3D Learning in anatomical and surgical education in relation to visual-spatial abilities

Bogomolova, K.

5

Stereoscopic video visualization of spatially-complex procedures increases performance of surgical novices with high visual-spatial abilities:

a randomized controlled trial

Katerina Bogomolova, Jeroen J.G. van Merriënboer, Jan E. Sluimers,
Jeroen Donkers, Theo Wiggers, Steven E.R. Hovius, Jos A. van der Hage

The American Journal of Surgery 2021; 222:739-745

ABSTRACT

Background

The effect of three-dimensional (3D) versus two-dimensional (2D) video on performance of a spatially complex procedure and perceived cognitive load were examined among residents in relation to their visual-spatial abilities (VSA).

Methods

In a randomized controlled trial, 108 surgical residents performed a 5-Flap Z-plasty on a simulation model after watching the instructional video either in a 3D or 2D mode. Outcomes included perceived cognitive load measured by NASA-TLX questionnaire, task performance assessed using Observational Clinical Human Reliability Analysis and the percentage of achieved safe lengthening of the scar.

Results

No significant differences were found between groups. However, when accounted for VSA, safe lengthening was achieved significantly more often in the 3D group and only among individuals with high VSA (OR=6.67, 95%CI: 1.23–35.9, p=.027).

Conclusions

Overall, 3D instructional videos are as effective as 2D videos. However, they can be effectively used to enhance learning in high VSA residents.

INTRODUCTION

Surgical residents experience difficulties with learning and performing spatially-complex procedures that require spatial and conceptual understanding.[1] Consequently, residents tend to feel less confident about performing procedures with increasing complexity. In particular, individuals with lower visual-spatial abilities (VSA) experience difficulties in learning spatially complex procedures.[2,3] VSA is defined as the ability that allows individuals to construct visual-spatial, i.e., three-dimensional (3D), mental representations of two-dimensional (2D) images and to mentally manipulate these representations.[4,5] VSA has been found to be positively associated with both subjective and objective assessments of surgical performance, especially among novices.[6,7] In anatomical education, VSA has been widely explored and showed repeatedly its positive association with anatomical knowledge.[8]

The differences in performance between low and high VSA individuals are explained within the cognitive load theory (CLT)[9]. According to CLT, the capacity of human working memory involved in processing new information is severely limited. Types of cognitive load include intrinsic (caused by performing the task itself), extraneous (caused by the way learning material is presented), and germane (caused by actual learning) load[9]. When the sum of the three types of load exceeds working memory capacity, a cognitive overload occurs which impairs learning.[10] Since low VSA individuals devote more cognitive resources to performing a spatially complex task, their intrinsic cognitive load increases during learning. Subsequently, they are left with less available resources that they can allocate to learning which leads to decreased performance in comparison with high VSA individuals. One way to compensate for the increased cognitive load in low VSA individuals, is to decrease extraneous load by improving the instructional method.

Instructional videos are one of the most used and effective ways of preparing for surgeries among residents.[11,12] When presented in a segmented rather than continuous format, using a step-by-step approach, video-based learning can be even more effective.[13] However, procedures are viewed monoscopically on a 2D screen without real perception of visual depth. This can make learning spatially-complex procedures more challenging, especially for residents with lower levels of VSA, as they have less ability to transform 2D images into 3D mental representations. However, according to the compensating hypothesis within the CLT,[10] this mental transformation can be assisted by presenting images stereoscopically, in real 3D.[14,15] In other words, stereoscopic visualization can compensate for low VSA by providing depth cues. In this case, the mental 3D model is already built and provided. This can eventually lead to a decreased cognitive load and improved learning.[15,16] A binocular vision of the viewer, though, is required to perceive spatial visual depth that is obtained with the use of this technology.

Compared to a monoscopic visualization, stereoscopic visualization has been shown to improve performance in laparoscopic surgery training.[17,18] However, it is yet unknown whether this applies to video-based learning and would probably depend on the "spatial" nature of the procedure. Roach et al. evaluated the effectiveness of a stereoscopic 3D instructional video of a four-flapped Z-plasty and a rhomboid flap and found no significant differences in performance when compared to a monoscopic 2D video of the same procedures.[19] As stated by the authors, both procedures were "essentially two dimensional in their design and performance, leaving their complexity to appear more conceptually geared rather than spatially." This suggests that watching an instructional video in real 3D that is eventually performed in a 2D plane does not offer extra gain in terms of knowledge and performance skill. This also suggests that video demonstration of a spatially complex procedure performed in a 3D plane can benefit from stereoscopic visualization, given that the complexity of the procedure is based more on spatial thinking and understanding.

Therefore, the primary aim of this study was to evaluate whether a 3D instructional video compared to a 2D video of a 5-flap Z-plasty would improve performance of surgical residents. The secondary aim was to evaluate the perceived cognitive load both during watching the instructional video and performing the procedure. The outcomes were evaluated in relation to the VSA of participants.

METHODS

Study design and population

A randomized controlled trial was conducted at the Leiden University Medical Center, The Netherlands (Figure 1). The experiment took place during a hands-on session as part of a special educational surgery program in September 2019. The study protocol was approved by the Netherlands Association for Medical Education (NVMO) Ethical Review Board (NERB case number: 2019.6.10).

Participants

Participants were first-year surgical residents from various teaching hospitals in The Netherlands. Participation in this study was voluntary, and written consent was obtained from all enrolled residents.

Randomization

Participants were randomly allocated to either a 2D or 3D video group using an Excel Random Group Generator (Microsoft Excel for Office 365 MSO, version 2012). Randomization was stratified by sex to ensure an equal female-to-male ratio in both groups.

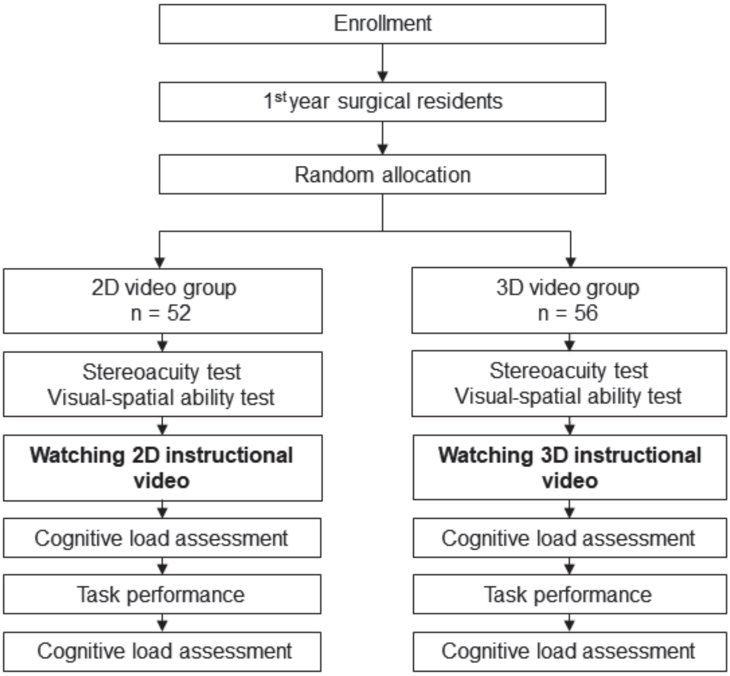

Figure 1. Flowchart of study design. 2D, two-dimensional; 3D, three-dimensional; n, number of participants.

Stereovision

Stereovision of participants was measured by a Random Dot 3 - LEA SYMBOLS® Stereoacuity Test [Vision Assessment Corp., Elk Grove Village, IL, USA] prior to the experiment to identify and exclude individuals with absent stereovision.

Assessment of VSA

VSA was measured by a validated Mental Rotation Test (MRT).[20,21] MRT is the gold standard to assess VSA that has been associated with anatomical knowledge and surgical skills assessment.[7,22] The duration of the test was 10 min, and the test scores ranged from 0 to 24 points. High VSA was defined as a mean score above the average; and low VSA was defined as a mean score below the average.

Surgical procedure

A 5-flap Z-plasty, also referred to as the jumping man flap, is a spatially complex procedure that involves two different types of tissue movements in a 3D plane. Its goal is to simultaneously provide lengthening and deepening of a skin contracture. The lengthening is achieved by two Z-plasties, while the deepening is achieved by a V-Y advancement flap (*Supplementary material 8).* Spatial understanding of the problem and solution is required to perform the procedure correctly.

Instructional videos

The instructional video of a 5-flap Z-plasty was developed using a validated step-by-step approach.[12] The video consisted of a step-by-step demonstration of the procedure on a simulation model accompanied by auditory narration. The duration of the video was 4 min.

The 2D video group watched the instructional video on a large flat screen. The 3D video group watched the same video on a large flat screen with stereoscopic projection and the use of active 3D glasses. All participants watched the video twice to stimulate active processing. Participants were asked to mentally rehearse the surgical steps while watching the video the second time without auditory narration.

Cognitive load assessment

Perceived cognitive load was measured twice by a validated NASA-TLX questionnaire.[23] The participants filled out the questionnaire immediately after watching the instructional video and for the second time, after performing the task. The cognitive domains included mental, physical, and temporal demands; performance; effort; and frustration. The total score was a product of the six domains and ranged between 0 and 10 points.

Task performance

Participants performed a 5-flap Z-plasty on a skin contracture in the simulation model (Figure 2) after watching the instructional video. The task was to execute the most favorable advanced 5-flap Z-plasty given the initial length of contracture and maximal slack of the skin to achieve maximal lengthening and deepening, taking into account the viability of the skin flaps. Participants were given 15 min to complete the procedure.

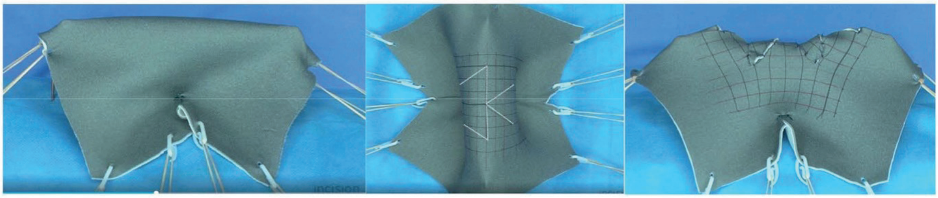

Figure 2. Simulation model of a skin contracture. Left: front view of the skin model, middle: above view of the skin with a 5-flap Z-plasty being drawn, right: the final result.

OCHRA checklist

The performed procedures were assessed independently by two experts (JS and KB) in a blinded manner using the OCHRA checklist (*Supplementary material 9*). The OCHRA is a procedure-specific, step-by-step skills assessment checklist that is characterized by a breakdown of a procedure into tasks or substeps.[24] The 5-flap Z-plasty comprises 10 substeps, each of which can attain scores of 0.5, 1, or 1.5 points if performed correctly. The total maximum possible score was 10 points.

5

Safe lengthening

The safe range of minimal and maximal gain in lengthening that can be achieved, if performed correctly, was calculated for each performed procedure given the initial length of the contracture and the chosen size of the angles of the Z-plasty flaps. The calculations were performed blindly by two experts (KB and JS) after the experiment. Safe lengthening was achieved (yes/no) if the calculated value felt in the permitted range of lengthening without compromising blood flow of the created flaps.

Outcome measures

The primary outcome measure was defined as the difference in mean OCHRA scores between 2D and 3D video groups. Secondary outcomes were defined as the difference in proportions of achieved safe lengthening and mean cognitive load scores of instructional videos, and task performance between the two groups. Additionally, the outcomes were evaluated for different levels of VSA.

Statistical analysis

Owing to the novelty of the study, including the type of procedure and assessment tool, no previous data were available to calculate the sample size. A minimal sample size of 100 participants was assumed to be appropriate. Participants' baseline characteristics were summarized using descriptive statistics. The differences in baseline measurements were assessed using an independent *t*-test for differences in means and chi-squared test for differences in proportions. The differences in mean OCHRA and cognitive load scores were assessed with an independent *t*-test, and the proportions of safe lengthening, with a chi-squared test. To assess the possible modifying effect of VSA on outcomes, regression analyses were performed. In ANCOVA, the OCHRA and cognitive load scores were included as dependent variables, intervention group as the fixed factor, VSA score as the covariate, and 'intervention x VSA' as the interaction term. In the logistic regression model, safe lengthening was included as the dependent variable; the remaining factors were identical as for the ANCOVA test. Odd ratios were calculated based on the values of predictors and interaction term from the logistic regression model. Analyses were performed using SPSS statistical software package version 23.0 for Windows (IBM Corp., Armonk, NY). Statistical significance was determined at the level of $p < .05$.

RESULTS

A total of 108 participants were included (Table 1). All participants could perceive spatial-visual depth as measured by the stereoacuity test.

The differences between the 2D and 3D video groups are presented in Table 2. The performance on the task, as measured by the OCHRA checklist, was not significantly different between the two groups ($t(106) = .813$; $p = .487$). Although safe lengthening was achieved more often by the participants in the 3D group, these differences were not significant (56.8% vs. 43.2%; $c^2(1) = .602$; $p = .555$). The perceived cognitive load was similar in both groups.

The effect of VSA

The mean OCHRA scores remained similar for all levels of VSA, as measured by the interaction term in ANCOVA ($F(1, 102)=0.43$, $p = .513$). However, regardless of intervention, VSA was significantly and positively associated with performance only among individuals with initially low levels of VSA ($F(1,48) = 5.37$, partial $\eta^2 = 0.11$, $p = .025$) (Figure 3a). This association was not found among individuals with high VSA ($F(1, 43) = 0.26$, partial $\eta^2 = 0.006$, $p = .610$) (Figure 3b).

Table 1. Baseline characteristics of the included participants.

	2D video	3D video	p value
	n = 52	n = 56	
Sex, n (%)			
Male	25 (48.1)	26 (46.4)	.622
Female	27 (51.9)	30 (53.6)	
Age, mean ±SD in years	29.8 ± 2.09	30.5 ± 1.5	.053
Residency, n (%			
General	26 (50.0)	34 (60.7)	.571
Orthopedic	13 (25.0)	10 (17.9)	
Plastic and Reconstructive Surgery	4 (7.7)	6 (10.7)	
Urology	9 (17.3)	6 (10.7)	
Clinical experience, n (%)			
Yes	50 (96.2)	56 (100)	.210
No	2 (3.8)	0 (0)	
Clinical experience, mean ± SD in months	21.3 ± 11.3	20.8 ± 8.8	.801
OR Hours, median (IQR)	1.5 (4.8)	1.0 (6.4)	.588
Performed a Z or V-Y plasty before, n (%)			
Yes	6 (11.5)	4 (7.1)	.414
No	46 (88.5)	52 (92.9)	
Mental Rotation Test score, mean ± SD	13.2 ± 5.4	13.8 ± 5.5	.596

2D, two-dimensional; 3D, three-dimensional; n, number of participants; SD, standard deviation; OR, operating room; IQR, interquartile range

Table 2. Differences in outcomes between 2D and 3D instructional video groups.

	2D video	3D video	p value
	n = 52	n = 56	
OCHRA score, mean ± SD	8.0 (1.7)	7.8 (2.0)	.487
Safe lengthening, n (%)			
Achieved	19 (43.2)	25 (56.8)	.555
Not achieved	32 (50.8)	31 (49.2)	
Cognitive load video, mean ± SD	6.1 ± 1.2	6.0 ± 1.4	.738
Cognitive load task, mean ± SD	6.5 ± 1.0	6.1 ± 1.2	.142

2D, two-dimensional; 3D, three-dimensional; n, number of participants; SD, standard deviation.

5

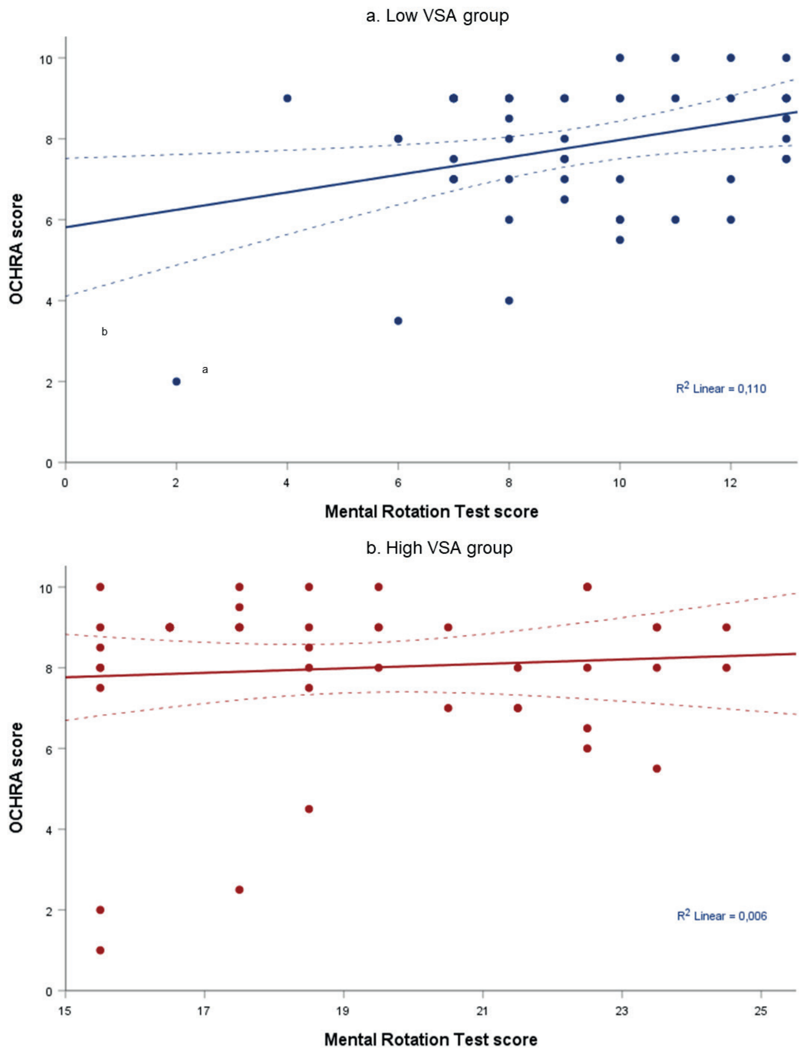

Figure 3. The relationship between Mental Rotation Test (MRT) scores and OCHRA scores in the low VSA (a) and high VSA (b) groups. The continuous lines represent the prediction values of the OCHRA scores; the dashed lines represent the 95% confidence Interval of these values.

The proportions of achieved safe lengthening were significantly different for participants with low and high VSA, as measured by the interaction term in logistic regression (OR = 6.67, 95% CI: 1.23–35.9, p = .027). Safe lengthening was achieved significantly more often in the 3D group and only among individuals with high VSA than in the 2D group (OR = 8.8). The corresponding odds ratios (OR) are illustrated in Figure 4. The coefficients' estimates for the model are provided in *Supplementary material 10*.

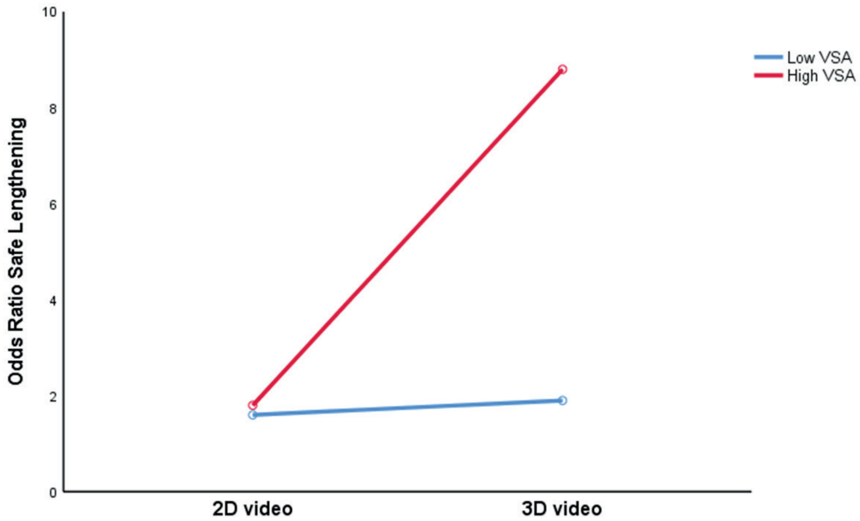

Figure 4. The interaction between intervention (2D versus 3D video group) and visual-spatial abilities (low versus high VSA) for the outcome "safe lengthening." The OR was calculated based on the values of predictors and interaction term in the logistic regression model.

The perceived cognitive load scores remained similar for all levels of VSA in both study groups, as measured by the interaction term in ANCOVA (*video:* $F(1, 100) = 1.81$, $p = .182$; *task:* $F(1,97) = 0.78$, $p = .379$). Regardless of intervention, VSA was not associated with the perceived cognitive load.

DISCUSSION

In this study, the effectiveness of a 3D versus 2D instructional video of a spatially complex procedure was evaluated in terms of performance and perceived cognitive load among surgical residents. Additionally, the outcomes were evaluated in relation to VSA. The beneficial effect of the stereoscopic video visualization hypothesized to be the greatest for individuals with lower VSA was not observed. On the contrary, the learning effect of the 3D mode was greater among individuals with higher VSA than in those with lower VSA.

These findings are in contrast to the proposed compensating hypothesis assuming that presenting images stereoscopically would compensate for low VSA, as the mental 3D model is already built and provided.[15,16] This eventually would lead to a decreased cognitive load for low VSA learners and improve learning. However, the opposite effect was observed with high VSA learners benefiting most from 3D visualization. This interaction could have been caused by the so-called "expertise reversal effect" that occurs between different levels of learner's expertise and instructional techniques.[25,26] Novices with minimal prior knowledge must process many novel elements of information that can easily overload their working memory capacity. Accordingly, instructional methods that are effective for experienced learners may become less effective or even disadvantageous for novice learners, and vice versa. It has been reported that learners with high prior knowledge benefit more from continuous dynamic animations, while novices with no or low prior knowledge can have difficulties with processing high degrees of transitivity of visual presentations.[27,28] Because the participants of this study were novices in terms of most surgical procedures with an average of 1.5 hours spent in OR a week, the instructional video could have been perceived as very challenging, which was reflected by the relatively large proportion of participants (58.3%) that did not achieve safe lengthening and the relatively high perceived cognitive load. The direction of this interaction is in line with the co-existing mechanism, the ability-as-enhancer hypothesis, within the cognitive load theory.[29,30] This hypothesis predicts that high levels of VSA are required to benefit from improved instructional method, while low VSA levels are hindered by it. In other words, high levels VSA enhance learning with 3D visualization, while low VSA levels do not. This is because high VSA individuals have still sufficient number of cognitive resources after processing spatially complex procedure. Low VSA individuals, on the other hand, are depleted in their cognitive resources and are not able to benefit from learning with 3D visualization.

Notably, the interaction effect of VSA and performance, as measured by the OCHRA checklist, was not observed. This can be explained by the choice of the assessment tool. In research and practice, surgical performance is effectively assessed both by global rating scales and task-specific checklists.[31] A task-specific OCHRA checklist was preferred in the current study for several reasons. First, the OCHRA checklist allowed individual assessment of each performed step of the 5-flap Z-plasty. This was essential, because each step of the procedure requires spatial understanding to be performed correctly, and this can affect the final result. Second, global rating scales include items that are much less relevant in performing a 5-flap Z-plasty, such as knowledge and use of instruments, and respect for other tissues.[32]. However, the association between VSA and performance appears to be most prominent when performance is measured by a global rating scale. Wanzel et al. evaluated the effect of VSA, as measured by the MRT, on performance of a spatially complex procedure.[3] Performance was measured both by a global rating scale and task-specific checklist. The MRT scores were significantly associated with the scores on the global rating scale ($r = 59$, $p = .0013$), but not with the checklist scores ($r = 0.36$, $p = .068$). This can explain why the interaction effect between MRT and OCHRA scores was not observed in the current study.

Regardless of intervention, VSA was associated with performance only among individuals with lower VSA levels. The specific VSA scores of learners higher on the VSA continuum did not affect their surgical performance. A similar association was previously described by Roach et al. evaluating the effect of stereoscopic visualization on laparoscopic skills in relation to VSA.[33] This suggests that VSA may impair performance if not well developed, but VSA does not affect performance when a certain VSA level has been reached. Previous research has shown that VSA can be trained by repeated practice and learning.[34-36] Therefore, training of VSA skills in novices with initially low VSA could be beneficial.

Limitations

This study has some limitations. First, all participants watched the instructional video simultaneously on a large computer screen. Watching it at their own pace was therefore not possible. It has been reported that pauses provide learners more time to process, consolidate, and transfer information to long-term memory before moving forward to the next step. Moreover, they can also replay the steps and better process the difficult parts of the procedure.[37] As our study participants were novices, those subjects with low VSA could have benefited from self-paced control and perhaps, even benefited from 3D visualization. Second, the included participants were first-year surgical residents with little to no prior surgical knowledge and experience. As more complex instructional videos can be more effective for more experienced learners, better effects of 3D visualization could have been obtained in the later stages of surgical training. Third, 5-flap Z-plasty was performed on a simulated model that optimally, but not fully, resembled normal skin. The used material could have permitted a larger amount of stretch of the flaps than normal skin. This could have led to an unintended increase in the achieved lengthening and deepening of the contracture. To account for the possible bias, the achieved safe lengthening was calculated for each participant. The assessment of achieved deepening was not possible owing to variability in height of the contracture. Last, all participants could complete the procedure within the given time frame of 15 min. Differences in time to complete the procedure between groups were not evaluated. However, by including time as an outcome measure, more insight could have been gained into the effects of VSA on performance.

Future implications

The findings of this study have implications for both research and practice. When designing a new study, it is instrumental to consider VSA as a potential effect modifier. Additionally, a potential interaction between VSA and instructional design should be considered when performing statistical analysis. In surgical practice, an individualized approach could be helpful for residents with lower levels of VSA. Individuals with low VSA can benefit from deliberate practice and feedback and achieve a comparable level of competency as those with high VSA.[2] Because spatial skills are malleable, VSA-based training is recommended in the early stages of surgical training to reinforce the development of surgical skills among novices.[34-36] For individuals with low VSA, 3D instructional videos seem to initially have no advantage over conventional 2D videos. Further research is encouraged to evaluate whether instructional videos are more effective in the later stages of surgical training and whether low VSA learners will then benefit from 3D visualizations.

CONCLUSIONS

Overall, watching an instructional video of a spatially complex procedure in 3D is as effective as watching it in a 2D traditional format. However, when considering VSA of the learners, 3D visualization is more favorable for individuals with high VSA. Authors hypothesized that the educational effect of 3D visualization possibly depends not only on VSA but also on the expertise level of the learner and/or the complexity of the procedure. Future research should focus on the effect of 3D instructional videos of simpler vs. complex procedures for residents with low VSA in relation to their level of expertise.

REFERENCES

1. Klebanoff JS, Marfori CQ, Vargas MV, et al. Ob/Gyn resident self-perceived preparedness for minimally invasive surgery. BMC Med Educ. 2020;20(1):185.
2. Wanzel KR, Hamstra SJ, Anastakis DJ, et al. Effect of visual-spatial ability on learning of spatially-complex surgical skills. Lancet. 2002;359(9302):230-231.
3. Wanzel KR, Hamstra SJ, Caminiti MF, et al. Visual-spatial ability correlates with efficiency of hand motion and successful surgical performance. Surgery. 2003;134(5):750-757.
4. Gordon HW. The cognitive laterality battery: tests of specialized cognitive function. Int J Neurosci. 1986;29(3-4):223-244.
5. Kozhevnikov M, Hegarty M. A dissociation between object manipulation spatial ability and spatial orientation ability. Mem Cognit. 2001;29(5):745-756.
6. Maan ZN, Maan IN, Darzi AW, Aggarwal R. Systematic review of predictors of surgical performance. Br J Surg 2012;99(12):1610-1621.
7. Langlois J, Bellemare C, Toulouse J, Wells GA. Spatial abilities and technical skills performance in health care: a systematic review. Med Educ. 2015;49(11):1065-1085.
8. Langlois J, Bellemare C, Toulouse J, Wells GA. Spatial Abilities Training in Anatomy Education: A Systematic Review. Anat Sci Educ. 2020;13(1):71-79.
9. van Merriënboer JJ, Sweller J. Cognitive load theory in health professional education: design principles and strategies. Med Educ. 2010;44(1):85-93.
10. Sweller J, van Merrienboer JJG, Paas F. Cognitive architecture and instructional design: 20 years later. Educ Psychol Rev. 2019;31(2):261-292.
11. Green JL, Suresh V, Bittar P, et al. The utilization of video technology in surgical education: A systematic review. J Surg Res. 2019;235:171-180.
12. Ahmet A, Gamze K, Rustem M, Sezen KA. Is video-based education an effective method in surgical education? A systematic review. J Surg Educ. 2018;75(5):1150-1158.
13. Nazari T, van de Graaf FW, Dankbaar MEW, et al. One step at a time: step by step versus continuous video-based learning to prepare medical students for performing surgical procedures. J Surg Educ. 2020;7(4):779-787
14. Hoffler TN. Spatial ability: Its influence on learning with visualizations-a meta-analytic review. Educ Psychol Rev. 2010;22(3):245-269.
15. Mayer RE. Multimedia Learning. Cambridge University Press: Cambridge, 2001.
16. Lee DY, Shin DH. An empirical evaluation of multi-media-based learning of a procedural task. Comput Hum Behav. 2012;28(3):1072-1081.
17. Fergo C, Burcharth J, Pommergaard HC, et al. Three-dimensional laparoscopy vs 2-dimensional laparoscopy with high-definition technology for abdominal surgery: a systematic review. Am J Surg. 2017;213(1):159-170.
18. Sorensen SM, Savran MM, Konge L, Bjerrum F. Three-dimensional versus two-dimensional vision in laparoscopy: a systematic review. Surg Endosc. 2016;30(1):11-23.
19. Roach VA, Brandt MG, Moore CC, Wilson TD. Is three-dimensional videography the cutting edge of surgical skill acquisition? Anat Sci Educ. 2012;5(3):138-145.
20. Shepard RN, Metzler J. Mental rotation of three-dimensional objects. Science. 1971;171(3972):701-703.
21. Vandenberg SG, Kuse AR. Mental rotations, a group test of three-dimensional spatial visualization. Percept Mot Skills. 1978;47(2):599-604.

5

22. Langlois J, Bellemare C, Toulouse J, Wells GA. Spatial abilities and anatomy knowledge assessment: A systematic review. Anat Sci Educ. 2017;10(3):235-241.
23. Hart S, Staveland L. Development of NASA-TLX (Task Load Index): Results of empirical and theoretical research. Adv Psychol. 1988;52:139-183.
24. Nazari T, Vlieger EJ, Dankbaar MEW, et al. Creation of a universal language for surgical procedures using the step-by-step framework. BJS Open. 2018;2(3):151-157.
25. Chen OH, Kalyuga S, Sweller J. The expertise reversal effect is a variant of the more general element interactivity effect. Educ Psychol Rev. 2017;29(2):393-405.
26. Kalyuga S, Ayres P, Chandler P, Sweller J. The expertise reversal effect. Educ Psychol-US. 2003;38(1):23-31.
27. Kalyuga S. Relative effectiveness of animated and static diagrams: An effect of learner prior knowledge. Comput Hum Behav, 2008;24(3):852-861.
28. Kalyuga S, Rikers R, Paas F. Educational implications of expertise reversal effects in learning and performance of complex cognitive and sensorimotor skills. Educ Psychol Rev. 2012;24(2):313-337.
29. Huk T. Who benefits from learning with 3D models? The case of spatial ability. J Comp Assisted Learning. 2006;22(6):392-404.
30. Mayer RE, Sims VK. For whom is a picture worth 1000 words - extensions of a dual-coding theory of multimedia learning. J Educ Psychol. 1994;86(3):389-401.
31. van Hove PD, Tuijthof GJM, Verdaasdonk EGG, et al. Objective assessment of technical surgical skills. Brit J Surg. 2010;97(7):972-987.
32. Martin JA, Regehr G, Reznick R, et al. Objective structured assessment of technical skill (OSATS) for surgical residents. Br J Surg. 1997;84(2):273-278.
33. Roach VA, Mistry MR, Wilson TD. Spatial visualization ability and laparoscopic skills in novice learners: evaluating stereoscopic versus monoscopic visualizations. Anat Sci Educ. 2014;7(4):295-301.
34. Uttal DH, Meadow NG, Tipton E, et al. The malleability of spatial skills: a meta-analysis of training studies. Psychol Bull. 2013;139(2):352-402.
35. Langlois J, Bellemare C, Toulouse J, Wells GA. Spatial abilities training in anatomy education: A systematic review. Anat Sci Educ. 2020;13(1):71-79.
36. Bogomolova K, Hierck BP, van der Hage JA, Hovius SER. Anatomy dissection course improves the initially lower levels of visual-spatial abilities of medical undergraduates. Anat Sci Educ. 2020; 13(3):333-342.
37. Schwan S, Riempp R. The cognitive benefits of interactive videos: learning to tie nautical knots. Learn Instr. 2004;14(3):293-305.