereas its dependence on the ability to maintain and transform spatial information persists.

The present results inform the interpretation of previous studies showing correlations between spatial ability and surgical skill among relative novices (Keehner et al., 2004; Wanzel et al., 2003). It now appears possible that such relationships found early in training might in fact originate in a covarying *general* factor, rather than in spatial ability *per se*. It is therefore important that future research efforts assess the effects of general ability separately from domain specific abilities.

A limitation of this study is that because Experiment 1 participants were novices, it was not possible to examine transfer from the simulation to real surgery. This could be addressed in future studies examining the effects of practice in a simulation on the training of new surgeons. This would be an important step to further investigate the potential of virtual environments for surgical training.

Of course, use of the angled laparoscope is only one surgical skill, albeit a challenging one. In our previous study with practicing surgeons (Keehner et al., 2004) we assessed a much wider set of skills, such as handling and suturing tissues, planning movements of surgical instruments, and giving verbal direction to assistants. However in that study measures of performance were based on rating scales of expert observers, which had limited reliability. In future research it will be important to use the methodology employed in the present study to examine a wider set of surgical skills. This would involve developing virtual reality simulations of other aspects of surgery that could provide objective measures of performance along the learning curve. In fact, the development of virtual simulations is well under way for common procedures such as the laparoscopic cholecystectomy (removal of the gall bladder), and simulations such as these could be used in future controlled studies. In these studies, it will also be important to examine correlations with a wider set of abilities including cognitive, psychomotor and perceptual speed.

In conclusion, our findings have implications for those involved in recruiting and training of surgeons. To optimize performance in surgery, should the focus be on using ability tests to select the most able students, or the development of virtual environments to train all students to proficiency? The possibility of pre-screening to assess aptitude has been discussed frequently in the surgical profession (Graham & Deary, 1991). On the one hand, our data suggest that pre-screening may not be warranted, as even initially low-performing learners can acquire skill with the angled laparoscope with practice. On the other hand, our data suggest that high-ability learners reach proficiency faster and are still somewhat faster at the skill at the end of training. For educators considering the issue of whether to pre-screen, an important question is whether this variability in time to train and final performance has practical significance in the context of real surgery.

With respect to training, this study shows promise for the use of virtual environments for acquiring surgical skills. Currently, the dominant method for training surgical skills is the apprenticeship model, in which trainees practice their skills *in vivo* (i.e. on real patients), supported by experienced practitioners. The present study showed that at least some perceptual-motor aspects of laparoscopy can be learned in a virtual environment. In particular, virtual environments could be used to supplement in vivo training, and might be especially useful for low-spatial individuals who need more time to train. In fact, a comparison of Experiments 1 and 2 suggested that novice performance after 12 sessions with our simulation was equivalent to that of experienced laparoscopic surgeons, suggesting that the simulation develops at least some skills equivalent to those acquired through laparoscopic practice. Thus, virtual reality technology has potential for improving the effectiveness and safety of surgical training in the future.

ACKNOWLEDGEMENTS

This research was supported by National Science Foundation grants 9980122 and 0313237. We thank Gunnar Proppe and David Feygin for software and hardware assistance, Daniel Haun and Toru Ishikawa for help in pilot testing, and Tijana Vujovic, and Anoushka Yekazarian for testing participants.

REFERENCES

Ackerman, P. L. (1986). Individual differences in information processing: an investigation of intellectual abilities and task performance during practice. *Intelligence, 10*, 109–139.

Ackerman, P. L. (1987). Individual differences in skill learning: an integration of psychometric and information processing perspectives. *Psychological Bulletin, 102*, 3–27.

Ackerman, P. L. (1988). Determinants of individual differences during skill acquisition: cognitive abilities and information processing. *Journal of Experimental Psychology: General, 117*(3), 288–318.

Ackerman, P. L. (1992). Predicting individual differences in complex skill acquisition: dynamics of ability determinants. *Journal of Applied Psychology, 77*(5), 598–614.

Ackerman, P. L., & Cianciolo, A. T. (2002). Ability and task constraint determinants of complex task performance. *Journal of Experimental Psychology: Applied, 8*(3), 194–208.

Adams, J. A. (1987). Historical review and appraisal of research on the learning, retention, and transfer of human motor skills. *Psychological Bulletin, 101*(1), 41–74.

Anderson, J. R. (1982). Acquisition of cognitive skill. *Psychological Review, 89*, 369–406.

Bennet, G. K., Seashore, H. G., & Wesman, A. G. (1981). *Differential aptitude tests.* Sidcup, UK: The Psychological Corporation.

Dailey, R. J. (1994). *A re-analysis of the relationship of psychomotor and perceptual skills to performance in dental education*. Los Angeles: University of Southern California.

Ekstrom, R. B., French, J. W., Harman, H. H., & Dermen, D. (1976). Manual for kit of factor-referenced cognitive tests. Princeton, NJ: Educational Testing Service.

Eliot, J., & Macfarlane-Smith, I. (1983). *An international directory of spatial tests*. Windsor, Berks: Nfer-Nelson.

Engle, R. W., Kane, M. J., & Tuholski, S. W. (1999). Individual differences in working memory capacity and what they tell us about controlled attention, general fluid intelligence and functions of the prefrontal cortex. In P. Shah (Ed.), *Models of working memory: Mechanisms of active maintenance and executive control* (pp. 102–134). New York: Cambridge University Press.

Eyal, R., & Tendick, F. (2001). Spatial ability and learning the use of an angled laparoscope in a virtual environment. *Stud Health Technol Inform, 81*, 146–152.

Fitts, P. (1964). Perceptual-motor skill learning. In A. W. Melton (Ed.), *Categories of human learning* (pp. 243–285). New York: Academic Press.

Fleishman, E. A., & Rich, S. (1963). Role of kinesthetic and spatial-visual abilities in perceptual-motor learning. *Journal of Experimental Psychology, 66*(1), 6–11.

Gibbons, R. D., Baker, R. J., & Skinner, D. B. (1986). Field articulation testing: a predictor of technical skills in surgical residents. *Journal of Surgical Research, 41*, 53–57.

Gibbons, R. D., Gudas, C., & Gibbons, S. W. (1983). A study of the relationship between flexibility of closure and surgical skill. *Journal of the American Podiatry Association, 73*(1), 12–16.

Gilligan, J. H., Welsh, F. K., Watts, C., & Treasure, T. (1999). Square pegs in round holes: has psychometric testing a place in choosing a surgical career? A preliminary report of work in progress. *Ann R Coll Surg Engl, 81*(2), 73–79.

Gordon, H. (1986). The cognitive laterality battery: tests of specialized cognitive functions. *International Journal of Neuroscience, 29*, 223–244.

Graham. (1972). Substitution of perceptual-motor ability test for chalk carving in dental admissions testing program. *Journal of Dental Education, 36*, 9–14.

Graham, K. S., & Deary, I. J. (1991). A role for aptitude testing in surgery? *Journal of the Royal College of Surgeons of Edinburgh, 36*(April), 70–74.

James, W. (1890). *The principles of psychology* (Vol. 1). New York: Holt.

Keehner, M., Tendick, F., Meng, M., Anwar, H., Hegarty, M., Stoller, M., et al. (2004). Spatial ability, experience, and skill in laparoscopic surgery. *American Journal of Surgery, 188*(1), 71–75.

Lohman, D. F. (1996). Spatial ability and g. In P. Tapsfield (Ed.), *Human abilities: Their nature and measurement* (pp. 97–116). Mahwah, NJ: Lawrence Erlbaum Associates.

Marshalek, B., Lohman, D. F., & Snow, R. E. (1983). The complexity continuum in the radex and hierarchical models of intelligence. *Intelligence, 7*, 107–127.

Miyake, A., Rettinger, D. A., Friedman, N. P., Shah, P., & Hegarty, M. (2001). Visuospatial working memory, executive functioning and spatial abilities. How are they related? *Journal of Experimental Psychology: General, 130*, 621–640.

Miyake, A., & Shah, P. (Eds.). (1999). *Models of working memory: Mechanisms of active maintenance and executive control*. New York: Cambridge University Press.

Murdoch, J. R., Bainbridge, L. C., Fisher, S. G., & Webster, M. H. C. (1994). Can a simple test of visual-motor skill predict the performance of microsurgeons? *Journal of the Royal College of Surgeons Edinburgh, 39*, 150–152.

Raven, J. C., Court, J. H., & Raven, J. (1977). *Standard progressive matrices*. London: H. K. Lewis & Co.

Reed, H. B. (1931). The influence of training on changes in variability in achievement. *Psychological Monographs, 41*(2), Whole No. 185.

Risucci, D., Geiss, A., Gellman, L., Pinard, B., & Rosser, J. (2001). Surgeon-specific factors in the acquisition of laparoscopic surgical skills. *The American Journal of Surgery, 181*, 289–293.

Schueneman, A. L., Pickleman, J., Hesslein, R., & Freeark, R. J. (1984). Neuropsychologic predictors of operative skill among general surgery residents. *Surgery, 96*(2), 288–293.

Shiffrin, R. M., & Schneider, W. (1977). Controlled and automatic human information processing: II. Perceptual learning, automatic attending, and a general theory. *Psychological Review, 84*, 127–190.

Steele, R. J. C., Walder, C., & Herbert, M. (1992). Psychomotor testing and the ability to perform an anastamosis in junior surgical trainees. *British Journal of surgery, 79*(October), 1065–1067.

Steiger, J. H. (1980). Tests for comparing elements of a correlation matrix. *Psychological Bulletin, 87*, 245–251.

Tendick, F., Downes, M., Goktekin, T., Cavusoglu, M. C., Feygin, D., Wu, X., Eyal, R. Hegarty, M, & Way, L. W. (2000). A virtual environment testbed for training laparoscopic surgical skills. *Presence, 9*, 236–255.

Thorndike, E. L. (1908). The effect of practice in the case of a purely intellectual function. *American Journal of Psychology, 19*, 374–384.

Vandenberg, S. G., & Kuse, A. R. (1978). Mental rotations, a group test of three-dimensional spatial visualization. *Perceptual & Motor Skills, 47*, 599–604.

Wanzel, K. R., Hamstra, S. J., Caminiti, M. F., Anastakis, D. J., Grober, E. D., & Reznick, R. K. (2003). Visual-spatial ability correlates with efficiency of hand motion and successful surgical performance. *Surgery, 134*(5), 750–757.

Williams, E. J. (1959). The comparison of regression variables. *Journal of the Royal Statistical Society (Series B), 21*, 396–399.

Witkin, H. A., Oltman, P. K., & Karp, S. A. (1971). *A manual for the embedded figures test*. Palo Alto, CA: Consulting Psychologists' Press.