White Paper

BENEFITS OF USING IMMERSIVE TRAINING IN INTERVENTIONAL MEDICINE

https://number9vr.com/vr9/
info@number9vr.com

CONTENTS

INTRODUCTION

Training in the operating room (OR) is an essential aspect in the education of a wide range of healthcare professionals. It is the main venue for surgeons in training, and the ideal setting for the specialized training of OR nurses, technicians and other perioperative personnel. Training in the OR, however, presents important challenges. The OR is a highly complex environment, marked by the interactions of personnel fulfilling different roles, the presence of complex medical equipment and the dynamics of situational necessity. Additionally, training in the OR may be constrained by workplace and logistical demands. Recurrent limitations include lack of surgeries, limited available OR time, and physical space for only a small number of trainees per surgery. Training in the OR can be costly if it requires that trainees travel to attend training and need to cover transportation and lodging expenses. Institutions, health professionals and trainees may find it difficult to coordinate schedules and resources, and the possibility of sudden surgery rescheduling increases the monetary risk.

Over the last few years, institutions and companies have resorted to digital technology to enhance medical education, and to counter the limitations and difficulties of in-person, on-location training. Technical advancements have made it possible to realistically reproduce the view of a human observer in a given environment. It is now common in healthcare to use various technologies such as digital simulations, 3D modeling of anatomy or medical devices, and immersive audiovisual experiences for educational or therapeutic purposes. The increased capacity and lower cost of available computing power have made it feasible to create and deploy immersive experiences on a mass scale, while retaining high fidelity as an audiovisual medium.

Among digital technologies, immersive digital video techniques such as virtual reality (VR) have been used in healthcare for more than 20 years (see e.g. Mc Cloy, 2001). Today's immersive technology affords the possibility of recording in video, then playing back, any environment in such a way that the viewer experiences a sense of presence, i.e. the feeling of being physically present in the environment where the video was recorded. This perception facilitates attention and knowledge retention to a similar or equal degree as being physically present in the original environment.

Since 2018, VR9 has recorded training sessions in the OR, during real surgical procedures. The intervening surgeons perform the surgery and provide training in front of a

360-degree camera. After postproduction, a trainee can view the training session as if standing where the camera was located.

This white paper describes this modality of training, with a special focus on using training sessions recorded during surgical interventions. An overview is provided of the advantages and disadvantages of this modality in comparison to on-location training, as well as budgetary considerations and the possibilities of integrating immersive content with other training materials.

BRIEF OVERVIEW OF VR AND IMMERSIVE VIDEO

The term "virtual reality" or VR is often employed informally. The "range of relatively varied and heterogeneous definitions of VR that can be found in the literature" has caused at least one group of healthcare professionals to issue a call to unify definitions (Kardong-Edgren et al, 2019). The term is sometimes used to designate 360-degree immersive video recorded in real physical environments. While VR and 360-degree video share key attributes, specially regarding visualization, in this white paper we consider them as different media formats.

The term "virtual reality" usually refers to a simulated environment composed of computer-generated images (CGI). This environment contains virtual objects and entities that simulate the behavior of real-world entities. VR media and devices may provide interfaces for real-time user interaction, such as hand-held devices with haptic feedback. The user can interact with elements in the virtual world, for example by manipulating hand controls to manipulate virtual objects. In VR, the totality of the virtual world, or at least all elements that allow for user interaction, are generated by CGI.

Unlike VR, immersive video is recorded in a real environment using a camera with an omnidirectional field of view, called a 360-degree camera. The camera uses at least two lenses to capture the totality of the environment around it. Each lens captures a portion of the visible environment around the camera. In postproduction, the images from each lens are assembled to conform a final 360-degree spherical image. As with VR, a user viewing the finished video in a head-mounted display (HMD) has the illusion of being physically present within the environment, that is to say where the camera was located when the video was recorded. The user can turn to look at any part of the environment just as in the real world. The experience, however, is not interactive, except if interactive elements were added after the actual filming.

Degrees of Immersion

In the context of VR research, immersion is defined as the objective degree to which an immersive system conveys a naturalistic portrayal in all sensory modalities (Slater, 2003). The sense of "presence" is a human response to the sensory input provided by the immersive system. The degree of immersion can be determined by objective criteria such as visual and auditory fidelity. For instance, visual fidelity is determined, in part, by image and color resolution, field of view and field of regard.

Immersive videos can be played back on 2D screens, but more effective playback devices are head-mounted displays (HMDs) such as VR headsets. In this case, immersion, and the ensuing sense of presence, are so effective that they engage the user's proprioceptive senses, enhancing spatial awareness and memory (Krokos et al., 2018).

IMMERSION AND LEARNING

Numerous studies have documented that human memory is essentially dependent on context (Essoe et al., 2022, Smith & Vela, 2001, Smith, 1979). Contextual cues can aid knowledge retention and recall, even if they are not related to the knowledge itself; conversely, their absence can hinder recall (Tulving & Thompson, 1973).

The "method of loci" or "memory palace" is a well-known mnemonic technique that relies heavily on context. It uses a mental representation of a physical environment to spatially organize information within that environment. The environment itself may be a recollection of an actual location, or entirely imaginary. During recall, the subject mentally traverses the imaginary environment—e.g. different rooms in a palace—and "retrieves" the memorized items of information. Recall is activated by the context in which the items are "stored," i.e. the specific place within the environment. The technique was described in classical Greek and Roman treatises on oratory. The Roman orator Cicero is believed to have employed it to visualize portions of his speeches in locations within the auditorium he was in.

Recent research in cognitive psychology supports the view that the mind is embodied, i.e. that bodily states or the body's interaction with the surrounding environment contribute to cognition (Barsalou 2008, Shapiro, 2007). Sensory and motor systems influence the way we construct concepts and recall information. Information about the environment is stored and recalled by the memory system known as spatial memory. Spatial memory and embodied navigation are closely connected to memory (Leutgeb et al., 2005). Immersive environments engage our vestibular system and proprioceptive senses, providing superior spatial awareness. Compared to traditional 2D display systems, they al-

low viewers to better recall information (Krokos et al., 2018). Many studies indicate that display systems with better immersion capabilities aid in episodic memory performance (Smith, 2019).

How exactly learning in an immersive environment compares to real-world learning is not readily apparent. Several studies indicate that memory retention in immersive virtual environments is inferior than in real-world learning (Flannery & Walles, 2003). This apparent deficiency, however, may be offset by the possibility of users spending more time in the immersive experience (Waller et al., 1998), e.g. by repeated viewings.

The usefulness of VR training depends on the possibility of transfer of learning from the VR setting to the real-world environment that was simulated (Smith, 2019). Knowledge gained in an OR inside an immersive experience would not be useful if it could not be precisely recalled during actual surgical interventions. In this regard, many studies have demonstrated that information acquired in a virtual environment can be reliably recalled in an analogous real-life environment (Smith, 2019, Ragan et al., 2010).

Most importantly, numerous studies show positive results when implementing immersive environments for training, including performance improvements for beginner surgeons in the OR (Lan et al., 2023). A systematic review of 17 articles covering 307 participants in four disciplines found that participants using immersive VR "demonstrated improvement of surgical skills through iVR" including "faster procedural completion times, greater task completion scores, and improved accuracy of surgical device implantation" (Mao et al., 2021). Compared to in-person training, VR systems afford the possibility of training in a risk-free environment, and of repeating training lessons as many times as is desired.

IMMERSIVE TECHNOLOGIES IN HEALTHCARE

Immersive technologies have been in use in healthcare for years. They are often employed in simulations—e.g. for training, equipment evaluation or interacting with patients—and in managing patient health and rehabilitation (Dhar E. et al. 2023).

This section focuses on immersive videos of classes and clinical cases, recorded during surgical interventions in the OR. This modality of training is especially used to expand the reach of traditional training methods. Immersive technology allows more medical professionals to access training provided by interventionists in the OR. Common subjects of training sessions include surgical procedures, evaluation or usage of medical equipment, and demonstrations of specific clinical cases. —

Relative to traditional training in the OR, some of the main advantages of immersive training are:

- Possibility of reaching a much larger audience
- Reduced costs
- Reduced logistic demands on workplace and personnel
- Possibility of repeated viewings of training sessions
- Possibility of integrating immersive training or VR simulations into wider training curricula

Employing immersive training allows for facilitating an experience very similar to training in the OR to a far greater number of trainees than would be possible with in-person training. It also allows trainees to access training from anywhere in the world, as many times as needed.

For the company or institution producing training sessions, one of the clearer advantages is the lower cost associated with providing training. Depending on a number of factors, cost reduction can reach 85%.

The below sections examine some usage possibilities, as well as advantages and disadvantages, of immersive training.

Example Use Case: Surgical Intervention

Medical Lesson in the OR

In the OR, an interventional cardiologist (IC) and his team are ready to begin a surgical procedure. The patient is fully sedated. The IC is standing next to the operating table. On the other side of the operating table is a 360-degree camera, about the size of a smartphone, set on a tripod.

The IC talks directly to the camera, as if it were a trainee. He briefly describes the clinical case: a 74-year-old man, with a history of coronary disease and chronic kidney disease, will receive a cardiac device for percutaneous left atrial appendage closure. The patient has been imaged with MRI during pre-procedural planning.

The IC begins the procedure. For the next hour he and his team will explain the proceedings as they perform the intervention. The camera remains fixed in a stationary position on its tripod, while a single technician monitors the video recording. There are two additional cameras recording closeups of the IC's hands and the output of medical instrument monitors.

Occasionally as he works, the IC comments on the procedure. He outlines the steps followed in the intervention, describes best practices, remarks on the readings of medical instruments, and gives such practical indications as where to place the instrument monitors and what to watch out for during catheter insertion.

The procedure lasts 55 minutes. Once it is finished, the IC confirms that everything went as planned, and thanks his colleagues.

The Experience in the VR Headset

For this specific clinical case, it was decided that the most effective approach was to begin the training video not with the surgical intervention, but with an overview of the pre-operative studies conducted on the patient. Thus the immersive video begins with the MRI images of the patient studies, and a verbal explanation provided by the IC.

As the immersive video begins, the trainee has the illusion of standing on an abstract white space, facing a large, finely detailed screen of cardiac MRI images. In voice-over, the IC describes the clinical case and procedural planning criteria that determined the use of MRI images. The MRI images zoom and scroll as the voice-over explains how to use them as aids for device implantation, covering subjects such as identifying the entry point and determining the angle and landing zone for the device.

This first section of the immersive experience runs for about five minutes. The trainee is then transported into the OR. He or she has the sense of standing exactly where the camera was located during the intervention, next to the operating table and facing the IC. The IC talks directly to the trainee.

The intervention begins. The procedure has been edited to remove redundant portions, such as most of the catheter insertion. In this case, the client—the biotechnology engineering firm that manufactures the cardiac implant—requested several edited versions, including a 30-minute version and a summarized version of 12 minutes.

The trainee can look directly at the IC, or down at the patient, or up and to the sides to view enlarged versions of the instrument monitors. After the first viewing, the trainee can watch the experience as many times as desired.

BUDGETARY CONSIDERATIONS

This section contains a comparison of costs between traditional in-person training in the OR and training in immersive video. Costs and time estimates are based on a US company providing training sessions in the US East Coast, Latin America and Europe.

Data for this section is derived from VR9's collaboration with clients in the United States during the years 2022 and 2023.

Training in the OR

Materializing a training session in the OR requires fulfilling a series of steps, of which some are subject to the influence of adverse external factors such as scheduling conflicts or lack of available surgeries. Roughly outlined, these steps are:

- Coordinating with interventional surgeons and the healthcare institution to select the relevant surgical interventions
- Coordinating attendance and travel schedules with each trainee, and each trainer when required
- Covering travel and lodging expenses for each trainee, and each trainer when required

The time estimates for one trainee's attendance to 10 surgical cases are:

- **Organization**: 0.5 to 2 days depending on surgery and resource availability
- **Attendance to 10 surgical cases**: At least 9 days per trainee, including travel

(The above estimates assume that surgeries are programmed back-to-back, minimizing transportation and lodging costs.)

Budget Estimate

The average budget estimate for one trainee's attendance to 10 surgical cases is:

- **Surgical cases within the US**: $2500-$3000
- **Surgical cases outside the US**: $4500-$6000

(These numbers do not include the cost of trainers who are sometimes required to provide support en route.)

Real-world Example

In order to provide a training course comprising 10 surgeries, one of our clients estimated an average cost of $3720 per trainee. Since the client needed to provide training to 90 company representatives, the overall cost was estimated at $335,000. The minimum required time to provide the training course for all 90 representative was three years.

The required number of surgeries to provide training to all representatives was difficult to predict. Variable factors such as the number of trainees allowed to witness each surgery in the OR, and the possible need to repeat surgical procedures in certain cases, caused the final estimate to fall at anywhere between 300 and 900 surgical interventions.

Disadvantages of Training in the OR

- Adapting to varying surgical demand/lack of surgeries
- Training is time-consuming to organize and implement
- Quality of surgeries and clinical cases can be inconsistent
- Attendant risks such as surgery rescheduling
- Need to repeat process for every new trainee
- Each procedure may be performed and seen only once by each trainee
- Limited number of trainees in the OR, each with a limited view

Advantages of Training in the OR

- Presence in the actual operating theater will likely bring trainee attention to its highest level
- If the situation allows, the interventionist may interact with trainees, for example by answering questions

Training with Immersive Experiences

Filming a training session with immersive video in the OR requires the same steps as taking one trainee to witness the surgical intervention. However, these steps are performed only once per surgery or set of surgeries. The presence of a videographer supplants that of a trainee. Once filmed and edited, each training session can be seen any number of times by any number of trainees.

After filming, the training session is assembled in postproduction. Editing makes it possible to introduce modifications to the training session: It is possible to add audio commentary, explanatory captions, additional medical instrument monitors, or graphs such as patient scans from pre-operative studies. The final training session can include content filmed outside the OR, as in the use case outlined earlier.

Budget Estimate

The final budget will be divided into three stages: Filming, editing, and visualization.

Filming and Editing

Since only one videographer is required to film a training session in the OR, the logistics of filming cost the same as the logistics of providing training for one trainee. To this must be added a flat fee for each day of filming. The cost of editing is usually covered by a flat fee per minute of edited video or as part of a multiple-video package.

The client mentioned above, who had estimated $335,000 for training 90 company representatives, was able to film almost all surgeries back-to-back, for a final cost of under $7000. Including postproduction, final spend was under $26,000 for all completed immersive videos. The client was then free to choose when and how to provide the immersive training to its representatives.

Type of training	Number of surgeries required	Cost in USD
In-person in the OR	300-900	335,000
Immersive	10	26,000

Table 1. Cost comparison, 90 company representatives, 10 surgeries each.

The above table does not include the cost of visualization by trainees. The cost can be negligible if the viewing equipment is already available, or can increase if equipment needs to be bought or rented. Some of the available visualization options are discussed below.

Visualization

There is no fixed cost per trainee for accessing an immersive training session, since each session can be viewed repeatedly, in different contexts and viewing platforms.

If trainees will access the training session outside company or institutional premises, it may be necessary to send them hardware (ideally, a VR headset). A training session may be visualized without a headset—in a computer, tablet or phone—but the provided level of immersion will decrease.

An institution can designate a single VR headset for use by multiple trainees, whether simultaneously (as in a headset available on-premises) or sequentially (a headset successively loaned to different trainees as required).

The estimated average cost of buying and delivering a headset (Meta Quest 3) within the US is $350. In the case outlined above, if the company chose to purchase headsets for half of its representatives, the total cost of purchases and deliveries would be $15,750. The total cost of providing training, including filming, editing and visualization, would be $41,750.

Even after viewing the training sessions and returning the headset, a trainee can still access training sessions for training reinforcement on a desktop computer, tablet or smartphone, if the sessions are made available online.

Training sessions may also be exhibited at institutional or company events, in full or abbreviated form, at stands provided with headsets or via other equipment such as phones or tablets. Sessions may be integrated, totally or in part, into traditional 2D videos, e-learning platforms or Metaverse conferences.

ENHANCING IMMERSIVE TRAINING VIDEOS

The digital medium offers almost endless possibilities for recording, editing and interacting with training videos.

An immersive training video can be recorded in the OR and streamed live in a multi-user environment, such as the Metaverse. In this scenario, as many trainees as desired can view the intervention live, in immersive video, from their remote locations. Live streaming offers trainees the possibility of direct interaction with the medical specialist in the OR, e.g. by asking questions.

Once recorded, immersive training videos can incorporate additional audiovisual and even interactive elements. In the use case outlined earlier, MRI images and the interventionist's voice-over explanation were used as an introduction to the intervention in the OR. Other examples of additional audiovisual material include captions and graphs, and CGI renderings of human anatomy or medical instruments.

Traditional editing techniques offer a wide range of narrative possibilities to enhance the educational potential of a training session. New user interface technologies provide the additional capability of creating videos that are both immersive and interactive, with the possibility of including tests or exams directly within the training session.

For example, immersive training videos in the OR frequently include captions to indicate or highlight a particular event or action. Captions can be static or interactive. In the latter case, the training session pauses while the trainee answers multiple-choice questions.

Immersive training videos themselves can be integrated into new contexts. For instance, they can be streamed live during events in the Metaverse, or integrated into e-learning platforms and learning management systems (LMS). They can be converted to traditional 2D video for use in non-360-capable platforms, for institutional or product presentations, or as part of additional training modules.

CONCLUSION

The available evidence indicates that immersive training is a more effective learning tool than traditional 2D-based remote training. How exactly it compares to real-world learning is not yet fully understood. While some studies suggest the superiority of real-world learning, this apparent deficiency may be offset by the possibility of repeated viewings of the immersive experience. Based on our own experience, we add that the possibilities for enhancement offered by editing and interactive technologies may also contribute to bringing immersive learning to the level of real-world learning.

From a practical standpoint, three important advantages of immersive learning are readily apparent. The first is the enormous reduction in cost. The second is the possibility of virtually unlimited reach, for any number of trainees and for any amount of time. The third is the ductility of immersive training sessions to be re-used and/or integrated with other learning systems and contexts, beyond their original scope.

For these reasons, immersive training offers a cost-effective solution to the need of providing training in the OR for health professionals. Immersive videos use the power of the human sense of presence to leverage context as a knowledge acquisition and retention tool. They are a significant contribution to the democratization of medical learning, and a valuable tool to complement healthcare education.

REFERENCES

Barsalou, L. W. (2008). Grounded Cognition. *Annual Review of Psychology 59*, 617–645. https://doi.org/10.1146/annurev.psych.59.103006.093639

Dhar, E., Upadhyay, U., Huang, Y., Uddin, M., Manias, G., Kyriazis, D., Wajid, U., AlShawaf, H., Abdul S.S. (2023). A scoping review to assess the effects of virtual reality in medical education and clinical care. *DIGITAL HEALTH 9*. https://doi.org/10.1177%2F20552076231158022

Essoe, J. KY., Reggente, N., Ohno, A.A. et al. (2022). Enhancing learning and retention with distinctive virtual reality environments and mental context reinstatement. *npj Science & Learning 7*, 31. https://doi.org/10.1038/s41539-022-00147-6

Flannery, K. A., & Walles, R. (2003). How does schema theory apply to real versus virtual memories? *Cyberpsychology & Behavior, 6*(2), 151–159. https://doi.org/10.1089/109493103321640347

Kardong-Edgren S., Farra, S., Alinier, G., Young, H.M. (2019). A Call to Unify Definitions of Virtual Reality. *Clinical Simulation in Nursing 31*, 28-34. https://doi.org/10.1016/j.ecns.2019.02.006

Krokos, E., Plaisant, C. y Varshney, A. (2018). Virtual memory palaces: immersion aids recall. *Virtual Reality 23*, 1–15. https://doi.org/10.1007/s10055-018-0346-3

Lan L., Mao, R. Qui, R. Kay, J., & de Sa, D. (2023). Immersive Virtual Reality for Patient-specific Preoperative Planning: A Systematic Review. *Surgical Innovation 30*(1), 109-122. https://doi.org/10.1177%2F15533506221143235

Leutgeb, S., Leutgeb, J., Moser, M-B., Moser, E. (2005). Place cells, spatial maps and the population code for memory. *Current Opinion in Neurobiology, 15*(6), 738-746. https://doi.org/10.1016/j.conb.2005.10.002

Mao R., Lan, L., Kay, J., Lohre, R., Ayeni, O., Goel, P., de Sa, D. (2021). Immersive Virtual Reality for Surgical Training: A Systematic Review. *Journal of Surgical Research 268*, 40-58. https://doi.org/10.1016/j.jss.2021.06.045

McCloy, R. (2001). Virtual reality in surgery. *BMJ 323*, 912. https://doi.org/10.1136/bmj.323.7318.912

Ragan, E. D., Sowndararajan, A., Kopper, R., & Bowman, D. A. (2010). The effects of higher levels of immersion on procedure memorization performance and implications for educational

virtual environments. *Presence: Teleoperators and Virtual Environments 19*(6), 527–543. https://doi.org/10.1162/pres_a_00016

Shapiro, L. (2007). The Embodied Cognition Research Programme. *Philosophy Compass 2/2*, 338-349, https://doi.org/10.1111/j.1747-9991.2007.00064.x

Slater, M. (2003). A Note on Presence Terminology. *Presence Connect 3*(3), 1-5. Retrieved from https://www.academia.edu/download/46463849/presenceterminology.pdf

Smith, S. A. (2019) Virtual reality in episodic memory research: a review. *Psychonomic Bulletin & Review 26*, 1213–1237. https://doi.org/10.3758/s13423-019-01605-w

Smith, S. M. (1979). Remembering in and out of context. *Journal of Experimental Psychology: Human Learning and Memory, 5*(5), 460–471. https://doi.org/10.1037/0278-7393.5.5.460

Smith, S. M., & Vela, E. (2001). Environmental context-dependent memory: a review and meta-analysis. *Psychonomic Bulletin & Review 8*, 203–220. https://doi.org/10.3758/bf03196157

Tulving, E., & Thomson, D. M. (1973). Encoding specificity and retrieval processes in episodic memory. *Psychological Review, 80*(5), 352–373. https://doi.org/10.1037/h0020071

Waller, D., Hunt, E., & Knapp, D. (1998). The transfer of spatial knowledge in virtual environment training. *Presence: Teleoperators and Virtual Environments, 7*(2), 129–143. https://doi.org/10.1162/105474698565631