Computerized feedback during colonoscopy training leads to improved performance: a randomized trial

Vilmann, Andreas Slot; Norsk, David; Svendsen, Morten Bo Søndergaard; Reinhold, Rasmus; Svendsen, Lars Bo; Park, Yoon Soo; Konge, Lars

Published in:
Gastrointestinal Endoscopy

Link to article, DOI:
[10.1016/j.gie.2018.07.008](https://doi.org/10.1016/j.gie.2018.07.008)

Publication date:
2018

Document Version
Peer reviewed version

[Link back to DTU Orbit](#)

Citation (APA):
Vilmann, A. S., Norsk, D., Svendsen, M. B. S., Reinhold, R., Svendsen, L. B., Park, Y. S., & Konge, L. (2018). Computerized feedback during colonoscopy training leads to improved performance: a randomized trial. *Gastrointestinal Endoscopy*, 88(5), 869-876. <https://doi.org/10.1016/j.gie.2018.07.008>

General rights

Copyright and moral rights for the publications made accessible in the public portal are retained by the authors and/or other copyright owners and it is a condition of accessing publications that users recognise and abide by the legal requirements associated with these rights.

- Users may download and print one copy of any publication from the public portal for the purpose of private study or research.
- You may not further distribute the material or use it for any profit-making activity or commercial gain
- You may freely distribute the URL identifying the publication in the public portal

If you believe that this document breaches copyright please contact us providing details, and we will remove access to the work immediately and investigate your claim.

Accepted Manuscript

Computerized feedback during colonoscopy training leads to improved performance: a randomized trial

Andreas Slot Vilmann, MD, David Norsk, M.Sc.Eng, Morten Bo Søndergaard Svendsen, M.Sc.Eng, Ph.d., Rasmus Reinhold, M.Sc.Eng, Ph.d., Lars Bo Svendsen, Prof., MD, DMSc, Yoon Soo Park, Ph.d., Lars Konge, Prof., MD, Ph.d.

PII: S0016-5107(18)32849-9

DOI: [10.1016/j.gie.2018.07.008](https://doi.org/10.1016/j.gie.2018.07.008)

Reference: YMGE 11149

To appear in: *Gastrointestinal Endoscopy*

Received Date: 13 December 2017

Accepted Date: 11 July 2018

Please cite this article as: Vilmann AS, Norsk D, Svendsen MBS, Reinhold R, Svendsen LB, Park YS, Konge L, Computerized feedback during colonoscopy training leads to improved performance: a randomized trial, *Gastrointestinal Endoscopy* (2018), doi: 10.1016/j.gie.2018.07.008.

This is a PDF file of an unedited manuscript that has been accepted for publication. As a service to our customers we are providing this early version of the manuscript. The manuscript will undergo copyediting, typesetting, and review of the resulting proof before it is published in its final form. Please note that during the production process errors may be discovered which could affect the content, and all legal disclaimers that apply to the journal pertain.

Cover page

Title: *Computerized feedback during colonoscopy training leads to improved performance: a randomized trial*

Andreas Slot Vilmann MD¹, David Norsk M.Sc.Eng², Morten Bo Søndergaard Svendsen M.Sc.Eng, Ph.d.¹, Rasmus Reinhold M.Sc.Eng, Ph.d.², Lars Bo Svendsen Prof., MD, DMSc⁴, Yoon Soo Park Ph.d.³, Lars Konge Prof., MD, Ph.d.¹

1. CAMES-Rigshospitalet, Copenhagen, Denmark
2. Technical University of Denmark, Kgs. Lyngby, Denmark
3. Department of the Medical Education, University of Illinois, Chicago, USA
4. Department of surgical gastroenterology C-Tx, Rigshospitalet, Copenhagen, Denmark

Corresponding author:

Andreas Slot Vilmann, MD.

CAMES-Rigshospitalet

Blegdamsvej 9

2100 Copenhagen

Denmark

Email: andreas.vilmann@gmail.com

Phone: +45 51 36 51 85

Title: Computerized feedback during colonoscopy training leads to improved performance: a randomized trial

Abstract

Background and Aims: Simulation-based training in colonoscopy is increasingly replacing the traditional apprenticeship method to avoid patient-related risk. Mentoring during simulation is necessary to provide feedback and to motivate, but expert supervisors are a scarce resource. We aimed to determine whether computerized feedback in simulated colonoscopy would improve performance, optimize time spent practicing, and optimize the pattern of training.

Methods: Forty-four participants were recruited and randomized to either a feedback group (FG) or a control group (CG). Participants were allowed 2 hours of self-practice where they could practice as they saw fit on 2 different cases: one easy and one difficult. The CG practiced without feedback but the FG was given a score of progression every time they reached the cecum. All participants were tested on a different case after end of training. The primary outcome was progression score in the final case and secondary outcomes were time spent practicing and training pattern.

Results: Regression analysis adjusting for sex was done due to an uneven sex distribution between groups ($P = 0.026$) and significantly higher performance scores by men (37.6, SD 25.9) compared with women (19.7, SD 18.7), $P = 0.012$. The FG outperformed the CG in the final case (FG scoring 14.4 points (95% CI, 1.2 - 27.6) more than the CG, $P = 0.033$) and spent more time practicing (FG practicing 25.8 minutes [95% CI, 11.6 - 39.9] more than the CG; $P = 0.001$). The FG practiced more on the easy case and reached the cecum 3.2 times more (95% CI, 2 - 4.5) during practice ($P < 0.001$).

Conclusions: Our findings of this study revealed that an automatic, computerized score of progression during simulated colonoscopy motivates the novices to improve performance, optimize time spent practicing, and optimize their pattern of training (Clinical trial registration number: NCT03248453).

Introduction

Colonoscopy is the criterion standard for diagnosing colorectal diseases. Novices handling the colonoscope impair patient-related safety, diagnostic accuracy, and cecal intubation rates (1,2). Traditionally, the training is done bedside but simulation-based training is increasingly replacing the initial introduction to avoid patient-related risks (3). Virtual reality colonoscopy simulators can provide the user with feedback (3,4) but lack tactile sensory realism. In contrast, physical simulators have been found to be more realistic but lack the feedback mechanism. Feedback is central in motivating and guiding the learner and is essential for all involved in education. How and when feedback is applied affects and motivates with different results (5). Evidence supports findings that feedback given by instructors compared with feedback from a simulator helps to increase performance (5), but drawbacks such as time spent by instructors and observer bias have led to the need for less-expensive, automated and objective systems.

Predefined proficiency level based on expert performances and not a fixed training time is currently the end goal for novices (6,7). Previous studies have explored how skills gained in a simulation-based setting transfer to the clinic but have not focused on how feedback affects performance during training (3,4,8). Currently, no one has investigated the influence on performance of receiving computerized feedback during simulated colonoscopy on a physical model.

The colonoscopy progression score (CoPS) is an automated and objective computerized score. CoPS was created to assess the progression of the tip of the endoscope by tracking it through the colon, calculating the progression score using a patented algorithm. The system has been validated in both a simulation-based and a clinical setting (9,10) and found to be positively correlated with patient-related pain during a colonoscopy (11).

We aimed to determine whether feedback from a computerized system could motivate and improve the novices' performance during self-practice in simulated colonoscopy on a physical model. We hypothesized that CoPS as feedback on a physical model versus no feedback motivates novices to improve performance, optimize time spent practicing, and optimize their pattern of training.

Methods

Participants:

Forty-four volunteer participants (interns and junior residents) were recruited for inclusion into this study. Participants were identified from mailing lists of graduate students within the last 2 years and advertisements on websites at the Copenhagen Academy for Medical Education and Simulation. Participants were included if they had no former experience with simulated or clinical colonoscopy. Demographics are shown in **Table 1**. All participants gave oral and written informed consent before enrolment. Participants were not compensated for their participation.

The simulator and equipment:

The Kyoto Kagaku Colonoscopy Training Model, a realistic physical phantom model, was used as a model for the human colon. The phantom can be configured into 6 standard cases, each with a different formation and difficulty. Two cases were chosen for training, one easy (Case 1, no loops) and one difficult (Case 4, alpha loop with a very flexible sigmoid), and the final case for testing (Case 3) had a more classic fixed alpha loop in the sigmoid. We used the Magnetic Endoscope Imaging (MEI, ScopeGuide, Olympus), to record the route of the colonoscope (CF-H180DL, Evis Exera II video center CV-180, Olympus Medical System Ltd, Tokyo, Japan) when participants performed colonoscopies. The established CoPS was measured after each colonoscopy (9,10). The CoPS is based on the movement of the tip of the colonoscope as recorded by the MEI displayed on a monitor (Figure 1). The score is reduced if the tip returns to a position already visited and increased if the tip continuously visits new positions (indicating progression along the colon).

Intervention and randomization:

Participants had no former experience with colonoscopy and were introduced by the same instructor to the colonoscope and the simulator, and were taught basic skills to handle the colonoscope correctly. All participants were given a standardized introduction of 30 minutes and randomized (by the instructor) in 1:1 ratio to either the feedback group (FG) or the control group (CG). One investigator prepared numbered, sealed and opaque envelopes for the randomization. Both groups were allowed a maximum of 2 hours of practice without supervision. Participants were free to switch between the easy and the difficult case or stop the practice session and proceed to the final test whenever they wanted.

Participants in the FG were given the CoPS as the only feedback each time they reached the cecum. They were encouraged to score as high as possible, and for comparison a leaderboard presenting high scores from experts in colonoscopy was present. The CG was not given any feedback during practice. All participants performed once on the final case where they were allowed a maximum of 30 minutes to reach cecum. The CoPS in the final case was registered as zero if the participants did not reach the cecum within the time limit. Neither of the groups was aware of the primary outcome measure in the final case.

The participants in the FG were asked to complete a questionnaire on motivation after the test. Figure 2 gives a summary of the study design.

Sample Size

The sample size was calculated based on data from a previous trial (9). We assumed that participants in the FG would have a mean score of 75 points and in the CG 40 points, with a standard deviation estimated to be 36 points. A significance level of 5% and a power of 0.9 required 22 participants in each group.

Performance measures:

The primary outcome was the actual CoPS in the final case, and the secondary outcomes were as follows: (1) Time spent during practice measured from the insertion of the colonoscope in the physical phantom model and until the participants voluntarily stopped or the time limit was reached and (2)

Training pattern measured as times the cecum was reached during practice on the easy and the difficult case, respectively.

Ethics

The regional committee of ethics evaluated and approved the study (no. H-16039935). All participants were provided with oral and written information regarding the trial. Participation was voluntary; no material goods were donated to the participants. The trial was registered (October 13, 2016) at clinicaltrials.gov with trial identification number: NCT03248453.

Statistics

Statistical analysis was done with IBM SPSS statistics (PASW, version 22; SPSS Inc, Chicago, Ill). Univariate analyses were conducted to examine the randomisation of participants in the control and feedback groups. Comparison between groups was made using independent sample t-test for continuous data and the Pearson chi-squared test for categorical data. To control for confounding factors, multiple regression analyses were conducted to adjust for confounding factors including gender, to estimate the main effect of intervention. The correlation coefficient was calculated and tested for significance to determine the probability of factors being related. Logistic regression was conducted to examine the effect of case difficulty between feedback and control groups. The level of statistical significance was set at $p < 0.05$ for all tests.

Results:

The data collection was conducted over a period of four months in 2016-2017. Forty-four physicians participated; there were no incomplete data sets in the final case. One participant in the CG ended the practice session early due to orthostatic hypotension, and one participant in the FG ended training early due to pregnancy-related nausea, but both were able to complete the final case. All participants reached cecum at least once during the practice session. Men scored significantly higher; men 37.6 (SD 25.9) and women 19.7 (SD 18.7), $P = 0.012$ (fig. 3). Although randomization of participants should have minimized confounding effects, we found an uneven sex distribution between groups (FG = 4 men and CG = 11 men, $P = 0.026$). To account for this issue, we used multiple regression analysis to adjust for sex and to estimate the intervention effect.

Multiple regression analysis: Three regression analyses adjusting for sex were conducted to examine the effect of the intervention with the following: (1) CoPS in the final case, (2) time spent practicing, and (3) number of times the learner reached cecum during practice, as outcome variables. For CoPS in the final case, the difference in CoPS was significant with the FG scoring 14.4 points (95% CI, 1.2 - 27.6) greater than the CG, $P = 0.033$. For time spent practicing, a t-test revealed a significant difference before adjusting for sex, with the FG practicing 103 minutes (SD 19) and CG practicing 74 minutes (SD 25), $P = 0.001$. Men (76 minutes, SD 25) spent less time during practice compared with women (95 minutes, SD 25) (Fig. 4). When adjusting for sex, the difference remained significant, with the FG practicing 26 minutes (95% CI, 11.6 - 39.9) more than the CG, $P = 0.001$; the trainees in the FG and the CG reached the cecum during practice 5.5 (SD = 2.4) times and 2.6 times (SD = 1.6), respectively, $P < 0.001$. Adjusting for sex, trainees in the FG were able to reach the cecum 3.2 more times (95% CI,

2 - 4.5) during training than the CG, $P < 0.001$. Interaction effects between intervention and sex were not significant, confirming the reported intervention effects.

Logistic regression: Logistic regression adjusted for sex was conducted to examine the likelihood of performing easy or difficult cases between the feedback and control group. Results show that the odds of the FG to perform the easy case in the first 5 cases was 2.1 times greater (95% CI, 1.1 - 4.1) than the CG, $P = 0.022$; the differences diminished when accounting for all 10 cases. Figure 5 shows the number of times participants reached cecum during practice.

Number of times the cecum was reached was positively correlated to the time spent practicing ($r = 0.4$, $P = 0.005$) and CoPS in the final case ($r = 0.4$, $P = 0.017$). Figure 6 summarizes the choice of case-difficulty during practice.

Nineteen out of 22 participants in the FG completed the questionnaire (86%). Responses revealed that 95% of the participants completing the questionnaire were motivated by the CoPS, but only 58% (11/19) felt motivated by the expert leaderboard because the expert scores felt unreachable. Nine participants (47%) believed they were affected by the CoPS when choosing case difficulty during practice.

Discussion:

Worldwide the numbers of simulation centers and simulation based training are increasing and it seems essential to investigate how to optimize the time spent on simulation based training without compromising the performance. Within the field of simulated endoscopy, training programs such as the

fundamentals of endoscopic surgery (FES) increasingly rely on simulation based practice before the trainees take the tests. The FES consist of 3 parts; didactic material in a web-based format, online written examination and a hands-on skill test on a simulator (12). Unsatisfactory pass-rates on the hands-on skill tests have led to the creation of simulation training systems and curricula (13). Self-practice represents a major part of the preparation before the trainee is ready to take the examinations, and our study represents a method to optimize the time spent when novices engage the initial steps of learning a colonoscopy, the progression from anus to cecum. This randomized controlled trial provides evidence that a simple score of progression during simulated colonoscopy self-practice leads novices to improve performance, optimize time spent practicing, and change their training pattern to a more favorable learning approach compared with a control group not receiving feedback.

In a study investigating the efficacy of a structured training program on insertion skills on the *Kyoto Kagaku Colonoscopy Training model* the participants spent 18 one-hour sessions spread across 9 weeks during which no performance related feedback was given (14). The performance related outcomes improved and the investigators concluded that future unsupervised training on the *Colonoscopy Training Model* is worthwhile. However, our study indicates that a feedback system has the potential to further improve the efficacy of training. Furthermore, a meta-analysis investigating the effect of different types of feedback in simulation-based medical education found that feedback given at the end of a training session increased long-term skill retention in novices compared with feedback given during a training session (5,15). Traditionally, supervisors are needed to provide feedback when training on realistic physical models using real colonoscopes. Supervisors are a scarce resource and our study shows that automatic, computerized feedback also has a positive effect. This finding aligns with

previous studies examining the effects of feedback in simulated training that demonstrated similar performance related outcomes regardless of the type of feedback (5).

Without feedback and the opportunity of comparing with expert scores, participants in the CG were not aware of the gap between their own and the expert's performance. The gap represented a reference-scale and gave the participants a goal but also forced a judgment of one's capability to progress within the time limit based on previous CoPS.

Time spent during practice was significantly lower in the CG compared with the FG. This could be overcome by having trainees practice until a predefined proficiency level but a performance measure is also necessary for this approach and reaching the predefined goals may not be cost-beneficial when comparing time spent with skills gained. A randomized trial investigated instructor feedback versus no instructor feedback in a virtual reality simulated laparoscopy and found reduced time spent reaching a predefined proficiency level when given standardized instructor feedback (16). They concluded that feedback given by the instructor helped the participants to be more time-efficient and focus their learning and thereby reaching the proficiency goals faster. The control group was not given feedback by an instructor, but the simulator gave them performance related information at the end of each completed scenario. Our primary focus was not how we make the trainees practice more, but rather how we optimize the time spent and still increase performance scores. One interpretation of this might be that without feedback or goals given by a simulator or instructor participants are not motivated to practice longer as seen in our CG.

Spending more time practicing seems a reasonable explanation for why the FG outperformed the CG in the final case, but the regression analysis examining the relationship between the intervention and time spent practicing revealed that time spent practicing in itself did not lead to the outperformance of the

CG. Participants in the FG repeated practicing on the easy case significantly more times than the CG. Participants in the FG deliberately chose to repeat practicing on the easy case (Figure 6 “Decision tree” generalizes decisions made based on group) as a consequence of getting feedback. Whereas the CG believed reaching the cecum on the easy case was the only success criteria and continued practicing on the difficult case instead. The expert leaderboard for score comparison made the FG aware of the performance gap and as a natural consequence repeated practicing on the easy case before taking a more technical challenge, the difficult case. A balance between the complexity and difficulty of a task has an impact on learner’s performance and learning (17). Progression from easy to difficult case scenarios during practice has previously shown improved performances, which partly explain why the FG outperformed the CG (17,18).

Our results show that men outperformed women. Evidence suggests that men have an initial advantage compared with women when learning new technical skills such as colonoscopy but that the difference (at least) evens out throughout following training programs (19). Reasons suspected are the ability to comprehend objects in a three-dimensional space (visual-spatial ability) and translate that into movement, and interest related to a simulator task such as video games and ball games during childhood (19). Former studies indicate that one-on-one training favors the performance related outcomes for women compared with self-practice (16,19). First, this partly explains why the women didn’t gain as much when given a computerized feedback as compared with men; second, participants had none or very little experience (average time since graduation was less than 1 year), and thus men had an advantage based on technical abilities. Performance and training related results at the level of residency is more complicated and do not show a clear difference between sexes, although men do have a tendency to take more risks and women are more thorough (20,21). A large retrospective study

investigating adenoma detection rate (ADR) characteristics among physicians found a difference of 4.2% ($P = 0.02$) between men and women, favoring the latter (20). These results support the theory that training and experience (at least) evens out the initial advantage that men may possess when learning a new technical skill.

Hand size has been proposed to be part of the explanation for men to perform better. In a survey by Cohen et al(22) investigating the relationship between hand size and endoscopic learning, 227 of 1295 gastroenterology fellows in the United States responded. Forty-one percent (93/227) considered their hand size too small for standard endoscope handle and 62% felt that small hand size affect performing endoscopy whereas 78% felt that learning was affected. They concluded that hand size is a factor when learning endoscopy and females, due to smaller hand size, would more likely be affected.

Limitations to this study are that the CoPS is a score of progression from anus and to cecum, and the score only assess the technical performance of the colonoscopy during insertion. Correct insertion skills of the colonoscope is more than just progression of the tip, but also relevant use of suction, inflation, loop management, and body position changes during the procedure (23). One might fear that trainees get bad habits when competing against a computerized score of progression, such as forceful progression, but a recent study exploring the correlation between pain measured on a visual analog scale on patients and CoPS revealed a positive correlation, meaning that increasing CoPS represents a smooth and even progression of the colonoscope and not forceful progression (11). However, it is important to acknowledge that a good insertion is not enough to ensure a good procedure. One of the major goals when performing a colonoscopy is to see as much mucosa as possible thereby increasing the chances of seeing a potential disease – CoPS does not measure this clinically important issue.

Currently, the feedback system gives a final score based on the movement of the tip but real-time feedback regarding the whole length of the colonoscope could give the opportunity of assessing every step of the procedure and identifying specific issues that the learner can improve. We did not ask participants of their former gaming experience and because a link between video games and increased technical abilities have been postulated, even though results are conflicting (24), this might have affected our results.

The sample size was large compared with other randomized training trials (8,17,25,26). All participants were interns or junior residents which should have reduced random errors such as performance linked interest and experience in endoscopy. However, including complete novices in endoscopy might have favored men as previously discussed. One participant in the FG had former experience with esophagogastroduodenoscopy (n=16), which might have improved performance in the final test. Eliminating this participant from data still reveals a significant difference between groups with the FG scoring 13.8 ($P = 0.043$; 95% CI, 0.45 - 27.25) greater the CG. Our randomization was not balanced regarding sex distribution which should be considered in future studies. The similarities of the difficult case and the final case might have favored participants spending most time practicing on the difficult case, but this was not apparent in our data analysis. When introducing automatic feedback systems it is important to acknowledge that it is necessary to master more than just the technical skills to be competent in colonoscopy (27). An important limitation is that we did not test transferability to the clinical world and our data are limited by the lack of validated metrics for the clinical setting. However, a systematic review and meta-analysis studying the relationship between simulations based assessment and patient-related outcome has revealed a positive correlation meaning that simulation leads to improved clinical performance (28).

In conclusion, this study revealed that a simple computerized score of progression during self-practice in simulated colonoscopy leads the novices to, increase time spent practicing, changes the training pattern to a more rewarding approach, and improve the performance. Applying the Colonoscopy Progression Score as feedback may diminish the financial and personnel cost in colonoscopy training.

References

1. Rabeneck L, Paszat LF, Hilsden RJ, Saskin R, Leddin D, Grunfeld E, et al. Bleeding and Perforation After Outpatient Colonoscopy and Their Risk Factors in Usual Clinical Practice. *Gastroenterology* 2008;135:1899–1906.e1.
2. Sedlack RE, Coyle WJ. Assessment of competency in endoscopy: Establishing and validating generalizable competency benchmarks for colonoscopy. *Gastrointest Endosc* 2016;83: 516–523.e1.
3. Haycock A, Koch AD, Familiari P, van Delft F, Dekker E, Petruzzello L, et al. Training and transfer of colonoscopy skills: a multinational, randomized, blinded, controlled trial of simulator versus bedside training. *Gastrointest Endosc* 2010;71:298–307.
4. McIntosh KS, Gregor JC, Khanna N V. Computer-based virtual reality colonoscopy simulation improves patient-based colonoscopy performance. *Can J Gastroenterol Hepatol* 2014;28:203–6.
5. Hatala R, Cook DA, Zendejas B, Hamstra SJ, Brydges R. Feedback for simulation-based procedural skills training: a meta-analysis and critical narrative synthesis. *Adv Heal Sci Educ* 2014;19:251–72.
6. Hashimoto DA, Petrusa E, Phitayakorn R, Valle C, Casey B, Gee D. A proficiency-based virtual reality endoscopy curriculum improves performance on the fundamentals of endoscopic surgery examination. *Surg Endosc Other Interv Tech* 2017;32:1–8.
7. Mueller CL, Kaneva P, Fried GM, Feldman LS, Vassiliou MC. Colonoscopy performance correlates with scores on the FES??? manual skills test. *Surg Endosc Other Interv Tech*.

2014;28:3081–5.

8. Grover SC, Garg A, Scaffidi MA, Yu JJ, Plener IS, Yong E, et al. Impact of a simulation training curriculum on technical and nontechnical skills in colonoscopy: a randomized trial. *Gastrointest Endosc* 2015;82:1072–9.
9. Nerup N, Preisler L, Svendsen MBS, Svendsen LB, Konge L. Assessment of colonoscopy by use of magnetic endoscopic imaging: Design and validation of an automated tool. *Gastrointest Endosc* 2015;81:548–54.
10. Preisler L, Bo M, Svendsen S, Søndergaard B, Brink L, Nordentoft T. Automatic and unbiased assessment of competence in colonoscopy : exploring validity of the Colonoscopy Progression Score (CoPS). *Endosc Int open* 2016;4:1238–43.
11. Preisler L, Bulut M, Svendsen MS, Svendsen LB, Konge L. An automatic measure of progression during colonoscopy correlates to patient experienced pain. *Scand J Gastroenterol* 2018;53:345–9.
12. Vassiliou MC, Dunkin BJ, Marks JM, Fried GM. FLS and FES: Comprehensive models of training and assessment. *Surg Clin North Am* 2010;90:535–58.
13. Gardner AK, Ujiki MB, Dunkin BJ. Passing the fundamentals of endoscopic surgery (FES) exam: linking specialty choice and attitudes about endoscopic surgery to success. *Surg Endosc Other Interv Tech* 2018;32:225–8.
14. Plooy AM, Hill A, Horswill MS, Cresp ASG, Karamatic R, Riek S, et al. The efficacy of training insertion skill on a physical model colonoscopy simulator. *Endosc Int Open*

2016;4:E1252–60.

15. Walsh CM, Ling SC, Wang CS, Carnahan H. Concurrent versus terminal feedback: it may be better to wait. *Acad Med* 2009;84(10 Suppl):S54–7.
16. Strandbygaard J, Bjerrum F, Maagaard M, Winkel P, Larsen CR, Ringsted C, et al. Instructor Feedback Versus No Instructor Feedback on Performance in a Laparoscopic Virtual Reality Simulator. *Ann Surg* 2013;257:839–44.
17. Grover SC, Scaffidi MA, Khan R, Garg A, Al-Mazroui A, Alomani T, et al. Progressive Learning in Endoscopy Simulation Training Improves Clinical Performance: A Blinded Randomized Trial. *Gastrointest Endosc* 2017;86:881-9.
18. Brydges R, Carnahan H, Rose D, Rose L, Dubrowski A. Coordinating progressive levels of simulation fidelity to maximize educational benefit. *Acad Med* 2010;85:806–12.
19. Ali A, Subhi Y, Ringsted C, Konge L. Gender differences in the acquisition of surgical skills: a systematic review. *Surg Endosc Other Interv Tech Springer US*; 2015;29:3065–73.
20. Mehrotra A, Morris M, Gourevitch RA, Carrell DS, Leffler DA, Rose S, et al. Physician characteristics associated with higher adenoma detection rate. *Gastrointest Endosc* 2018;87:688–694.e2.
21. Wallis CJ, Ravi B, Coburn N, Nam RK, Detsky AS, Satkunasivam R. Comparison of postoperative outcomes among patients treated by male and female surgeons: a population based matched cohort study. *British Med J* 2017;4366:j4366.
22. Cohen DL, Naik JR, Tamariz LJ, Madanick RD. The perception of gastroenterology fellows

- towards the relationship between hand size and endoscopic training. *Dig Dis Sci*. 2008;53:1902–9.
23. Kaminski M, Thomas-Gibson S, Bugajski M, Bretthauer M, Rees C, Dekker E, et al. Performance measures for lower gastrointestinal endoscopy: a European Society of Gastrointestinal Endoscopy (ESGE) Quality Improvement Initiative. *Endoscop*. 2017;49:378–97.
24. Lynch J, Aughwane P, Hammond TM. Video Games and Surgical Ability: A Literature Review. *J Surg Educ* 2010;67:184–9.
25. Ende A, Zopf Y, Konturek P, Naegel A, Hahn EG, Matthes K, et al. Strategies for training in diagnostic upper endoscopy: a prospective, randomized trial. *Gastrointest Endosc* 2012;75:254–60.
26. Snyder CW, Vandromme MJ, Tyra SL, Hawn MT. Proficiency-Based Laparoscopic and Endoscopic Training With Virtual Reality Simulators: A Comparison of Proctored and Independent Approaches. *J Surg Educ* 2009;66:201–7.
27. Sedlack RE, Shami VM, Adler DG, Coyle WJ, Degregorio B, Dua KS, et al. Colonoscopy core curriculum. *Gastrointest Endosc* 2012;76:482–90.
28. Brydges R, Hatala R, Zendejas B, Erwin PJ, Cook DA. Linking Simulation-Based Educational Assessments and Patient-Related Outcomes. *Acad Med* 2015;90:246–56.

Tables:**Table 1. Demographics and experience among participants**

Variable	Control Group (n = 22)	Feedback Group (n = 22)	P value
Mean age (range)	29 (25-38)	28.3 (26-32)	0.364
Sex (Female/Male)	F = 11 (50%) / M = 11 (50%)	F = 18 (82%) / M = 4 (18%)	0.026
Mean years since graduation (range)	0.8 (0-4)	0.9 (0-7)	0.912
Mean colonoscopies performed (range)	0 (0)	0 (0)	
Mean gastroscopies performed (range)	0.2 (0-3)	1.9 (0-16)	0.252

Figure legends:

Figure 1 – Two different procedures on the final case (case 3, alfa loop), exemplified by a colonoscopy progression map. Left: a low Colonoscopy Progression Score (CoPS) from a novice

(CoPS = 15) with uneven progression. Right: a high score from an expert (CoPS = 8556) with even progression.

Figure 2 - Summary of the study design.

Figure 3 - Colonoscopy Progression Score in the final case between groups and sex.

Figure 4 - Time spent practicing between groups and sex.

Figure 5 - x-axis: Number of times a participant reached cecum during practice. y-axis: Number of participants.

Figure 6 – Decision-tree of the training pattern showing how the participants in the two groups decided to practice. The intervention group performed the easy case several times before proceeding to the difficult case whereas the control group very quickly began training on the difficult case.

Andreas Slot Vilmann MD¹, David Norsk M.Sc.Eng², Morten Bo Søndergaard Svendsen M.Sc.Eng, Ph.d.¹, Rasmus Reinhold M.Sc.Eng, Ph.d.², Lars Bo Svendsen Prof., MD, DMSc⁴, Yoon Soo Park Ph.d.³, Lars Konge Prof., MD, Ph.d.¹

1. CAMES-Rigshospitalet, Copenhagen, Denmark
2. Technical University of Denmark, Kgs. Lyngby, Denmark
3. Department of the Medical Education, University of Illinois, Chicago, USA
4. Department of surgical gastroenterology C-Tx, Rigshospitalet, Copenhagen, Denmark

Corresponding author:

Andreas Slot Vilmann, MD.

CAMES-Rigshospitalet

Blegdamsvej 9

2100 Copenhagen

Denmark

Email: andreas.vilmann@gmail.com

Phone: +45 51 36 51 85

Disclosure Statement:

The intellectual property rights for the algorithm for analysing scope progression are protected by patent by Professor Lars Konge, Morten Bo Søndergaard Svendsen, and Professor Lars Bo Svendsen. Andreas Slot Vilmann, David Norsk, Rasmus Reinhold Paulsen and Yoon Soo Park disclosed no financial relationship relevant to this publication.

Author Contributions:

Study concept and design: Andreas Slot Vilmann and Lars Konge

Acquisition of Data: Andreas Slot Vilmann, Morten Bo Søndergaard Svendsen and David Norsk

Drafting of the article: Andreas Slot Vilmann

Statistical analysis: Yoon Soo Park, Andreas Slot Vilmann, and Lars Konge

Critical revision for important intellectual content: All authors

Final approval of the article: All authors

ACCEPTED MANUSCRIPT

Acronyms

CoPS = Colonoscopy Progression Score

3D-CoPS = 3D-Colonoscopy Progression Score

3D-CoRS = 3D-Colonoscopy Retractions Score

FG = Feedback Group

CG = Control Group

SD = Standard Deviation

P = alpha

95% CI = 95% Confidence Interval

MEI = Magnetic Endoscope Imaging