

ARTICLE TYPE

Interactive Mixed Reality Anatomy Learning of Normal and Malformed Inner Ears

Zhiyue Wu^{1,2} | Zheng Xia³ | Xin Chen³ | Dongliang Guo³ | Yu Wang⁴ | Shaoxing Zhang^{*4} | Liang Zhou^{*1,2}

¹Institute of Medical Technology, Peking University Health Science Center, Beijing, China

²National Institute of Health Data Science, Peking University, Beijing, China

³School of Information Science and Engineering, Department of Software Engineering, Yanshan University, Qinhuangdao, China

⁴Department of Otolaryngology, Peking University Third Hospital, Beijing, China

Correspondence

*: Shaoxing Zhang and Liang Zhou contributed equally to this article.

Corresponding author Liang Zhou,
Email: zhouling@pku.edu.cn

Present address

38 Xueyuan Road, Beijing, 100191, China.

Abstract

The objective of this study is to evaluate the effectiveness of interactive Mixed Reality (MR) learning and compare it to traditional methods for independently learning inner ear anatomy through two experiments. The evaluation was conducted using HololInnerEar, our new interactive MR tool for inner ear anatomy learning. In the first controlled experiment (N=32), we compared independent learning techniques for normal anatomy using the interactive 3D+2D features of HololInnerEar to the traditional method using a video course combined with a mobile medical imaging application. Results show that interactive MR learning is significantly better than the traditional method for the interest of learning and knowledge confidence, and yields better knowledge test scores on multiple tasks although not reaching statistical significance. In a second experiment, the malformation learning module of multiple-linked views of HololInnerEar was assessed with otolaryngologists (N=7). Experts agreed that this module fills a gap in inner ear malformation learning which is important yet overlooked in the current curriculum. Positive feedback is acquired through a customized questionnaire, and NASA Task Load Index (NASA TLX). Overall, we believe that interactive MR learning is a promising complementary approach to learning the anatomy of normal and malformed inner ears alike.

KEYWORDS

Mixed reality, visualization, inner ear anatomy, congenital inner ear malformations, medical imaging, learning technology

1 | INTRODUCTION

The standard method of learning inner ear anatomy and malformation involves lectures and practical sessions¹. Lectures involve explanations from experts to understand the basic concepts of inner ear structure and 2D medical images². In practical sessions, students explore the 3D structure of the inner ear using anatomical specimens³. However, the small size and intricate structure of the inner ear make it challenging for the traditional approach. For centuries, dissecting the human body has been the main method for teaching and learning gross anatomy⁴. The anatomical complexity of the inner ear stems from both its miniature scale (the cochlear is of 5-8 mm in height) and intricate functional structures, including the auditory transduction organ of Corti and vestibular sensory crista ampullaris. These small and delicate structures, combined with the unique embedded position in the bone, impose challenges for gross anatomical dissection and education⁵. On top of that, the use of

saws and drills to remove the temporal bone and internal bony structures hinders the study of intact anatomy. This often restricts a thorough, accurate, and comprehensive understanding of the anatomical features of the inner ear.

Inner ear malformations are an important cause of congenital hearing loss⁶. Characterized by complex pathological structures and diverse clinical manifestations, inner ear malformations present significant challenges for anatomical education and clinical diagnosis⁷. The classification of inner ear malformation primarily relies on the identification of abnormalities in the bony labyrinth structure, with malformations such as Incomplete Partition Type II (Mondini deformity) and Enlarged Vestibular Aqueduct being among the more commonly encountered types in clinical practice^{6,8}. At present, literature review is the primary method for learning about inner ear malformations. However, this approach is limited in providing comprehensive spatial understanding as the literature typically presents only static, two-dimensional images and lacks case variety.

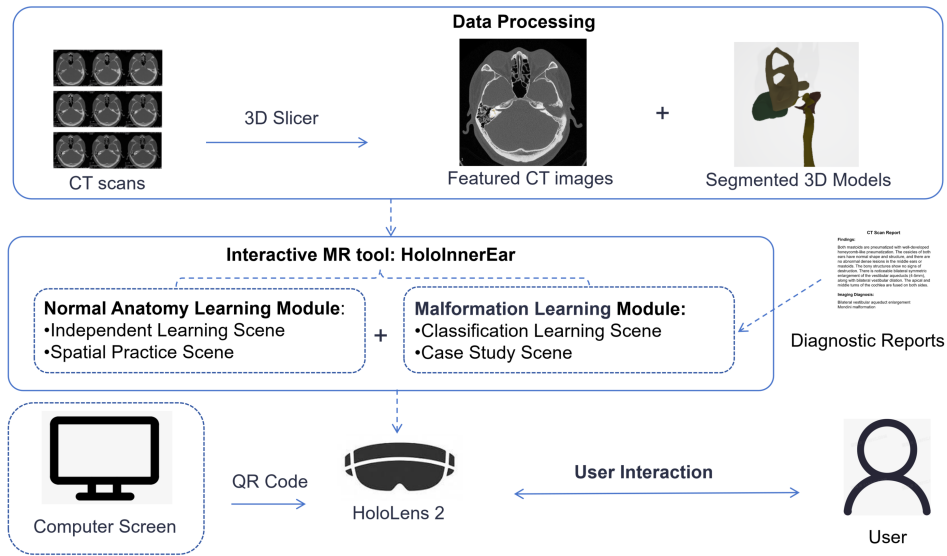


FIGURE 1 The architecture of the interactive MR tool for inner ear anatomy learning-HoloInnerEar.

Computer-assisted techniques, especially visualization in virtual reality (VR), augmented reality (AR), and MR have emerged as crucial complementary tools in anatomical education^{9–11}. Visualization and MR technologies bring new possibilities to anatomy education of inner ears. Nicholson et al.¹² use a web-based 3D inner ear anatomy learning system for teaching. This study demonstrates that computer-based 3D anatomical models can enhance the learning of inner ear anatomy. Zariwny et al.¹³ design a system that integrates optical glyphs with AR for interactive inner ear anatomy learning. This tool combines visual and haptic feedback to improve the understanding of complex anatomical structures, however, the user's 3D perception depends on the physical model. Gnanasegaram et al.¹⁴ compare the effectiveness of lectures, computer programs, and 3D holographic models in MR for inner ear anatomy learning. The results show that compared to traditional lectures and computer software, mixed reality has better teaching effects and is more popular among students. However, this study only allows passive observation without interactions.

In clinical practice, medical imaging such as computational tomography (CT) and magnetic resonance imaging (MRI) are typically used for inner ear diagnosis. Much proficiency is required for identifying key imaging features¹⁵ and relating them to 3D structures¹⁶ in the process. However, the substantial spatial and dimensional disparities between 3D anatomical structures and 2D images pose considerable challenges for reconstructing 3D models from medical image scans and correlating these with their 2D representations^{17–20}. In practice, medical students and doctors typically learn the associations using video lectures and practice on medical imaging software. Studies on MR anatomy learning of 3D structures combined with medical images are available but few²¹. Brun et al.²² find evidence of superior learning performance for 3D visualization in MR compared to 2D CT slices and 3D printing for cardiac morphology. Ho et al.²³ combine MRI with the study of the

physical structure of the brain to enhance understanding of the spatial relationship between 2D images and 3D structures.

Existing studies on interactive MR learning of inner ear anatomy have, to our knowledge, not explored the integration of 3D models with real-world medical scans. Furthermore, none of the studies addresses the learning of inner ear malformations.

In this study, we conducted experiments to compare interactive MR visualization with standard independent learning methods for inner ear anatomy. We used a new MR tool, HoloInnerEar, designed to address the aforementioned limitations of MR learning of normal and malformed inner ear anatomy. We compared the independent learning experiences of normal inner ear structures using interactive MR (integrating 3D models with image slices) against the typical approach of video lectures and medical imaging applications. The performances were assessed using knowledge tests and questionnaire surveys. The learning of inner ear malformations in HoloInnerEar was evaluated through a usability study with experienced otolaryngologists. Results suggest that interactive MR-based learning has comparable (better yet not statistically significant) learning effectiveness compared to the standard independent learning technique, while significantly enhancing both interest and confidence in learning normal inner ear anatomy. Expert feedback indicates that HoloInnerEar provides a valuable reference for congenital inner ear malformation classification with real cases and bridging the gap in the teaching of malformations.

2 | MATERIALS AND METHODS

This study is based on our newly devised MR interactive inner ear learning tool, HoloInnerEar. The study consists of two experiments: a controlled experiment of 32 participants (Section 2.3) and an expert evaluation by 7 otolaryngologists (Section 2.4). The experiments have

been approved by the Institutional Review Boards of Peking University (approval number IRB00001052-23206) and Qinhuangdao First Hospital (approval number 2024Y001).

2.1 | Datasets

As shown in Fig. 1, CT scans and associated diagnostic reports are used in our study. The CT data used in our study were scanned, anonymized, and provided by the Department of Otolaryngology, Head and Neck Surgery at the Peking University Third Hospital, Beijing, China. It is a leading institute in the nation for treating otologic diseases with a particular expertise in cochlear implantation and temporal bone surgery. The data were high-resolution thin-slice CT scans of the temporal bone, with a slice thickness of 0.625 mm. A total of 21 temporal bone CT scans were obtained, each accompanied by a radiological diagnostic reports. Among these scans, one represented a normal inner ear, while the remaining 20 were associated with various inner ear malformations. Specifically, the abnormal group included 16 cases of Incomplete Partition Type II (Mondini deformity), 2 cases of Enlarged Vestibular Aqueduct, 1 case of Common Cavity Malformation, and 1 case of Rudimentary Otolocyst. The data were segmented into vestibular system organs and cochlea using 3D Slicer (<http://www.slicer.org>)²⁴. All inner ear structures were manually segmented slice-by-slice under the guidance of experienced otolaryngologists, ensuring anatomical accuracy in the 3D reconstructions prior to mesh conversion for MR analysis.

2.2 | HololInnerEar

We devised an MR inner ear anatomy learning tool-HololInnerEar, implemented as a Unity application for HoloLens2, to address identified shortcomings of existing learning tools for inner ear anatomy. As illustrated in Figure 1, we used 3D Slicer to process CT scans and generated featured 2D images and segmented 3D models, which were subsequently integrated with diagnostic reports to form the dataset. The Mixed Reality Toolkit (MRTK, <https://github.com/microsoft/MixedRealityToolkit-Unity>) was used to aid the implementation of hand interaction and data visualization. Using QR-based spatial registration, our tool provides immersive virtual visualization while maintaining the real-world context for diagnosis. HololInnerEar consists of two complementing study modules for 1. normal inner ear anatomy, and 2. inner ear malformations.

2.2.1 | Normal Anatomy Learning Module

The normal anatomy learning module comprises two scenes. The first is the independent learning scene as shown in Fig 2, which features a static 3D inner ear model that users can slice interactively by adjusting a cutting plane. The CT image corresponding to the current cutting plane of the inner ear model is shown on the left. When the cutting plane intersects the inner ear model, the corresponding area is highlighted in color

in the CT image. Annotation lines appear when the cutting plane intersects the model, connecting the inner ear model to the corresponding CT image position. In this scene, users can save the current CT image, inner ear model, and cutting plane information when necessary. All saved information is organized as snapshots on a back learning wall to assist learning. The second scene focuses on practicing spatial positioning of

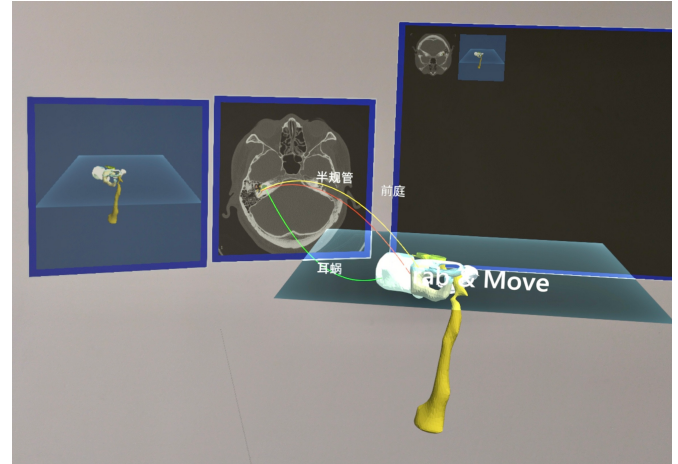


FIGURE 2 The independent learning scene shows the current position of the cutting plane and the corresponding CT image (left), the 3D inner ear model and the cutting plane (front), and the learning wall (back-right).

the inner ear as shown in Fig 3. In this scene, the 3D inner ear model is movable and rotatable, while maintaining a fixed size. It is accompanied by CT cross-sectional, coronal, and sagittal images. Users are required to manipulate the 3D model and place it in the correct position that corresponds to the CT images. During this process, users can freely drag the CT images along the X, Y, and Z axes to observe them from different angles. They can also hide specific parts of the inner ear model to better assess the correct spatial position of the inner ear structures. Once users believe they have positioned the model correctly, they can submit to confirm their choice, and the tool reveals the correct position of the model in the space.

2.2.2 | Malformation Learning Module

The malformation learning module contains a classification learning scene and a case study scene using situated visualization^{25,26}. Situated visualization brings data visualizations into their context of use, connecting data with the physical environment. As shown in Fig. 4, the classification scene includes a malformation classification view and a case visualization view. The classification of inner ear malformations is presented through an interactive icicle plot that effectively visualizes the hierarchical structure of these categories. The case visualization

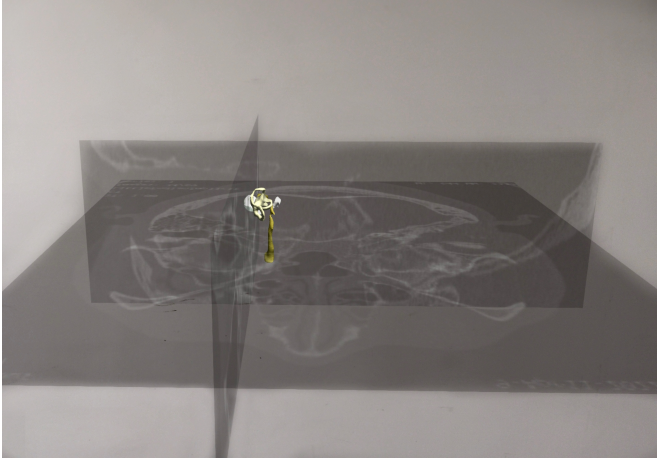


FIGURE 3 The spatial practice scene features the 3D inner ear model with 2D CT images for axial, coronal, and sagittal views.

component consists of five parts: malformation description panel, malformation case studies, the interactive 3D model slicing area, CT images, and diagnostic reports. Users can select a case of interest from the malformation case shelf on the right, and the corresponding 3D model is displayed in the slicing area at the center. By adjusting the position of the cutting plane, the CT image display area shows the corresponding cross-sectional CT image, while the text report area displays the diagnostic report for reference. Key elements in the diagnostic report are highlighted and visually connected to the corresponding regions in the 3D model with curves, enhancing user comprehension.

The malformation classification plot and the case visualization views are connected through brushing and linking – a coordinated multi-view technique that selections (brushing) of a user in one view dynamically highlight related data (linking) across all views. For example, when a malformation is selected in the plot, the case visualization is automatically updated including the corresponding description, 3D model, CT images, and diagnostic reports. A typical case of the “IP-II (Mondini deformity)” is shown in Fig. 4. Since some malformations are relatively rare conditions, CT data for these cases are not available, and CT slice images from a classification study⁸ are used instead.

The situated-visualization case study scene facilitates the understanding of malformations in a typical clinical setting where a doctor is working on a case of a patient using a computer. When HololnnerEar detects a QR code with a diagnostic report displayed on the computer screen, the corresponding interactive 3D model of the inner ear malformation and its diagnostic report are integrated into the mixed reality space. For CT images on the screen along with a QR code, the associated 3D model and diagnostic report are triggered and visualized in the mixed reality environment. The functionality was realized with spatial tracking aided by Vuforia Engine (<https://developer.vuforia.com/>).

2.3 | Controlled Experiment on Normal Anatomy Learning

To test the effectiveness of interactive MR learning for normal anatomy features using HololnnerEar, we designed a controlled experiment.

2.3.1 | Study Design

A between-subjects design was used for the experiment. The experiment compared two groups: the mixed reality group using HololnnerEar (MR group), and the control group using video lectures and imaging software on mobile phones (IS group). The videos were selected from recorded courses publicly available online (Systemic Anatomy, Sichuan University, https://www.bilibili.com/video/BV1xs411v7nV/?p=44&share_source=copy_web&vd_source=97afdd705c79d21ade46315ec14c032c) as shown in Fig. 6(a). The imaging software used was the Anatomical Imaging Atlas (<https://apps.apple.com/us/app/imaging-anatomy-atlas/id6449857896>), a professional tool for learning human anatomical imaging—developed in cooperation with major medical schools and hospitals—and features numerous annotated imaging atlases (Fig. 6(b)).

We formulated our hypotheses as follows:

- **H1:** The MR method performs better than the IS method in identifying relevant anatomical features in CT images.
- **H2:** The MR method performs better than the IS method in establishing the spatial correspondence between the 3D structure and the CT images of inner ears.
- **H3:** The MR method is more effective in enhancing knowledge confidence and learning interest compared to the IS method.

2.3.2 | Tasks

To test these hypotheses, we formulated three tasks focusing on the comprehension of 3D models, 2D CT images, and their corresponding relationships. The experiment contains two parts: a knowledge test and a questionnaire survey, as summarized in Table 1.

TABLE 1 Tasks in the User Study.

Hypothesis	Task
H1	T1. Identify the corresponding inner ear structures in 2D CT images.
H2	T2. Establish spatial correspondence between 3D inner ear structures and 2D CT images.
H3	T3. Complete the questionnaire on knowledge confidence and learning interest.

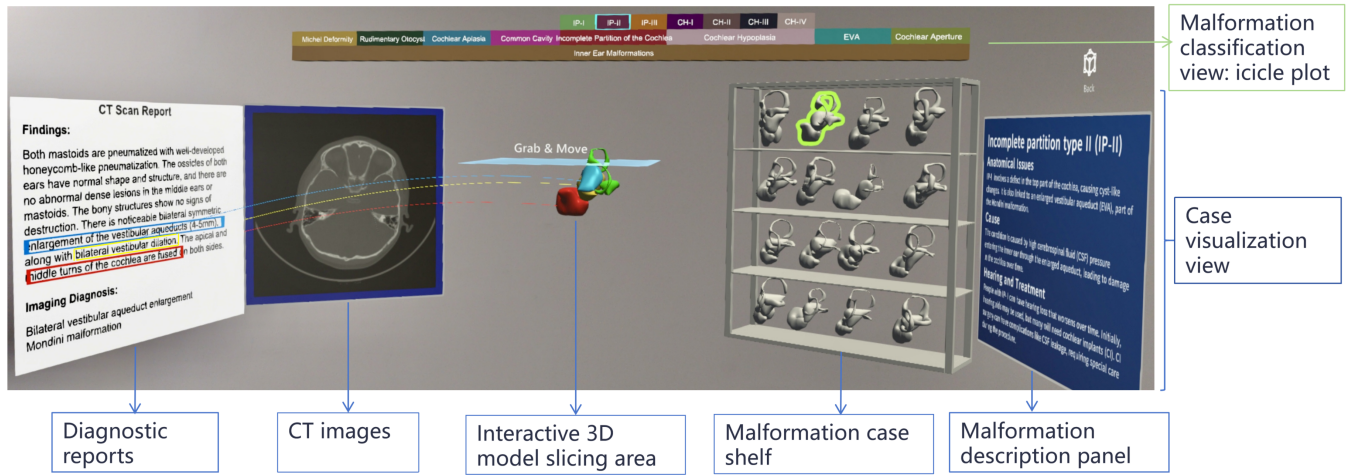


FIGURE 4 The classification scene includes a malformation classification view as an interactive icicle plot, and a case visualization view, including a malformation description panel, a malformation case shelf, the interactive 3D model slicing area, CT images, and diagnostic reports.

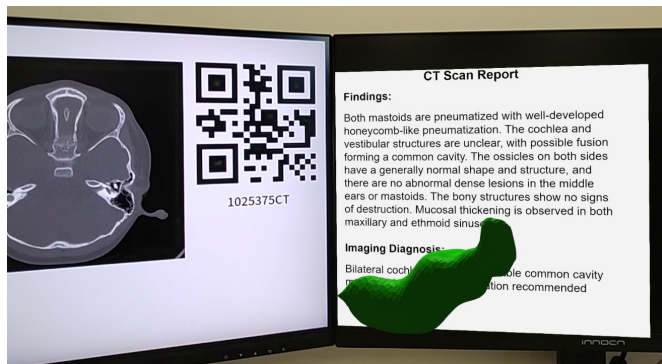


FIGURE 5 The situated visualization case study scene links a virtual 3D model and its associated diagnostic report or CT images to the physical computer screen through QR code tracking.

The knowledge test includes tasks **T1** (paper-based test) and **T2** (MR test), while the questionnaire survey is designated as **T3**. **T1** consists of 5 single-choice questions to evaluate users' comprehension of the inner ear in CT images. **T2** is an MR test conducted in HoloLens 2, comprising 10 questions divided into two parts. **T3** is a 6-question survey employing a 5-point Likert scale concerning students' knowledge confidence and learning interest. Note that the paper-based knowledge test and the questionnaire can be found in the supplemental material.

Participants have to answer questions through interaction in the MR test (**T2**). The first part of the questions features a fixed-position inner ear model and a horizontal cross-section as shown in Fig. 7(a). The participant has to select the correct answer from the four CT images on the left based on the relationship between the 3D model and the horizontal cross-section. The tool reveals the correct answer and proceeds to the next question after the participant commits her/his answer. This part consists of 5 questions, and the correctness of each question are recorded. The second part requires the participant to move the image

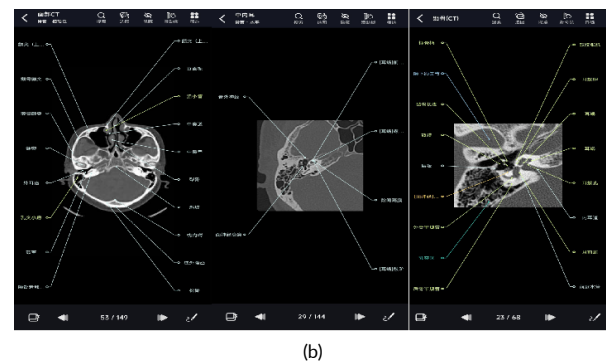
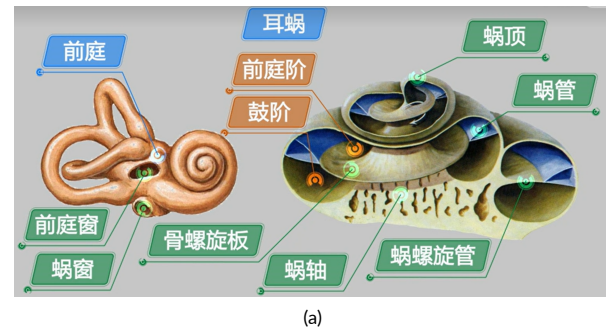


FIGURE 6 The video lecture (a) and the imaging software (b) used in the control group. Note that both sources have annotations (in Chinese) for inner ear structures.

slice along the vertical axis of the 3D model to match the CT image on the left as shown in Fig. 7(b). This part also consists of 5 questions, with the difference between the placed and correct positions recorded.

The survey (**T3**) used 5-point Likert scales (1=strongly disagree, 5=strongly agree) to assess participants' experiences on the following aspects:

- I learned about the anatomical structure of the inner ear;

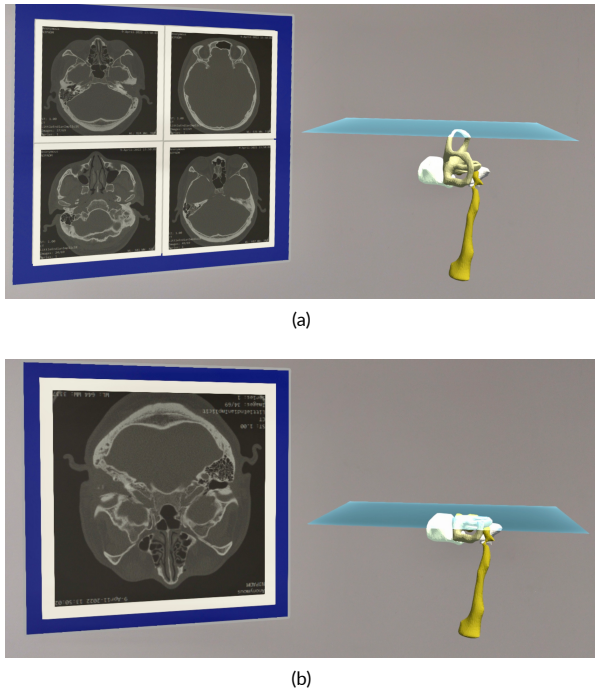


FIGURE 7 The knowledge test in MR contains two parts: (a) choosing the correct CT slice on the left for the cross-section of the 3D model, and (b) matching the position of the cross-section on the 3D model to the CT slice.

- I learned how to interpret inner ear structures in CT images;
- I learned about the correspondence between anatomical structures and CT images;
- I have always been proactive throughout the learning process;
- I find this method very engaging;
- I hope to continue learning in the same way in the future.

2.3.3 | Participants

The study included 32 participants (12 females), spanning from undergraduates to doctoral candidates. Among them, 9 participants had prior experience using MR-related devices. They were required to have no formal knowledge of inner ear anatomy; to be free from severe motion sickness; and to have no chronic or acute major health conditions that could interfere with their participation in the study. We employed a two-tailed power analysis to determine the minimum sample size. Previous research has demonstrated that the effect size relationship between VR/AR learning and knowledge test performance ranges from 0.38 to 1.99, with an average effect size of 0.64 and an average standard deviation of 0.6 in these studies, which were used for the power analysis ^{23,27}.

2.3.4 | Procedure

The procedure of the experiment is outlined in Fig. 8. Participants were paired for the experiment, and each participant was randomly assigned to either the MR or IS group in a pair. The 32 participants were divided into 16 pairs, and the experimental duration of each pair was approximately 95 minutes.

First, a 15-minute explanation of the experiment was provided. The experimenters introduced the purpose of the study, learning objectives, and procedures to the two participants. They were encouraged to ask questions during this session. After ensuring that all participants fully understood the information provided, they were given informed consent forms along with the experimental instructions. Participants were informed that they could withdraw from the experiment at any time; however, none chose to do so.

Once written consent was granted, a 20-minute pre-experiment preparation was conducted. Participants learned how to use the HoloLens, focusing on clicking, dragging, remote clicking, and remote dragging. The experimenters then demonstrated the use and tips for the medical imaging application.

Next, a 30-minute learning session was performed. Participants were randomly assigned to the experimental group or the control group and engaged in independent learning based on the learning objectives. Experimenters could answer any question asked by participants at any time during the learning process.

After the learning session, participants took a 30-minute knowledge test. They first took the paper test (T1) and then proceeded to the HoloLens-based MR test (T2, regardless of their assigned groups). Experimenters were available to answer any non-academic questions throughout the process.

Finally, a 5-minute questionnaire survey of 6 questions using 5-point Likert scales was conducted (T3). Experimenters also talked to participants to understand their feedback on the tools they used and the overall experiment.

2.4 | Expert Assessment on Malformation Learning

The classification and diagnosis of inner ear malformations is a specialized area within otorhinolaryngology, which is typically not included in the core curriculum for medical students. Therefore, we invited 7 otolaryngologists to evaluate the malformation learning module of HoloInnerEar. The expert panel comprised 4 females and 3 males, with years of clinical experience and substantial expertise in otorhinolaryngology as summarized in Table 2. In addition, they had a strong understanding of medical education technologies.

The entire expert evaluation process lasted approximately 25 minutes. After obtaining informed consent and receiving a system description (5 minutes), experts assessed the observability and interactivity of

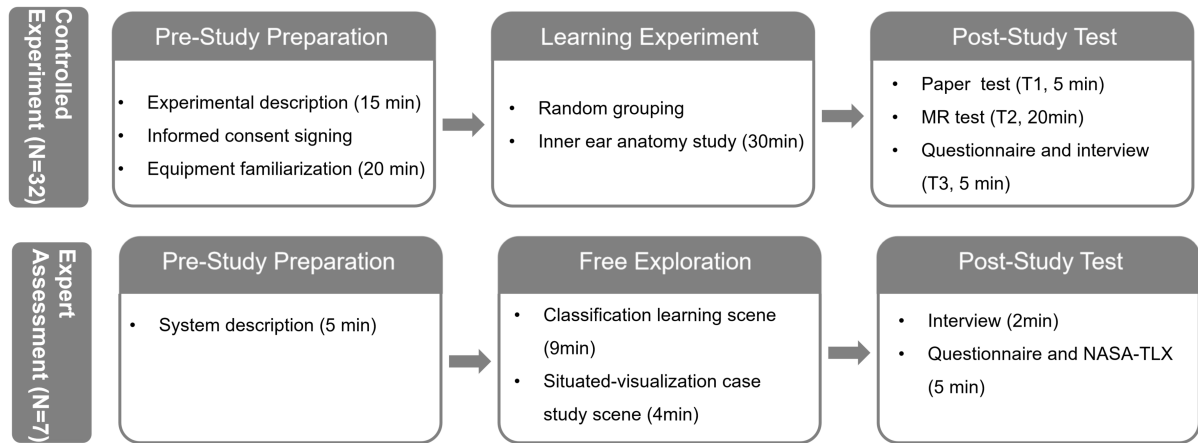


FIGURE 8 The flowchart of the controlled experiment and the expert assessment.

TABLE 2 Basic Information of the Experts

Gender	Age	Clinical experience (yrs)	MR experience
Female	37	12	Yes
Male	26	4	Yes
Female	30	5	Yes
Female	24	2	Yes
Female	24	2	No
Male	37	10	No
Male	40	16	No

the 3D models, 2D CT images, and diagnostic reports in the malformation learning module. They evaluated the effectiveness of the tool in enhancing the understanding of inner ear malformations.

The interactive analysis session lasted about 13 minutes, 9 minutes for the classification learning scene and 4 minutes for the situated-visualization case study scene. During the session, the experts freely interacted with and provided feedback on the two scenes for malformation learning under the guidance of the research team. An interview followed to collect further feedback and suggestions for improvement. At the end of the assessment, experts completed the NASA-TLX²⁸ questionnaires to evaluate the system's usability and perceived task load. Throughout the process, a think-aloud protocol was employed, and feedback from both the interactive evaluation session and the interview was documented.

3 | RESULTS

We report on the results of the two experiments in this section. More details can be found in the supplemental material.



FIGURE 9 An expert is evaluating the situated-visualization of the malformation learning module. When HoloLens scans the QR code on the computer screen, she can simultaneously see the 2D medical image on the screen (physical object), a 3D model, and a diagnostic report (virtual components).

3.1 | Controlled Experiment

We analyze performances of participants of the controlled experiment to test hypotheses. IBM SPSS 26 and R was used to analyze and process data, and data was evaluated with a significance level of 0.05.

3.1.1 | Knowledge Test

The reliability of the knowledge test questions was assessed using Cronbach's alpha²⁹, yielded a value of 0.654, indicating a reasonable level of

TABLE 3 Results of the controlled experiment.

Task	Group	Mean	SD	Method	p Value
T1-CT Comprehension	MR Group	0.80	0.20	Mann-Whitney U test	0.273
	IS Group	0.68	0.29		
T2-Relationship between Model and CT	MR Group	0.57	0.13	t-test	0.180
	IS Group	0.50	0.16		
T3-Knowledge Confidence	MR Group	0.92	0.12	Mann-Whitney U test	0.001*
	IS Group	0.70	0.18		
T3-Learning Interest	MR Group	0.93	0.16	Mann-Whitney U test	<0.001*
	IS Group	0.61	0.24		

reliability for these questions. To facilitate data analysis and presentation, all scores were normalized. The results of the knowledge test are summarized in Table 3 and Fig. 10 as violin plots.

For **T1** (2D CT structure identification), the mean score of the MR group (0.80 ± 0.20) exceeded that of the IS group (0.68 ± 0.29). The Mann-Whitney U test showed no significant difference between the groups ($U = 100.00$, $Z = -1.096$, $p = 0.273$), indicating comparable performances between the groups.

For **T2** (correspondence between 3D models and 2D CT), the mean score of the MR group (0.57 ± 0.13) was also higher than that of the IS group (0.50 ± 0.16). With homogeneous variances according to Levene's test, the t-test showed no significant difference between the groups ($t(30) = 1.372$, $p = 0.180$), indicating comparable performance.

3.1.2 | Questionnaire

The questionnaire in **T3** was evaluated for reliability with a Cronbach's alpha of 0.903 indicating a high level of internal consistency among the questionnaire items. The six questions are divided into two dimensions: knowledge confidence (questions 1–3) and learning interest (questions 4–6). The results for these two dimensions are shown in Table 3 and Fig. 10.

For knowledge confidence, the Mann-Whitney U test revealed a significant difference between the MR group (0.92 ± 0.12) and the IS group (0.70 ± 0.18) ($U = 42.50$, $Z = -3.301$, $p = 0.001 < 0.05$), indicating that the interactive MR learning tool can effectively enhance learning confidence compared to the traditional method. For learning interest, the variance was homogeneous, and the Mann-Whitney U test indicated a statistically significant difference between the MR group (0.93 ± 0.16) and the IS group (0.61 ± 0.24) ($U = 35.00$, $Z = -3.665$, $p < 0.001$), suggesting that the MR method can significantly enhance the interest. Therefore, Hypothesis **H3** can be accepted.

3.2 | Feedback of Expert Users

We collect expert feedback on the malformation learning module of our tool to assess its usability and effectiveness. The mean responses from the customized questionnaire as summarized in Table 4 reflect

the general acceptance of the experts on the capabilities and user-friendliness of the tool. Among others, the classification functionality and the multiple linked view received highest mean responses.

TABLE 4 Questionnaire Results of Expert Users

Question	Mean	SD
1. Effectiveness of MR visualization in understanding inner ear structures.	4.14	0.69
2. Intuitiveness and ease of use of the interactive design.	3.71	0.76
3. Effectiveness in understanding inner ear malformation classification.	4.57	0.53
4. Effectiveness of situated visualization in understanding inner ear malformation cases.	4.14	0.69
5. The ability to relate CT data, diagnostic reports, and 3D models of malformed inner ears.	4.86	0.38

The NASA-TLX index was analyzed to assess the perceived workload during the free exploration of the tool. The average NASA-TLX score was 37.85 ± 20.48 , indicating a moderate level of perceived workload. The variability in scores suggests that some experts experienced higher cognitive and physical demands, particularly during complex tasks like manipulating 3D models while simultaneously interpreting CT scans and diagnostic reports.

Some common preferences of participants are as follows after summarizing useful comments from the exploration sessions and interviews. First, they appreciated the interactive features that allowed them to manipulate 3D models of the inner ear, which helped in understanding the relationship between 3D structures and 2D CT images. Second, the malformation classification module was particularly helpful in highlighting deformity types, which facilitated a deeper understanding of clinical diagnoses. Third, the situated visualization that augments the computer screen was particularly engaging and enabled the acquisition of information from multiple dimensions.

In terms of criticisms, one expert, who had never used MR devices before, initially found the virtual interaction interface challenging, as the virtual components were not well-aligned with the physical environment, making them difficult to interact with. However, after a brief

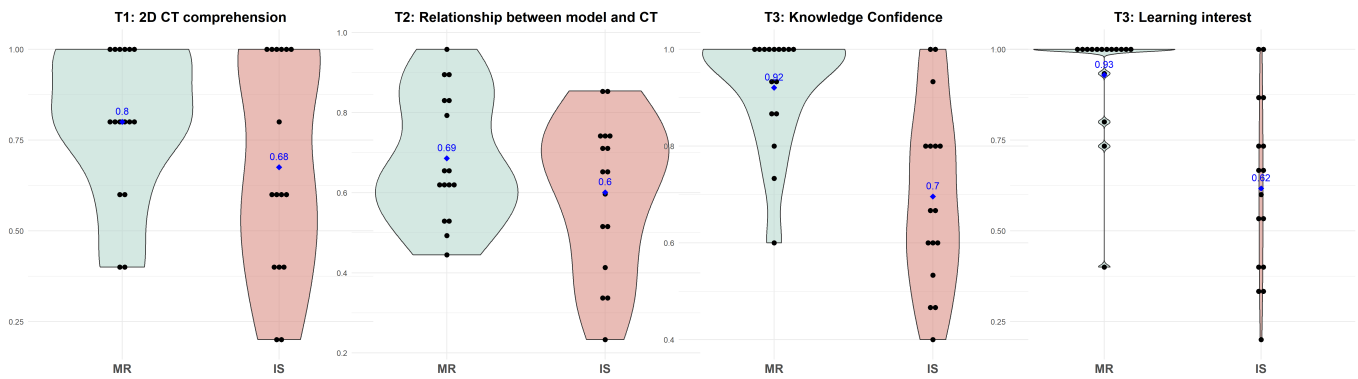


FIGURE 10 Violin Plots of Knowledge Test and Questionnaire Results. The numbers in blue indicate the mean values.

explanation and demonstration, the confusion was resolved. Five experts reported that it was challenging to simultaneously view the CT scan and the diagnostic report clearly while slicing a 3D inner ear model due to the limited field of vision. This issue was resolved after the researchers demonstrated the remote operation function for the cutting slides to the experts.

Overall, all experts agreed that the system improved the understanding of inner ear anatomy and malformations, particularly through its interactive and immersive approach. They also noted that the system's integration of 3D models and 2D CT images was effective in teaching complex anatomical concepts, with some recommendations for further refinements in user interface clarity and tutorial support.

4 | DISCUSSION

In the controlled experiment and expert assessment, both inexperienced students and experienced otolaryngologists agreed that:

1. The combination of 3D models and 2D CT images significantly improved their understanding of the spatial relationship between the two modalities compared to traditional methods.
2. The use of real clinical medical data offered valuable insights into the clinical characteristics of rare malformations, enhancing the educational experience.
3. The combination of virtual and real elements in MR created an immersive and interactive environment, fostering active exploration and improving engagement with the learning material.

Overall, the study results demonstrate better performances and preferences of the interactive MR method compared to the traditional method for independent learning of inner ear anatomy. All participants agreed that MR technology provides a more engaging and interactive learning experience compared to traditional literature-based methods. The integration of real clinical cases further enriches the learning process by exposing learners to diverse scenarios and addressing the shortage of cases in conventional clinical training. The experts suggested that

the malformation learning module not only enhances the understanding of inner ear malformations but also provides detailed, context-driven diagnostic guidance which is very helpful in the clinical setting.

4.1 | Linked 3D+2D Inner Ear Anatomy Learning

Participants in the controlled experiment noted that the 3D +2D MR inner ear anatomy learning tool was particularly beneficial in understanding the spatial relationship between the inner ear and surrounding anatomical structures. Similarly, in the expert assessment, the experts praised the interactive features of the module, especially the integration of 3D models and CT imaging, for facilitating a deeper understanding of malformation types and their clinical implications. Expert evaluations confirmed that the method meets educational standards, highlighting its potential as a valuable tool for medical training related to both normal and malformed inner ear structures.

4.2 | Benefits and Limitations of Interactive MR

One benefit of MR is that it addresses the limitations of visualizations on computer screens by offering 3D representations in stereo vision. This enables learners to walk around and interact with 3D models as if they were real objects, allowing for a better understanding of spatial relationships and complex structures. Another benefit of the interactive MR tool is that the integration of virtual and real objects allows users to simultaneously observe their real-world environment, enabling uninterrupted communication with colleagues, fostering collaboration, and minimizing disruptions to daily workflows. A third benefit is that the MR headset is rather lightweight and can work independently of a computer, which enhances the adaptability of the MR tool, allowing its potential use in clinical settings.

However, the limitations of MR negatively affect the results of the study. For those who are new to MR, the 3D visualizations and unfamiliar user interactions—compared to standard mouse and keyboard interactions—pose challenges that can hinder performance. We note that in the controlled study, despite the MR group being exposed to a 3D scene, several participants still perceived it as a 2D representation as if it was on a screen. One expert in the free exploration experienced difficulty clicking on the items in the classification plot as he was not aware of the depth of the items.

4.3 | Learning Curve of the MR Device and Future Plans

In this study, HoloLens 2 was chosen as the MR device for HoloInnerEar for its robust hand-tracking capabilities, spatial mapping accuracy, and established use in medical applications. While the device presents a learning curve, most participants adapted quickly within the 20-minute familiarization session with the guidance of the researcher. Notably, even first-time MR users achieved proficiency with our tool, indicating that our workflow-specific design and tailored introduction allow users to learn the tool quickly. Furthermore, our MR interface provides visual cues, such as "Click" labels below interactive buttons, to support intuitive interactions.

The discontinuation of HoloLens development necessitates a hardware-agnostic approach for long-term viability. Future iterations of HoloInnerEar would take advantage of the cross-platform OpenXR standards to provide cross-device functionalities. Based on Unity, our tool can be ported to emerging MR devices, e.g., Meta Quest Pro, Apple Vision Pro, or Varjo XR-4, with ease, as these platforms support OpenXR and similar interaction paradigms to MRTK used by HoloLens2.

4.4 | Potential Explanations of the Controlled Experiment Results

In the controlled experiment, the mean scores of T1 and T2 in the MR group were higher than those in the IS group, although the differences were not statistically significant. Except for the rather small sample size, several factors may explain this outcome.

One factor contributing to the learning outcomes is that video lectures, as a passive learning method, allow participants to quickly absorb and memorize important contents. This approach facilitates the mastery of complex concepts in a short period. Similarly, the medical imaging software provides a predefined learning protocol, where participants adhere to a fixed sequence of steps to acquire information. In contrast, the MR group engaged in active exploration during the learning process, which resulted in spending additional time spent on content not directly aligned with the learning objectives. This increased exploration time resulted in a less efficient allocation of time towards core learning activities, which may have impacted their knowledge test results. To improve these results, providing more structured and targeted guidance

within the MR environment, as well as optimizing the learning schedule to ensure better time allocation, could further enhance the MR group's performance.

Another factor is familiarity of participants to interactions in the medical imaging software. The software used in the IS group is among the most advanced and widely adopted in China, and users demonstrate greater proficiency when interacting with the software via smartphones than with HoloLens. However, these factors bring the control group closer to a realistic scenario, which enhances the external validity of the study.

4.5 | Suggestions for Improvements

Participants, in particular the experts, made several suggestions regarding the system design. Currently, the tool only offers a 3D model of the inner ear, which does not visually indicate its position within the head. This can be confusing for beginners. It is recommended to incorporate an interactive model of the entire head to improve spatial awareness and understanding. Second, the resolution of the existing model is not sufficient, making it difficult to observe the smaller and more complex internal structures of the inner ear. Third, the interference of the augmented reality background has resulted in the indistinct display of CT images due to a lack of contrast. It is suggested to optimize the background settings to enhance contrast. Moreover, experts see potential applications for HoloInnerEar in collaborative clinical training and distance education, which could broaden its utility in medical education.

5 | CONCLUSIONS

Our study evaluates interactive MR technology for learning inner ear anatomy and malformations using HoloInnerEar—a specialized interactive MR learning tool. In a between-subjects controlled experiment, we compared interactive MR learning to the state-of-the-art independent-learning technique using video lectures and a widely used medical imaging software for anatomy learning. The MR group demonstrated better yet not statistically significant knowledge test scores compared to the IS group, suggesting a modest advantage for the MR approach; significantly better learning interest and knowledge confidence are shown for the MR group. Participants also expressed a preference for the MR method. In the second experiment, an expert assessment with otolaryngologists confirmed the effectiveness of interactive MR in improving the identification and classification of inner ear malformations. The experts gave positive feedback on the interactive features and integration of real and virtual objects. All participants liked the ability of the method to integrate 3D models and 2D CT images, which facilitated a deeper understanding of normal and malformed anatomical structures of inner ears. Overall, interactive MR learning is a promising new method that complements the traditional anatomical learning of inner ears.

ACKNOWLEDGMENTS

This work was supported in part by the National Natural Science Foundation of China (grant 62372012), the Clinical Medicine Plus X - Young Scholars Project of Peking University (PKU2024LCXQ028), and the Fundamental Research Funds for the Central Universities.

AUTHOR CONTRIBUTIONS

Zhiyue Wu: Conceptualization, Methodology, Software, Validation, Investigation, Data curation, Writing - original draft, Writing - review & editing, Visualization. Zheng Xia: Software, Validation, Investigation, Data curation, Writing - original draft. Xin Chen: Software, Investigation. Dongliang Guo: Conceptualization, Methodology. Yu Wang: Resources, Data curation. Shaoxing Zhang: Conceptualization, Methodology, Resources, Data curation, Supervision, Project administration, Funding acquisition. Liang Zhou: Conceptualization, Methodology, Resources, Writing - original draft, Writing - review & editing, Supervision, Project administration, Funding acquisition.

REFERENCES

- Duarte ML, Santos LR, Guimarães Júnior JB, Peccin MS. Learning anatomy by virtual reality and augmented reality. A scope review. *Morphologie*. 2020;104(347):254–266. Available from: <https://www.sciencedirect.com/science/article/pii/S1286011520300813>.
- Mansilla L, Milone DH, Ferrante E. Learning deformable registration of medical images with anatomical constraints. *Neural Networks*. 2020 Apr;124:269–279. Available from: <http://dx.doi.org/10.1016/j.neunet.2020.01.023>.
- Zargarán A, Turki M, Bhaskar J, Spiers H, Zargarán D. The Role of Technology in Anatomy Teaching: Striking the Right Balance. *Advances in Medical Education and Practice*. 2020;11:259 – 266.
- Triepels CPR, Smeets CFA, Notten K, Kruitwagen R, Futterer J, Vergeldt T, et al. Does three-dimensional anatomy improve student understanding? *Clinical Anatomy* (New York, Ny). 2019;33:25 – 33.
- Alenzi S, Dhanasingh A, Alanazi H, Alsanosi A, Hagr A. Diagnostic Value of 3D Segmentation in Understanding the Anatomy of Human Inner Ear Including Malformation Types. *Ear, Nose & Throat Journal*. 2020;100:675S – 683S.
- Sennaroglu L. In: DeSaSouza S, editor. *Classification and Management of Inner Ear Malformations*. Singapore: Springer Nature Singapore; 2022. p. 245–264. Available from: https://doi.org/10.1007/978-981-19-0452-3_11.
- Jackler R, Luxfor WM, House W. Congenital malformations of the inner ear: A classification based on embryogenesis. *The Laryngoscope*. 1987;97.
- Sennaroglu L, Bajin MD. Classification and Current Management of Inner Ear Malformations. *Balkan Medical Journal*. 2017;34(5):397–411. Available from: <https://doi.org/10.4274/balkanmedj.2017.0367>.
- Moro C, Štromberga Z, Raikos A, Stirling A. The effectiveness of virtual and augmented reality in health sciences and medical anatomy. *Anatomical Sciences Education*. 2017;10(6):549–559. Available from: <https://anatomypubs.onlinelibrary.wiley.com/doi/abs/10.1002/ase.1696>.
- Taylor L, Dyer T, Al-Azzawi M, Smith C, Nzeako O, Shah Z. Extended reality anatomy undergraduate teaching: A literature review on an alternative method of learning. *Annals of Anatomy - Anatomischer Anzeiger*. 2022;239:151817. Available from: <https://www.sciencedirect.com/science/article/pii/S0940960221001436>.
- Hu M, Wattchow D, de Fontgalland D. From ancient to avant-garde: a review of traditional and modern multimodal approaches to surgical anatomy education. *ANZ Journal of Surgery*. 2018;88.
- Nicholson DT, Chalk C, Funnell WR, Daniel SJ. Can virtual reality improve anatomy education? A randomised controlled study of a computer-generated three-dimensional anatomical ear model. *Medical Education*. 2006 Nov;40(11):1081–1087.
- Zariwny A, Stewart P, Dryer M. Visuo-haptic learning of the inner ear: using the optical glyphs and augmented reality of the InvisibleEar™. *SIGCAS Comput Soc*. 2014 Jul;44(2):5–7. Available from: <https://doi.org/10.1145/2656870.2656871>.
- Gnanasegaram JJ, Leung R, Beyea J. Evaluating the effectiveness of learning ear anatomy using holographic models. *Journal of Otolaryngology - Head & Neck Surgery*. 2020;49.
- Yu FF, Feltrin FS, Bathla G, Raj K, Agarwal A, Lee WC, et al. Imaging Guide to Inner Ear Malformations: An Illustrative Review. *Current Problems in Diagnostic Radiology*. 2023;52(6):576–585. Available from: <https://www.sciencedirect.com/science/article/pii/S0363018823000981>.
- Assiri M, Khurayzi T, Almuhawwas F, Schlemmer K, Hagr A, Dhanasingh A. Cochlear nerve visualization in Normal anatomy and inner ear malformations. *Laryngoscope Investigative Otolaryngology*. 2024;9(6):e70023. Available from: <https://onlinelibrary.wiley.com/doi/abs/10.1002/liot.70023>.
- Renna JM, Sondereker KB, Cors CL, Chaszeyka SN, Keenan KN, Corigliano MR, et al. From 2D slices to a 3D model: Training students in digital microanatomy analysis techniques through a 3D printed neuron project. *Anatomical Sciences Education*. 2024;17(3):499–505. Available from: <https://anatomypubs.onlinelibrary.wiley.com/doi/abs/10.1002/ase.2396>.
- Labranche L, Wilson TD, Terrell M, Kulesza RJ. Learning in Stereo: The Relationship Between Spatial Ability and 3D Digital Anatomy Models. *Anatomical Sciences Education*. 2022;15(2):291–303. Available from: <https://anatomypubs.onlinelibrary.wiley.com/doi/abs/10.1002/ase.2057>.
- Roach VA, Mi M, Mussell J, Van Nuland SE, Luffler RS, DeVeau KM, et al. Correlating Spatial Ability With Anatomy Assessment Performance: A Meta-Analysis. *Anatomical Sciences Education*. 2021;14(3):317–329. Available from: <https://anatomypubs.onlinelibrary.wiley.com/doi/abs/10.1002/ase.2029>.
- Crossingham JL, Jenkinson J, Woolridge N, Gallinger S, Tait G, Moulton C. Interpreting three-dimensional structures from two-dimensional images: a web-based interactive 3D teaching model of surgical liver anatomy. *HPB : the official journal of the International Hepato Pancreato Biliary Association*. 2009;11:523–8.
- Bergman EM. Discussing dissection in anatomy education. *Perspectives on Medical Education*. 2015;4:211 – 213.
- Brun H, Lippert M, Langø T, Sanchez-Margallo J, Sanchez-Margallo F, Elle OJ. Comparing assisting technologies for proficiency in cardiac morphology: 3D printing and mixed reality versus CT slice images for morphological understanding of congenital heart defects by medical students. *Anatomical Sciences Education*; Available from: <https://anatomypubs.onlinelibrary.wiley.com/doi/abs/10.1002/ase.2530>.
- Ho S, Liu P, Palombo DJ, Handy TC, Krebs C. The role of spatial ability in mixed reality learning with the HoloLens. *Anatomical Sciences Education*. 2022;15(6):1074–1085. Available from: <https://anatomypubs.onlinelibrary.wiley.com/doi/abs/10.1002/ase.2146>.

24. Kikinis R, Pieper SD, Vosburgh KG. In: Jolesz FA, editor. 3D Slicer: A Platform for Subject-Specific Image Analysis, Visualization, and Clinical Support. New York, NY: Springer New York; 2014. p. 277–289. Available from: https://doi.org/10.1007/978-1-4614-7657-3_19.
25. White S, Feiner S. SiteLens: situated visualization techniques for urban site visits. In: Proceedings of the SIGCHI Conference on Human Factors in Computing Systems. CHI '09. New York, NY, USA: Association for Computing Machinery; 2009. p. 1117–1120. Available from: <https://doi.org/10.1145/1518701.1518871>.
26. Bressa N, Korsgaard H, Tabard A, Houben S, Vermeulen J. What's the Situation with Situated Visualization? A Survey and Perspectives on Situatedness. *IEEE Transactions on Visualization and Computer Graphics*. 2022;28(1):107–117.
27. Bogomolova K, van der Ham IJM, Dankbaar MEW, van den Broek WW, Hovius SER, van der Hage JA, et al. The Effect of Stereoscopic Augmented Reality Visualization on Learning Anatomy and the Modifying Effect of Visual-Spatial Abilities: A Double-Center Randomized Controlled Trial. *Anatomical Sciences Education*. 2020;13(5):558–567. Available from: <https://anatomypubs.onlinelibrary.wiley.com/doi/abs/10.1002/ase.1941>.
28. Hart SG. NASA-task load index (NASA-TLX); 20 years later. In: Proceedings of the human factors and ergonomics society annual meeting. vol. 50. Sage publications Sage CA: Los Angeles, CA; 2006. p. 904–908.
29. Ponsiglione AM, Amato F, Cozzolino S, Russo G, Romano M, Improta G. A Hybrid Analytic Hierarchy Process and Likert Scale Approach for the Quality Assessment of Medical Education Programs. *Mathematics*. 2022;10(9). Available from: <https://www.mdpi.com/2227-7390/10/9/1426>.