Education
Presented at the Academic Surgical Congress 2019

Screening surgical residents' laparoscopic skills using virtual reality tasks: Who needs more time in the sim lab?

Hossein Mohamadipanah, PhD[a,‡], Kenneth H. Perrone, MD[a,‡], Jay Nathwani, MD[b], Chembian Parthiban, PhD[b], Katherine Peterson, MS[b], Brett Wise, BS[a], Anna Garren, BS[b], Carla Pugh, MD, PhD, FACS[a,*]

[a] Department of Surgery, Stanford University School of Medicine, Stanford, CA
[b] Department of Surgery, School of Medicine and Public Health, University of Wisconsin-Madison, Madison, WI

ARTICLE INFO

ABSTRACT

Background: This study investigated the possibility of using virtual reality perceptual-motor tasks as a screening tool for laparoscopic ability. We hypothesized that perceptual-motor skills assessed using virtual reality will correlate with the quality of simulated laparoscopic ventral hernia repair.

Materials and Methods: Surgical residents ($N = 37$), performed 2 virtual reality perceptual-motor tasks: (1) force matching and (2) target tracking. Participants also performed a laparoscopic ventral hernia repair on a simulator and final product quality score, and endoscopic visualization errors were calculated. Correlational analysis was performed to assess the relationship between performance on virtual reality tasks and laparoscopic ventral hernia repair.

Results: Residents with poor performance on force matching in virtual reality—"peak deflection" ($r = -0.34$, $P < .05$) and "summation distance" ($r = -0.36$, $P < .05$)—had lower final product quality scores. Likewise, poor performance in virtual reality–based target tracking—"path length" ($r = -0.49$, $P < .05$) and "maximum distance" ($r = -0.37$, $P < .05$)—correlated with a lower final product quality score.

Conclusion: Our findings support the notion that virtual reality could be used as a screening tool for perceptual-motor skill. Trainees identified as having poor perceptual-motor skill can benefit from focused curricula, allowing them to hone personal areas of weakness and maximize technical skill.

© 2019 Elsevier Inc. All rights reserved.

Introduction

During the past several decades, general surgery has shifted from open approaches to minimally invasive approaches. Minimally invasive techniques, including endoscopy, thoracoscopy, and laparoscopy, have extended across multiple specialties in surgery. This shift from open to minimally invasive techniques can be seen in the dramatic increase in laparoscopic cases that surgery residents include in their operative logs.[1,2] As a result, laparoscopic skills are becoming increasingly critical for trainees to master.

Laparoscopy presents a unique set of technical challenges compared with open surgery. In addition to procedural knowledge, success depends on strong visual-spatial and psychomotor ability. Changes in duty hours mean that these components must be learned in a condensed amount of time.[3] The decrease in duty hours to 80 hours per week from 90 to 100 hours has effectively reduced residents' total clinical time by 6 months to 1 year when compared with 5 years of 100-hour work weeks.[4] This decrease in hours has resulted in a need to utilize clinical time more efficiently to build sufficient skill for independent practice. Simulation has been demonstrated as a safe and effective way to develop fundamental skills that transfer to the operating room.[5] Simulation provides an opportunity for deliberate practice of necessary skills and focus on areas of weakness.[6]

Trainees enter residency with different levels of innate ability that can have a significant effect on skill acquisition.[7] Two areas of innate ability that contribute to laparoscopic skill include psychomotor skill and visual-spatial ability. Psychomotor skill has been "defined as a stable and reliable link between perception of body and environment, and execution of goal-directed motor actions,
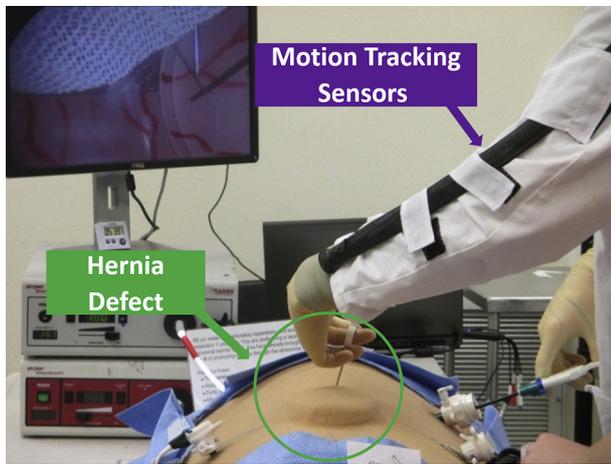
Fig 1. A participant performing an LVH repair on the simulator.

which is both consistent across repeated performances of the action, and can be flexibly adapted to changes in task constraints."[3] The visual-spatial skill refers to the ability to mentally reconstruct the 3D space from the 2D image of the operating field on a screen.[8] It is expected that training will allow for compensation of the many differences in innate ability.[9,10] Given the limited training time and differences in innate ability, there is a need to identify residents who would benefit from additional practice time in the simulation lab to build their fundamental laparoscopic skills.

This study investigated the possibility of establishing virtual reality (VR) perceptual-motor tasks as a screening tool for laparoscopic ability. This study uses laparoscopic ventral hernia (LVH) repair as an archetypical procedure. The goal was to investigate whether the quality of an LVH repair correlated with the psychomotor skills assessed in a VR setting. We hypothesized that perceptual-motor skills assessed using VR will correlate with the quality of an LVH repair.

Methods

Participants

A total of 37 participants were recruited from general surgery training programs across the United States. Participants completed a simulated partial LVH repair and 2 VR tasks. Basic demographic data including gender, training institution, post-graduate year and years of clinical experience were collected for each participant. Because the participants were analyzed based on performance, data on laparoscopic experience, including case logs, were not collected. By analyzing based on performance, each participant would be acting as his or her own control. A participant's performance on the LVH repair would be compared with his or her own performance on the 2 VR tasks. This study was approved by the University of Wisconsin-Madison Institutional Review Board, Madison, WI.

Laparoscopic ventral hernia repair

Participants were surgical residents who were not assumed to have extensive laparoscopic experience. As such, they were invited to perform a partial LVH repair on a simulator with previously established validity evidence.[11,12] The simulator contained a 10-cm × 10-cm hernia, 5-cm inferior to the umbilicus. The abdomen was assumed to be insufflated, 4 laparoscopic ports were placed, and 2 of 4 sets of the mesh anchoring sutures were pulled through the abdominal wall when participants began the simulation (Fig 1). All essential tools required to complete the LVH repair were provided, with the exception of cautery, and participants were allotted 20 minutes to complete the task.

Participants were provided with the following directions before beginning the repair:

> 69 y/o male s/p exploratory laparotomy and bowel resection 1 year ago. You are performing a laparoscopic incisional hernia repair. You have already brought up 2 of the 4 of anchoring sutures through the abdominal wall. You also know: The hernia defect is 10 × 10 cm, abdomen is insufflated, ports are in place, and mesh (16 × 16 cm, with anchoring sutures attached) has been inserted. Please complete the procedure by bringing up the remaining anchoring sutures and placing your first 5 tacks.

Performance during the repair was graded on 4 visualization errors: (1) inserting an instrument without camera visualization, (2) losing instrument in the field of view inside the abdomen, (3) failing to properly visualize the hernia defect with the camera, and (4) placing tacks without visualization of the tacking location.[11] Additionally, completed hernia repairs were given a final product quality score (FPQS), using a checklist with validity evidence that included the following 10 metrics (totaling 24 points):

- Suture sets brought through the abdominal wall (2),
- Suture sets tied down (4),
- Suture tails cut (4),
- Suture sets brought up with a bridge of peritoneum (2),
- Mesh placed flat (1),
- Alignment with previously placed suture sets (2),
- Mesh coverage of the hernia defect (2),
- Maintaining 2 cm of mesh beyond the hernia defect (1),
- Number of tacks placed to secure the mesh (5), and
- Minimal excess suture between mesh attachment and incision (1).

This resulted in a potential score from 0 to 24.[13,14]

Participants completed a survey both before and after data collection. They were asked to report their self-assessed confidence level and perceived difficulty in both visualizing the hernia defect and performing the entire procedure. Confidence levels were reported on a Likert scale from 1 to 5, with 1 being "Not Confident" and 5 being "Extremely Confident." Similarly, perceived procedure difficulty levels were reported from 1 to 5, with 1 being "Not Difficult" and 5 being "Extremely Difficult."

Virtual reality tasks

Participants completed two different VR task modules, with the goal of examining psychomotor skills using the Force Dimension Omega 6 haptic device (Nyon, Switzerland), a motorized haptic device that applies force feedback on the user's hand via a stylus (Fig 2).

Force matching in VR

The first module, initially developed by Parthiban et al,[15] tested participants' ability to counteract distracting forces to maintain the haptic interface stylus at the center of a motionless sphere (Fig 3). Participants underwent 20 trials, with the magnitude of force randomly ranging from 2 N to 6.5 N and randomly changing direction.[16] Correct stylus location was indicated by the color of the motionless sphere: *red* when outside the sphere and *blue* when inside (Figs 2 and 3).

The 2 following metrics were analyzed: (1) "Peak Deflection upon Force Release," defined as the maximum stylus deflection
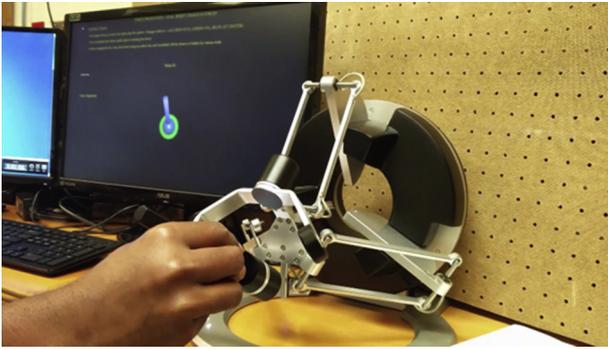
Fig 2. This motorized haptic device applies force feedback on the user's hand via the black stylus. The stylus is currently centered in the motionless sphere as indicated by a blue color of the inner sphere and a green color of the outer sphere.
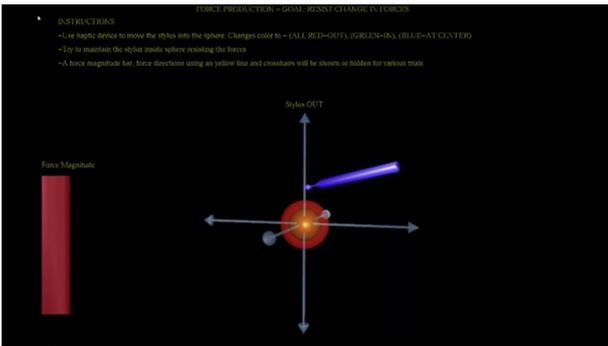

Fig 3. A screenshot of the force matching task module in which the virtual blue stylus should be kept at the sphere center while resisting the applied force from the haptic device. A red color on the motionless spheres indicates the styles is outside of the sphere.
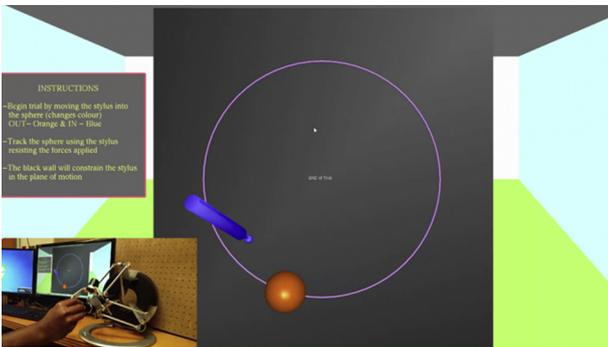

Fig 4. The target tracking module in which a participant tracks the rotating sphere with the tip of the stylus.

from the center of the sphere immediately after the force was released and (2) "Summation of Distances from Sphere," defined as the summation of distances between the tip of the stylus and the center of the sphere. Both metrics were averaged independently across 20 trials.

Target tracking in VR

The second module assessed participants' ability to track a moving sphere while encountering distracting forces (Fig 4). An orange sphere traveled around a defined circular trajectory in either a clockwise or a counterclockwise direction. While the sphere moved around the track, distracting forces, with

Table I
Participants' demographic information

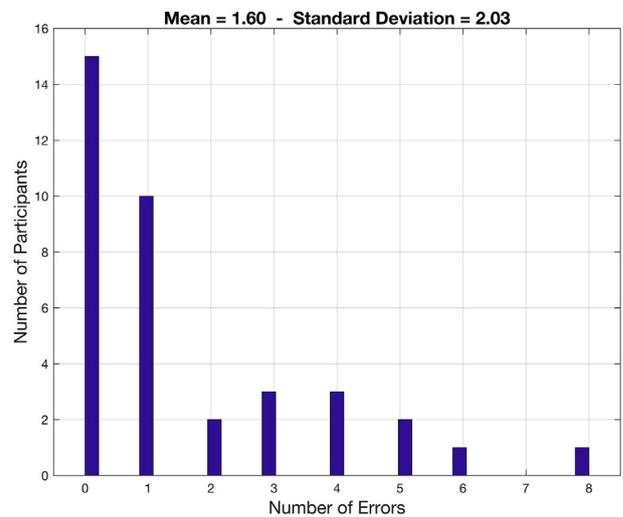|	Participants $n = 37$ (100%)
Male sex (%)	17 (45.9)
Institution	
1	8 (21.6)
2	4 (10.8)
3	2 (5.4)
4	4 (10.8)
5	1 (2.7)
6	10 (27.0)
7	9 (24.3)
Postgraduate year	
1	2 (5.4)
2	14 (37.8)
3	13 (35.1)
4	6 (16.2)
5	2 (5.4)
Clinical years completed	
1	2 (5.4)
2	27 (73.0)
3	7 (18.9)
4	1 (2.7)
Research years completed	
0	23 (62.2)
1	7 (18.9)
2	6 (16.2)
3	1 (2.7)

Fig 5. A plot of the visualization errors committed by the participants during LVH repair.

magnitudes from 2 N to 6 N, are applied by the haptic device. Participants were instructed to attempt to resist the force and continue to follow the sphere. Different magnitudes of force and direction were given to participants throughout 15 distinct and mathematically randomized trials.[16]

A total of 2 metrics were calculated in this module: (1) "Path Length," defined as average of the total traveled distance of the stylus tip throughout the 15 trials and (2) "Maximum Distance from Sphere," defined as the maximum distance between the tip of the stylus and the center of the sphere averaged throughout the 15 trials.

Data analysis

Spearman's rho correlation coefficient was used to assess correlations between visualization errors, final product score, survey metrics, and VR metrics. Alpha was set at 0.05 to determine
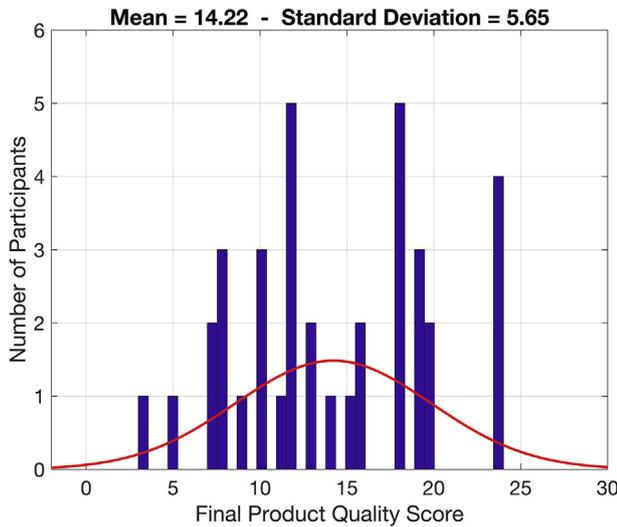
Fig 6. The distribution of scores associated with the quality of the final product in the LVH repair. The fitted normal distribution is depicted by the *red line*.

statistical significance. Statistical analysis was preformed using MATLAB (MathWorks, Natick, MA).

Results

The demographic information of the 37 surgical residents who participated are presented in Table I.

When analyzing the LVH performance data, participants committed a mean of 1.6 ± 2.0 visualization errors (Fig 5) and had a mean FPQS of 14.22 ± 5.65 out of a total possible score of 24 (Fig 6). A total of 4 residents achieved a perfect FPQS. Because of the small number of senior residents, performance was not stratified by training level. Assessment of the relationship between the number of LVH visualization errors and the LVH FPQS revealed a moderately strong, negative correlation (r = −0.52, P = .001; Fig 7). Meaning, a lower number of visualization errors correlated with a higher FPQS.

When comparing VR performance with LVH repair performance data, we found significant correlations. In the VR force matching module, the metric "Peak Deflection upon Force Release" and the metric "Summation of Distances from Sphere" showed significant negative correlations to the FPQS (r = −0.34, P = .04 and r = −0.36, P = .03, respectively). This indicates that
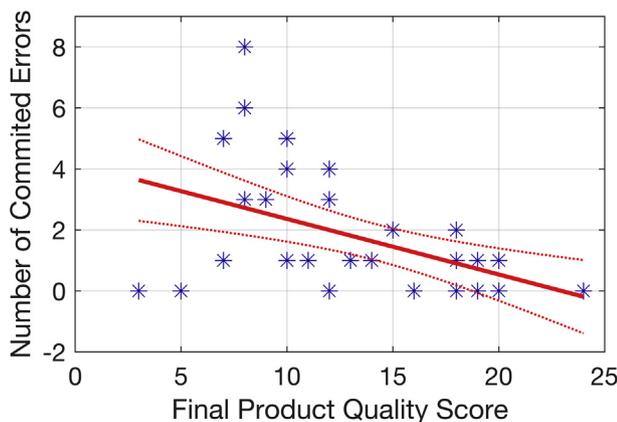

Fig 7. The linear relationship between the number of committed visualization errors of participants and their corresponding final product quality score in performing LVH repair.

better performance during the force matching module in VR corresponded to a higher quality repair in the simulated LVH procedure. Similarly, in the VR target tracking module, the metrics "Path Length" and "Maximum Distance from Sphere" showed significant negative correlations to the FPQS of the LVH repair (r = −0.49, P < .01 and r = −0.37, P = .03, respectively; Table II). This indicates that better performance in target tracking is related to a higher quality LVH repair.

In addition, the metric "Path Length" in the Target Tracking VR module had a significant positive correlation with the LVH visualization errors (r = +0.38, P = .02). Meaning, lower performance on the target tracking module in VR correlated with more visualization errors while performing the LVH repair. No correlation was identified between the other VR metrics and number of visualization errors (Table II).

Both preprocedure and postprocedure survey responses regarding confidence and perceived procedure difficulty showed significant correlations to the FPQS (Table III). Self-reported confidence in "completely visualizing the hernia defect" and "successfully performing the entire surgical task" were positively correlated with FPQS on both the presurvey results (r = +0.37, P=0.02 and r = +0.43, P < .01, respectively) and postsurvey results (r = +0.37, P = .02 and r = +0.53 P < .01, respectively). Perceived procedure difficulty negatively correlated with FPQS for "completely visualize the hernia defect" on the presurvey (r = −0.41, P = .02) and "successfully perform the entire surgical task" on the postsurvey (r = −0.48, P < .01). However, survey metrics on confidence and perceived difficulty did not correlate with the number of visualization errors committed during the LVH procedure (Table III).

Discussion

This study sought to understand whether performance on specific VR tasks could be used as a screening tool for LVH repair performance. We hypothesized that assessment using a motorized haptic device has the potential to be used as an assessment tool for identifying the psychomotor skills that contribute to the ability to perform a high-quality LVH repair. Our results identified four motion metrics that correlate with FPQS, suggesting that these metrics could be used to screen residents.

To address the psychomotor and visual-spatial challenges unique to laparoscopic surgery, sequential training of these skills has been proposed. Nickel et al[17] found that medical students who initially underwent a sequential training approach focused on psychomotor skill acquisition in a transparent shoebox required less time on the laparoscopic box trainer compared with those who underwent the traditional simultaneous approach to training psychomotor skills. This finding suggests that trainees with higher psychomotor skill may need less time to achieve proficiency. Therefore, the ability to screen residents' psychomotor skill, using VR would allow for identification of residents who would need less time in the simulation laboratory. Similarly, identification of residents with weaker psychomotor skill may see greater benefit from a sequential training method.

The number of visualization errors committed during the simulated LVH repair was found to be a predictor of FPQS. In contrast, with the exception of "Path Length" metric on the VR target tracking module commission of visualization errors did not correlate with performance on VR. We believe that visualization errors did not correlate with VR outcomes because visualization errors may reflect visual-spatial ability and VR outcomes reflect psychomotor ability. Similarly, commission of visualization errors did not correlate with confidence or perceived procedural difficulty as self-reported in the survey.

Table II
Correlation of the VR metrics to the final product quality score and number of visualization errors committed

Virtual reality metrics		Final product score		Visualization errors	
		R	P value	R	P value
Force matching	Peak deflection upon force release	−0.344	.037[*]	+0.032	.848
	Summation of distances from sphere	−0.359	.029[*]	+0.090	.597
Target tracking	Path length	−0.488	.002[*]	+0.375	.022[*]
	Maximum distance from sphere	−0.365	.026[*]	+0.132	.437

[*] $P < .05$.

Table III
Correlation of the survey metrics to the final product quality score and number of visualization errors committed

Survey metrics			Final product score		Visualization errors	
			R	P value	R	P value
Presurvey	Confidence	Completely visualize hernia defect	+0.370	.024[*]	−0.046	.788
		Successfully perform the entire surgical task	+0.430	.008[*]	−0.061	.719
	Difficulty[†]	Completely visualize hernia defect	−0.408	.015[*]	+0.076	.663
		Successfully perform the entire surgical task	−0.284	.099	−0.013	.941
Postsurvey	Confidence	Completely visualize hernia defect	+0.371	.024[*]	+0.019	.911
		Successfully perform the entire surgical task	+0.529	.001[*]	−0.095	.576
	Difficulty	Completely visualize hernia defect	−0.254	.129	−0.090	.595
		Successfully perform the entire surgical task	−0.482	.003[*]	+0.017	.919

[*] $P < .05$.
[†] These correlations reported using 35 participants as 2 participants' survey data were missed during data collection.

The ability of VR tasks to separate psychomotor ability from visual-spatial ability and procedural knowledge may indicate VR's role as a diagnostic tool. The ability to break performance into domain-specific scores would allow for the identification of specific weaknesses and individualized curriculum. For example, a trainee with a low FPQS on the simulated LVH and poor performance on the VR tasks may benefit from focusing on psychomotor skills. In contrast, a trainee with a low FPQS and high performance on the VR tasks may benefit more from focusing on visual-spatial skills. We believe additional work on these relationships could shine more light on the utility of VR in shaping training. Although not identified in this work, a different VR task could be developed to focus on visual-spatial ability. We would expect a visual-spatial focused VR task to correlate with visualization errors. Additional work could also focus on developing a VR task to improve our ability to screen trainees to determine their learning needs.

This study was limited in terms of number of participants and range of experience. We did not have sufficient numbers of more-senior residents ($n = 8$) to have sufficient power to stratify based on training level. Increasing the range of experience to include more-senior residents and attendings would allow for a better understanding of the robustness of the correlation between VR performance and quality of the hernia repair. It would also allow for further investigation into the effect of experience on skills acquisition.

In conclusion, this study investigated the possibility of screening surgical residents using VR. Our findings indicate that assessing residents using VR-based psychomotor skill assessment correlates with the quality of an LVH repair. The results suggest that screening residents with VR could potentially be used to inform overall resident competency in LVH repair performance and used to guide training.

Conflict of interest

The authors have indicated that they have no conflict of interest regarding the content of this article.

References

1. Malangoni MA, Biester TW, Jones AT, Klingensmith ME, Lewis FR. Operative experience of surgery residents: Trends and challenges. *J Surg Educ*. 2013;70:783–788.
2. Alkhoury F, Martin JT, Contessa J, Zuckerman R, Nadzam G. The impact of laparoscopy on the volume of open cases in general surgery training. *J Surg Educ*. 2010;67:316–319.
3. White C, Rodger MW, Tang T. Current understanding of learning psychomotor skills and the impact on teaching laparoscopic surgical skills. *Obstet Gynaecol*. 2016;18:53–63.
4. Lewis FR, Klingensmith ME. Issues in general surgery residency training—2012. *Ann Surg*. 2012;256:553–559.
5. de Montbrun SL, MacRae H. Simulation in surgical education. *Clin Colon Rectal Surg*. 2012;25:156–165.
6. Han JJ, Brooks AD. How can we be more "deliberate" in training surgeons? *Am J Surg*. 2018;216:359–360.
7. Stefanidis D, Korndorffer Jr JR, Black FW, et al. Psychomotor testing predicts rate of skill acquisition for proficiency-based laparoscopic skills training. *Surgery*. 2006;140:252–262.
8. Henn P, Gallagher AG, Nugent E, et al. Visual spatial ability for surgical trainees: Implications for learning endoscopic, laparoscopic surgery and other image-guided procedures. *Surg Endosc*. 2018;32:3634–3639.
9. Wanzel KR, Hamstra SJ, Caminiti MF, Anastakis DJ, Grober ED, Reznick RK. Visual-spatial ability correlates with efficiency of hand motion and successful surgical performance. *Surgery*. 2003;134:750–757.
10. Wanzel KR, Hamstra SJ, Anastakis DJ, Matsumoto ED, Cusimano MD. Effect of visual-spatial ability on learning of spatially-complex surgical skills. *Lancet*. 2002;359:230–231.
11. D'Angelo A-LD, Law KE, Cohen ER, et al. The use of error analysis to assess resident performance. *Surgery*. 2015;158:1408–1414.
12. Pugh C, Plachta S, Auyang E, Pryor A, Hungness E. Outcome measures for surgical simulators: Is the focus on technical skills the best approach? *Surgery*. 2010;147:646–654.
13. Nathwani JN, Wise BJ, Garren ME, et al. Residents' surgical performance during the laboratory years: An analysis of rule-based errors. *J Surg Res*. 2017;219:226–231.
14. Mohamadipanah H, Nathwani J, Peterson K, et al. Shortcut assessment: Can residents' operative performance be determined in the first five minutes of an operative task? *Surgery*. 2018;163:1207–1212.
15. Parthiban C, Ray R, Rutherford D, Zinn M, Pugh C. Development and analysis of psychomotor skills metrics for procedural skills decay. *Stud Health Technol Inform*. 2016;220:285–288.
16. Mohamadipanah H, Parthiban C, Law K, et al. Hand smoothness in laparoscopic surgery correlates to psychomotor skills in virtual reality. In: *2016 IEEE 13th International Conference on Wearable and Implantable Body Sensor Networks (BSN)*. San Francisco, CA: IEEE; 2016:242–246.
17. Nickel F, Hendrie JD, Kowalewski K-F, et al. Sequential learning of psychomotor and visuospatial skills for laparoscopic suturing and knot tying—A randomized controlled trial "The Shoebox Study" DRKS00008668. *Langenbecks Arch Surg*. 2016;401:893–901.