

# One step at a time: step by step versus continuous video-based learning to prepare medical students for performing surgical procedures

Tahmina Nazari, MD<sup>1,2</sup>

Floyd W. van de Graaf, MD<sup>1</sup>

Mary E.W. Dankbaar, PhD<sup>3,4</sup>

Johan F. Lange, MD, PhD<sup>4,5</sup>

Jeroen J.G. van Merriënboer, PhD<sup>5</sup>

Theo Wiggers, MD, PhD, FRCS<sup>2</sup>

*Journal of Surgical Education (2020) 77(4), 779-787*

<sup>1</sup> Department of Surgery, Erasmus University Medical Center, Rotterdam, the Netherlands

<sup>2</sup> Incision Academy, Amsterdam, the Netherlands

<sup>3</sup> The institute of Medical Education Research Rotterdam (iMERR), the Netherlands

<sup>4</sup> Department of Education, Erasmus University Medical Center, Rotterdam, the Netherlands

<sup>5</sup> Department of Surgery, IJsselland Hospital, Capelle aan den IJssel, The Netherlands

<sup>6</sup> Department of Educational Development and Research, Faculty of Health, Medicine & Life Sciences, Maastricht University, Maastricht, the Netherlands

## ABSTRACT

### Objective

The objective of this study was to compare the effects of cognitive load and surgical performance in medical students that performed the open inguinal hernia repair after preparation with step-by-step video-demonstration versus continuous video-demonstration. Hypothetically, the step-by-step group will perceive lower extraneous load during the preparation of the surgical procedure compared to the continuous group. Subsequently, fewer errors will be made in the surgical performance assessment by the step-by-step group, resulting in better surgical performance.

### Design

In this prospective study, participants were randomly assigned to the step-by-step or continuous video-demonstration. They completed questionnaires regarding perceived cognitive load during preparation (10-point Likert scale). Their surgical performance was assessed on a simulation hernia model using the Observational Clinical Human Reliability Assessment.

### Setting

Erasmus University Medical Center, Rotterdam, the Netherlands.

### Participants

Participants included medical students who were enrolled in extracurricular anatomy courses.

### Results

Forty-three students participated; 23 students in the step-by-step group and 20 in the continuous group. As expected, the step-by-step group perceived a lower extraneous cognitive load ( $2.92 \pm 1.21$ ) compared to the continuous group ( $3.91 \pm 1.67$ ,  $p = 0.030$ ). The surgical performance was not statistically significantly different between both groups; however, in subanalyses on a selection of students that prepared for 1 to 2 hours, the step-by-step group made less procedural errors,  $1.67 \pm 1.11$ , compared to the continuous group,  $3.06 \pm 1.91$ ,  $p = 0.018$ .

### Conclusions

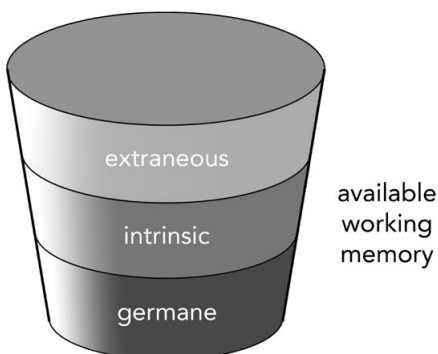
Our results suggest that preparation using step-by-step video-based learning results in lower extraneous cognitive load and subsequently fewer procedural errors during the surgical performance. For learning purposes, demonstration videos of surgical procedures should be presented in a segmented format.

## INTRODUCTION

Since the dawn of the digital age, surgical education has undergone an immense evolution, from its initial “master and apprentice” model in which apprentices learned from observing in the operating room to a time in which the 21st-century learner has the availability to learn by observing a multitude of online resources, for example, medical apps, books and videos.<sup>1</sup> Online videos are used frequently by medical students and residents and are known to be excellent tools to build anatomical and surgical knowledge.<sup>1-3</sup>

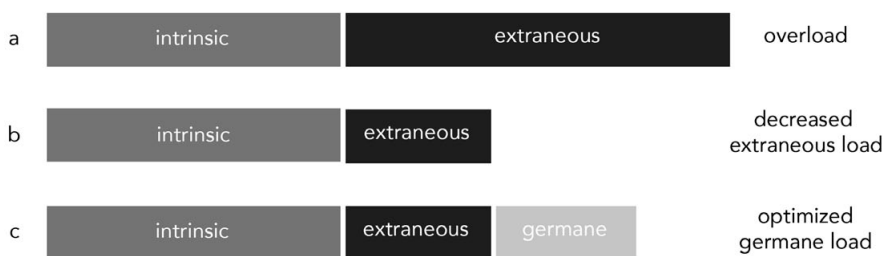
To understand how a trainee learns surgical procedures from observing videos, the limited cognitive capacity of the human brain must be taken into account. The cognitive capacity can be burdened when new and complex information is presented in a dynamic and transient format, as in a video-demonstration of a surgical procedure. To grasp the entire surgical procedure video-demonstration, the cognitive load can be high as disappearing information from the video needs to be retained and processed in working memory to understand the information that is presented in the video later.<sup>4</sup> Novices tend to learn better when this complex and transient information is presented in learner-paced segments, rather than as one continuous unit.<sup>5</sup> The learner-paced chunks result in lower perceived cognitive load and, subsequently, in potentially better learning.<sup>5,6</sup> In cognitive learning theory, this is referred to as the segmentation principle.<sup>7</sup>

The segmentation principle is an approach to prevent cognitive overload.<sup>5</sup> As shown in Figure 1, 3 types of cognitive load can be distinguished: intrinsic, germane, and extraneous cognitive load.<sup>8,9</sup> The complexity of new information determines intrinsic cognitive load. This type of cognitive load is higher for novices, and as the learner advances, the intrinsic cognitive load decreases. Germane cognitive load is determined by the construction and automation of cognitive schemas and is often categorized together with the intrinsic load.<sup>10</sup> Finally, extraneous cognitive load is determined by the suboptimal presentation of new information.<sup>11,12</sup>



**Figure 1.** Cognitive load types.

While processing new information, the total load of these 3 types of cognitive load cannot exceed the working memory available as the bucket in Figure 1 will overflow.<sup>8</sup> During simple tasks that yield low intrinsic cognitive load, the learner will be able to manage the task even if the extraneous cognitive load is high. On the contrary, during complex tasks, such as closely observing or performing a surgical procedure, the intrinsic load will be high. Therefore, the extraneous cognitive load should be reduced as much as possible so that learning and the corresponding germane load can still occur. Theoretically, as shown in Figure 2, unsegmented surgical procedure video-demonstration demands high extraneous load (Figure 2a). The application of the segmentation principle on video-based learning of surgical procedures would reduce the extraneous cognitive load because it provides additional processing time (Figure 2b). This extraneous load reduction gives more opportunity for germane processing (construction of cognitive schemas; Figure 2c), and subsequently improve the performance of the surgical procedure.<sup>13</sup>



**Figure 2.** Optimizing cognitive capacity: lowering extraneous load and providing opportunity for germane processing (adapted from Sweller 1998).

Segmenting surgical procedures into steps and substeps can be done in a standardized approach using our developed step-by-step framework.<sup>14</sup> A step is defined as a surgical goal that needs to be reached and evaluated before proceeding to the next step. A step consists of one or more substeps, a combination of anatomical structure with an action (for example, incise, transect, dissect, et cetera).

Surgical performance can be assessed using various methods. For a stepwise assessment, a validated option is the Observational Clinical Human Reliability Assessment (OCHRA).<sup>15</sup> The OCHRA is a systematic assessment checklist assessing errors on a substep level. Each substep could be assessed as “correct,” “procedural error,” or “executional error.” A substep is assessed as a “procedural error” when a substep was not performed, partially performed, repeated, or done out of sequence. Executional errors concern a substep performed with too much or too little force, speed, depth, or distance, or a substep executed in the wrong direction or on a wrong structure.

To investigate the effects of segmentation in video-based learning, the Lichtenstein open inguinal hernia repair (LOIHR) was chosen as an example surgical procedure as it is a complex procedure with multiple steps. Medical students prepared themselves using either a step-by-step video-demonstration or a continuous video-demonstration to perform the LOIHR surgery in a controlled environment using an open inguinal hernia repair simulation model.<sup>16</sup> The hypotheses are that the step-by-step group will perceive lower extraneous load during the preparation of the surgical procedure compared to the continuous group. Subsequently, fewer errors will be made in the surgical performance assessment by the step-by-step group, resulting in better surgical performance.

## MATERIAL AND METHODS

### Participants, Setting and Design

Medical students of Erasmus University Medical Center, Rotterdam, the Netherlands who were enrolled in extracurricular anatomy research courses, were approached for participation. The extracurricular anatomy research courses at Erasmus University Medical Center select their students on the grounds of significant interest and knowledge of surgical anatomy. Participation was voluntary, and written consent was gathered before the study. This study among medical students did not require institutional board review according to Dutch law.

During this prospective randomized trial, the participating medical students were randomly assigned to 2 groups; the step-by-step group (n = 23) or the continuous group (n = 20). Randomization was stratified per study year. Figure 3 shows the study design.

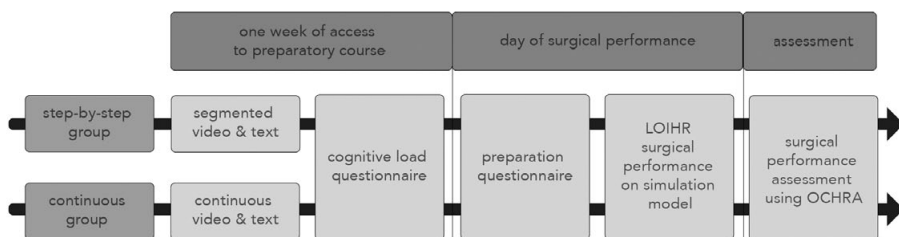


Figure 3. Study design.

### Step-by-Step Versus Continuous Preparatory Course

Before the participants performed the surgical procedure, they were granted 1 week of access to their assigned online preparatory course: the step-by-step or continuous online preparatory course.

The step-by-step group had access to the segmented video-demonstration alongside the associated textual description. The segmentation of the LOIHR video-demonstration and description consisted of 6 steps and 25 substeps, which were constructed using the step-by-step framework.<sup>14</sup> In this step-by-step course, the student was presented the video-demonstration one step at a time. After viewing the video-demonstration of one step (Figure 4a), the student had to press on the “next” button to continue to the next webpage to view the associated textual description of this step (Figure 4b). This process was repeated for all 6 steps (Table 1).

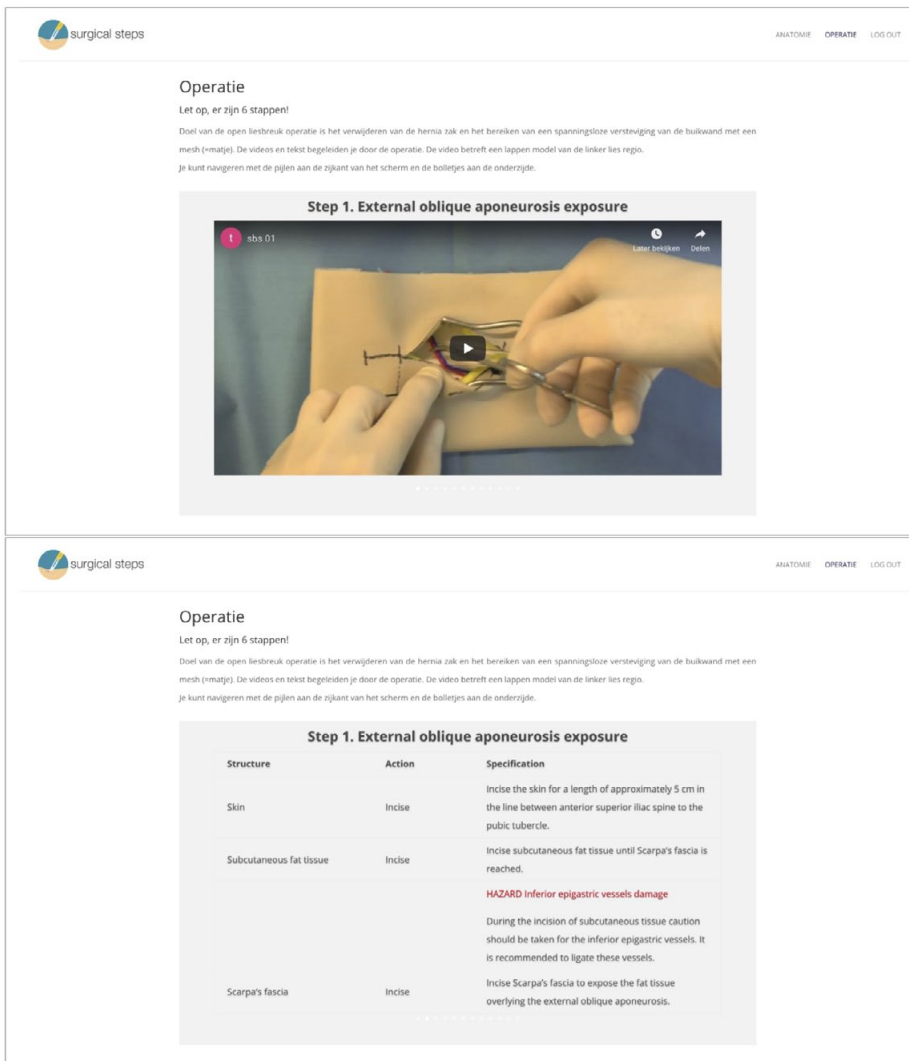


Figure 4. (a) Step by step video-demonstration and (b) textual description on the website.

**Table 1.** Duration Video-Demonstrations

Step-by-Step Video-Demonstration	Duration (mm:ss)
Step 1 External oblique aponeurosis exposure	01:38
Step 2 Inguinal canal exposure	00:30
Step 3 Spermatic cord mobilization	00:24
Step 4 Hernia sac removal	00:52
Step 5 Mesh placement	03:22
Step 6 Wound closure	01:02
Total duration	07:48

Continuous video-demonstration	Duration (mm:ss)
Total duration	07:30

The continuous group had access to a continuous video-demonstration of the LOIHR procedure and its associated textual description without segmentation. The continuous video-demonstration and textual description were displayed on separate webpages. After viewing the video-demonstration, the students could access the textual description of the procedure on a separate webpage in the online course by pressing on the “next” button.

The participants were allowed to study the online preparatory course at their own pace. The students could pause and rewatch the videos on demand. The content of the online courses (video-demonstrations and textual descriptions) were identical in both groups, with segmentation being the only difference.

### Cognitive Load Questionnaire

At the end of the online preparatory course, students were requested to fill out a questionnaire on their perceived cognitive load during the entire online course. A modified version of an existing questionnaire was used, composed of 12 statements assessing the intrinsic/germane cognitive load (8 statements) and the extraneous cognitive load (4 statements).<sup>11</sup> All statements were rated on a 10-point Likert type scale, ranging from 1 = totally disagree to 10 = totally agree.

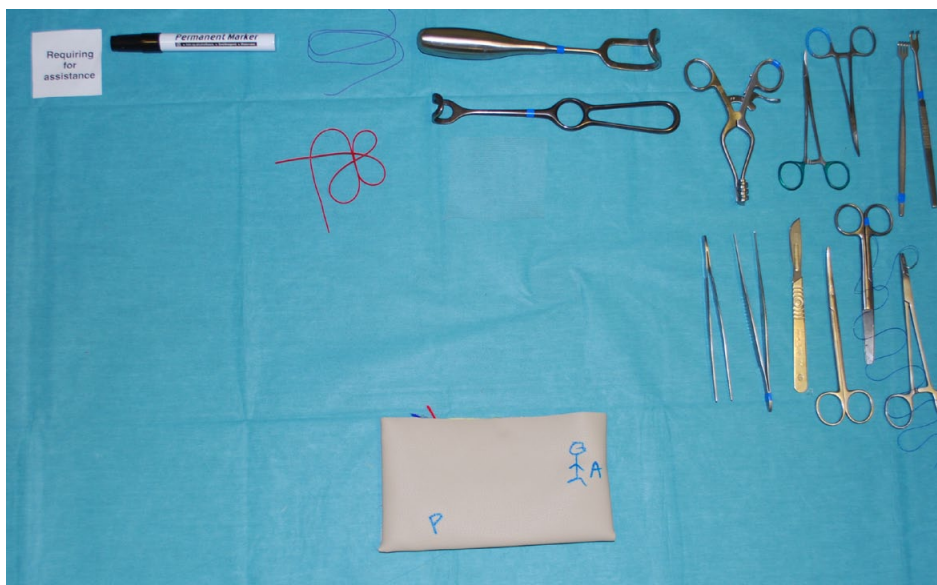
On the day of the surgical assessment, students were asked to fill out a questionnaire regarding their preparation (time spent on self-study during the online course in hours, use of other sources for self-study, and satisfaction during online preparation on a 10-point Likert scale, ranging from 1 = not at all to 10 = completely liking the teaching method).

### LOIHR Surgical Performance

All students performed the LOIHR surgical procedure on a simulation model.<sup>16</sup> This model mimicked the human abdominal wall anatomy, as each textile layer corresponded with a layer of the abdominal wall. The blood vessels, nerves (ilioinguinal, iliohypogastric, and geni-

tal branch of the genitofemoral nerve), the spermatic cord, and an indirect hernia sac were placed in the correct anatomical position within the textile layers. The simulation model used in the surgical performance assessment was identical to the model used in the preparatory video-demonstration.

To perform the LOIHR surgical procedure, each student received the necessary instruments and materials, such as a scalpel, forceps, scissors, retractor, mesh, needle driver, sutures, ligatures, marker, and a Penrose drain (Figure 5). The students had a maximum of 30 minutes to perform the LOIHR surgical procedure. The students were allowed to ask for help. Each time a student requested help regarding the execution or the correct order of the steps, this was flagged by one of the experimenters (TN or FvdG) as “requiring help.” Requests for an extra pair of hands by the students, such as cutting threads or holding retractors, were provided but not flagged as “requiring help.”



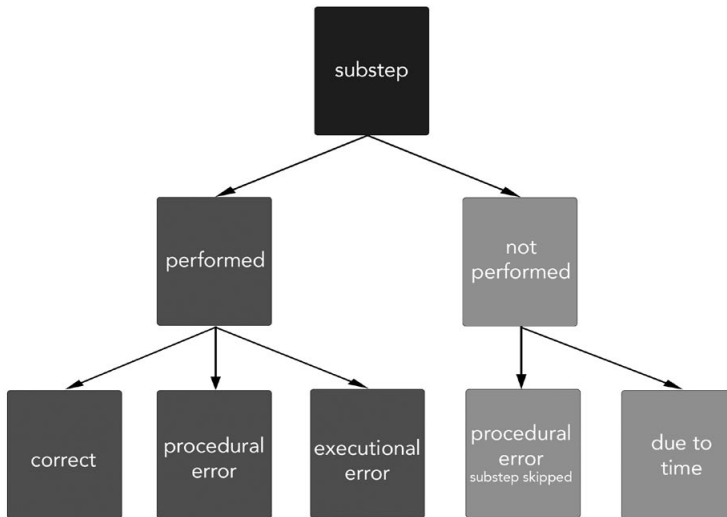
**Figure 5.** Set up operating table

### Surgical Performance Assessment

The LOIHR surgical procedures were video recorded using a head-mounted GoPro Hero 5 Black (GoPro Inc. San Mateo, California), with the following settings: resolution 720p, 60 frames per second; FOV: Narrow; White Balance 4000k; Locked exposure. The video recordings were anonymized and stored. Two trained assessors (TN, FvdG) were blinded for the randomization and reviewed the video recordings independently. Any discrepancies were discussed and reviewed by the 2 assessors and resolved through consensus. The assessment was done according to the principles of OCHRA.<sup>15</sup> As shown in Figure 6, a performed



substep could be assessed as “correct,” “procedural error” or “executorial error.” When the substep was not performed, this could be categorized as a “procedural error” if the students skipped this substep, or as “due to time” if it was caused by time constraints. The number of errors was registered for each medical student.



**Figure 6.** Assessment of a substep using observational clinical human reliability assessment.

### Statistical Analysis

Data were tested for normality using the Shapiro-Wilk test and presented as means and standard deviations, or as medians and interquartile ranges [Q1-Q3], according to their normality of distribution. If normal distribution was present, an independent samples *t*-test was used; otherwise, the Mann-Whitney *U* test was conducted. Categorical data were presented as numbers and percentages and compared using the Chi-square test. For the performance assessed by the OCHRA checklist, the mean of each category was presented. Subanalyses were performed on comparable subgroups of participants that spent 1 to 2 hours preparing the online course. A *p*-value of less than 0.05 was considered statistically significant.

Effect sizes were calculated using Cohen’s delta (*d*). Different formulas were used for parametric and nonparametric data.<sup>17</sup> Effect sizes of 0.20 were considered small,  $\geq 0.50$  were considered medium, and  $\geq 0.80$  were considered large.<sup>18</sup> The internal consistency was determined using Cronbach’s alpha ( $\alpha$ ). Data were analyzed with IBM SPSS Statistics for Windows (IBM Corp. Version 24.0, Armonk, New York).

## RESULTS

A total of 43 students participated in this study, of which 23 students were randomly assigned to the step-by-step group and 20 students to the continuous group. There were no statistically significant differences between the 2 groups regarding time spent during preparation, satisfaction during the preparation, and usage of other resources (Table 2).

**Table 2.** Total Group of Students – Demographics and Preparation

		Step by step (n=23)	Continuous (n=20)	p-value
Gender (n)	Female	13	9	.451 <sup>a</sup>
	Male	10	11	
Age in years (median [IQR])		20 [19-21]	20 [19-21]	.805 <sup>b</sup>
Year of study (n)	Year 1	6	5	.744 <sup>a</sup>
	Year 2	6	6	
	Year 3	8	7	
	Year 4	3	1	
	Year 5	0	1	
Time spent during preparation				
How much time did you spend studying the online course? (n)	0 - 1 hour(s)	5	2	.326 <sup>a</sup>
	1 - 2 hours	15	16	
	2 - 3 hours	1	2	
	3 - 4 hours	2	0	
Satisfaction during the preparation				
Over all, I appreciated the way the procedure was taught (median [IQR])	Scale 1-10	8 [7-9]	8 [6.25-8]	.053 <sup>b</sup>
I felt well prepared after watching the video and studying the text (median [IQR])	Scale 1-10	7 [6-8]	7 [4.50-8]	.487 <sup>b</sup>
Usage of other learning resources				
Did you, besides the online course, use other resources or materials to prepare for the surgery? (n)	Yes	12	12	.606 <sup>a</sup>
	No	11	8	
Which other different resources or materials did you use? (n)	Books	3	2	.758 <sup>a</sup>
	Other websites	4	6	
	Other videos	3	3	
	Other...	1 anatomy images	1 Google images	
How much time did you spend studying other resources or materials? (n)	0 - 1 hour(s)	10	12	.286 <sup>a</sup>
	1 - 2 hours	1	0	

IQR interquartile range [Q1 – Q3]

<sup>a</sup> analyzed using Chi square test

<sup>b</sup> analyzed using Mann Whitney U test

The perceived cognitive load and surgical performance are shown in Table 3. The mean (SD) extraneous cognitive load was perceived lower by the step-by-step group, 2.92 (1.21), than by the continuous group, 3.91 (1.67), with a medium effect size ( $t(41) = -2.24$ ,  $p = 0.030$ ,  $d = 0.68$ , Cronbach  $\alpha = 0.836$ ). The surgical performance was not significantly different between both groups on any of the measures. The median [Q1-Q3] satisfaction during preparation tended to be higher in the step-by-step group, 8 [7-9], than in the continuous group, 8 [6.25-8], with a small effect size ( $U = 153.00$ ,  $p = 0.053$ ,  $d = 0.09$ ).

Additional subanalyses were run on comparable subgroups that spent the same amount of time studying the preparatory course (1-2 hours). In this selection, gender, age, years of study, satisfaction during the preparation, and usage of other sources for preparation were not statistically significantly different between the groups (Table 4). As shown in Table 5, in the subanalyses, the step-by-step group perceived a lower level of extraneous cognitive load than the continuous group, with a medium effect size ( $t(29) = -2.091$ ,  $p = 0.045$ ,  $d = 0.75$ , Cronbach  $\alpha = 0.827$ ). Furthermore, the step-by-step group made less "performed – procedural errors," mean (SD) of 0.33(0.49), than the continuous group, 1.13 (1.09), with a small effect size ( $U = 65.00$ ,  $p = 0.018$ ,  $d = 0.15$ ).

**Table 3.** Total group of students – Performance and cognitive load

	Step by step (n=23)	Continuous (n=20)	$p$ - value
	Mean (SD)	Mean (SD)	
<b>Performance</b>			
Total performed - correct	7.30 (2.80)	7.75 (2.31)	.531 <sup>b</sup>
Total performed - incorrect	6.43 (2.00)	6.25 (1.91)	.759 <sup>a</sup>
Total performed - procedural error	0.39 (0.50)	0.90 (1.07)	.109 <sup>b</sup>
Total performed - executional error	6.00 (2.00)	5.25 (1.89)	.215 <sup>a</sup>
Total performed - procedural and executional error	0.04 (0.21)	0.10 (0.31)	.473 <sup>b</sup>
Total not performed	11.04 (3.23)	10.85 (2.13)	.763 <sup>a</sup>
Total not performed - due to procedural error (skipped)	1.48 (1.31)	1.70 (1.46)	.644 <sup>b</sup>
Total not performed - due to time	9.52 (3.18)	9.05 (2.31)	.109 <sup>b</sup>
Total procedural error combined (total performed – procedural error and tot not performed – procedural error)	1.87 (1.39)	2.60 (1.98)	.223 <sup>b</sup>
Total times asked for help	1.26 (1.57)	1.30 (1.63)	.868 <sup>b</sup>
<b>Cognitive load</b>			
Intrinsic/germane cognitive load Cronbach $\alpha = .807$	6.10 (1.17)	6.43 (1.10)	.351 <sup>a</sup>
Extraneous cognitive load Cronbach $\alpha = .836$	2.92 (1.21)	3.91 (1.67)	.030 <sup>a,*</sup>

<sup>a</sup> analyzed using independent samples t-test

<sup>b</sup> analyzed using Mann Whitney U test

\* statistically significant

**Table 4.** Students with 1-2 hours preparation – Demographics and preparation

		Step by step (n=15)	Continuous (n=16)	p-value
Gender (n)	Female	10	7	.200 <sup>a</sup>
	Male	5	9	
Age in years (median [IQR])		20 [19-21]	20 [19-21]	.896 <sup>b</sup>
Year of study (n)	Year 1	4	3	.764 <sup>a</sup>
	Year 2	4	6	
	Year 3	5	5	
	Year 4	2	1	
	Year 5	0	1	
Satisfaction during the preparation				
Over all, I appreciated the way the procedure was taught (median [IQR])	Scale 1-10	9 [8-9]	8 [7-8.75]	.090 <sup>b</sup>
I felt well prepared after watching the video and studying the text (median [IQR])	Scale 1-10	7 [7-8]	7 [4-8.75]	.340 <sup>b</sup>
Usage of other learning resources				
Did you, besides the online course, use other resources or materials to prepare for the surgery? (n)	Yes	7	10	.376 <sup>a</sup>
	No	8	6	
Which other different resources or materials did you use? (n)	Books	2	2	.752 <sup>a</sup>
	Other websites	2	4	
	Other videos	1	3	
	Other...	1 anatomy images	1 Google	
How much time did you spend studying other resources or materials? (n)	0 - 1 hour(s)	5	10	.182 <sup>a</sup>
	1 - 2 hours	1	0	

IQR interquartile range [Q1 – Q3]

<sup>a</sup> analyzed using Chi square test<sup>b</sup> analyzed using Mann Whitney U test

## DISCUSSION

Video-demonstrations create high extraneous cognitive load for managing the transiency of information as relevant information disappears quickly from the screen.<sup>6,9</sup> Segmentation provides smaller portions of information with pauses in between to reduce the extraneous load. In our study, this theory was affirmed as

the segmented step-by-step group showed a lower extraneous cognitive load compared to the continuous group. The intrinsic cognitive load was not statistically significantly different between the groups, as was expected since the complexity of the new information – the LOIHR surgical procedure for the medical students – was similar in both groups.

**Table 5.** Students with 1-2 hours preparation – Performance and cognitive load

	Step by step (n=15)		Continuous (n=16)		p-value
	Mean	SD	Mean	SD	
<b>Performance</b>					
Total performed - correct	7.80	2.43	7.19	2.20	.460 <sup>b</sup>
Total performed - incorrect	6.33	2.06	6.63	1.63	.667 <sup>a</sup>
Total performed - procedural error	0.33	0.49	1.13	1.09	.018 <sup>b,*</sup>
Total performed - executional error	6.00	2.17	5.44	1.63	.425 <sup>a</sup>
Total performed - procedural and executional error	0.00	0.00	0.06	0.25	.333 <sup>b</sup>
Total not performed	10.67	2.92	10.88	1.82	.815 <sup>a</sup>
Total not performed - procedural error (skipped)	1.33	1.18	1.94	1.48	.247 <sup>b</sup>
Total not performed - due to time	9.33	3.29	8.94	2.18	.286 <sup>b</sup>
Total procedural error combined (total performed – procedural error and tot not performed – procedural error)	1.67	1.11	3.06	1.91	.018 <sup>b,*</sup>
Total times asked for help	1.00	1.36	1.44	1.78	.531 <sup>b</sup>
<b>Cognitive load</b>					
Intrinsic/germane cognitive load Cronbach $\alpha = .827$	6.53	1.08	6.59	1.10	.879 <sup>a</sup>
Extraneous cognitive load Cronbach $\alpha = .827$	2.87	0.92	3.92	1.74	.045 <sup>b,*</sup>

<sup>a</sup> analyzed using independent samples T-test

<sup>b</sup> analyzed using Mann Whitney U test

\* statistically significant

When comparing students in our study with the same preparation time (1-2 hours), the step-by-step group made fewer procedural errors than the continuous group. Procedural errors are errors concerning the performance of the surgical procedure in the correct order and are determined by a trainee's procedural knowledge. A likely explanation for fewer procedural errors in the step-by-step group is that surgical knowledge was better learned while watching the segmented video leading to higher surgical performance compared to the continuous group. The executional errors were not significantly different between both groups. The executional errors concern surgical skills, such as knotting and suturing. Surgical skills are determined by repetitive practice and are therefore not solely dependable on video-based preparation.

To our knowledge, this is the first study to demonstrate the effects of segmentation of video-based surgical procedure learning on cognitive load and surgical performance. The findings of this study need to be viewed in light of several limitations. First, our prospective experimental design allowed students in both groups to pause and rewatch the video-demonstration on demand, similar to reality. The option to pause continuous videos effectively segments videos by providing smaller portions of information at a time. The continuous

group had thus the option to compensate for potential suboptimal teaching in this condition by investing more study time in preparation for the surgery (e.g., by pausing or rewatching the video, consulting other resources, et cetera). Additional subanalyses were therefore performed on the selection of students with the same preparation time of 1 to 2 hours in order to correct for potential compensation. This selection concerned the majority of the students, 31 of the 43 participating students.

In this study, the effects of segmentation were investigated in medical students as they form a homogeneous group with similar surgical experience and are more readily available compared to surgical residents. The next step is to investigate the segmentation effect in surgical residents. Finally, the segmentation in this study was performed using the step-by-step framework.<sup>14</sup> Further research is needed to investigate if the step-by-step framework offers the best way to define these segments.

## Conclusions

This study compared the effects of a step-by-step versus a continuous video-demonstration of a surgical procedure on perceived cognitive load and surgical performance. The step-by-step group perceived a lower extraneous cognitive load compared to the continuous group. Among students with the same preparation time (1-2 hours), the step-by-step group showed a lower extraneous cognitive load and higher performance, specifically, fewer procedural errors. Based on the findings in our study, we suggest presenting surgical video-demonstrations in a segmented format.

## REFERENCES

1. Evans CH, Schenarts KD. Evolving Educational Techniques in Surgical Training. *Surgical Clinics of North America*. 2016;96(1):71-88.
2. Glass NE, Kulaylat AN, Zheng F, et al. A national survey of educational resources utilized by the Resident and Associate Society of the American College of Surgeons membership. *The American Journal of Surgery*. 2015;209(1):59-64.
3. Barry DS, Marzouk F, Chulak-Oglu K, Bennett D, Tierney P, O’Keeffe GW. Anatomy education for the YouTube generation. *Anatomical Sciences Education*. 2015;9(1):90-96.
4. Mayer RE. Cognitive Theory of Multimedia Learning. *The Cambridge Handbook of Multimedia Learning*: Cambridge University Press; 2005:31-48.
5. Rey GD, Beege M, Nebel S, Wirzberger M, Schmitt TH, Schneider S. A Meta-analysis of the Segmenting Effect. *Educational Psychology Review*. 2019;31(2):389-419.
6. Spanjers IAE, van Gog T, van Merriënboer JGG. A Theoretical Analysis of How Segmentation of Dynamic Visualizations Optimizes Students’ Learning. *Educational Psychology Review*. 2010;22(4):411-423.
7. Mayer RE. Principles for Managing Essential Processing in Multimedia Learning : Segmenting, Pretraining, and Modality Principles. *The Cambridge Handbook of Multimedia Learning*: Cambridge University Press; 2005:169-182.
8. van Merriënboer JGG, Kirschner PA. Ten Steps to Complex Learning. *A systematic approach to four-component instructional design*: Taylor & Francis; 2013:22.
9. Sweller J. Element Interactivity and Intrinsic, Extraneous, and Germane Cognitive Load. *Educational Psychology Review*. 2010;22(2):123-138.
10. Sweller J, van Merriënboer JGG, Paas F. Cognitive Architecture and Instructional Design: 20 Years Later. *Educational Psychology Review*. 2019;31(2):261-292.
11. Leppink J, Paas F, van Gog T, van der Vleuten CPM, van Merriënboer JGG. Effects of pairs of problems and examples on task performance and different types of cognitive load. *Learning and Instruction*. 2014;30:32-42.
12. Leppink J, Gog T, Paas F, Sweller J. Cognitive load theory: researching and planning teaching to maximise learning. *Researching Medical Education*: John Wiley & Sons, Ltd; 2015:207-218.
13. Sweller J, Van Merriënboer JGG, Paas FGWC. Cognitive Architecture and Instructional Design. *Educational Psychology Review*. 1998;10(3):251-296.
14. Nazari T, Vlieger EJ, Dankbaar MEW, van Merriënboer JGG, Lange JF, Wiggers T. Creation of a universal language for surgical procedures using the step-by-step framework. *BJS Open*. 2018;2(3):151-157.
15. Tang B. Identification and Categorization of Technical Errors by Observational Clinical Human Reliability Assessment (OCHRA) During Laparoscopic Cholecystectomy. *Archives of Surgery*. 2004;139(11):1215.
16. Nazari T, Simons MP, Zeb MH, et al. Validity of a low-cost Lichtenstein open inguinal hernia repair simulation model for surgical training. *Hernia*. 2019.
17. Fritz CO, Morris PE, Richler JJ. Effect size estimates: Current use, calculations, and interpretation. *Journal of Experimental Psychology: General*. 2012;141(1):2-18.
18. Hojat M, Xu G. A Visitor’s Guide to Effect Sizes – Statistical Significance Versus Practical (Clinical) Importance of Research Findings. *Advances in Health Sciences Education*. 2004;9(3):241-249.