



Transfer of virtual reality endoscopy training to live animal colonoscopy: a randomized control trial of proficiency vs. repetition-based training

Douglas J. Cassidy¹ · Taylor M. Coe¹ · Kristen M. Jogerst^{1,2} · Sophia K. McKinley¹ · Naomi M. Sell¹ · Michael Sampson¹ · Yoon Soo Park³ · Emil Petrusa¹ · Robert N. Goldstone¹ · Daniel A. Hashimoto¹ · Denise W. Gee^{1,4}

Received: 3 October 2021 / Accepted: 9 December 2021 / Published online: 10 February 2022
© The Author(s), under exclusive licence to Springer Science+Business Media, LLC, part of Springer Nature 2022

Abstract

Background Low first-time pass rates of the Fundamentals of Endoscopic Surgery (FES) exam stimulated development of virtual reality (VR) simulation curricula for test preparation. This study evaluates the transfer of VR endoscopy training to live porcine endoscopy performance and compares the relative effectiveness of a proficiency-based vs repetition-based VR training curriculum.

Methods Novice endoscopists completed pretesting including the FES manual skills examination and Global Assessment of GI Endoscopic Skills (GAGES) assessment of porcine upper and lower endoscopy. Participants were randomly assigned one of two curricula: proficiency-based or repetition-based. Following curriculum completion, participants post-tested via repeat FES examination and GAGES porcine endoscopy assessments. The two cohorts pre-to-post-test differences were compared using ANCOVA.

Results Twenty-two residents completed the curricula. There were no differences in demographics or clinical endoscopy experience between the groups. The repetition group spent significantly more time on the simulator (repetition: 242.2 min, SD 48.6) compared to the proficiency group (proficiency: 170.0 min, SD 66.3; $p=0.013$). There was a significant improvement in porcine endoscopy (pre: 10.6, SD 2.8, post: 16.6, SD 3.4; $p<0.001$) and colonoscopy (pre: 10.4, SD 2.7, post: 16.4, SD 4.2; $p<0.001$) GAGES scores as well as FES manual skills performance (pre: 270.9, SD 105.5, post: 477.4, SD 68.9; $p<0.001$) for the total cohort. There was no difference in post-test GAGES performance or FES manual skills exam performance between the two groups. Both the proficiency and repetition group had a 100% pass rate on the FES skills exam following VR curriculum completion.

Conclusion A VR endoscopy curriculum translates to improved performance in upper and lower endoscopy in a live animal model. VR curricula type did not affect FES manual skills examination or live colonoscopy outcomes; however, a proficiency curriculum is less time-consuming and can provide a structured approach to prepare for both the FES exam and clinical endoscopy.

Keywords Endoscopy · Colonoscopy · Virtual reality · Simulation · Proficiency-based training · Repetition-based training

The Fundamentals of Endoscopic Surgery (FES) program, created by the Society of American Gastrointestinal and

Endoscopic Surgeons (SAGES), is an educational and assessment tool of knowledge and skill in flexible gastrointestinal endoscopy. The exam is a validated assessment

Douglas J. Cassidy and Taylor M. Coe are co-first authors.

✉ Douglas J. Cassidy
dcassidy3@partners.org

✉ Denise W. Gee
dgee@mgh.harvard.edu

¹ Department of Surgery, Massachusetts General Hospital, 55 Fruit St. GRB-425, Boston, MA 02114, USA

² Department of Surgery, Mayo Clinic Hospital, Phoenix, AZ, USA

³ Department of Emergency Medicine, Massachusetts General Hospital, Boston, MA, USA

⁴ Department of Surgery, Massachusetts General Hospital, 15 Parkman St. WAC-460, Boston, MA 02114, USA

tool that consists of both a cognitive knowledge and manual skills assessment, the latter of which involves five tasks assessing the following skills: scope navigation, loop reduction, mucosal inspection, retroflexion, and targeting [1]. The skills portion of the exam is hosted on the 3D Systems (formerly Symbionix) GI Mentor endoscopic simulator, a virtual reality (VR) endoscopy simulator that has been the subject of multiple studies assessing construct validity [2–5]. Virtual reality simulators allow for the utilization of proficiency-based curricula, which require the identification of construct valid tasks and performance-based rather than time- or repetition-based criteria for completion [6]. A curriculum for attainment of proficient performance accounts for different rates of learning in individuals and ensures that trainees are truly acquiring an acceptable level of skill prior to performing procedures [6].

Beginning in 2018, all surgical residents are required to pass the FES examination to graduate; thus, there is a need for an effective curriculum to prepare trainees to pass this exam. The urgency for effective endoscopy curriculum had been heightened by work that demonstrated poor baseline pass rates of surgical trainees, suggesting that standard endoscopy experience is insufficient [7]. Fortunately, a few proficiency-based curricula have been developed to help trainees prepare for the FES manual skills examination [8–11]. Despite improved FES exam pass rates, no studies have assessed the impact of a VR training curriculum on trainee's endoscopic skills in a live animal model or patients.

Simulation-based endoscopy training has been shown to improve performance on a porcine model and studies have shown a correlation between live colonoscopy and FES manual skills assessment performance [12, 13]. The ultimate goal of the FES curriculum and endoscopic training is to prepare surgeons to perform endoscopy on real patients, and thus, understanding the optimal VR training preparation for colonoscopy in a live animal model is relevant for all general surgery training programs. The purpose of this study is to evaluate the transfer of VR endoscopy training to live porcine colonoscopy and to compare the relative effectiveness of two different VR training paradigms in preparing for live colonoscopy: proficiency-based training vs repetition-based training.

Materials and methods

Participants

Novice endoscopists, defined as trainees who have completed fewer than 10 colonoscopies, were recruited via email from a single large academic surgical residency between January 2020 and June 2021. At our institution, residents complete a dedicated endoscopy rotation during their second postgraduate year; therefore, recruitment was limited to surgical interns

(categorical general surgery and general surgery preliminary interns) whose rotations do not generally include colonoscopies. Participants were excluded if they had previous participation in a simulation-based endoscopy training curriculum or had previously taken the FES examination.

Instrumentation

The GI Mentor (3D Systems, Rock Hill, SC) is a virtual reality endoscopic simulator that is designed to assist in the teaching and practice of both upper and lower gastrointestinal endoscopy. This platform was chosen by SAGES for the administration of the FES examination and as a result, was selected for use in this study [1].

Study design

Our Institutional Review Board (IRB) exempted this protocol from further review with regards to human subject protections (IRB Protocol: 2015P000522-AME4) as an amendment and extension of our research team's prior work on FES VR curricula [10]. The Partners Healthcare Institutional Animal Care and Use Committee (IACUC) approved the porcine endoscopy protocol for use of swine in surgical training (IACUC Protocol #2019N000140).

This study employed a randomized control design. Participants completed baseline demographics surveys and were provided with a one-on-one familiarization session with the simulator by an experienced operator, including a full demonstration and an opportunity to ask questions. All participants underwent pre-testing by completing the FES manual skills examination and performing a porcine endoscopy and colonoscopy assessed with the Global Assessment of GI Endoscopic Skills (GAGES) scoring system, a tool that has been shown to have validity evidence for technical skills in flexible endoscopy (Appendix A and B) [14]. Healthy male Yorkshire pigs (30–40 kg) underwent a bowel preparation consisting of a clear liquid diet and polyethylene glycol starting the day prior to the procedure. Endoscopy was performed in the left lateral decubitus position under general anesthesia and animals were euthanized at the conclusion of the experiment. A single channel colonoscope (KARL STORZ Endoscopy-America, USA) was used for both the upper and lower endoscopy.

Participants were then randomized in modified matched pairs to one of two VR curricula: proficiency-based training or repetition-practice based training. The proficiency group completed practice GI Mentor VR tasks with expert level benchmarks as described in Hashimoto et al. (Appendix C) [10]. Each participant in the proficiency group was required to meet the benchmarks for each task, as determined by expert performance, on two consecutive occasions. In comparison, the repetition group completed the same tasks as the proficiency group but for a set number of repetitions (10),

regardless of performance quality or benchmark achieved. For each group, a FES-certified instructor was available to provide coaching during the first repetition of each task. Subsequent repetitions in the two groups were self-directed by the participant with feedback provided by the simulator (e.g., total time, proficiency achievement) after each task completion.

All participants underwent a post-test consisting of a repeat FES manual skills exam and porcine endoscopy and colonoscopy with GAGES assessment. Faculty performing the GAGES assessment were expert endoscopists and blinded to the study arm to which each participant has been assigned. Participants completed repeat demographic surveys indicating how many endoscopies and colonoscopies they performed in clinical training during the study period.

Assessment

Performance on porcine endoscopy and colonoscopy was measured with the GAGES scoring system, a checklist with high inter-rater reliability adopted by the American Board of Surgery to assess endoscopic skills in both upper and lower endoscopy [14]. The endoscopy and colonoscopy GAGES checklists are a 5-item checklist with 5-point behaviorally anchored ratings assigned to each item, assessing skills such as scope navigation and instrumentation by a trainee or other user. The instrumentation portion of the GAGES checklists was excluded from this study, and participants were evaluated with a maximum possible score of 20. Porcine endoscopy and colonoscopy were assessed by FES-certified surgeons who were blinded to each participant's assigned study arm.

Data analysis

We assessed for balanced randomization based on pretest performances and baseline participant characteristics. Participant demographics, FES-scaled scores, and pre-test GAGES scores were compared between the two training arms using student's unpaired t-tests, Chi-square, and Wilcoxon Rank Sum. The two curricula cohorts were compared based on VR simulator use data using unpaired student's t-tests. To determine the potential transfer of skills attained during VR training to live animal colonoscopy, pre- to post-test differences within the two cohorts on both FES-scaled scores and GAGES performances on porcine endoscopy and colonoscopy were assessed using paired student's t-tests. After confirming the parallel slopes assumption, the post-tests between the proficiency- and repetition-trained participants were then compared using analysis of covariance (ANCOVA). Statistical analysis was performed using Stata (StataCorp. 2019. Stata Statistical Software: Release 16. College Station, TX: StataCorp LLC), and statistical significance was set at $p < 0.05$ for all tests.

Results

Participants

Twenty-four ($n = 24$) general surgery interns were recruited to participate in the study across two academic years (2019–2020, 2020–2021). There were two dropouts during the study, one from the repetition group and one from the proficiency group. In total, eleven ($n = 11$) inexperienced endoscopists were randomized into the repetition group and eleven ($n = 11$) into the proficiency group. There were no significant differences between the repetition and proficiency groups in sex, age, glove size, handedness, or clinical experience with endoscopy or colonoscopy at time of enrollment (Table 1). There was no difference in mean baseline scaled performance on FES manual skills pre-testing (proficiency = 253.5, SD 115.7, repetition = 288.2, SD 96.5; $p = 0.45$). There was also no difference in porcine endoscopy and colonoscopy GAGES scores between the two groups (Endoscopy: proficiency = 10.2, SD 2.7, repetition = 11.1, SD 2.8, $p = 0.92$; Colonoscopy: proficiency = 10.5, SD 2.9, repetition = 10.3, SD 2.6, $p = 0.96$).

Performance

The repetition group spent an average of 242.2 (SD 48.6) minutes on the VR simulator compared to the proficiency group's average of 170.0 (SD 66.3) minutes ($p = 0.013$; Table 2). The proficiency group spent significantly less time on the Endoscopic Navigation task (proficiency = 43.1 min, SD 25.0, repetition = 70.6 min, SD 17.0; $p = 0.010$) and the Advanced Mucosal Evaluation task (proficiency = 43.5 min, SD 20.9, repetition = 68.5 min, SD 22.4; $p = 0.019$). There was no difference in time spent on completion of the upper endoscopy bleeding modules or the colonoscopy modules.

There was a significant difference in live porcine endoscopy (pre-endoscopy mean 10.6, SD 2.8, post-endoscopy mean 16.6, SD 3.4; $p < 0.001$) and colonoscopy (pre-colonoscopy mean 10.4, SD 2.7, post-colonoscopy 16.4, SD 4.2; $p < 0.001$) performance as measured by the GAGES score

Table 1 Comparison of the demographics of novice endoscopists participating in the proficiency vs. repetition curricula

	Proficiency group ($n = 11$) n /mean (%/SD)	Repetition group ($n = 11$) n /mean (%/SD)	p -value
Age	27.7 (2.3)	27.2 (1.2)	0.49
female sex	9 (75%)	5 (45.5%)	0.66
Right handedness	9 (81.8%)	9 (81.8%)	> 0.99
Glove size	6.7 (0.8)	6.9 (0.7)	0.48
Prior endoscopy or colonoscopy experience	5 (22%)	3 (14%)	0.67

Table 2 Comparison of time spent (in minutes) on VR simulator tasks between the proficiency and the repetition groups

	Proficiency group (n=11) Mean (SD)	Repetition group (n=11) Mean (SD)	p-value
FES curriculum	170.0 (66.3)	242.2 (48.6)	0.013
Endoscopic Navigation	43.1 (25.0)	70.6 (17.0)	0.010
Advanced Mucosal Evaluation I	43.5 (20.9)	68.5 (22.4)	0.019
Bleeding Module 1, Case 5	11.6 (8.9)	10.0 (3.3)	0.611
Bleeding Module 1, Case 6	10.6 (10.2)	14.0 (5.5)	0.318
Colonoscopy Module 1, Case 9	39.9 (21.4)	48.0 (21.2)	0.394
Colonoscopy Module 2, Case 10	21.3 (12.9)	30.5 (8.1)	0.074

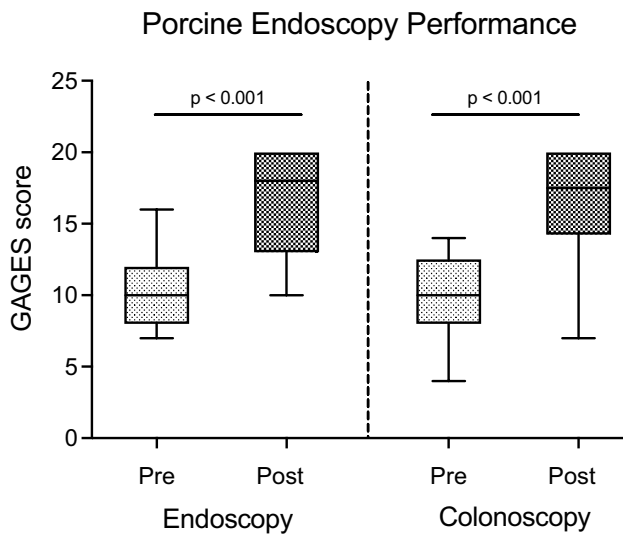


Fig. 1 Comparison of GAGES performance in pre- and post-curriculum porcine endoscopy and colonoscopy testing

before and after curriculum completion (Fig. 1). There was no difference in post-test performance between the proficiency and repetition groups, after accounting for prior pre-test performance (Table 3).

There was a significant different in FES manual skills performance (pre-curriculum mean 270.9, SD 105.5, post-curriculum mean 477.4, SD 68.9; $p < 0.001$) before and after curriculum completion (Fig. 2). However, no difference was seen in performance on the post-curriculum exam between the proficiency and repetition groups (proficiency = 465.3, SD 58.6, repetition = 489.6, SD 78.7; $p = 0.658$, Table 4). Likewise, there was no difference in performance between the two curricula on subset analysis of each component of the FES exam. All participants ($n = 22$) in both the proficiency and repetition groups had a 100% FES manual skills pass rate.

Table 3 GAGES scores presented as means plus or minus standard deviations. Scores are reported before and after completing the assigned curricula. After confirming parallel slopes assumption,

the scores between groups compared with Analysis of Covariance (ANCOVA). There was no significant difference between the two groups

GAGES skill	Repetition curriculum			Proficiency curriculum			p-value
	Pre-test	Post-test	Improvement	Pre-test	Post-test	Improvement	
EGD intubation	2.09 ± 1.14	3.55 ± 1.75	1.45 ± 2.38	1.91 ± 1.37	3.64 ± 1.80	1.73 ± 2.37	0.959
EGD navigation	2.91 ± 0.83	4.18 ± 0.87	1.27 ± 1.10	2.64 ± 0.50	4.36 ± 0.81	1.73 ± 0.90	0.546
EGD clear field	3.18 ± 0.87	4.18 ± 0.87	1.00 ± 1.18	3.00 ± 1.00	4.27 ± 0.90	1.27 ± 1.42	0.826
EGD exam quality	2.91 ± 0.83	4.18 ± 0.75	1.27 ± 1.01	2.64 ± 0.67	4.27 ± 0.90	1.64 ± 1.29	0.844
EGD total score	11.09 ± 2.91	16.36 ± 2.84	5.27 ± 4.24	10.18 ± 2.75	16.73 ± 4.03	6.55 ± 5.43	0.917
Colonoscopy navigation	2.27 ± 0.79	4.00 ± 1.00	1.73 ± 1.01	2.54 ± 0.69	3.82 ± 1.40	1.27 ± 1.62	0.669
Colonoscopy strategies	2.54 ± 0.52	4.09 ± 0.94	1.55 ± 0.82	2.36 ± 0.67	4.18 ± 1.25	1.82 ± 1.47	0.786
Colonoscopy clear field	2.91 ± 0.94	4.27 ± 1.01	1.36 ± 0.92	2.91 ± 0.83	4.36 ± 1.12	1.45 ± 1.63	0.847
Colonoscopy exam quality	2.54 ± 0.82	4.00 ± 0.89	1.45 ± 1.13	2.64 ± 0.92	4.09 ± 1.30	1.45 ± 1.81	0.828
Colonoscopy total score	10.27 ± 2.61	16.36 ± 3.64	6.09 ± 3.45	10.45 ± 2.91	16.45 ± 4.89	6.00 ± 6.29	0.964
GAGES total	21.36 ± 5.26	32.73 ± 5.59	11.36 ± 6.56	20.64 ± 4.52	33.18 ± 8.22	12.55 ± 9.97	0.876

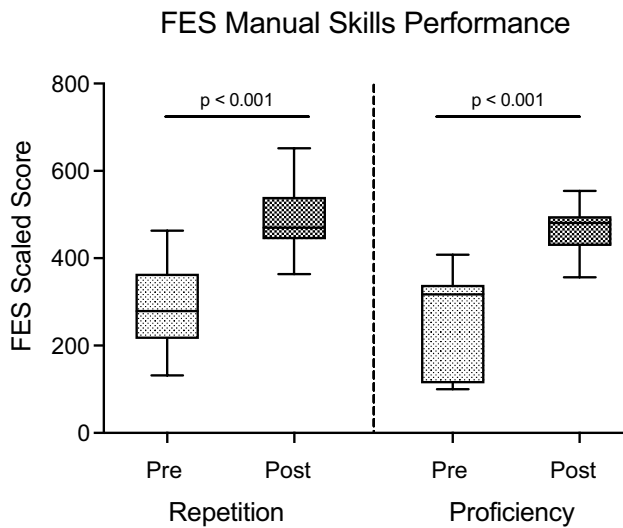


Fig. 2 Comparison of FES manual skills exam performance in pre- and post-curriculum testing between repetition- and proficiency-based curriculum groups

Discussion

This study demonstrates the utility of using a VR curriculum in training novice endoscopists prior to their clinical exposure to endoscopy. A VR curriculum for endoscopy results in significantly improved performance in both upper and lower endoscopy as measured by the GAGES scoring system in a live animal model. Following curriculum completion, novice endoscopists demonstrated significant improvement in all four of the GAGES parameters assessed for both endoscopy and colonoscopy as well as significant overall improvement.

Skills transfer following completion of simulation-based training is critical for novices to make the transition from skill development to safe clinical practice. Trainees who reached simulation-based skill proficiency before undergoing patient-based assessments have previously demonstrated improved

performance in both an animal model and in the operating room for laparoscopic procedures, with higher global assessment scores and fewer errors than their counterparts [15, 16]. Trainees completing simulator-based training for endoscopy perform at a similar level in a clinical setting compared to trainees who participated in patient-based training, suggesting that the skills learned on a simulator are comparable to clinical learning [17]. Our study suggests that the endoscopic skills obtained through completion of a VR endoscopy curriculum transfer to a live animal model, an important pre-clinical model representative of clinical practice.

Recommended procedure numbers for general surgery residents are 35 and 50 for upper endoscopy and colonoscopy, respectively [18]. Prior studies have suggested that these recommendations may not represent the experience needed to achieve proficiency, as the learning curve for upper endoscopy and colonoscopy, based on GAGES scoring, begins to plateau around 50 and 75 cases, respectively [19]. Given the nuanced balance between trainee skill acquisition and patient safety, it is critical trainees reach a baseline level of proficiency prior to clinical exposure to maximize within-case learning. With trainee time limits and competing clinical demands, efficient training modalities are essential for residents to reach proficiency prior to graduation. Completion of our VR curriculum demonstrated high FES pass rates and significant clinical improvement, serving as a sufficient preparation for a clinical endoscopy rotation.

Similar to our group's prior work, completion of the VR curriculum resulted in a high pass rate (100% for both proficiency and repetition curricula) and high scores (465.3 for proficiency curriculum, 489.6 for repetition curriculum) on the FES skills exam [10]. Several other groups have adopted their own endoscopy curricula using either the GI Mentor II or an alternative model to improve institutional FES scores [8, 9, 11, 20–22]. These findings continue to advocate for the use of simulation-based training for endoscopy training and preparation for the FES examination. The results of our study in conjunction with our prior work suggests that our

Table 4 FES-scaled scores presented as means plus or minus standard deviations

FES tasks	Repetition curriculum			Proficiency curriculum			p-value
	Pre-test	Post-test	Improvement	Pre-test	Post-test	Improvement	
Task 1	47.2 ± 19.4	90.6 ± 10.7	43.4 ± 8.7	37.7 ± 20.5	81.2 ± 12.4	43.5 ± 8.0	0.091
Task 2	11.7 ± 17.2	59.6 ± 34.8	47.9 ± 17.6	30.2 ± 28.1	65.5 ± 21.7	35.3 ± 6.4	0.594
Task 3	70.0 ± 22.5	76.6 ± 19.3	6.5 ± 3.2	59.3 ± 22.5	76.1 ± 16.1	16.8 ± 6.4	0.783
Task 4	57.1 ± 15.7	49.8 ± 10.7	7.3 ± 5.0	47.2 ± 16.8	48.2 ± 14.0	0.9 ± 2.7	0.994
Task 5	67.9 ± 18.6	98.4 ± 9.9	30.5 ± 8.8	50.0 ± 22.5	89.3 ± 11.9	39.3 ± 10.6	0.106
FES total	50.8 ± 11.6	75.0 ± 9.5	24.2 ± 2.1	44.9 ± 16.7	72.1 ± 7.0	27.2 ± 9.7	0.720
FES scaled	288.2 ± 96.5	489.6 ± 78.7	201.3 ± 17.8	253.5 ± 115.7	465.3 ± 58.6	211.8 ± 57.1	0.658

Scores are reported before and after completing the assigned curricula. After confirming parallel slopes assumption, the scores between groups compared with Analysis of Covariance (ANCOVA). There was no significant difference between the two groups

institutional VR curriculum and performance benchmarks can be utilized by programs with access to the GI Mentor II in order to best prepare their residents for the FES manual skills examination [10].

Our study sought to compare the relative effectiveness of two different VR training paradigms, proficiency-based and repetition-based training, in preparing trainees for live colonoscopy. Between the two curricula, there was no difference in performance on the post-training FES manual skills examination or live colonoscopy. Despite the similar performance, there were significant differences in time spent completing the curriculum. The participants in the proficiency-based group spent on average 72 less minutes completing the curriculum (170.0 min vs. 242.2 min; $p=0.013$) and less time completing the Endoscopic Navigation and Advanced Mucosal Evaluation tasks. All participants in the repetition group met proficiency standards in both tasks, suggesting that the additional repetitions in this group were unnecessary with an additional time cost. This is similar to laparoscopic skill acquisition in novices, where criterion-based training reduces overall training time without impacting training outcome and overtraining, despite a faster learning curve, has no long-term effect on skill retention and no additional time benefit [23, 24]. With the time saved and proven comparable outcomes with proficiency-based curriculum for skill acquisition, there does not appear to be any benefit to the use of a repetition-based curriculum.

Interestingly, 2 participants (18.2%) in the repetition group did not meet proficiency standards in Task 9 and 4 participants (36.4%) did not meet proficiency standards in Task 10. Both tasks are designed to practice the skill of loop reduction, notoriously the most difficult task for the FES examination [25, 26]. Despite this, there was not a difference in performance on the Loop Reduction task between groups in the FES manual skills examination (proficiency=65.5, SD 21.7, repetition=59.6, SD 34.8; $p=0.594$), suggesting that proficiency benchmarks of expert performance for these tasks may be too strict.

Trainees begin with different levels of fundamental ability, experience, and skill. If a standard number of hours or number of tasks is prescribed to all trainees, the outcome will be variable performance levels based on individuals' manual skill learning curves [27]. One of the benefits of a proficiency-based curriculum is that all trained individuals perform at a pre-determined benchmark level of competence. This allows for flexibility in training, and as we found, a reduction in overall time on the simulator to reach proficiency standards with similar outcomes to repetition-based training. This is critical given there are currently 73 institutions within the USA that are FES test centers with GI Mentor II access [28]. Given there was no difference in clinical or examination performance between a proficiency-based and a repetition-based curriculum, the recommendation to use a proficiency-based curriculum is secondary to the problem of limited simulator access. We want to help

ensure residents can complete an appropriate curriculum in a timely and efficient manner.

This study has several limitations that must be considered. First, this study was conducted at a single academic institution with residents from a single general surgery residency. Due to costs and equipment availability, including both the animal laboratory and the GI Mentor, it was difficult to expand the sample size outside of a single institution. As a consequence of these limitations, the study was not powered in order to see a difference between the two curricula. Power calculations would reveal a sample size of $n=59$ necessary to detect a difference on FES performance between the proficiency-based and repetition-based curricula and a sample size of $n=1870$ to detect a difference in GAGES score, the latter suggesting essentially no difference between the curricula in clinical endoscopy performance.

The pig model also has its own anatomical considerations that differ significantly from human anatomy. For the upper endoscopy portion, successful endoscopic intubation is more difficult given the presence of a pharyngeal diverticulum. Normal pig anatomy has a pharyngeal diverticulum similar to the pathological Zenker's diverticulum that can be seen in a human. In the pig, the pharyngeal diverticulum is located posteriorly, at the level of the upper sphincter, and evident upon passing the pharyngeal sinus on either side of the glottis [29]. After entry into the diverticulum, slow withdrawal of the scope is required to allow viewing of both the septum and the esophageal lumen, facilitating safe passage. For the colonoscopy portion, normal pig anatomy includes a proximal spiral colon arranged in a series of centrifugal and centripetal coils [30]. This consists of the cecum, ascending colon, and transverse colon, whereas the anatomy of the left colon and rectum is similar to that of humans. Given these findings, participants were only required to reach the spiral colon but not completely traverse it to reach the cecum.

Finally, it is important to recognize that the original study design was interrupted by the COVID-19 pandemic during the first year of the study. Residents during the 2019–2020 academic year were pre-tested prior to the pandemic; however, due to a suspension in all research and simulation activities to maintain social distancing guidelines, the VR curriculum was completed during the beginning of their second clinical year. As a result, this group may have gained more clinical endoscopy exposure; however, no participant completed their dedicated endoscopy rotation prior to GAGES post-testing, likely minimizing this clinical education confounder.

Conclusion

Participation in a VR curriculum leads to both improved FES performance as well as improved clinical endoscopy performance for novice endoscopists. A VR curriculum and training

program can provide a structured approach for residents to acquire endoscopic skills. Completion of a VR curriculum leads to improved clinical performance in an animal model, maximizes the benefits of a clinical endoscopic experience, and optimally prepares residents for performing patient endoscopy.

Appendix A

Global Assessment of Gastrointestinal Endoscopic Skills (GAGES) Scoresheet for Upper Endoscopy

GAGES - UPPER GI ENDOSCOPY SCORESHEET

GLOBAL ASSESSMENT OF GASTROINTESTINAL ENDOSCOPIC SKILLS

INTUBATION OF THE ESOPHAGUS

SCORE

Reflects patient management, understanding of anatomy and sedation

- 5 Able to independently (successfully) intubate esophagus without patient discomfort
- 4
- 3 Requires detailed prompting and cues
- 2
- 1 Unable to properly intubate requiring take over

SCOPE NAVIGATION

SCORE

Reflects navigation of the GI tract using tip deflection, advancement/withdrawal and torque

- 5 Expertly able to manipulated the scope in the upper GI tract autonomously.
- 4
- 3 Requires verbal guidance to completely navigate the upper GI tract
- 2
- 1 Not able to achieve goals despite detailed verbal cues, requiring take over

ABILITY TO KEEP A CLEAR ENDOSCOPIC FIELD

SCORE

Utilization of insufflation, suction and/or irrigation to maximize mucosal evaluation

- 5 Uses insufflation, suction, and irrigation optimally to maintain clear view of endoscopic field
- 4
- 3 Requires moderate prompting to maintain clear view
- 2
- 1 Inability to maintain view despite extensive verbal cues

INSTRUMENTATION (if applicable; leave blank if not applicable)

SCORE

Random biopsy: targeting is assessed by asking the endoscopist to take another biopsy from the identical site. Targeted instrumentation: evaluation is based on ability to direct the instrument to the target.

- 5 Expertly directs instrument to desired target
- 4
- 3 Requires some guidance and/or multiple attempts to direct instrument to target
- 2
- 1 Unable to direct instrument to target despite coaching

QUALITY OF EXAMINATION

SCORE

Reflects attention to patient comfort, efficiency, and completeness of mucosal evaluation

- 5 Expertly completes the exam efficiently and comfortably
- 4
- 3 Requires moderate assistance to accomplish a complete and comfortable exam
- 2
- 1 Could not perform a satisfactory exam despite verbal and manual assistance requiring takeover of the procedure

Appendix B

Global Assessment of Gastrointestinal Endoscopic Skills (GAGES) Scoresheet for Colonoscopy

GAGES - COLONOSCOPY SCORESHEET GLOBAL ASSSMENT OF GASTROINTESTINAL ENDOSCOPIC SKILLS

SCOPE NAVIGATION

SCORE

Reflects navigation of the GI tract using tip deflection, advancement/withdrawal and torque

- 5 Expertly able to manipulate the scope in the GI tract autonomously
- 4
- 3 Requires verbal guidance to completely navigate the lower GI tract
- 2
- 1 Not able to achieve goals despite detailed verbal guidance requiring takeover

USE OF STRATEGIES

SCORE

Examines use of patient positions, abdominal pressure, insufflation, suction and loop reduction to comfortably complete the procedure

- 5 Expert use of appropriate strategies for advancement of the scope while optimizing patient comfort
- 4
- 3 Use of some strategies appropriately, but requires moderate verbal guidance
- 2
- 1 Unable to utilize appropriate strategies for scope advancement despite verbal assistance

ABILITY TO KEEP A CLEAR ENDOSCOPIC FIELD

SCORE

Utilization of insufflation, suction and/or irrigation to maximize mucosal evaluation

- 5 Used insufflation, suction, and irrigation optimally to maintain clear view of endoscopic field
- 4
- 3 Requires moderate prompting to maintain clear view
- 2
- 1 Inability to maintain view despite extensive verbal cues

INSTRUMENTATION (if applicable; leave blank if not applicable)

SCORE

Random biopsy: targeting is assessed by asking the endoscopist to take another biopsy from the identical site. Targeted instrumentation: evaluation is based on ability to direct the instrument to the target.

- 5 Expertly directs instrument to desired target
- 4
- 3 Requires some guidance and/or multiple attempts to direct instrument to target
- 2
- 1 Unable to direct instrument to target despite coaching

QUALITY OF EXAMINATION

SCORE

Reflects attention to patient comfort, efficiency, and completeness of mucosal evaluation

- 5 Expertly completes the exam efficiently and comfortably
- 4
- 3 Requires moderate assistance to accomplish a complete and comfortable exam
- 2
- 1 Could not perform a satisfactory exam despite verbal and manual assistance requiring takeover of the procedure

Appendix C

Endoscopy curriculum and established proficiency benchmarks

GI mentor module	FES skill practiced	Metrics	Benchmark
Endoscopic navigation	Scope navigation, targeting	Total time	9 min 19 s
Advanced mucosal evaluation I	Mucosal inspection	Total time, % lesions	11 min 11 s; 100% lesions
Colonoscopy module 1, case 9	Loop reduction	Time to cecum	4 min 12 s
Colonoscopy module 2, case 10	Loop reduction	Time to cecum	2 min 30 s
Bleeding module 1, case 5	Retroflexion, targeting	Total time	1 min 14 s
Bleeding module 1, case 6	Retroflexion, targeting	Total time	1 min 21 s

Acknowledgements The authors would like to thank SAGES for funding for this project and KARL STORZ Endoscopy-America for their generous support and equipment donation. In addition, we would like to thank the staff of the Knight Surgical Research Laboratory, in particular, Michael Duggan, DVM, for their support and care of the animals. In addition, we would like to thank Michelle Lynn Sanchez for her administrative work and assistance.

Author contributions DC and TC were directly involved in the design and implementation of the study, collection of data, analysis and interpretation of data, writing the manuscript, and approving the final version for submission. KJ was directly involved in the collection of data, analysis, and interpretation of data, providing revisions and approving the final version for submission. SM and DH were directly involved in the conception and design of the study, providing revisions, and approving the final manuscript. NS, MS, and YSP were involved in the collection of data, analysis and interpretation of data, and approval of the final manuscript. EP was involved in the design of the study, analysis and interpretation of data, providing revisions and approval of the final manuscript. RG and DG were directly involved in the oversight of the project as well as providing critical revisions and approval of the final manuscript for submission.

Funding This work was supported by a 2019 SAGES General Research Grant awarded to Dr. Denise Gee. The funding source had no role in the study design; the collection, analysis and interpretation of data; the writing of the report; and in the decision to submit the article for publication.

Declarations

Disclosures Drs. Cassidy, Coe, Goldstone, Jogerst, McKinley, Sell, Park, and Petrusa and Mr. Sampson have no conflicts of interest or financial ties to disclose. Dr. Hashimoto is a consultant for Google/

Verily Life Sciences and Johnson & Johnson Institute. He is a member of the Advisory Board of Activ Surgical and has research support from Olympus Corporation and Intuitive Foundation. Dr. Gee is a consultant for New View Surgical, Boston Scientific, Ethicon, and Medtronic. None of these commitments or relationships impacted the design or implementation of the study.

References

- Vassiliou MC, Dunkin BJ, Fried GM et al (2014) Fundamentals of endoscopic surgery: creation and validation of the hands-on test. *Surg Endosc* 28:704–711
- Fayez R, Feldman LS, Kaneva P, Fried GM (2010) Testing the construct validity of the Symbionix GI Mentor II virtual reality colonoscopy simulator metrics: module matters. *Surg Endosc* 24:1060–1065
- Felsher JJ, Olesevich M, Farres H et al (2005) Validation of a flexible endoscopy simulator. *Am J Surg* 189:497–500
- Grantcharov TP, Carstensen L, Schulze S (2005) Objective assessment of gastrointestinal endoscopy skills using a virtual reality simulator. *JLS J Soc Laparoendosc Surg*. 9:130
- Koch AD, Buzink SN, Heemskerk J et al (2008) Expert and construct validity of the Symbionix GI Mentor II endoscopy simulator for colonoscopy. *Surg Endosc* 22:158–162
- Aggarwal R, Grantcharov TP, Darzi A (2007) Framework for systematic training and assessment of technical skills. *J Am Coll Surg* 204:697–705
- Gardner AK, Scott DJ, Willis RE et al (2017) Is current surgery resident and GI fellow training adequate to pass FES? *Surg Endosc* 31:352–358
- Gearhart S, Marohn M, Ngamruengphong S et al (2018) Development of a train-to-proficiency curriculum for the technical skills component of the fundamentals of endoscopic surgery exam. *Surg Endosc* 32:3070–3075
- Guzzetta AA, Weis JJ, Hennessy SA et al (2018) Proficiency-based preparation significantly improves FES certification performance. *Surg Endosc* 32:4451–4457
- Hashimoto DA, Petrusa E, Phitayakorn R, Valle C, Casey B, Gee D (2018) A proficiency-based virtual reality endoscopy curriculum improves performance on the fundamentals of endoscopic surgery examination. *Surg Endosc* 32:1397–1404
- Mizota T, Anton NE, Huffman EM et al (2020) Development of a fundamentals of endoscopic surgery proficiency-based skills curriculum for general surgery residents. *Surg Endosc* 34:771–778
- Mueller C, Kaneva P, Fried G, Feldman L, Vassiliou M (2014) Colonoscopy performance correlates with scores on the FES™ manual skills test. *Surg Endosc* 28:3081–3085
- Telem DA, Rattner DW, Gee DW (2014) Endoscopic simulator curriculum improves colonoscopy performance in novice surgical interns as demonstrated in a swine model. *Surg Endosc* 28:1494–1499
- Vassiliou MC, Kaneva PA, Poulou BK et al (2010) Global Assessment of Gastrointestinal Endoscopic Skills (GAGES): a valid measurement tool for technical skills in flexible endoscopy. *Surg Endosc* 24:1834–1841
- Dawe SR, Windsor JA, Broeders JA, Cregan PC, Hewett PJ, Madern GJ (2014) A systematic review of surgical skills transfer after simulation-based training: laparoscopic cholecystectomy and endoscopy. *Ann Surg* 259:236–248
- Martin JR, Anton N, Timsina L, Whiteside J, Myers E, Stefanidis D (2019) Performance variability during training on simulators is associated with skill transfer. *Surgery* 165:1065–1068

17. Haycock A, Koch AD, Familiari P et al (2010) Training and transfer of colonoscopy skills: a multinational, randomized, blinded, controlled trial of simulator versus bedside training. *Gastrointest Endosc* 71:298–307
18. Education ACfGM (2019) Defined category minimum numbers for general surgery residents and credit role
19. Vassiliou MC, Kaneva PA, Poulou BK et al (2010) How should we establish the clinical case numbers required to achieve proficiency in flexible endoscopy? *Am J Surg* 199:121–125
20. Dyke C, Franklin BR, Sweeney WB, Ritter EM (2021) Early implementation of Fundamentals of Endoscopic Surgery training using a simulation-based mastery learning curriculum. *Surgery*. 169(5):1228–1233
21. Franklin BR, Placek SB, Gardner AK et al (2018) Preparing for the American Board of Surgery Flexible Endoscopy Curriculum: development of multi-institutional proficiency-based training standards and pilot testing of a simulation-based mastery learning curriculum for the Endoscopy Training System. *Am J Surg* 216:167–173
22. Van Sickle KR, Buck L, Willis R et al (2011) A multicenter, simulation-based skills training collaborative using shared GI mentor II systems: results from the Texas association of surgical skills laboratories (TASSL) flexible endoscopy curriculum. *Surg Endosc* 25:2980–2986
23. Brinkman WM, Buzink SN, Alevizos L, de Hingh IH, Jakimowicz JJ (2012) Criterion-based laparoscopic training reduces total training time. *Surg Endosc* 26:1095–1101
24. Kolozsvari NO, Kaneva P, Brace C et al (2011) Mastery versus the standard proficiency target for basic laparoscopic skill training: effect on skill transfer and retention. *Surg Endosc* 25:2063–2070
25. Lineberry M, Park YS, Hennessy SA, Ritter EM (2020) The Fundamentals of Endoscopic Surgery (FES) skills test: factors associated with first-attempt scores and pass rate. *Surg Endosc* 34:3633–3643
26. Lineberry M, Ritter EM (2017) Psychometric properties of the Fundamentals of Endoscopic Surgery (FES) skills examination. *Surg Endosc* 31:5219–5227
27. Gallagher AG, Ritter EM, Champion H et al (2005) Virtual reality simulation for the operating room: proficiency-based training as a paradigm shift in surgical skills training. *Ann Surg* 241:364
28. (FES) TFoES. FES Test Centers (2021) Society of American Gastrointestinal and Endoscopic Surgeons (SAGES)
29. Seaman DL, de la Mora LJ, Gostout CJ, Rajan E, Herman L, Knipschild M (2007) An animal training model for endoscopic treatment of Zenker's diverticulum. *Gastrointest Endosc* 65:1050–1053
30. Swindle MM, Smith AC, Hepburn BJ (1988) Swine as models in experimental surgery. *J Invest Surg* 1:65–79

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.