

## ORIGINAL RESEARCH

# Optimizing Residents' Performance of Lumbar Puncture: An RCT Comparing the Effect of Preparatory Interventions on Performance and Self-Confidence

Mikael Johannes Vuokko Henriksen, MD<sup>1,2</sup>, Troels Wienecke, PhD<sup>2,3</sup>, Helle Thagesen, MD<sup>3</sup>, Rikke Borre Vita Jacobsen, PhD<sup>2,4</sup>, Yousif Subhi, MD<sup>2,5</sup>, Ryan Brydges, PhD<sup>6,7,8</sup>, Charlotte Ringsted, PhD<sup>9</sup>, and Lars Konge, PhD<sup>1,2</sup>

<sup>1</sup>Copenhagen Academy for Medical Education and Simulation, Copenhagen, The Capital Region of Denmark, Denmark; <sup>2</sup>Faculty of Health and Medical Sciences, University of Copenhagen, Copenhagen, Denmark; <sup>3</sup>Department of Neurology, Zealand University Hospital, Roskilde, Denmark; <sup>4</sup>Center for Head and Orthopedic/UFU 4231 Anesthesiology, Rigshospitalet, Copenhagen, Denmark; <sup>5</sup>Department of Ophthalmology, Zealand University Hospital, Roskilde, Denmark; <sup>6</sup>The Wilson Center, University Health Network, Toronto, Canada; <sup>7</sup>Allan Waters Family Simulation Centre & Professorship in Technology-Enabled Education, St. Michael's Hospital, Toronto, Canada; <sup>8</sup>Department of Medicine, University of Toronto, Toronto, Canada; <sup>9</sup>Centre for Health Science Education, Faculty of Health, Aarhus University, Aarhus, Denmark.

**BACKGROUND:** Lumbar puncture is often associated with uncertainty and limited experience on the part of residents; therefore, preparatory interventions can be essential. There is growing interest in the potential benefit of videos over written text. However, little attention has been given to whether the design of the videos impacts on subsequent performance.

**OBJECTIVE:** To investigate the effect of different preparatory interventions on learner performance and self-confidence regarding lumbar puncture (LP).

**DESIGN:** Randomized controlled trial in which participants were randomly assigned to one of three interventions as preparation for performing lumbar puncture: 1) goal- and learner-centered video (GLV) presenting procedure-specific process goals and learner-centered information; 2) traditional video (TV) providing expert-driven content, but no process goals; and 3) written text (WT) with illustrations.

**PARTICIPANTS AND MAIN MEASURES:** Participants were PGY-1 doctors without LP experience. After the preparatory intervention, participants performed an LP in a simulated setting with a standardized patient and an assistant. Two content experts, blinded to participant group allocation, assessed video recordings of the performance using the Lumbar Puncture Assessment Tool (LumPAT) and an overall global rating. Participants rated their self-confidence immediately prior to performing the procedure. The primary outcome was the difference in LumPAT scores among groups.

**KEY RESULTS:** A total of 110 PGY-1 doctors were included. Results demonstrated significant differences in LumPAT mean scores among the three groups: GLV, 42.8; TV, 40.6; WT, 38.1 ( $p = 0.01$ ). The global rating scores were highest in the GLV group ( $p = 0.026$ ). Self-confidence scores

differed significantly among the three groups ( $p = 0.003$ ), with the TV group scoring the highest. There were no significant correlations between self-confidence scores and performance scores in any of the groups.

**CONCLUSION:** A video designed with procedure-specific process goals and learner-centered information resulted in better subsequent LP performance than a traditionally designed video or written text. Participants' self-confidence was not predictive of their actual performance.

**KEY WORDS:** medical education; medical education-clinical skills training; medical education-instructional design.

J Gen Intern Med 33(2):148-54

DOI: 10.1007/s11606-017-4207-7

© Society of General Internal Medicine 2017

## INTRODUCTION

Challenges in providing sufficient clinical instruction and supervision of invasive procedures have led to an increased use of supplementary multimedia sources, such as instructional videos.<sup>1,2</sup> Indeed, free online video-sharing websites have become increasingly popular among medical students and residents.<sup>3</sup> This trend will likely continue as video recording equipment and editing software become less expensive, and anyone will be easily able to produce videos and distribute them on the internet. Concerns have been raised regarding this widespread production and use of videos. First, studies assessing videos available on YouTube have identified generally low-quality and even dangerously misleading content.<sup>4,5</sup> Second, video design often fails to address pedagogical effectiveness,<sup>6</sup> and the principles underlying the design choices are rarely described in studies. Third, studies evaluating the effectiveness of video instruction in preparation for performing an invasive procedure have demonstrated conflicting results.<sup>7-10</sup> The Association of American Medical Colleges (AAMC) Institute for Improving Medical Education reported that research in educational technology is characterized by

**Electronic supplementary material** The online version of this article (<https://doi.org/10.1007/s11606-017-4207-7>) contains supplementary material, which is available to authorized users.

?

Received April 21, 2017

Revised August 30, 2017

Accepted October 11, 2017

Published online November 13, 2017

deficiencies in either methodological approach, conceptual framework, or both.<sup>11</sup> That report recommended that researchers investigate the selection, sequencing and presentation of information in instructional videos.

The formats of preparatory interventions (PIs) based on text and video are by definition a one-way presentation of information, without provision of feedback. These limitations underscore the need for accurate and unambiguous choices on the content presented. PIs are often produced by content experts. However, because performing the procedure is so perfunctory for these experts, they may omit key elements that are essential for novice learners, thus compromising the effectiveness of the intervention for student learning.<sup>12</sup>

The challenges with respect to pedagogical design can be addressed by two evidence-based principles for optimizing the pedagogical potential of the PI for clinical procedures. First, preparatory interventions including goal orientation have been found to be effective.<sup>13</sup> Models of goal setting broadly distinguish two strategies: a process-goal orientation focusing on the strategies used to complete a task, and outcome-goal orientation focusing on the product or end results.<sup>14</sup> Presenting task-specific process goals to novices prior to their performance of a task is reported to be more effective than presenting outcome-specific goals,<sup>15</sup> which indicates that the use of task-specific process goals may optimize the pedagogical potential of instructional videos intended for invasive procedure preparation.

Second, the content incorporated in the PI should present only the knowledge needed to improve novices' performance<sup>16</sup> and should avoid presentation of extraneous information.<sup>17</sup>

Lumbar puncture (LP) is a complex procedure<sup>18</sup> which is crucial for the diagnosis of serious medical conditions, and frequently must be performed on an urgent basis.<sup>19</sup> Junior doctors express uncertainty about performing the LP,<sup>20,21</sup> and perform below stakeholder expectations.<sup>22</sup> As residents often use video and text to prepare for performing procedures on patients, the effectiveness of different formats should be investigated. Although instructional videos have been reported to improve operator self-confidence,<sup>7,23</sup> the correlation with actual clinical performance demonstrates conflicting results.<sup>7,23</sup> Hence, there is a need to explore how different instructional designs affect novice self-confidence, and how their confidence correlates with performance. We hypothesized that the performance of participants who prepared using a goal-setting/learner-centered video would be superior to that of groups using a traditionally designed video or instructional text. The aims of this study were to investigate the effect of three different formats of preparatory instructional material—a video including goal-setting and learner-centered information, a traditionally designed video, and written instructional text—on novice doctors' LP performance and self-confidence in a simulated clinical setting, and whether the different interventions influenced the relation between self-confidence and performance.

## METHODS

### Study Design

This study was an interventional, single-blinded randomized controlled trial with three arms. A CONSORT diagram is provided in Figure 1.

### Participants

We recruited postgraduate-year 1 (PGY-1) doctors with less than 6 months of postgraduate clinical experience and with no previous experience in performing LP.

Participants were recruited over a 12-month period by email invitation to new graduates from the medical school at the University of Copenhagen and to junior doctors enrolled in first-year postgraduate training in one of the ten hospitals in eastern Denmark. As part of the enrollment, participants used a web-based portal to provide demographic information and booked a date for participation. Throughout the inclusion period, as participants approached their designated date, they were cluster-randomized by LK upon enrollment using ([Random.org](http://Random.org)).<sup>24</sup>

### Interventions

For the study, we designed two instructional videos and a written instructional text. As participants in our study were non-English speaking, we found it most fair to design both videos in the local language and according to local standards for performing LP. Participants prepared individually with one of the interventions for 15 min. Participants were allowed to repeat the video at their discretion within the time limit.

### Goal- and Learner-Centered Video (GLV)

The design of the goal- and learner-centered video (GLV) was based on a previous study of learner-requested information and the differences between experts and novices with respect to their approach to performing LP.<sup>25</sup> The results demonstrated that expert performers contextualized the procedure by integrating both the technical and non-technical aspects related to the patient, the clinical surroundings, the equipment, and the doctors' own preparedness to perform. By contrast, novices express a fear of causing damage and focused on the outcome goal of getting the sample,<sup>25,26</sup> and they requested information about anatomy related to needle insertion.

Based on these results, MJVH, YS and CR designed the video, structured as follows: anatomy and needle insertion, process goals for the procedure, attention to positioning of the patient and communication during the procedure, equipment, the procedure, and problem-solving. The video was 7 min long.

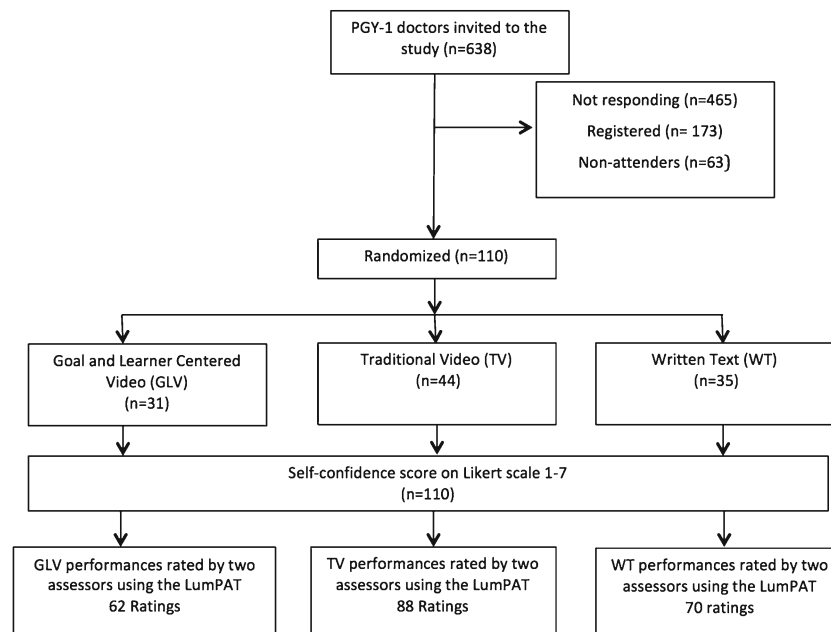


Figure 1 CONSORT flow diagram of the study.

## Traditional Video (TV)

The design of the traditional video was inspired by the content and structure of a frequently cited instructional video for the LP procedure from *The New England Journal of Medicine (NEJM)*.<sup>27</sup> The *NEJM* video focuses on the technical aspects of performing the procedure and does not address any process goals.

We developed the TV to include the following elements: indications, contraindications, equipment, positioning, the procedure, and anatomy. No elements of patient communication or strategy for performing the procedure were presented. This video was six minutes long.

An example of the difference between process goals and outcome goals in video design is exemplified by how each video presents the identification of the needle insertion point. The TV instructs the learner to identify the L4 level and subsequently the interspaces between the superior and inferior vertebrae. In contrast, the GLV instructs the learner to palpate the iliac crest, and the L4 level is then identified by bringing the thumbs together. The optimal needle insertion point is identified by palpating the spinous processes in a caudal to sacral direction, maintaining awareness of the patient's back and the spaces between the vertebrae.

## Written Text (WT)

The written text described the steps for performing the procedure with three illustrations,<sup>28</sup> and was six pages long (pocket size).

## Assessment of Self-Confidence

After the preparatory intervention, the participants rated their self-confidence on a seven-point Likert-scale, based on the

following question: "How confident do you feel in the procedure you are about to perform?" (1 = very unconfident to 7 = very confident).

## Study Setting

We established a hybrid simulated setting. The participants performed the LP individually and were instructed to act as they would in a real clinical setting, including engaging with the patient. We chose to limit participant performance to a single attempt of the procedure to mimic the clinical context where trainees access instructional materials and subsequently perform the procedure on a patient. We provided the participants with authentic medical records including laboratory results and computed tomography (CT) brain results. We created a ward-like setting including a standardized patient (SP), a standardized assistant, and an LP phantom (Kyoto Kagaku Lumbar Puncture Simulator II; Kyoto Kagaku Co., Ltd., Kyoto, Japan). The phantom simulated the lumbar anatomy including landmarks for palpation, lifelike skin and tissue resistance, and the possibility of obtaining a fluid sample. The SP was instructed to express moderate fear and being uninformed about the procedure. The phantom was introduced when participants had positioned and marked their insertion point(s) on the SP. The procedure continued with the SP having the phantom strapped to her back, and thus communication continued during the remainder of the procedure. Participants who failed to obtain fluid had to inform the SP and terminate the session.

## Assessment

For assessment of LP performance we used the previously developed Lumbar Puncture Assessment Tool (LumPAT),

which has demonstrated acceptable inter-rater reliability and positive correlation with other measures of procedural performance in a similar population of residents.<sup>25</sup> The LumPAT is a global rating scale that assesses both technical aspects of performance and non-technical aspects, such as patient communication. The LumPAT scale ranges from 11 to 55 points, with a previously established standard of 44 points required to pass (see online appendix for the LumPAT). In addition, the raters assigned a global judgment score for the entire performance based on a seven-point Likert scale. All procedures were video-recorded. The anonymized videos, as well as the assessment tool, were distributed to the two raters through a web-based system.<sup>29</sup> The raters were content experts—consultants and associate professors in anesthesia (RVBJ) and neurology (HT). We further documented whether participants were successful in obtaining liquid from the mannequin.

## Statistical Analysis

All continuous variables were evaluated for normal distribution. Normally distributed variables were reported as a mean and standard deviation (SD), and were tested using parametric tests. Non-normally distributed variables were reported with median and interquartile range (IQR), and tested using non-parametric tests.

We compared baseline characteristics among the three groups. Age was compared using one-way analysis of variance (ANOVA), and gender distribution was compared using the chi-square test. Due to a lack of normal distribution, the number of observed LP procedures was compared using the Kruskal–Wallis test.

Inter-rater agreement was evaluated with the intraclass correlation coefficient (absolute agreement definition).

To explore the effect of the three different PIs, we compared performance scores among the three groups using the mean of the raters' scores on all the LumPAT items, and subsequently the mean of the raters' global ratings on the seven-point Likert scale, using one-way ANOVA. As post hoc tests to identify differences between groups, we used independent-samples *t* tests. Effect sizes for differences between groups were calculated using Cohen's *d*, based on mean values for LumPAT items and global ratings.

The chi-square test was used to compare the difference between success and failure in obtaining liquid among the groups. To assess the relative effect of passing the established standard of the LumPAT, we defined WT as the reference for calculating the relative risk (RR) of a passing score for the other groups using a chi-square test.

Self-confidence scores were not normally distributed, and thus we used nonparametric tests. To explore whether age and the number of observed LP procedures correlated with self-confidence scores, we used Spearman's correlation. Self-confidence scores were compared by gender using the Mann–Whitney *U* test. Differences in self-confidence among

groups were compared using the Kruskal–Wallis test. Individual groups were compared with one another using the Mann–Whitney *U* test.

We used Spearman's correlation to explore the correlation between self-confidence scores and both the mean LumPAT score of the two raters and the mean of the raters' global judgment scores for each of the three groups.

Statistical analysis was performed using SPSS version 22.0 software (IBM Corp., Armonk, NY, USA).

## Ethics

This study contained no patient participation or collection of biological material. The project was presented to the local ethics committee of Capital Region Denmark, who waived the need for further approval (journal number: H-15018242). All participants were informed verbally and in writing regarding the purpose, that the study was voluntary and entirely anonymous, and that it would not impact their future clinical appointments. All participants completed written informed consent.

## RESULTS

A total of 110 participants completed the study. Data on baseline characteristics of participants are presented in Table 1, and demonstrate no differences among the three groups.

Intraclass correlation coefficients for the LumPAT mean score were 0.71 and 0.69 for the raters' global ratings.

There were significant differences in the LumPAT mean scores among the three study groups ( $p = 0.01$ ; Table 2). Post hoc analysis demonstrated that the GLV group had a significantly higher mean LumPAT score than the WT group, with a medium effect size ( $d = 0.73$ ,  $p = 0.02$ ). There was also a significant difference among the global ratings of the three groups ( $p = 0.026$ ). Post hoc analysis demonstrated that the GLV group scored higher than both the TV and WT groups.

There was a significant difference in self-confidence scores among the three groups ( $p = 0.003$ ). Post hoc analysis demonstrated that the TV group had a significantly higher self-confidence score than the WT group, with a high effect size

Table 1 Participant Characteristics

	Goal- and learner-centered video ( $n = 31$ )	Traditional video ( $n = 44$ )	Written text ( $n = 35$ )	<i>p</i> -value
Age, years, mean (SD)	32.8 (4.1)	31.8 (2.3)	32.5 (3.1)	0.36
Gender, no. (%)				0.14
Female	25 (81)	27 (61)	21 (60)	
Male	6 (19)	17 (39)	14 (40)	
No. of observed LP procedures, median (IQR)	2 (1–3)	3 (1–4)	2 (1–4)	0.44

LP lumbar puncture, IQR interquartile range



**Table 2 Results of Performance Tests for LumPAT Score, Rater Global Judgment, Participant Self-Confidence Scores and Number of Participants Successfully Obtaining Liquid**

	Goal- and learner-centered video (n = 31)	Traditional video (n = 44)	Written text (n = 35)	p-value
LumPAT total score, mean (SD)	42.8 (6.3)	40.6 (5.8)	38,1 (6.7)	0.01
Rater global judgment score	4.4 (1.2)	3.8 (1.2)	3.6 (1.2)	0.026
Self-confidence score, mean (SD)	3.8 (1.7)	4.3 (1.4)	3.1 (1.4)	0.003
Number of participants successfully obtaining liquid	20	21	20	0.34

LumPAT score is distributed from 11 to 55 points. Rater global judgment scores are based on a seven-point (1–7) Likert scale. Participant self-confidence scores are rated on a seven-point Likert scale (1 = not confident, 7 = very confident). Number of successful LP procedures is given in absolute numbers

( $d = 0.85$ ,  $p = 0.002$ ). The GLV group had a higher median score than the WT group ( $p = 0.046$ ).

The percentage of participants who successfully obtained liquid was highest for the GLV group and lowest for the TV group, but the difference was non-significant (see Tables 2 and 3).

The relative effect of the different formats revealed that the GLV group was 2.6 times as likely to obtain a passing score as the WT group ( $p = 0.02$ ). There was no difference between TV and WT (see Table 4).

There were no significant correlations between participants' self-confidence scores and age ( $\rho = -0.02$ ;  $p = 0.85$ ; Spearman's correlation), number of previously observed LP procedures ( $\rho = 0.07$ ;  $p = 0.44$ ; Spearman's correlation), or gender ( $p = 0.22$ ; Mann–Whitney  $U$  test).

**Table 3 Results of Post Hoc Test Demonstrating the Differences Between Groups in the LumPAT Scores, Rater Global Judgment and Participant Self-Confidence Scores**

	Mean difference (95% CI of difference)	Cohen's $d$	p-value
LumPAT score			
GLV vs. TV	2.23 (−0.56 to 5.04)	0.49	0.12
GLV vs. WT	4.75 (1.53 to 7.97)	0.73	0.004
TV vs. WT	2.52 (−0.28 to 5.31)	0.40	0.08
Rater global judgment score			
GLV vs. TV	0.56 (0.02 to 1.10)	0.48	0.04
GLV vs. WT	0.78 (0.19 to 1.38)	0.64	0.01
TV vs. WT	0.22 (−0.31 to 0.76)	0.19	0.41
Self-confidence score			
GLV vs. TV	−0.48 (−1.19 to 0.24)	0.31	0.19
GLV vs. WT	0.72 (−0.041 to 1.48)	0.46	0.06
TV vs. WT	1.19 (0.56 to 1.82)	0.85	<0.001

GLV goal- and learner-centered video, TV traditional video, WT written text

LumPAT score is distributed from 11 to 55 points. Rater global judgment score is based on a seven-point (1–7) Likert scale. Participant self-confidence score is based on a seven-point Likert scale (1 = not confident, 7 = very confident)

**Table 4 Consequences of Different Formats of Pre-Training in Meeting the Standard for Lumbar Puncture**

	Passed, no. (%)	RR	CI 95%	p-value
Goal- and learner-centered video	14 (45)	2.6	tbcolw30pt1.2–6.0	0.02
Traditional video	10 (23)	1.3	0.5–3.3	0.54
Written text	6 (17)	Ref	Ref	Ref

Chi-square test showed a significant difference between groups in the distribution of participants who passed ( $p = 0.026$ ). Control: written text is used as the reference to calculate the RR, e.g. pre-training with goal- and learner-centered video is 2.6 times as likely to result in a passing score as the control: written text

The correlations between self-confidence scores and mean LumPAT scores were non-significant for all three groups: GLV group correlation coefficient = 0.14,  $p = 0.46$ ; TV group correlation coefficient = 0.015,  $p = 0.93$ ; WT group correlation coefficient = 0.19,  $p = 0.27$ . The correlations between self-confidence scores and mean rater global judgment scores were non-significant for all three groups: GLV group correlation coefficient = 0.19,  $p = 0.31$ ; TV group correlation coefficient = −0.12,  $p = 0.43$ ; WT group correlation coefficient = 0.05,  $p = 0.77$ .

## DISCUSSION

This study demonstrates that a video based on a learner-centered design, including a demonstration of procedure-specific process goals, significantly improved immediate performance of a simulated LP compared to written text instruction. However, such benefits were not found for a video based on a traditional design. We further demonstrated that learners' self-confidence did not correlate with their performance, irrespective of the format used for preparation. Participants preparing with the GLV were 2.6 times as likely as those using the written text to meet the established standard for readiness to perform the procedure in clinical practice. As instructional videos are often used as just-in-time preparation for clinical procedures, the present study suggests that doctors using the GLV would be better prepared for performing standard accepted LP techniques and patient-centered care than those using traditionally designed videos or text instructional material.

The strength of this study is that participants were residents with limited procedural experience, as this group frequently uses either text- or video-based instruction as preparation for performing clinical procedures.<sup>30</sup> The study demonstrated sufficient inter-rater reliability<sup>31</sup> and used an assessment tool specific to the procedure with an established standard for readiness to perform the procedure independently.<sup>25</sup>

The outcomes for the GLV group were significantly superior to those for the WT group, whereas those of the TV group were not. Previous studies on the effect of videos have not provided rich descriptions of the choices made when determining the means of delivering the content of their videos. Our study shows that reporting such details is necessary, as these

design and content decisions can have a significant impact on participant performance. These factors, in turn, may explain the conflicting results among previous studies exploring the effect of instructional videos.<sup>7-10</sup>

The significant difference in rater judgment between the GLV and TV groups suggests that important clinical contextualization can be provided by the instructional video format. The integration of non-technical and technical aspects of a procedure has been found to be beneficial for simulation-based training.<sup>32</sup> In our study, this contextualization, along with the presentation of the process goals, may explain the 2.6-fold greater chance of meeting the established performance standard with the GLV versus WT approach. However, we found no significant difference in the ability to obtain liquid.

The identification of a significant difference in self-confidence scores between the TV and WT groups, but not between the GLV and WT groups, demonstrates that the design and content of the video format influence learners' self-confidence but do not correlate with performance. Several studies have described a lack of congruence between objective performance scores obtained from simulation-based assessment and self-ratings of confidence,<sup>33,34</sup> including among residents performing infant LP.<sup>7</sup>

There are several clinical implications of this study. We find that the concept of presenting procedure-specific process goals can be easily integrated into the design of instructional videos, and we therefore recommended this approach. This additional effort is believed to be cost-effective, as the learners subsequently have a greater chance of achieving the standard for performing the procedure. Further, we recommend that developers of instructional videos examine differences in procedure-related perceptions and strategies for performance between expert and novice performers. Identifying this gap is important, because content requests from novices frequently differ from those of experts. As a result, there is a risk that content experts could unintentionally overlook important needs in their teaching strategies for clinical procedures.<sup>12</sup> Another implication of the study is that a video developer's evaluation of PI should not rely solely on learner self-rating of confidence, but rather on objective assessment of performance.

There are some notable limitations to the study and in the interpretation of the results. The differences between the GLV and TV groups in LumPAT scores and in the percentage successfully obtaining liquid were non-significant, which could be due to a type II error. It was not possible to perform a power calculation for the present study, but we felt that the sample sizes were large enough that any significant differences would be clinically relevant. Novice performers, as those in the present study, tend to have a high level of variance in their performance,<sup>35</sup> and therefore an additional study limitation is that we only evaluated a single one-off performance after the intervention. Further, the long-term effect of the different formats was not explored with a retention test. Nonetheless, studies comparing the instructional design of outcome versus process goals for simulation-based training found no differences between groups

for 1-week retention,<sup>36</sup> and no differences were found between post-training and a 6-week retention test for dyad versus solo training.<sup>37</sup> The GLV included presentation of procedure-specific process goals and learner-centered information, making it difficult to discern the specific mechanism for the observed effect. However, our aim was to optimize the pedagogical effectiveness of PIs, and hence the GLV should be considered as a single entity. It is also important to note that this study did not explore the effect of the intervention on clinical performance. However, this limitation is minimized by the fact that the LumPAT was previously developed and validated for assessing residents' readiness to perform LP independently, and the established standard reflects this aim.<sup>25</sup> Moreover, meta-analysis of the correlation between simulation-based assessments and patient-related outcomes has shown that these measures correlate well if based on rigorously established assessment tools.<sup>38</sup>

## CONCLUSION

We have identified that an instructional video with a learner-centered approach and procedure-specific process goals was associated with significantly better participant performance in comparison to written instructions. A video based on a traditional design was associated with less benefit with respect to participant performance, but corresponded to higher participant self-confidence than a written text intervention. No correlation was found between self-confidence scores and objective performance scores in any of the groups. More rigorous evaluation of preparatory interventions for invasive procedures should be considered in order to optimize patient safety.

---

**Acknowledgements:** Funding for this study was provided by TrygFonden (grant no. 105112), a non-medical non-governmental organization. The funding source had no role in the design or conduct of the study; collection, management, analysis or interpretation of the data; preparation, review or approval of the manuscript; or the decision to submit the manuscript for publication.

The study was presented at the Association for Medical Education in Europe (AMEE) 2016 conference, Barcelona, August 2016.

**Corresponding Author:** Mikael Johannes Vuokko Henriksen, MD; Copenhagen Academy for Medical Education and Simulation, Copenhagen, The Capital Region of Denmark, Denmark (e-mail: mikael.johannes.vuokko.henriksen@regionh.dk).

**Compliance with Ethical Standards:**

**Conflict of Interest:** The authors declare that they do not have a conflict of interest.

## REFERENCES

1. **Dong C, Goh PS.** Twelve tips for the effective use of videos in medical education. *Med Teach.* 2015;37(2):140-145.
2. **Forbes H, Oprescu FI, Downer T, et al.** Use of videos to support teaching and learning of clinical skills in nursing education: a review. *Nurse Educ Today.* 2016;42:53-56.

3. **Rapp AK, Healy MG, Charlton ME, Keith JN, Rosenbaum ME, Kapadia MR.** YouTube is the most frequently used educational video source for surgical preparation. *J Surg Educ.* 2016;73(6):1072-1076.
4. **Duncan I, Yarwood-Ross L, Haigh C.** YouTube as a source of clinical skills education. *Nurse Educ Today.* 2013;33(12):1576-80.
5. **Rössler B, Lahner D, Schebesta K, Chiari A, Plöchl W.** Medical information on the Internet: quality assessment of lumbar puncture and neuroaxial block techniques on YouTube. *Clin Neurol Neurosurg.* 2012;114:655-658.
6. **Borgersen NJ, Henriksen MJV, Konge L, Sorensen TL, Thomsen ASS, Subhi Y.** Direct ophthalmoscopy on YouTube: analysis of instructional YouTube videos' content and approach to visualization. *Clin Ophthalmol.* 2016;10:1535-1541.
7. **Srivastava G, Roddy M, Langsam D, Agrawal D.** An educational video improves technique in performance of pediatric lumbar punctures. *Pediatr Emerg Care.* 2012;28(1):12-16.
8. **Davis JS, Garcia GD, Wyckoff MM, et al.** Use of mobile learning module improves skills in chest tube insertion. *J Surg Res.* 2012;177(1):21-26.
9. **Yeung P, Justice T, Pasic RP.** Comparison of text versus video for teaching laparoscopic knot tying in the novice surgeon: a randomized, controlled trial. *J Minim Invasive Gynecol.* 2009;16:411-415.
10. **Salina L, Ruffinengo C, Garrino L, et al.** Effectiveness of an educational video as an instrument to refresh and reinforce the learning of a nursing technique: a randomized controlled trial. *Perspect Med Educ.* 2012;1(2):67-75.
11. **AAMC Institute for Improving Medical Education.** Effective use of educational technology in medical education - colloquium on educational technology: recommendations and guidelines for medical educators. Washington DC; 2007:1-19.
12. **Sullivan ME, Ortega A, Wasserberg N, Kaufman H, Nyquist J, Clark R.** Assessing the teaching of procedural skills: can cognitive task analysis add to our traditional teaching methods? *Am J Surg.* 2008;195(1):20-23.
13. **Mesmer-Magnus J, Viswesvaran C.** The role of pre-training interventions in learning: a meta-analysis and integrative review. *Hum Resour Manag Rev.* 2010;20:261-282.
14. **Cleary TJ, Durning SJ, Artino AR.** Microanalytic assessment of self-regulated learning during clinical reasoning tasks: recent developments and next steps. *Acad Med.* 2016;91(11):1516-1521.
15. **Cleary TJ, Zimmerman BJ.** Self-regulation differences during athletic practice by experts, non-experts, and novices. *J Appl Sport Psychol.* 2001;13(2):185-206.
16. **Campbell J, Tirapelle L, Yates K, et al.** The effectiveness of a cognitive task analysis informed curriculum to increase self-efficacy and improve performance for an open cricothyrotomy. *J Surg Educ.* 2011;68(5):403-407.
17. **Mayer RE (ed).** *The Cambridge handbook of multimedia learning*, 2nd edn. New York, NY: Cambridge University Press; 2005.
18. **Roos KL.** Lumbar puncture. *Semin Neurol.* 2003;23(1):105-114.
19. **Glimaker M, Johansson B, Grindborg O, Bottai M, Lindquist L, Sjolín J.** Adult bacterial meningitis: earlier treatment and improved outcome following guideline revision promoting prompt lumbar puncture. *Clin Infect Dis.* 2015;60(8):1162-1169.
20. **Promes SB, Chudgar SM, Grochowski COC, et al.** Gaps in procedural experience and competency in medical school graduates. *Acad Emerg Med.* 2009;16(SUPPL. 2):S58-62.
21. **Dehmer JJ, Amos KD, Farrell TM, Meyer AA, Newton WP, Meyers MO.** Competence and confidence with basic procedural skills: the experience and opinions of fourth-year medical students at a single institution. *Acad Med.* 2013;88(5):682-687.
22. **Lammers RL, Temple KJ, Wagner MJ, Ray D.** Competence of new emergency medicine residents in the performance of lumbar punctures. *Acad Emerg Med.* 2005;12(7):622-628.
23. **Hansen M, Oosthuizen G, Windsor J, et al.** Enhancement of medical interns' levels of clinical skills competence and self-confidence levels via video iPods: pilot randomized controlled trial. *J Med Internet Res.* 2011;13(1):e29.
24. **Randomness and Integrity Services Ltd.** [www.random.org](http://www.random.org). Available at: <http://www.random.org>. Accessed March 2013 to March 2014.
25. **Henriksen MJV, Wienecke T, Thagesen H, et al.** Assessment of residents readiness to perform lumbar puncture: a validation study. *J Gen Intern Med.* 2017;32(6):610-618.
26. **Brydges R, Hatala R, Mylopoulos M.** Examining residents' strategic mindfulness during self-regulated learning of a simulated procedural skill. *J Grad Med Educ.* 2016;8(3):364-371.
27. **Miles E, Tegtmeier K, Lai S, Braner DAW.** Lumbar puncture — NEJM. *N Engl J Med.*
28. **Finn J, Bogeskov M, Gustaffson A.** *Praktisk Procedure- Og Diagnoseguide (Practical Procedure and Diagnoseguide)*. Copenhagen: Munksgaard Denmark; 2008.
29. **Subhi Y, Todsén T, Konge L.** An integrable, web-based solution for easy assessment of video-recorded performances. *Adv Med Educ Pract.* 2014;5:103-105.
30. **Jang HW, Kim K-J.** Use of online clinical videos for clinical skills training for medical students: benefits and challenges. *BMC Med Educ.* 2014;14(1):56.
31. **Downing SM.** Reliability: on the reproducibility of assessment data. *Med Educ.* 2004;38(9):1006-1012.
32. **Brunckhorst O, Shahid S, Aydin A, et al.** The relationship between technical and nontechnical skills within a simulation-based ureteroscopy training environment. *J Surg Educ.* 2015;72(5):1039-1044.
33. **Liaw SY, Scherpbier A, Rethans J-J, Klainin-Yobas P.** Assessment for simulation learning outcomes: a comparison of knowledge and self-reported confidence with observed clinical performance. *Nurse Educ Today.* 2012;32(6):e35-e39.
34. **Barnsley L, Lyon PM, Ralston SJ, et al.** Clinical skills in junior medical officers: a comparison of self-reported confidence and observed competence. *Med Educ.* 2004;38(4):358-367.
35. **Ward P, Williams M, Hancock P.** Simulation for performance and training. In: Ericsson KA, Charness N, Feltovich P, Hoffman RR, eds. *The Cambridge handbook of expertise and expert performance*. New York: Cambridge University Press; 2006:243-265.
36. **Brydges R, Carnahan H, Safir O, Dubrowski A.** How effective is self-guided learning of clinical technical skills? It's all about process. *Med Educ.* 2009;43(6):507-515.
37. **Shanks D, Brydges R, den Brok W, Nair P, Hatala R.** Are two heads better than one? Comparing dyad and self-regulated learning in simulation training. *Med Educ.* 2013;47(12):1215-1222.
38. **Brydges R, Hatala R, Zendejas B, Erwin PJ, Cook DA.** Linking simulation-based educational assessments and patient-related outcomes: a systematic review and meta-analysis. *Acad Med.* 2015;90(2):246-256.