International Journal of Computer Assisted Radiology and Surgery (2024) 19:1213–1222
https://doi.org/10.1007/s11548-024-03138-7

ORIGINAL ARTICLE

Stand in surgeon's shoes: virtual reality cross-training to enhance teamwork in surgery

Benjamin D. Killeen[1] · **Han Zhang**[1] · **Liam J. Wang**[1] · **Zixuan Liu**[1] · **Constantin Kleinbeck**[1,3] · **Michael Rosen**[1] · **Russell H. Taylor**[1] · **Greg Osgood**[2] · **Mathias Unberath**[1]

Received: 4 March 2024 / Accepted: 28 March 2024 / Published online: 20 April 2024

Abstract

Purpose Teamwork in surgery depends on a shared mental model of success, i.e., a common understanding of objectives in the operating room. A shared model leads to increased engagement among team members and is associated with fewer complications and overall better outcomes for patients. However, clinical training typically focuses on role-specific skills, leaving individuals to acquire a shared model indirectly through on-the-job experience.

Methods We investigate whether virtual reality (VR) cross-training, *i.e*let@tokeneonedotexposure to other roles, can enhance a shared mental model for non-surgeons more directly. Our study focuses on X-ray guided pelvic trauma surgery, a procedure where successful communication depends on the shared model between the surgeon and a C-arm technologist. We present a VR environment supporting both roles and evaluate a cross-training curriculum in which non-surgeons swap roles with the surgeon.

Results Exposure to the surgical task resulted in higher engagement with the C-arm technologist role in VR, as measured by the mental demand and effort expended by participants ($p < 0.001$). It also has a significant effect on non-surgeon's mental model of the overall task; novice participants' estimation of the mental demand and effort required for the surgeon's task increases after training, while their perception of overall performance decreases ($p < 0.05$), indicating a gap in understanding based solely on observation. This phenomenon was also present for a professional C-arm technologist.

Conclusion Until now, VR applications for clinical training have focused on virtualizing existing curricula. We demonstrate how novel approaches which are not possible outside of a virtual environment, such as role swapping, may enhance the shared mental model of surgical teams by contextualizing each individual's role within the overall task in a time- and cost-efficient manner. As workflows grow increasingly sophisticated, we see VR curricula as being able to directly foster a shared model for success, ultimately benefiting patient outcomes through more effective teamwork in surgery.

Keywords Computer-assisted interventions · Surgical training · X-ray · Image-guided

Benjamin D. Killeen and Han Zhang have equally contributed to this work.

✉ Benjamin D. Killeen
killeen@jhu.edu

Han Zhang
hzhan206@jhu.edu

Liam J. Wang
wwang136@jhu.edu

Zixuan Liu
zliu189@jhu.edu

Constantin Kleinbeck
constantin.kleinbeck@fau.de

Michael Rosen
mar@jhu.aru

Russell H. Taylor
rht@jhu.edu

Greg Osgood
gosgood2@jhmi.edu

Mathias Unberath
unberath@jhu.edu

1 Johns Hopkins University, Baltimore, MD 21218, USA

2 Department of Orthopaedic Surgery, Johns Hopkins Medicine, Baltimore, MD 21218, USA

3 Friedrich-Alexander-Universität, Erlangen, Germany

Introduction

Teamwork is a critical component of surgical success. Poor teamwork leads to an increase in minor problems, such as difficulty finding instruments due to unpreparedness or team members' disengagement. Although seemingly innocuous, these minor problems create the opportunity for adverse events and are associated with worse outcomes for patients [1]. Effective teamwork, on the other hand, contributes to improved performance and shorter operating times [2], which has motivated numerous interventions aiming to improve teamwork in surgery. These include non-technical skills training, implementation of checklists, and structured communication protocols [3]. The mixed success of these interventions, though, suggests that complementary approaches are needed [4]. Targeting the specific learning goal of an "architectural knowledge about how different team roles and tasks fit together," i.e., a shared mental model for effective teamwork, is necessary so that all team members stay "focused and engaged, even when not directly involved in the action" [3].

Cross-training is another intervention aimed at enhancing teamwork, which it does by providing subjects with some exposure to other team members' specialized roles [5]. Despite the recognized benefits of cross-training [6], including for medical teams [7], non-surgeons have no direct exposure to the surgeon's role as a part of their required curriculum. Surgical training requires significant investment in time and resources, involving scarce resources such as cadaveric specimens. Providing the same level of training to non-surgeons is impossible from a cost and priority perspective, and the time required to gain even a base level of proficiency would be prohibitive for many [8]. Thus, there is an inherent imbalance in the information available to surgeons and non-surgeons about each others' roles.

Virtual reality (VR) presents the opportunity to design interventions to improve teamwork in surgery through swapping roles, the most direct form of cross-training. To show the potential value of this approach, we conduct a user study with two interventions in which participants carry out the "surgeon task" or an "enabling task" in VR. We focus in particular on percutaneous fracture fixation because it entails frequent communication between the surgeon, who inserts a K-wire through a narrow bony corridor under image guidance, and an X-ray technologist, who manipulates a C-arm to obtain necessary images. To conduct this study, we implement a VR environment that supports realistic X-ray image guidance for surgical workflows, including tool-to-tissue interactions and C-arm positioning, as shown in Fig. 1. Using the NASA Task Load Index as our primary measurement, we quantify individuals' engagement based on their self-reported mental demand and effort expended to achieve their level of performance in the task. For novice participants, cross-training on the surgeon task increases the engagement and reduces the frustration experienced when performing the enabling task. Moreover, we find that this cross-training has a significant effect on novice participants' shared mental model, changing their perception of the difficulty of the other task. A professional technologist with five years of experience in the OR reacted similarly in a study with their regular attending physician, remarking "Apparently your job is a lot more difficult" than previously understood.

Related work

Cross-training has long been recognized as an effective method for developing shared mental models [5, 6]. In the context of medical teams, [7] examines cross-training for physicians and pharmacists with a physical simulation-based practice scenario involving intubation. However, physical simulation of surgical procedures, which requires realistic phantoms or cadaveric specimens, is impractical due to cost. Computer simulation and virtual reality are an increasingly viable option for acquisition of surgical knowledge [9, 10]. Notable successes in this area focused on the acquisition of procedural knowledge [11–14], i.e., the steps to complete a surgery, which benefits from enhanced recall due to learning in an immersive environment [15]. [11] in particular compares remote instruction via virtual reality to traditional videoconferencing, while [14] explores the ability of VR to mitigate time constraints for apprenticeship-style training in interventional radiology. For skill-based knowledge, [16] show that endovascular interventional skills learned in VR transferred to a porcine model. More recently, [17] find in a randomized controlled trial that immersive VR is a superior teaching mechanism for glenoid exposure, while [18] demonstrates how VR training can provide rich data for training deep neural networks (DNNs).

For X-ray image-guided surgery in particular, much work has been proposed to improve C-arm placement. These include assistance systems in augmented reality (AR) [19–21], radiation-free physical simulation of X-ray images [22, 23], and computer simulation of C-arms [24]. Much of the work in this area originates with the desire to simulate images for training DNNs for C-arm positioning [25–27] and closed-loop control [28]. However, as in [18], VR simulators present opportunities for human-centered data collection, providing the basis for collaborative assistance systems in X-ray guided surgery [29].

Methods

To study the effect of VR cross-training for non-surgeon roles, we conduct a user study with two interventions: the

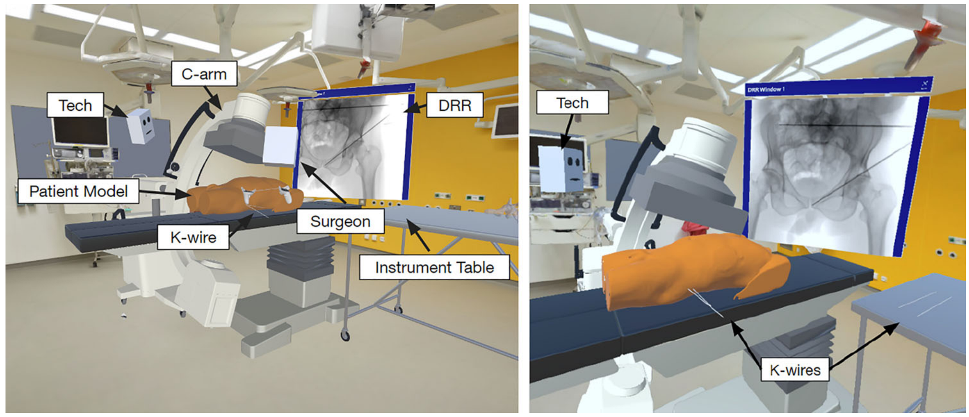


Fig. 1 Virtual reality simulation of an X-ray image-guided procedure. The simulator supports real-time multiplayer with interactive components, including a mobile C-arm, surgical instruments, and patient model. Basic haptics and a simplified tool-to-tissue interactions encourage the replication of real workflows

surgeon intervention and the enabling intervention. During the **surgeon intervention**, the participant carries out the surgical task in VR, placing a virtual K-wire along a narrow bony corridor under X-ray image guidance. The participant communicates with an investigator, who performs the enabling task as the C-arm technologist, manipulating a virtual C-arm to acquire simulated X-ray images under the participant's direction. During the **enabling intervention**, the participant performs the enabling task (manipulating the virtual C-arm) and the investigator performs the surgeon task (placing the virtual K-wire). Through relevant mailing lists and word of mouth, we recruited 30 participants with no prior surgical experience, ensuring that their impressions of the surgical and enabling tasks were based solely on these interventions. After the enabling intervention, these users are a reasonable analog for real technologists, having observed the surgical task while successfully performing the enabling task. We then repeated the same study with a professional technologist and attending physician in orthopedic surgery.

Figure 2 summarizes the study design. After a 5-minute tutorial on VR, participants were divided into two groups of 15 participants each, which differed only in the order of interventions. Group A experienced the enabling intervention first, while group B experienced the surgeon intervention first. The initial intervention was limited to 15 min due to time constraints, although investigators expressed that successful task completion was not necessary. Rather, they asked participants to balance minimizing the number of shots with the number of repeat attempts, that is inserting the K-wire along an unsuitable trajectory and subsequently retracting it. After the initial intervention, users reported their experienced task load using the standard NASA Task Load Index (TLX) on a scale of 1 (very low) to 7 (very high). To measure participants' shared mental model of the overall task, we also asked them to estimate the task load of the other role using a modified TLX survey. So, for example, group A estimated the task load of the surgeon task after performing the enabling task. To prevent any confusion, we explicitly distinguished between the task load of the investigator, who is skilled at the surgical task just as a real surgeon would be, and the estimated task load that individuals would experience performing the alternate role themselves. (See 7 for a full list of survey questions.) Afterward, participants experienced the second intervention and repeated the standard TLX for the second task. After both interventions, we asked individuals to list their major concerns, reflecting on the challenges they experienced during each intervention To enable a fair comparison, the same investigator conducted all studies.

Virtual reality environment for image-guided surgery

For this study, we implement a VR environment for simulating both the surgeon and C-arm technologist roles, shown in Fig. 1. The environment is built in Unity and deployed on the Meta Quest 2 headset, with interpupillary distance set according to user's preference. It includes a virtual mobile C-arm based on the GE Healthcare OEC One and simulated X-ray images displayed on a virtual monitor. A simplified physics model recreates the challenges of real K-wire placement in soft and hard tissue, with haptic feedback provided through physical joysticks. Users occupy the same virtual environment together in real time, enabling live demonstration and simulation of teamwork.

In designing a virtual C-arm, we focus on providing intuitive controls that emulate a physical system. X-ray technologists typically manipulate mobile C-arms by unlocking specific degrees of freedom (DOFs) and grasping physical handles on the arm itself (for rotation) or the base (for translation). Previous work has simulated mobile C-arm movement in a digital environment [24], but the user interface was limited to control single axes with a keyboard and mouse. By contrast, we enable users with minimal training to manipulate the C-arm by grabbing virtual handles with physical controllers and position them in space, so that even novice users can perform the enabling task successfully after a short

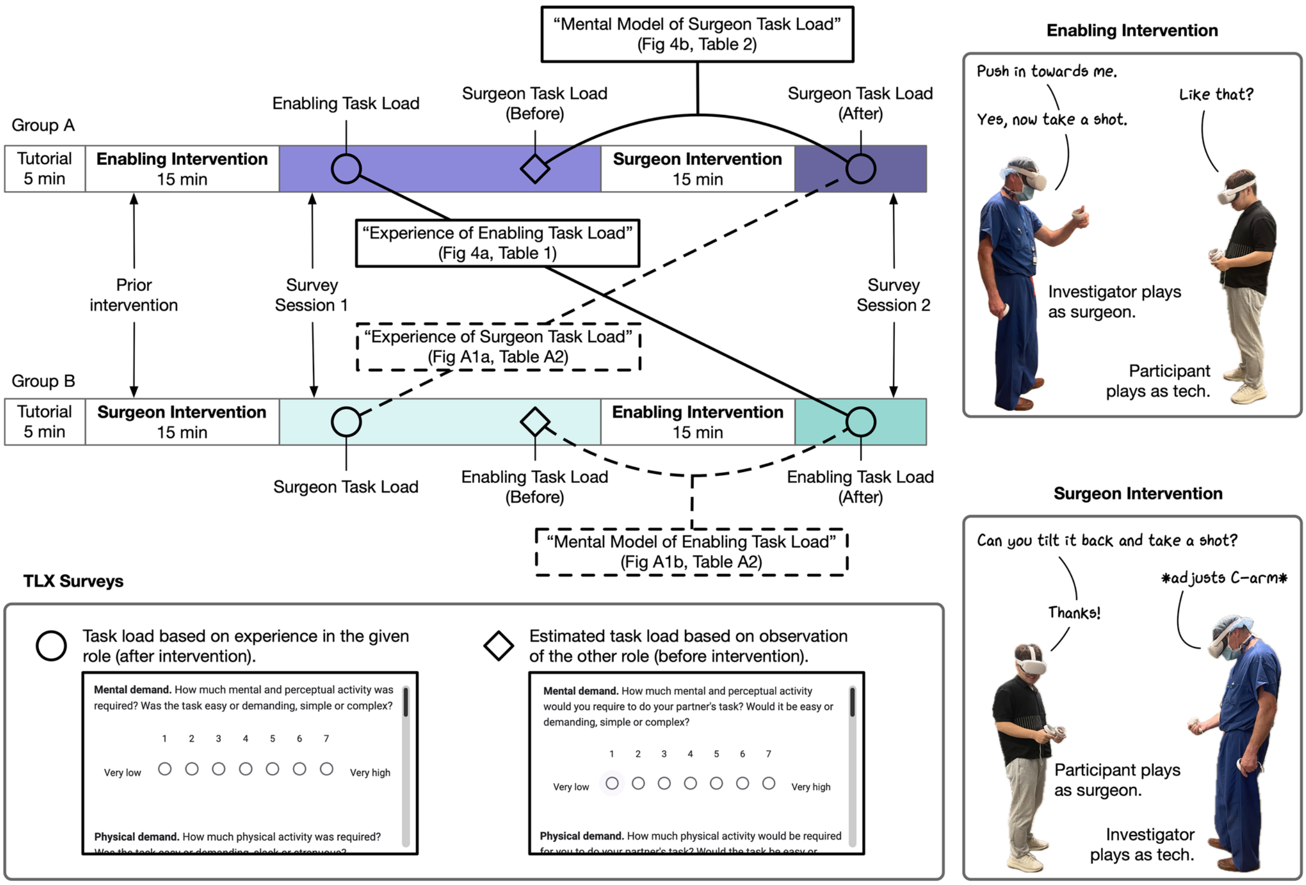


Fig. 2 Our study design investigates the effect of VR cross-training for non-surgeons. A session consists of a short tutorial followed by an initial intervention (the "prior"), a survey session, a second intervention, and a final survey. Participants belonged either to group A, which experienced the enabling intervention first, or group B, which experienced the surgeon intervention followed by the enabling. We consider the effect of cross-training first by comparing the actual experienced task load for the enabling intervention between the two groups (Fig. 4a, Table 1) and second by comparing Group A's evaluation of the surgeon task load, as imagined based on observation, to that same evaluation as based on first-hand experience (Fig. 4b, Table 2). For completeness, the corresponding comparisons for Group B (dashed lines) are available in the appendix

tutorial. An attending physician confirmed that this interface mirrors the physical interactions they have with real-world C-arm experience, making it intuitive and effective in a virtual setting.

We simulate X-ray images of a patient model, including inserted K-wires, in real time as digitally reconstructed radiographs (DRRs). For novice users, the patient model consists of a bony cylinder randomly oriented in a soft tissue cube, which makes it feasible for non-surgeons to interpret images while still maintaining the spatial reasoning aspect of the surgeon task. For expert users, the patient model consists of a segmented CT image from [30], from which we derive corresponding surface meshes for rendering. A server uses the original CT image to render DRRs based on the current pose of the patient model, instruments, and C-arm in the HMD environment. We use a modified version of DeepDRR [28, 31] to render 500×500 images with a pixel size of 1.1mm and a source-to-detector distance of 1020 mm.[1] The DRR server also relays the shared state of all objects and images among multiple HMDs, facilitating multiplayer interactions.

To approximate the physical challenges of placing wires, we implement a tool-to-tissue interaction suited for controllers in free space. When a user picks up a wire, it moves freely in space until inserted into the patient model. There, it continues to move in the direction of the controller but with 30% velocity. Additionally, the wire's center of rotation is continuously updated to the midpoint $\mathbf{p}_{\text{mid}}$ of the entry point $\mathbf{p}_{\text{entry}}$ and the tip $\mathbf{p}_{\text{tip}}$, with a maximum rotation of 30° from the initial trajectory at insertion (see Fig. 3). When $\mathbf{p}_{tip}$ makes contact with the bone surface, it becomes the center of rotation, as of a sharp point gaining purchase. In order to insert the wire into bone, the user activates a drill button and

[1] Available at github.com/arcadelab/deepdrr.

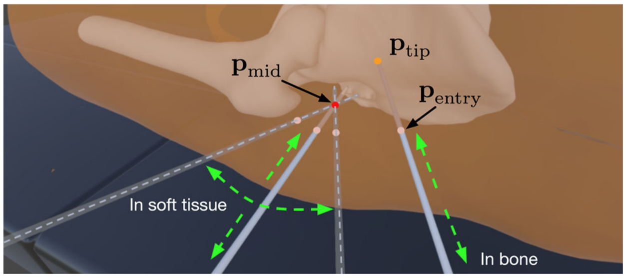


Fig. 3 K-wire interaction simulation. The K-wire pivot point $\mathbf{p}_{\text{mid}}$ in tissue is defined by the midpoint of the $\mathbf{p}_{\text{entry}}$ and $\mathbf{p}_{tip}$. The simulated resistance from the tissue changes as it goes deeper. The possible motions in the body are depicted in green arrows

advances their controller. During this motion, the K-wire's movement is constrained to the single degree of freedom along its current trajectory, as in Fig. 1. We further enhance this interaction with haptic feedback via controller vibration. A half-strength impulse indicates entry into soft tissue, while a full-strength impulse indicates contact and entry into bone. Altogether, these constraints force the user to fully retract and re-place the wire in order to change trajectories, following the real surgical workflow, despite having no physical constraints on free-hand controller movement. When the wire is fully inserted inside the bone, passing the designated endpoint, the environment indicates successful task completion and resets. Users were not penalized for incorrect placements but rather tried again until success was achieved or 15 min elapsed.

Results

The study design allows us to answer two important questions. First, what is the effect of VR cross-training on non-surgeon's experience of their own role, *i.e*let@tokeneonedot the enabling task? If cross-training enhances a shared mental model, then we would expect to see a change in users' task load, perhaps due to a greater understanding of their contribution to the overall objectives. Second, does cross-training change users' mental model of the surgical task? Participants estimate of the task load they themselves would experience in the surgeon's role reveals their understanding based on observation of the potential difficulty and challenges of the surgical task, which in this case depend on 3D spatial reasoning from 2D images. Given space constraints, the main text here focuses on data pertaining to these questions, but we make the full results available in the Appendix.

Experience of Enabling Task Load Cross-training on the surgeon task has a significant effect on participants' enabling task load. As shown in Fig. 4a, it increases the mental demand and effort that participants expend by 1.47 (on a 7-point scale) and 1.87, respectively ($p < 0.001$), indicating increased engagement with the task. This is reflected in the same tendency for the perception of temporal demand (+0.47 with $p < 0.001$), although we did not explicitly impose time constraints on task completion. Table 1 describes this comparison in more detail. A similar effect is present for the experience of the surgeon task load (see Table 4), indicating a possible learning effect where either prior intervention increases engagement with the subsequent intervention.

We observe, however, that no such learning effect is present for the frustration experienced during the enabling intervention. For Group A, the prior surgeon intervention is associated with an average decrease in frustration of 2.20 ($p < 0.001$) for the enabling intervention. That is, non-surgeon participants who had undergone VR cross-training were significantly less frustrated while performing the enabling task, obtaining X-ray images, but there was no change in frustration with the task order reversed. By contrast, for Group B, the prior enabling intervention is not associated with any significant change in frustration. This indicates that the decrease in frustration for the enabling task is not simply due to increasing familiarity as the study progresses but reflects the information imbalance between the two tasks. We conjecture that performing the surgeon task, which can actually produce the outcome of placing a K-wire, offers a more complete understanding of how exactly the enabling task contributes to overall task success. On the other hand, the enabling task can be performed successfully (although, we would argue, less effectively) simply by following the surgeon's instructions without requiring first-hand understanding.

Mental Model of Surgeon Task Load Indeed, our second analysis reveals just such a gap in understanding, as shown in Fig. 4b and Table 2. We observe that participants in group A underestimate the mental demand (by 1.07 on a 7-point scale), physical demand (0.73), and effort (1.27) that they would experience performing the surgeon task while also overestimating the overall performance they would achieve by 1.53. Together, these results indicate that the enabling intervention does not convey the difficulty participants would face performing the task themselves. This is not to say participants are necessarily confident they can perform the surgeon task successfully, but the degree to which they underestimate the challenges of that task may explain the increased engagement and decreased frustration found above.

The gap in understanding for the surgeon task is not reflected in group B's understanding of the enabling task, pointing toward a real imbalance between the two tasks. If anything, group B *over*estimates the difficulty of the enabling task in terms of physical demand, temporal demand, and frustration (see Table 5), although in this case the use of novice participants does not reflect the understanding surgeons nec-

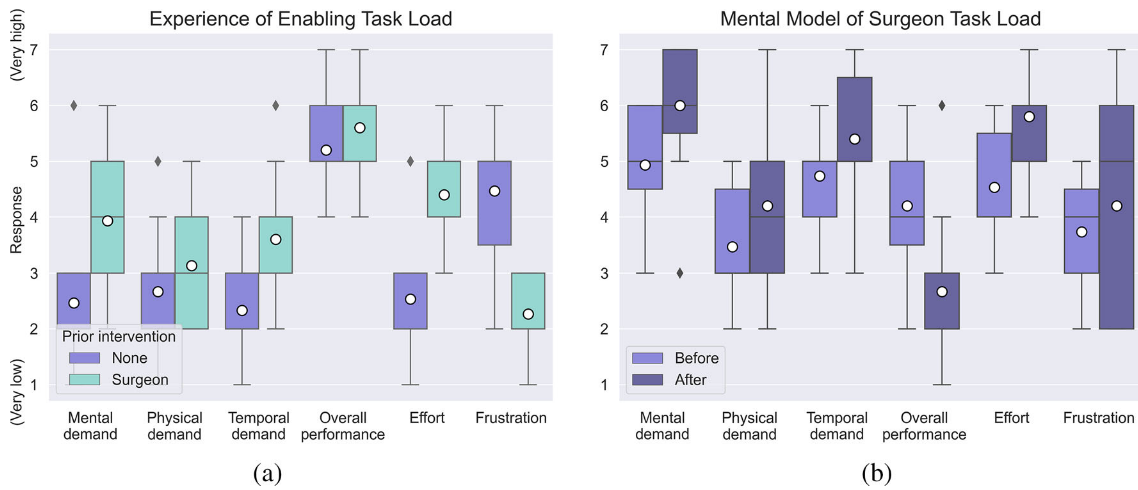


Fig. 4 **a** Participants' actual experience of the enabling task load, compared between group A (no prior intervention) and group B (prior: surgeon). Cross-training on the surgeon task reduces frustration and increases engagement with the enabling task, as reflected by mental demand and effort. **b** Group A's mental model of the enabling task, as quantified by the difference in their estimate based on observation and their actual experienced task load. Participants reliably underestimate the load they themselves will experience performing the surgeon task in terms of mental demand, physical demand, overall performance (which declines), and effort

Table 1 Experience of enabling task load

Question	Prior: None ($\mu \pm \sigma$)	Prior: Surgeon ($\mu \pm \sigma$)	Δ (μ)	t-statistic	p-value ($\mathbf{p < 0.05}$)
Mental demand	2.47 ± 1.13	3.93 ± 1.10	$+1.47$	3.61	**0.001**
Physical demand	2.67 ± 1.11	3.13 ± 1.06	$+0.47$	1.18	0.249
Temporal demand	2.33 ± 0.82	3.60 ± 1.06	$+1.27$	3.68	$<$ **0.001**
Overall performance	5.20 ± 0.86	5.60 ± 0.74	$+0.40$	1.37	0.183
Effort	2.53 ± 1.19	4.40 ± 0.91	$+1.87$	4.83	$<$ **0.001**
Frustration	4.47 ± 1.36	2.27 ± 0.59	-2.20	-5.76	$<$ **0.001**

Table 2 Mental model of surgeon task load

Question	Before ($\mu \pm \sigma$)	After ($\mu \pm \sigma$)	Δ ($\mu \pm \sigma$)	t-statistic	p-value ($\mathbf{p < 0.05}$)
Mental demand	4.93 ± 1.03	6.00 ± 1.13	$+1.07 \pm 1.61$	2.48	**0.027**
Physical demand	3.47 ± 1.13	4.20 ± 1.52	$+0.73 \pm 1.06$	2.58	**0.022**
Temporal demand	4.73 ± 0.88	5.40 ± 1.35	$+0.67 \pm 1.70$	1.47	0.164
Overall performance	4.20 ± 1.21	2.67 ± 1.29	-1.53 ± 1.93	-2.98	**0.010**
Effort	4.53 ± 1.13	5.80 ± 0.86	$+1.27 \pm 1.65$	2.87	**0.012**
Frustration	3.73 ± 0.96	4.20 ± 1.86	$+0.47 \pm 1.82$	0.96	0.354

essarily have of C-arm manipulation, acquired as a part of surgical training.

We find similar results in the evaluation of the surgeon task by a professional X-ray technologist. With an attending orthopedic surgeon acting as the investigator in VR, the technologist performed the enabling intervention followed by the surgeon intervention. They underestimated the difficulty of the surgeon task in terms of temporal demand (+ 2) and overall performance (− 2), and experienced higher frustration than anticipated. These results suggest that the findings from our study with novice participants may be present in professional surgical staff. Table 3 details these measurements.

Discussion

VR opens up new avenues for training surgical teams, involving exposure to multiple roles. The increasing realism and accessibility of VR has already yielded significant benefits for surgeons' skill acquisition [9]. Thus far, however, applica-

Table 3 Mental model of surgeon task load for professional technologist

Question	Before	After	Δ
Mental demand	5	5	0
Physical demand	4	4	0
Temporal demand	4	6	+ 2
Overall performance	5	3	− 2
Effort	5	5	0
Frustration	3	4	+ 1

tions for non-surgeons have focused on learning the steps of a procedure. Here, we have shown how experiential knowledge gained in VR can align the shared mental model for non-surgeons, increasing engagement and reducing frustration for individuals performing an enabling task in the OR. Before the surgeon intervention, for example, one participant noted, "I do not know whether the image I acquired is what the surgeon wants," focusing on satisfying short-term directives. After the surgeon intervention, participants' focus broadened to include overarching goals of the procedure, with one remarking, "It is hard to match the X-ray image to the real patient and K-wire position." Moreover, without experiencing the surgeon task firsthand, participants reliably underestimate the difficulty of completing the procedure. That is to say, experiencing the surgeon's role improved their understanding of the surgeon task, compared to understanding gained from observation. This is important in the context of teamwork because it allows technologists and other operating room staff to engage with long-term clinical goals and anticipate the related needs of their teammates.

Of course, the enabling intervention implemented here cannot fully capture on-the-job experience, and further study is needed to determine whether swapping roles in VR has a similar effect for professional technologists and other support staff. For instance, we found that VR cross-training had no significant effect on the number of shots taken or the likelihood of successful K-wire insertion, as innate spatial reasoning skills dominated performance outcomes for novice users. Nevertheless, our results indicate a significant effect on the shared mental model of non-surgeons, enabling a study of sufficient power involving professional support staff, surgeons, and surgical residents, a population for which performance can be meaningfully measured. They also point toward a very real informational asymmetry between novice participants and investigators is analogous to that of technologists and surgeons, but real surgeons already have exposure to the enabling task, including in the case of C-arm X-ray technologists. As surgical workflows grow more sophisticated, for instance by incorporating augmented reality [20] or artificial intelligence [21, 27], firsthand exposure to novel enabling tasks may prove valuable for surgeons' mental model.

Conclusion

In conclusion, we find that VR training may improve the effectiveness of teamwork in surgery by augmenting the shared mental model of non-surgeons. It has a significant effect on the engagement of study participants, as measured by their increased mental demand and effort, while also lowering the frustration experienced in the enabling task. Future work may measure participants engagement more directly, for instance by using newly available headsets with built-in gaze trackers. Moreover, participants underestimate the difficulty they themselves will experience performing the surgeon task, showing the positive effect of VR cross-training on a shared mental model. In future, we hope these results pave the way for some training in multiple roles to become a standard part of clinical curricula for surgeons and non-surgeons alike.

Appendix A: Study design details

The standard NASA TLX consists of the following questions, on a seven-point scale, where 1 is "Very low" and 7 is "Very high".

- **Mental demand.** How much mental and perceptual activity was required? Was the task easy or demanding, simple or complex?
- **Physical demand.** How much physical activity was required? Was the task easy or demanding, slack or strenuous?
- **Temporal demand.** How much time pressure did you feel due to the pace at which the tasks or task elements occurred? Was the pace slow or rapid?
- **Overall performance.** How successful were you in performing the task? How satisfied were you with your performance?
- **Effort.** How hard did you have to work (mentally and physically) to accomplish your level of performance?
- **Frustration.** How irritated, stressed, and annoyed versus content, relaxed, and complacent did you feel during the task?

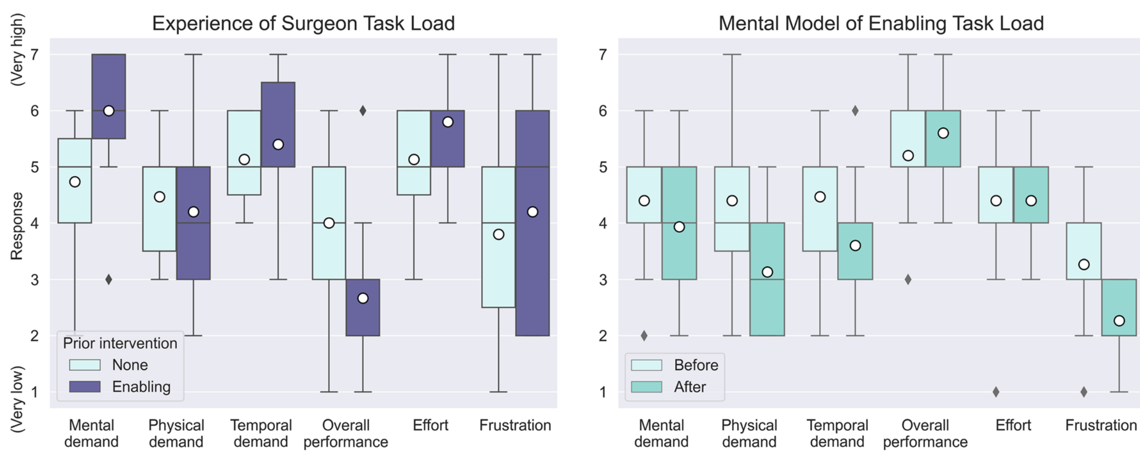
The observational TLX modifies these questions, so that participants estimate the task load they themselves would experience performing the task. Answers are given on the same scale.

- **Mental demand.** How much mental and perceptual activity would you require to do your partner's task? Would it be easy or demanding, simple or complex?
- **Physical demand.** How much physical activity would be required for you to do your partner's task? Would the task be easy or demanding, slack or strenuous?
- **Temporal demand.** How much time pressure would you feel doing your partner's task due to the pace at which the tasks or task elements occurred? Would the pace be slow or rapid?
- **Overall performance.** How successful would you be in performing your partner's task? How satisfied would you be with your performance?
- **Effort.** How hard would you have to work (mentally and physically) to accomplish your level of performance in your partner's task?
- **Frustration.** How irritated, stressed, and annoyed versus content, relaxed, and complacent would you feel doing your partner's task?

Participants included English-speaking non-medical students and Johns Hopkins employees over 18 with no disabilities and no history of motion sickness. Participants were consented verbally, and no further demographic information was recorded.

See Fig. 5.

See Tables 4 and 5.



(a) (b)

Fig. 5 **a** The task load of the surgeon intervention for group B, which had no prior intervention, and group A, which experienced the enabling intervention. **b** The difference between the estimated task load of the enabling, measured *before* the enabling intervention, and the task load measured after

Table 4 Experience of surgeon task load

Question	Prior: None ($\mu \pm \sigma$)	Prior: Tech ($\mu \pm \sigma$)	Δ (μ)	t-statistic	p-value ($\mathbf{p < 0.05}$)
Mental demand	4.73 ± 1.16	6.00 ± 1.13	+ 1.27	3.02	**0.005**
Physical demand	4.47 ± 1.13	4.20 ± 1.52	−0.27	−0.55	0.590
Temporal demand	5.13 ± 0.83	5.40 ± 1.35	+ 0.27	0.65	0.521
Overall performance	4.00 ± 1.46	2.67 ± 1.29	−1.33	−2.65	**0.013**
Effort	5.13 ± 0.99	5.80 ± 0.86	+ 0.67	1.97	0.059
Frustration	3.80 ± 1.61	4.20 ± 1.86	+ 0.40	0.63	0.534

Table 5 Mental model of enabling task load

Question	Before ($\mu \pm \sigma$)	After ($\mu \pm \sigma$)	Δ ($\mu \pm \sigma$)	t-statistic	p-value ($\mathbf{p < 0.05}$)
Mental demand	4.40 ± 1.06	3.93 ± 1.10	−0.47 ± 1.45	−1.20	0.250
Physical demand	4.40 ± 1.35	3.13 ± 1.06	−1.27 ± 1.98	−2.39	**0.031**
Temporal demand	4.47 ± 1.25	3.60 ± 1.06	−0.87 ± 1.15	−2.83	**0.013**
Overall performance	5.20 ± 1.08	5.60 ± 0.74	+ 0.40 ± 1.02	1.47	0.164
Effort	4.40 ± 1.30	4.40 ± 0.91	+ 0.00 ± 1.75	0.00	1.000
Frustration	3.27 ± 1.16	2.27 ± 0.59	−1.00 ± 1.03	−3.62	**0.003**

Supplementary Information The online version contains supplementary material available at https://doi.org/10.1007/s11548-024-03138-7.

Funding This work was supported by the Link Foundation Fellowship for Modeling, Training, and Simulation; the NIH under Grant No. R21EB028505, the NSF under Award No. 2239077, and Johns Hopkins University Internal Funds.

Declarations

Conflicts of interest The authors declare that they have no Conflict of interest.

Ethical approval The study was conducted in accordance with the Declaration of Helsinki, and the protocol was approved by the Johns Hopkins University Institutional Review Board (HIRB00017105).

Informed consent All subjects gave their informed consent for inclusion before they participated in the study.

References

1. Mazzocco K, Petitti DB, Fong KT, Bonacum D, Brookey J, Graham S, Lasky RE, Sexton JB, Thomas EJ (2009) Surgical team behaviors and patient outcomes. Am J Surg 197(5):678–685. https://doi.org/10.1016/j.amjsurg.2008.03.002
2. Catchpole KR, Giddings AEB, Wilkinson M, Hirst G, Dale T, Leval MR (2007) Improving patient safety by identifying latent failures in successful operations. Surgery 142(1):102–110. https://doi.org/10.1016/j.surg.2007.01.033
3. Aveling E-L, Stone J, Sundt T, Wright C, Gino F, Singer S (2018) Factors influencing team behaviors in surgery: a qualitative study to inform teamwork interventions. Ann Thorac Surg 106(1):115–120. https://doi.org/10.1016/j.athoracsur.2017.12.045
4. Weaver SJ, Rosen MA, DiazGranados D, Lazzara EH, Lyons R, Salas E, Knych SA, McKeever M, Adler L, Barker M, King HB (2010) Does teamwork improve performance in the operating room? A multilevel evaluation. Jt Comm J Qual Patient Saf 36(3):133–142. https://doi.org/10.1016/S1553-7250(10)36022-3
5. Volpe CE, Cannon-Bowers JA, Salas E, Spector PE (1996) The impact of cross-training on team functioning: an empirical investigation. Hum Factors 38(1):87–100. https://doi.org/10.1518/001872096778940741
6. Marks MA, Sabella MJ, Burke CS, Zaccaro SJ (2002) The impact of cross-training on team effectiveness. J Appl Psychol 87(1):3
7. Hedges AR, Johnson HJ, Kobulinsky LR, Estock JL, Eibling D, Seybert AL (2019) Effects of cross-training on medical teams' teamwork and collaboration: use of simulation. Pharmacy 7(1):13. https://doi.org/10.3390/pharmacy7010013
8. Anastakis DJ, Regehr G, Reznick RK, Cusimano M, Murnaghan J, Brown M, Hutchison C (1999) Assessment of technical skills transfer from the bench training model to the human model. Am J Surg 177(2):167–170. https://doi.org/10.1016/S0002-9610(98)00327-4
9. Mao RQ, Lan L, Kay J, Lohre R, Ayeni OR, Goel DP, Sa DD (2021) Immersive virtual reality for surgical training: a systematic review. J Surg Res 268:40–58. https://doi.org/10.1016/j.jss.2021.06.045
10. Cate G, Barnes J, Cherney S, Stambough J, Bumpass D, Barnes CL, Dickinson KJ (2023) Current status of virtual reality simulation education for orthopedic residents: the need for a change in focus. Global Surg Educat-J Associat Surg Educat. https://doi.org/10.1007/s44186-023-00120-w
11. Rebol M, Pietroszek K, Ranniger C, Hood C, Rutenberg A, Sikka N, Li D, Gütl C: Mixed reality communication for medical procedures: teaching the placement of a central venous catheter. In: 2022 IEEE international symposium on mixed and augmented reality (ISMAR), pp 17–21. IEEE. https://doi.org/10.1109/ISMAR55827.2022.00050
12. Blumstein G, Zukotynski B, Cevallos N, Ishmael C, Zoller S, Burke Z, Clarkson S, Park H, Bernthal N, SooHoo NF (2020) Randomized trial of a virtual reality tool to teach surgical technique for tibial shaft fracture intramedullary nailing. J Surg Educ 77(4):969–977. https://doi.org/10.1016/j.jsurg.2020.01.002
13. Luca A, Giorgino R, Gesualdo L, Peretti GM, Belkhou A, Banfi G, Grasso G (2020) Innovative educational pathways in spine surgery: advanced virtual reality-based training. World Neurosurg 140:674–680. https://doi.org/10.1016/j.wneu.2020.04.102
14. Miller ZA, Amin A, Tu J, Echenique A, Winokur RS (2019) Simulation-based training for interventional radiology and opportunities for improving the educational paradigm. Tech Vasc Interv Radiol 22(1):35–40. https://doi.org/10.1053/j.tvir.2018.10.008
15. Essoe JK-Y, Reggente N, Ohno AA, Baek YH, Dell'Italia J, Rissman J (2022) Enhancing learning and retention with distinctive virtual reality environments and mental context reinstatement. Npj Sci Learn 7(31):1–14. https://doi.org/10.1038/s41539-022-00147-6
16. Berry M, Lystig T, Beard J, Klingestierna H, Reznick R, Lönn L (2007) Porcine transfer study: virtual reality simulator training compared with porcine training in endovascular novices. Cardiovasc Intervent Radiol 30(3):455–461. https://doi.org/10.1007/s00270-006-0161-1
17. Lohre R, Bois AJ, Athwal GS, Goel DP (2020) on behalf of the Canadian Shoulder and Elbow Society (CSES)∗: Improved Complex Skill Acquisition by Immersive Virtual Reality Training: A Randomized Controlled Trial. JBJS 102(6):26. https://doi.org/10.2106/JBJS.19.00982
18. Munawar A, Li Z, Kunjam P, Nagururu N, Ding AS, Kazanzides P, Looi T, Creighton FX, Taylor RH, Unberath M (2021) Virtual reality for synergistic surgical training and data generation. Comput Methods Biomech Biomed Eng: Imag Visualiz. https://doi.org/10.1080/21681163.2021.1999331
19. Unberath M, Fotouhi J, Hajek J, Maier A, Osgood G, Taylor R, Armand M, Navab N (2018) Augmented reality-based feedback for technician-in-the-loop C-arm repositioning. Healthcare Technol Lett 5(5):143–147. https://doi.org/10.1049/htl.2018.5066
20. Killeen BD, Winter J, Gu W, Martin-Gomez A, Taylor RH, Osgood G, Unberath M (2023) Mixed reality interfaces for achieving desired views with robotic X-ray systems. Comput Methods Biomech Biomed Eng: Imag Visualiz. https://doi.org/10.1080/21681163.2022.2154272
21. Killeen BD, Gao C, Oguine KJ, Darcy S, Armand M, Taylor RH, Osgood G, Unberath M (2023) An autonomous X-ray image acquisition and interpretation system for assisting percutaneous pelvic fracture fixation. Int J CARS 18(7):1201–1208. https://doi.org/10.1007/s11548-023-02941-y
22. De Silva T, Punnoose J, Uneri A, Mahesh M, Goerres J, Jacobson M, Ketcha MD, Manbachi A, Vogt S, Kleinszig G, Khanna AJ, Wolinksy J-P, Siewerdsen JH, Osgood G (2018) Virtual fluoroscopy for intraoperative C-arm positioning and radiation dose reduction. J Med Imaging. https://doi.org/10.1117/1.JMI.5.1.015005
23. Allen DR, Moore J, Joschko A, Clarke C, Peters TM, Chen ECS (2020) Miniature C-arm simulator using wireless accelerometer based tracking. In: Proceedings Volume 11315, Medical Imaging 2020: Image-Guided Procedures, Robotic Interventions, and Modeling (vol. 11315, p. 1131502). SPIE, https://doi.org/10.1117/12.2547388

24. Bott OJ, Dresing K, Wagner M, Raab B-W, Teistler M (2011) Informatics in radiology: use of a C-Arm fluoroscopy simulator to support training in intraoperative radiography. Radiographics
25. Killeen BD, Chakraborty S, Osgood G, Unberath M (2022) Toward perception-based anticipation of cortical breach during K-wire fixation of the pelvis. In: Proceedings Volume 12031, Medical Imaging 2022: Physics of Medical Imaging (vol. 12031, pp. 410–415). SPIE, https://doi.org/10.1117/12.2612989
26. Killeen BD, Cho SM, Armand M, Taylor RH, Unberath M (2023) In silico simulation: a key enabling technology for next-generation intelligent surgical systems. Prog Biomed Eng 5(3):032001. https://doi.org/10.1088/2516-1091/acd28b
27. Kausch L, Thomas S, Kunze H, Privalov M, Vetter S, Franke J, Mahnken AH, Maier-Hein L, Maier-Hein K (2020) Toward automatic C-arm positioning for standard projections in orthopedic surgery. Int J CARS 15(7):1095–1105. https://doi.org/10.1007/s11548-020-02204-0
28. Gao C, Killeen BD, Hu Y, Grupp RB, Taylor RH, Armand M, Unberath M (2023) Synthetic data accelerates the development of generalizable learning-based algorithms for X-ray image analysis. Nat Mach Intell 5(3):294–308. https://doi.org/10.1038/s42256-023-00629-1
29. Killeen BD, Zhang H, Mangulabnan J, Armand M, Taylor RH, Osgood G, Unberath M (2023) Pelphix: Surgical Phase Recognition from X-Ray Images in Percutaneous Pelvic Fixation. In: Medical Image Computing and Computer Assisted Intervention – MICCAI 2023, pp 133–143. Springer, Cham, Switzerland. https://doi.org/10.1007/978-3-031-43996-4_13
30. Edgar H, Daneshvari Berry S, Moes E, Adolphi N, Bridges P, Nolte K (2020) New Mexico Decedent Image Database. Office of the Medical Investigator, University of New Mexico. https://doi.org/10.25827/5s8c-n515
31. Unberath M, Zaech J-N, Lee SC, Bier B, Fotouhi J, Armand M, Navab N (2018) Deepdrr - a catalyst for machine learning in fluoroscopy-guided procedures. In: Frangi AF, Schnabel JA, Davatzikos C, Alberola-López C, Fichtinger G (eds) Medical Image Computing and Computer Assisted Intervention - MICCAI 2018. Springer, Cham, pp 98–106