IMMERSIVE MEDIA IN CONNECTED HEALTH

EDITED BY: Panagiotis Evaggelos Antoniou, Georgios Tsoulfas,
Daphne Economou and Alkinoos Athanasiou
PUBLISHED IN: Frontiers in Digital Health and Frontiers in Virtual Reality

frontiers Research Topics

ISSN 1664-8714
ISBN 978-2-88971-542-8
DOI 10.3389/978-2-88971-542-8

About Frontiers

Frontiers is more than just an open-access publisher of scholarly articles: it is a pioneering approach to the world of academia, radically improving the way scholarly research is managed. The grand vision of Frontiers is a world where all people have an equal opportunity to seek, share and generate knowledge. Frontiers provides immediate and permanent online open access to all its publications, but this alone is not enough to realize our grand goals.

Frontiers Journal Series

The Frontiers Journal Series is a multi-tier and interdisciplinary set of open-access, online journals, promising a paradigm shift from the current review, selection and dissemination processes in academic publishing. All Frontiers journals are driven by researchers for researchers; therefore, they constitute a service to the scholarly community. At the same time, the Frontiers Journal Series operates on a revolutionary invention, the tiered publishing system, initially addressing specific communities of scholars, and gradually climbing up to broader public understanding, thus serving the interests of the lay society, too.

Dedication to Quality

Each Frontiers article is a landmark of the highest quality, thanks to genuinely collaborative interactions between authors and review editors, who include some of the world's best academicians. Research must be certified by peers before entering a stream of knowledge that may eventually reach the public - and shape society; therefore, Frontiers only applies the most rigorous and unbiased reviews.
Frontiers revolutionizes research publishing by freely delivering the most outstanding research, evaluated with no bias from both the academic and social point of view.
By applying the most advanced information technologies, Frontiers is catapulting scholarly publishing into a new generation.

What are Frontiers Research Topics?

Frontiers Research Topics are very popular trademarks of the Frontiers Journals Series: they are collections of at least ten articles, all centered on a particular subject. With their unique mix of varied contributions from Original Research to Review Articles, Frontiers Research Topics unify the most influential researchers, the latest key findings and historical advances in a hot research area! Find out more on how to host your own Frontiers Research Topic or contribute to one as an author by contacting the Frontiers Editorial Office: frontiersin.org/about/contact

IMMERSIVE MEDIA IN CONNECTED HEALTH

Topic Editors:
Panagiotis Evaggelos Antoniou, Aristotle University of Thessaloniki, Greece
Georgios Tsoulfas, Aristotle University of Thessaloniki, Greece
Daphne Economou, University of Westminster, United Kingdom
Alkinoos Athanasiou, University General Hospital of Thessaloniki AHEPA, Greece

Citation: Antoniou, P. E., Tsoulfas, G., Economou, D., Athanasiou, A., eds. (2021). Immersive Media in Connected Health. Lausanne: Frontiers Media SA.
doi: 10.3389/978-2-88971-542-8

Table of Contents

frontiers in Digital Health

EDITORIAL
published: 08 September 2021
doi: 10.3389/fdgth.2021.697336

Editorial: Immersive Media in Connected Health

Panagiotis E. Antoniou[1*‡], *Daphne Economou*[2†], *Alkinoos Athanasiou*[1†] and *George Tsoulfas*[3†]

[1] *Laboratory of Medical Physics and Digital Innovations, Department of Medicine, School of Health Sciences, Aristotle University of Thessaloniki, Thessaloniki, Greece,* [2] *Department of Computer Science and Engineering, University of Westminster, London, United Kingdom,* [3] *First General Surgery Department, Papageorgiou General Hospital/Aristotle University of Thessaloniki, Thessaloniki, Greece*

Keywords: mixed reality, medical education, surgery, telepresence, 3D printing

Editorial on the Research Topic

Immersive Media in Connected Health

INTRODUCTION

According to Fortune Business Insights, the market size of extended reality (XR), a blanket term comprising augmented, virtual, and augmented reality (AR/VR/MR), is estimated to grow to 30.4 billion USD by 2027 from an estimated 1.56 billion USD (1). This simple figure outlines more succinctly than any literary review the explosive growth of XR in the healthcare sector. With XR being only one vista in the whole landscape of immersive media in connected health, the scope of this research study becomes readily apparent. Medical image fusion, educational immersive visualizations, web-based virtual scenarios, and XR interventions for mental and physical health are all parts of the research ecosystem of this topic.

This special issue aims at contemporary immersive media research study as it pertains to healthcare. From the whole spectrum of XR to a diverse landscape of image processing and fusion, educational simulations, and technical/usability issues in proliferating immersive media in the whole connected healthcare ecosystem, this special issue aims to explore both the technologies of immersive media and their impact on contemporary healthcare.

On the educational front, virtual patients (VPs), with current web-based rapid development and deployment cycles, are ubiquitously present in the healthcare curricula. Meaningfully combining immersive three-dimensional (3D) virtual environments with VPs leads medical training to new aspects of open social learning. Learners can absorb educational content at their own pace and engage in an experiential way and not only feel the presence but also affect the sense of presence in the training event (2). Developing such resources and repurposing them for virtual reality (VR) are a rather straightforward task, that is, one, however, that must adhere to sound game-informed design principles and narrative techniques to achieve measurable impact both in learning outcomes and engagement (3).

However, education is not the only healthcare field that XR applies. Preoperative surgical planning using mixed reality (4) and 3D prints for patient-specific surgical approaches are proliferating (5). Haptic controls in conjunction with XR visualization and novel medical image co-registration and fusion are becoming common practices in surgical training and preoperative preparation.

Edited and reviewed by:
Constantinos S. Pattichis, University of Cyprus, Cyprus

***Correspondence:**
Panagiotis E. Antoniou
pantonio@otenet.gr

† *These authors have contributed equally to this work*

‡ *This author share first authorship*

Specialty section:
This article was submitted to Connected Health, a section of the journal Frontiers in Digital Health

Received: *19 April 2021*
Accepted: *29 June 2021*
Published: *08 September 2021*

Citation:
Antoniou PE, Economou D, Athanasiou A and Tsoulfas G (2021) Editorial: Immersive Media in Connected Health. Front. Digit. Health 3:697336. doi: 10.3389/fdgth.2021.697336

Beyond "hard" medicine applications, the whole spectrum of immersive health is benefited through healthcare interventions. The general public and targeted sensitive groups of people are benefiting from physical and mental wellness interventions based on immersive media such as XR, m-health, or even virtual environment ecosystems (6).

OUTLINE OF CONTRIBUTIONS

In this context, our topic attracted a diverse collection of studies.

Starting from the core need for surgical guidance, Drakopoulos et al. combined non-rigid registration of preoperative operative data with mixed reality presentation in order to provide an affordable, fast, and immersive solution that is able to tackle the intraoperative brain deformation issue in brain tumor resections. During resection, especially of significant tumor volumes, navigation systems based on preoperative imaging lose significant portions of accuracy and guidance value due to dynamic brain shifts. The authors used an Adaptive Physics-Based Non-Rigid Registration method (A-PBNRR) and multi-tissue mesh adaptation to register preoperative and intraoperative MRI on glioma surgery data of real patients and evaluated registration accuracy using visual assessment, Hausdorff distance-based metric, and a landmark-based approach using anatomical points identified by a neurosurgeon. Their approach reported several positive outcomes: It significantly improved the accuracy of deformable registration compared to traditionally employed methods; it was sufficiently fast for intraoperative application; and when coupled with Mixed Reality (MR), it offers an affordable alternative solution to current neuronavigation systems.

Moving to core medical education, the study by Jivram et al. explored the use of Second Life virtual 3D environment as a more realistic space for deploying virtual scenarios. Collaborative learning through case-based or problem-based learning (PBL) is an established way to cultivate workplace knowledge associated with specific competencies. At the St George's University of London, an interactive online form of decision-based PBL (D-PBL) was developed using web-based VPs, which was subsequently transferred to the Second Life virtual environment. It is interesting to note that while a small number of students found the Second Life experience more engaging in terms of realism, most students favored a simpler more concise interaction of the web-based VPs. In that context, providing the learner with the essential topical knowledge in a simple engaging medium trumped the increased immersion of the virtual environment.

Deepening the exploration of educational themes, Pieterse et al. researched how augmented reality provides us with the necessary technology in order to enhance the experience of the real world with computer-generated information and a variety of auditory, haptic, and visual stimuli. This opens a whole new world of educational opportunities (among many others), which have the potential to transform education at all different levels. In this study, the authors developed an augmented reality application for the presentation of shortness of breath (dyspnea). This was subsequently successfully implemented in the medical curriculum, where its value was confirmed by medical students. The authors provide a detailed description of an elegant and elaborate process, which has the potential to provide us with potent, novel educational tools.

Finally concluding the topic, an important technical usability hurdle is tackled by Schmidt et al. The issues of the unpleasant side effects caused by VR headset to users such as cybersickness and sensory disorders have been well studied. However, there has been almost no exploration of these problems with individuals with autism or other cognitive disabilities. The study with the title "process-model for minimizing adverse effects when using head mounted display-based virtual reality for individuals with autism" is an initial attempt to address this gap by conducting a study in two separate, independent research sites—one in the United States and one in the United Kingdom. The results of the study assert that the proposed guidelines could provide clarity in terms of design and implementation of headset-based VR for individuals with autism spectrum disorders, guide implementations of future researchers and practitioners, and contribute to minimizing and controlling for potential adverse effects.

CONCLUDING REMARKS

While the field is quite extensive, it is auspicious that the scope of the collected studies has been diverse enough to traverse most of its admittedly wide scope. It is within this auspicious context that we invite the reader to traverse the nuances of an eclectic collection of research studies on this topic.

AUTHOR CONTRIBUTIONS

PA: general section authorship and review of contributions to the Research Topic. DE, AA, and GT: review of contributions to the Research Topic. All authors contributed to the article and approved the submitted version.

REFERENCES

1. *"Virtual Reality Healthcare Market Size" Fortune Business Insights VR Health*. (2021). Available online at: https://www.fortunebusinessinsights.com/industry-reports/virtual-reality-vr-in-healthcare-market-101679 (accessed April 15, 2021).
2. Antoniou PE, Athanasopoulou CA, Dafli E, Bamidis PD. Exploring design requirements for repurposing dental virtual patients from the web to second life: a focus group study. *J Med Internet Res.* (2014) 16:3343. doi: 10.2196/jmir.3343
3. Antoniou PE, Dafli E, Arfaras G, Bamidis PD. Versatile mixed reality medical educational spaces; requirement analysis from expert users. *Personal Ubiquit Comput.* (2017) 21:1015–24. doi: 10.1007/s00779-017-1074-5
4. Antoniou PE, Athanasiou A, Bamidis PD. Virtual and augmented reality in neuroscience. In: de Albuquerque VH, Athanasiou A, Ribeiro

S, editors. *Neurotechnology: Methods, Advances and Applications*. The IET (2020).

5. Bangeas P, Tsioukas V, Papadopoulos VN, Tsoulfas G. Role of innovative 3D printing models in the management of hepatobiliary malignancies. *World J Hepatol*. (2019) 11:574–85. doi: 10.4254/wjh.v11.i7.574
6. Antoniou PE, Ioannidis L, Sidiropoulos E, Bamidis PD. DISCOVER-ing opensim. Design guidelines and implementation of scenario based learning for carers of the elderly. In: *Proceedings of the 7th International Conference on Education and New Learning Technologies*. Barcelona (2015).

Conflict of Interest: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher's Note: All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

frontiers in Virtual Reality

ORIGINAL RESEARCH
published: 28 October 2020
doi: 10.3389/frvir.2020.577534

Check for updates

Design and Implementation of "*AugMedicine: Lung Cases,*" an Augmented Reality Application for the Medical Curriculum on the Presentation of Dyspnea

Arianne D. Pieterse[1], Beerend P. Hierck[2,3], Peter G. M. de Jong[3], Jelger Kroese[4], Luuk N. A. Willems[5] and Marlies E. J. Reinders[1,6]*

[1] Department of Internal Medicine, Leiden University Medical Center, Leiden, Netherlands, [2] Department of Anatomy & Embryology, Leiden University Medical Center, Leiden, Netherlands, [3] Center for Innovation in Medical Education, Leiden University Medical Center, Leiden, Netherlands, [4] Centre for Innovation, Leiden University, The Hague, Netherlands, [5] Department of Pulmonology, Leiden University Medical Center, Leiden, Netherlands, [6] Department of Nephrology, Leiden University Medical Center, Leiden, Netherlands

OPEN ACCESS

Edited by:
Georgios Tsoulfas,
Aristotle University of Thessaloniki, Greece

Reviewed by:
Randy Klaassen,
University of Twente, Netherlands
Alexandros Sigaras,
Weill Cornell Medicine, Cornell University, United States

***Correspondence:**
Arianne D. Pieterse
a.d.pieterse@lumc.nl

Specialty section:
This article was submitted to Virtual Reality in Medicine, a section of the journal Frontiers in Virtual Reality

Received: 29 June 2020
Accepted: 15 September 2020
Published: 28 October 2020

Citation:
Pieterse AD, Hierck BP, de Jong PGM, Kroese J, Willems LNA and Reinders MEJ (2020) Design and Implementation of "AugMedicine: Lung Cases," an Augmented Reality Application for the Medical Curriculum on the Presentation of Dyspnea. Front. Virtual Real. 1:577534. doi: 10.3389/frvir.2020.577534

Introduction: Augmented Reality is a technique that enriches the real-life environment with 3D visuals and audio. It offers possibilities to expose medical students to a variety of clinical cases. It provides unique opportunities for active and collaborative learning in an authentic but safe environment. We developed an Augmented Reality application on the clinical presentation of shortness of breath (dyspnea), grounded on a theoretical instructional design model.

Methods: A team of various stakeholders (including medical teachers and students) was formed to design the application and corresponding small group learning session, grounded on principles of instruction as described by Merrill. Evaluation was performed by an explorative questionnaire, consisting of open and closed questions (Likert scales), covering user experience, content and physical discomfort.

Results: Multiple interactive cases of dyspnea were designed. The application plays back audio samples of abnormal lung sounds corresponding to a specific clinical case of dyspnea and displays a 3D model of the related pulmonary pathologies. It was implemented in the medical curriculum as an obligatory small group learning session scheduled preceding clinical clerkships. Prior knowledge was activated prior to the learning session. New knowledge was acquired with the application by solving an authentic problem based on a real patient case. In total 110 students participated in the study and 104 completed the questionnaire. 85% of the students indicated that the virtually auscultated lung sounds felt natural. 90% reported that the augmented reality experience helped them to better understand the clinical case. The majority of the students (74%) indicated that the experience improved their insight in the portrayed illness. 94.2% reported limited or no physical discomfort.

Discussion: An Augmented Reality application on the presentation of dyspnea was successfully designed and implemented in the medical curriculum. Students confirm the value of the application in terms of content and usability. The extension of this Augmented Reality application for education of other healthcare professionals in currently under consideration.

Keywords: augmented reality, mixed reality, medical education, technology enhanced learning, active learning, collaborative learning, pulmonary medicine, internal medicine

INTRODUCTION

Medical students need exposure to a wide variety of authentic clinical cases to prepare for clinical activities. Until recently, the preparation of medical students for the clinical clerkships at Leiden University Medical Center (LUMC) comprised mainly e-learning modules, small group learning, physical examination of peer students, interpretation of 2D radiology images and practicing with simulated patients. However, authenticity in these learning environments is often limited. Learning activities with simulated patients cannot expose students to abnormalities on physical examination such as pathological lung sounds, which can be important clues in the diagnostic process (Cleland et al., 2009). E-learning can offer virtual patients with ample examples of authentic clinical cases, but has limitations in practicing communication (Edelbring et al., 2011) and psychomotor skills, such as auscultation. However, these competencies will also be necessary once the students encounter real patients. And finally, simulated patients nor e-learning can demonstrate the relationship between auscultated lung sounds and 3D anatomy very well, which is one of the major challenges for students.

A possible technology to help students bridging the gap between virtual or simulated patients and clinical practice is Augmented Reality (AR). AR is an innovative digital technique that enriches the real-life environment by adding computer generated perceptual information, such as 3D visuals, audio and/or (haptic) feedback. It can promote active learning by literally making students walk around a 3D hologram and stimulating relevant motion patterns, and it is suitable for collaborative learning by synchronizing experiences among users (Kaufmann and Schmalstieg, 2003; Barmaki et al., 2019; Bogomolova et al., 2019, 2020). In order to add virtual content to the real-life environment, various hardware devices can be used for AR ranging from handheld devices such as smart phones or tablets to head-mounted transparent displays (Kamphuis et al., 2014). Examples of the latter are Google Glass® and Microsoft HoloLens®. An advantage of AR applications with a head-mounted display is that, in contrast to virtual reality (VR) headsets, learners can continue interacting with each other and their teachers because of the transparency of the AR headset. In this way, an authentic and safe learning environment for small group learning can be created.

Not surprisingly, the field of AR in medicine is growing rapidly (Eckert et al., 2019). There are many publications on this topic, and it should be noted that one should be careful interpreting literature data as AR technology changed markedly over time. Many of the published AR applications are on the topics of surgery and radiology, and described prototypes that needed further development (Eckert et al., 2019). When zooming in on AR specifically for medical education, most studies focus on surgical or anatomy education (Zhu et al., 2014; Sheik-Ali et al., 2019; Tang et al., 2020; Uruthiralingam and Rea, 2020). Studies on other medical fields are underrepresented (Nischelwitzer et al., 2007; Rosenbaum et al., 2007; Nilsson and Johansson, 2008; Pretto et al., 2009; Sakellariou et al., 2009; Lamounier et al., 2010; Rasimah et al., 2011; Kotranza et al., 2012; von Jan et al., 2012; Gerup et al., 2020). Interestingly, only a few AR applications incorporated a collaborative aspect (e.g., Kaufmann and Schmalstieg, 2003; Barmaki et al., 2019; Bogomolova et al., 2019), perhaps because of technical reasons, such as a lack of widely available network coverage or the computational ability to allow for shared or multiuser applications. In many studies, both learners and experts responded positively toward AR and the possibility of implementing AR in training programs. The current quality and range of AR research in medical education may not yet be sufficient to recommend the adoption of AR technologies into medical curricula (Tang et al., 2020). There are relatively few studies focusing on AR with head-mounted displays, and specifically in the field of lung diseases and auscultation, there are no AR applications described yet.

In this study we designed and developed an AR application for head-mounted display in the field of pulmonary medicine, to complement the diagnostic process on the presentation of dyspnea (shortness of breath). This application provides users with a synchronized experience comprising a virtual stethoscope and positional lung sounds, and visualizes relevant anatomy and pathology in an interactive setup. It is grounded on a theoretical instructional design model based on the First principles of instruction as described by Merrill (2002). The first implementation in our medical curriculum was evaluated by an explorative questionnaire amongst learners.

MATERIALS AND METHODS

Learning Experience Design and Design Process of the AR Application

In November 2017, a collaboration between Leiden University Medical Center (LUMC) and Leiden University's Center for Innovation was formed in order to co-create the AR application and small group learning sessions with this application. The team consisted of various stake holders: clinical teachers, anatomy experts, medical students, developers, project managers, educational experts and a medical illustrator. The total design

Abbreviations: ACM, Audio Collision Model; AR, Augmented Reality; COPD, Chronic Obstructive Pulmonary Disease; HLAPI, High Level Application Programming Interface; LED, learning experience design; LUMC, Leiden University Medical Center; SDK, software development kit; UI, user interface; VR, Virtual Reality; WLAN, wireless local area network.

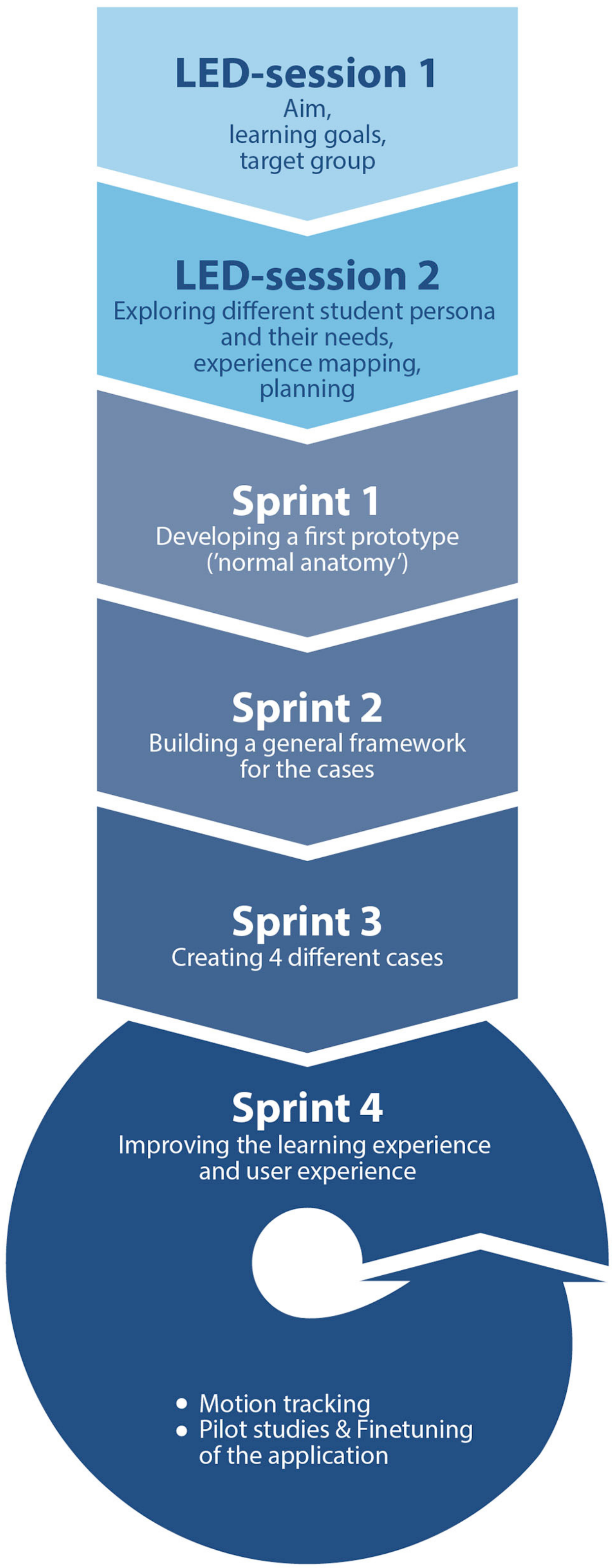

FIGURE 1 | Overview of the design process.

process had a duration of ∼22 months and is visualized in **Figure 1**. It was started by a so-called "Learning Experience Design (LED) session" which involved all stakeholders. During this session, the common goal was set, the target group was defined and the project process was outlined. The project was carried out using the SCRUM method (Schwaber and Beedle, 2008) working in short timeframes (sprints). Once every 2 months the project team held sprint meetings in which the progress was discussed and project activities were prioritized. This design process resulted in several functional requirements. The application should offer:

- An immersive authentic AR experience for use with a head-mounted display (the HoloLens®)
- A 3D thorax model that is anatomically correct, and several 3D thorax models with authentic pulmonary pathologies
- The possibility to auscultate lung sounds within the AR application corresponding to a clinical case in an authentic manner, and practice as long and as frequently as needed
- The possibility to visualize site-specific pathologies corresponding to localized lung sounds
- The possibility of augmentation of a real-life simulated patient with the digital models, so communication skills could be practiced in an authentic manner
- A synchronized experience by multiple users.

It was decided to use a modular design for the AR application, so extra cases could be easily added in the future.

The first LED session was followed by a second LED session and multiple sprints with cycles of design, evaluation and redesign (**Figure 1**). In the second LED session, the team focused on learning experience mapping. Specifically, distinct learner personas were created and the learner journey was outlined. A design goal was that the AR application would optimally serve the different student personas. One student persona is a motivated learner with affinity for ICT, and another persona that was less motivated requiring more assistance with ICT. The learner journey was translated into a series of modular steps in the learning process and within the AR application, which were further explored and developed in the sprints. In the first sprint, a first prototype of the 3D model was created and evaluated: a 3D thorax with normal anatomy and the possibility to listen to lung sounds. After testing the prototype with multiple team members (students and teachers), the feedback was used as input for the second sprint. The main suggestions for improvements were: a feature to move the model up or down (the default position was too low and could not yet be adjusted), a longer or resizable virtual stethoscope (by using a relative short stethoscope users had to come quite close to the model, which resulted in a limited overview of the model due to the decreased field of view). The second sprint resulted in a prototype with a general framework of a user interface with the possibility to adjust the position and the size of the model. Again, the prototype was evaluated by team members (both students and teachers). Their feedback included that the audio files did not start and stop in a natural way, which was fixed in a later sprint. In the third sprint, the content for four clinical cases was created and incorporated

TABLE 1 | Use of Merrill's First Principles of Instruction in the educational process with the AR application.

Principle:	Description of the use of this principle in the educational process:
Task/Problem-centered	The students perform a series of tasks simulating the diagnostic process of evaluating a patient with dyspnea at the Emergency Department, which is a real-world problem
Activation	Before entering the small group learning session, prior knowledge is activated by online assignments
Demonstration	Students are demonstrated the diagnostic process and problem-solving strategies by the teacher
Application	In the interactive small group learning session students collaboratively solve a patient case
Integration	Students are encouraged to integrate their knowledge and skills during their tasks as an intern

in the application. The 3D models corresponding to the cases were designed by a medical illustrator based on radiology images (mainly CT scans) of patients with corresponding diagnoses. Previously recorded audio samples of lung sounds of actual patients were used, and for each case the spatial positioning of the appropriate lung sounds was programmed to correspond to the location(s) of pulmonary abnormalities. The fourth sprint comprised further finetuning of the application to improve both learning and user experience and refining a system for motion tracking of the virtual stethoscope. After the fifth sprint, pilot studies with medical students were performed and evaluated. The prototype was tested by several medical students that were not involved in designing the application. The students were observed by several team members whilst using the application, and were asked to provide their feedback. One suggestion for improvement was providing feedback whether they listened on an appropriate or inappropriate location (e.g., it is not possible to listen to lung sounds if you place the virtual stethoscope on a bony surface such as the scapula). This feedback resulted in an extra feature: the virtual stethoscope changes in color from blue to orange if a student listens on an inappropriate location.

Instructional Design of the Small Group Learning Session

Merrill's First Principles of Instruction were used as theoretical framework for designing the small group learning session (Merrill, 2002). The incorporation of these five instructional design principles is shown in **Table 1**. The core principle is to engage learners in relevant, real-world problems. The presenting complaint *shortness of breath* was selected, because interns will encounter many patients with this symptom at the Emergency Department and Internal Medicine ward. The other principles of Merrill relate to consecutive phases of instruction. The second principle is that activation of previous knowledge or experience promotes learning. In order to stimulate this, students are provided with online assignments before entering the learning session, to recall knowledge relating to diagnosing patients with dyspnea. Because a simple recall of knowledge is rarely effective, the teacher refers to these assignments at the start of the session and allows opportunities for discussion. The third principle is that learning is promoted by demonstrating what needs to be learned. In order to do this, the teacher will demonstrate the consecutive steps of the diagnostic process and problem-solving strategies. The fourth principle is to promote learning by solving problems ("the application phase"). The vast part of the small group learning session is dedicated to this phase in which students collaboratively solve a patient case, including the parts in which the AR application is used. The fifth and final principle is the integration of knowledge. This is a natural step, because students will have ample opportunities to utilize the obtained knowledge and skills on new (real-world) patient cases during the Internal Medicine clerkship, which in our hospital starts within 2 weeks after the learning session. Because the students will be regularly assessed in the clinical clerkships, the subject of diagnosing cases with dyspnea was not formally assessed during this phase of the course.

Evaluation of the Implementation of the AR Application

The application and learning session have been developed for students in the preparative course of the Internal Medicine clerkship. In our curriculum, this clerkship is the first in a 3-years period of clinical rotations. The first experiences of students with the AR application in the small group learning session, were obtained by a voluntary questionnaire. This explorative questionnaire, as provided as **Supplementary Material**, covered the following topics:

- General information
- Use of AR application
- Content of the AR application
- Personal goal of the students
- Physical discomfort.

All medical students at LUMC participating in the obligatory small group learning sessions during the preparative course of the Internal Medicine clerkship were included in the study. Students that participated previously (e.g., students that had to repeat the clerkship including the preparative course) were excluded.

It was planned to study 10 groups of ∼10–12 students over a period of 5 months to reach at least 45 responses being a representative 15% sample of the full cohort of 300 students per year. If the number of 45 responses could not reached after 5 months, the study period would be expanded until the number of 45 responses has been reached.

Participants were invited by a team member that was not involved in assessing or grading them to fill out a paper questionnaire, immediately after the small group learning session. All the participants received a letter with detailed information about the study objectives and data safety, and signed an informed consent form. After collecting the questionnaires, the data were pseudonymized and provided as such to the research team. The data have been stored for further research purposes according to the "Nederlandse Gedragscode Wetenschapsbeoefening" of the VSNU (The

A B

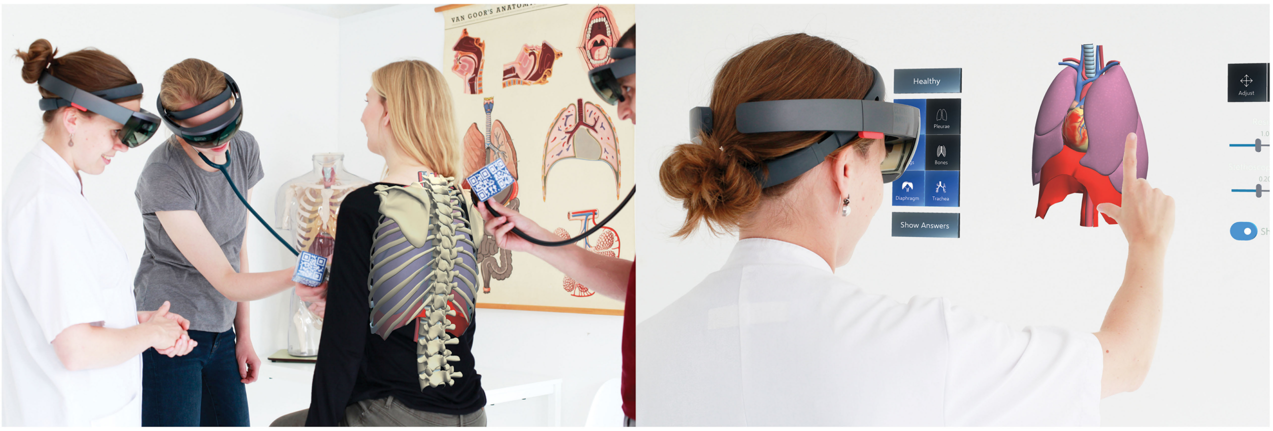

FIGURE 2 | Impressions of the use of the AR application. **(A)** Auscultation with the AR application on a standardized patient. **(B)** User interface of the AR application.

Association of Universities The Netherlands). The study protocol has been reviewed and approved by the Educational Research Review Board of the Leiden University Medical Center (OEC/ERRB/20191008/1).

The answers to the evaluative questions with a 4 or 5-point Likert scale were quantitatively analyzed. The answers to open ended questions were qualitatively analyzed by a thematic analysis. The answers were coded *in vivo*, overarching themes were selected in a inductive approach, and the quotations were assigned to overarching themes by the principal investigator (Miles et al., 2014). Illustrative quotations per theme were selected. The questions related to the type and extent of physical discomfort that students experienced during the use of the AR application, were descriptively analyzed.

RESULTS

Short Description of the AR Application

The design process resulted in an interactive multiuser AR application: *AugMedicine: Lung cases* [see online video's: Demonstration of AugMedicine: Lung Cases (2020), HoloLens App AugMedicine: Lung Cases (short) (2020)]. The application comprises a multilayer 3D anatomical model of a human thorax combined with 3D spatialized audio. When entering a session, users (either student or teacher) are assigned one of two possible roles: host or client. The host can control the application and select different cases of dyspnea (COPD + pneumonia, pulmonary embolism, pneumothorax or pleural effusion) or a 3D model with normal anatomy. The models are textured surface models which limits the amount of (synchronized) data, and improves performance. Since no cross sections are implemented in this application, solid models would have no added value. The device uses a coordinate system to place the model into the room, and the users can walk around to change their perspective on the 3D model. Users can manipulate a virtual stethoscope to listen to the lung sounds at different locations and from different angles, exactly as they would do when examining a real patient. The sounds are played back by the AR glasses and correspond to the selected case (see **Figures 2A,B**). The host can display the 3D model on a real person's body (e.g., a simulated patient) so the auscultation process becomes more authentic (**Figure 2A**). The host can also choose which anatomical layers are visible to the learners (bones, diaphragm, heart and blood vessels, trachea, lungs, pleurae; see **Figure 2B**). The host can eventually show a 3D visual representation of the actual diagnosis.

Technology Used for Developing the AR Application

Development Framework

The application was developed using the Unity game engine, with C# as scripting language. The open source Mixed Reality Toolkit was used for basic features, such as rotating, moving, resizing the 3D model, and basic user interface (UI) components such as buttons, sliders, input handlers, screen reticle, shaders. All other functionalities and app logic were custom developed for this application.

Multi-User Synchronizing

Synchronizing the application across multiple HoloLens® devices was done through connecting the devices over a wireless local area network (WLAN). To achieve this, custom networking functionality was developed for the application, using the UNet networking High Level Application Programming Interface (HLAPI) from Unity. Upon starting a group session, one of the users chooses to start the application in "host mode." This starts a new networked session in which this user is both host and server. Once the other users start the application, they are automatically added to the session as "clients." The HoloLens® running in "host mode" displays, additionally to the virtual torso, a user interface that can be used to control the session. The user interface includes the following options: changing the case, selecting which elements of the 3D model are visible, changing stethoscope length, resizing, moving and rotating the 3D model. Only the host can perform these actions. Changes in the application are automatically synchronized with the clients.

Aligned Spatial Positioning

It was aimed to achieve a multi-user experience that enables users to communicate about the virtual content in a way that is similar to communicating about a real world object, e.g.,: a user can direct the attention of other users toward a specific part of the 3D model by simply pointing a finger to it. To enable such an experience, it is required that the virtual content is displayed at the same position and orientation relative to the real-world space for all the users. This was achieved by implementing image recognition functionality in the application through the Vuforia® software development kit (SDK), and using it to enable the HoloLens® to recognize the position and orientation of a printed visual marker. Before the start of a session, a printed marker (a QR code) is placed on a non-moving surface, such as a table. During startup of the application, the user is asked to aim the HoloLens®' camera at the marker by looking at it. Once the marker is recognized, the application sets a spatial anchor on the position of the marker and displays the virtual on top of the marker for the rest of the session. Since all users make use of the same printed marker, the virtual content is aligned for all users. Over time, it may happen that a user's model loses alignment. A button was incorporated in the UI that enables the host of the session to trigger a realignment request for users. This prompts them with a message to rescan the printed marker for realigning the virtual content to the marker's position and orientation.

Virtual Stethoscope

Our aim was to simulate working with a stethoscope as accurately as possible. Therefore, the application enables users to control the position and orientation of the virtual stethoscope through operating a real stethoscope (or any other physical object). This requires 6-DOF tracking of the real stethoscope, which was achieved through making use of the image recognition functionality of the Vuforia® SDK. We created a printable custom "multi-target" cube, which is a foldable 3D object that can be recognized by the Vuforia® engine from all angles. Multiple shapes and sizes of the multi-target were tested: cube-, cylinder- and cone-shaped. A cube of 6 cm by 6 cm provided the best balance between tracking stability and size. Using tape, the cube can be attached to a real stethoscope (or any other object). Upon holding the stethoscope with attached cube in front of the HoloLens®' cameras, it is recognized and a virtual stethoscope is rendered over it. The tip of the virtual stethoscope consists of a 3D sphere with a collider, which is used to detect collisions with the virtual torso. The distance between the tip of the virtual stethoscope and the cube target, "stethoscope length", can be set from within the application. When set to a short length, users have to place the virtual stethoscope directly on the outside of the virtual torso to hear sounds, like in a real-world setting. Setting a longer "stethoscope length", allows users to also aim the virtual stethoscope at the virtual torso from a distance to auscultate. This was implemented for a better user experience. When working with larger groups, it enables multiple people to listen to the sounds while keeping personal distance. Given the HoloLens®' Field of View limitations, a longer stethoscope length also allows users to see the full torso when listening to the lung sounds.

Lung Sounds

A custom method was developed to simulate the positioning of sounds for an area of the torso that roughly corresponds with the torso area on which a sound can be heard in reality. This method makes use of audio recordings that are played back in mono. No audio spatialization effects are used, since a real stethoscope also simply provides mono audio. However, like in a real-world situation, the type of sound heard is determined by the position of the stethoscope. The application simulates this by using the location on which the virtual stethoscope collides with the virtual torso to control the type of audio sample that is played back and its volume. This is achieved in the following way.

An invisible 3D model that follows the general outline of the visible 3D torso (functioning as a sort of shell) was used to detect collisions between the torso and the virtual stethoscope. We call this invisible model "Audio Collision Model" (ACM). The ACM is divided into 50 segments that are used to distinguish between different collision areas of the model (see **Figure 3**). Each segment has its own "collider" to detect collisions with the virtual stethoscope. A custom script was written to map each segment to a sound file from our database of lung sounds. A Unique mapping was created for each of the clinical cases in the app. A mapping was made by first identifying, for each clinical case, the torso areas on which a collision with the virtual stethoscope should trigger playback of a particular sound. Then, the segments of the ACM covering that part of the torso area were selected and mapped to the corresponding audio file. This approach allowed us to reuse one ACM for each of the clinical cases in the application. Hence, when switching to a different case, the ACM remains unchanged, but the mapping between the ACM's segments and the sound files is updated. **Figure 3** shows differences in the mapping of audio files to the ACM between two clinical cases (**Figures 3A,B**: pneumothorax; **Figures 3C,D**: COPD + pneumonia). Each color corresponds to one unique sound file. This shows that multiple segments of the ACM can be mapped to the same audio file. Together with a medical specialist it was decided that dividing the ACM in 50 segments provided sufficient resolution for making mappings that simulate a real-world situation. When moving the virtual stethoscope from one segment of the ACM to another, a cross-fade between the volume of the previous and the new audio sample is used to create a smooth transition.

Description of the Small Group Learning Session

The small group learning session for 10–12 students is guided by a clinical teacher and is aimed at diagnosing one clinical case presenting with dyspnea. An illustrative overview of the session is shown in **Figure 4**. In two small subgroups (5–6 students each), the students follow sequential diagnostic steps: taking a patient history with a simulated patient, auscultating lung sounds using the AR application (displayed on the simulated patient), gathering and interpreting information of the physical

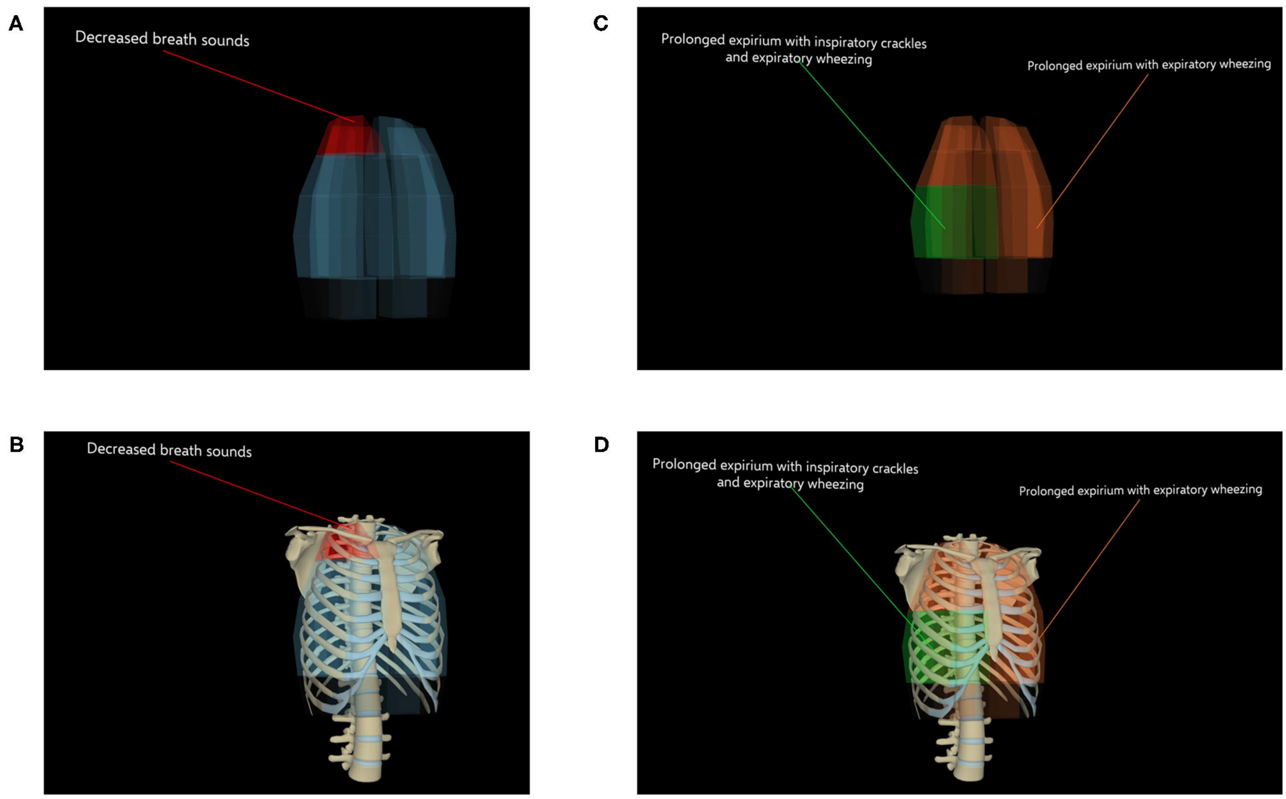

FIGURE 3 | Screenshots of the "Audio Collision Model" (ACM). **(A)** ACM with mapping of audio files of the case "Pneumothorax." **(B)** ACM with mapping of audio files of the case "Pneumothorax," with bones layer visible. **(C)** ACM with mapping of audio files of the case "COPD + pneumonia." **(D)** ACM with mapping of audio files of the case "COPD + pneumonia," with bones layer visible.

examination, ordering and interpreting relevant laboratory tests and radiology exams, and finally formulating a working diagnosis as well as a differential diagnosis. Students are encouraged to think-aloud and discuss their findings and thoughts with their peers at each step. After the last step, a 3D hologram of the underlying illness is displayed by the AR application and is discussed under guidance of a teacher. Because of the use of a simulated patient, students also train their communication skills. This offers the possibility of patient feedback, in addition to peer and teacher feedback. The duration of the session is ∼2 h, during which the students use the AR application twice: at the introduction phase while auscultating the model with the virtual stethoscope, and at the evaluation phase, each time for an estimated 10–15 min (total of 20–30 min per session). There are two short breaks during the session.

Implementation of the Small Group Learning Session With the AR Application

First, two pilot sessions were held in summer 2019. These sessions were used to refine logistics and to improve the study questionnaire. After these pilot studies, the small group learning session was implemented in our medical curriculum. The learning session is now integral part of the preparative course for the Internal Medicine clerkship. Every 4 weeks two new groups start (total of 20–24 new students) and participate in the obligatory small group learning session. During the study period, four different teachers guided the small group learning sessions. One of the teachers (AP) was involved in designing the AR application and in evaluating the AR experience. The other teachers received instructions about the AR application and evaluation prior to the session. Sometimes technical difficulties occurred during the learning sessions, e.g. problems connecting a HoloLens® to the group session, drifting of the model and crashing of the application. These technical problems appeared to increase when more students were using the AR application simultaneously, although we did not formally test this in this study. At the end of the study period, adjustments were made to prevent these technical problems in the future.

Evaluation of the AR Application

During the study period between October 2019 and February 2020, all students who participated in the obligatory small group learning session (n = 110) were asked to participate in the study. A total of 104 completed questionnaires was received (response rate 94.5 %). None of the students that were involved in designing the application participated in the study. The general information of the study participants is provided in **Table 2**. The average participant age was 23 years, with

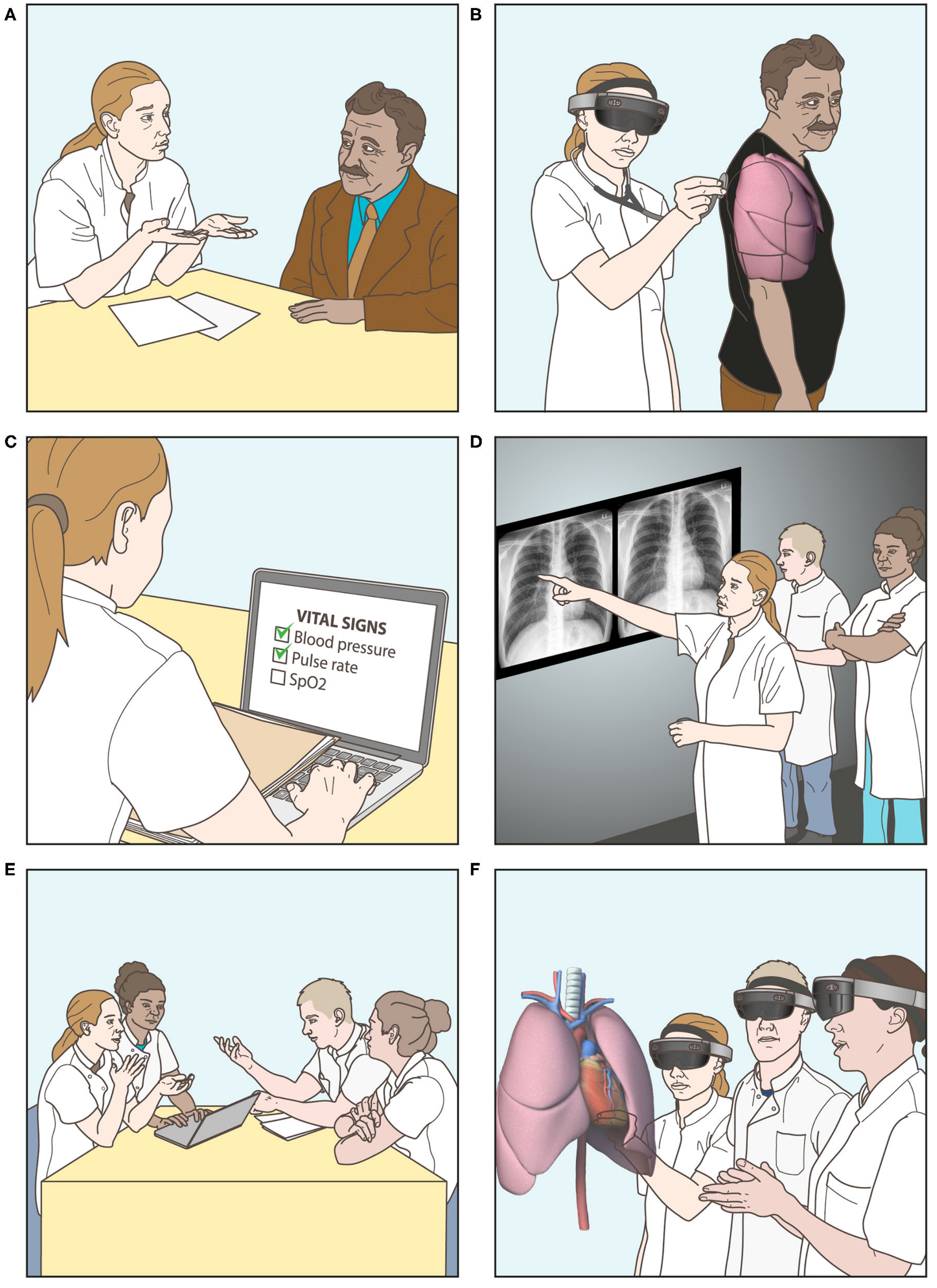

FIGURE 4 | Overview of small group learning session with the AR application. **(A)** Taking a patient history with a standardized patient. **(B)** Auscultating lung sounds using the virtual stethoscope. **(C)** Gathering and interpreting information of the physical examination. **(D)** Ordering and interpreting laboratory tests and radiology examinations. **(E)** Formulating a differential and working diagnosis. **(F)** Observing a 3D hologram of the underlying illness, displayed by the AR application, and discussing the case.

TABLE 2 | General information of students.

Characteristics	Total $n = 104$*
Age (y)†	23 ± 3.12
Sex (male)	44 (42.3)
Glasses or Contact lenses	Glasses 26 (25) Lenses 30 (29)
Wore glasses/contact lenses during assignment	Glasses 11 (11) Lenses 28 (27)
Hearing problems	4 (4)
No experience with headset	87 (84)
Any experience with AR headset	17 (16)
– 1–2 experiences	14 (13)
– 3–5 experiences	2 (2)
– >5 experiences	1 (1)

*Values in parentheses are percentages.
†Mean ± SD.

a range of 21–47 years. 42.3% of the participants was male. Many of the responders reported that they had glasses (25%) or contact lenses (29%). Only some of the students actually wore their glasses during the AR experience (11 out of 25 students), although it is no problem to wear glasses while using the AR application with the head-mounted display. Almost none of the students reported hearing problems. Most students did not have any prior experience with an AR headset (84%). Two different clinical cases were used in the study period: *Pneumothorax* and *COPD (Chronic Obstructive Pulmonary Disease) plus pneumonia.*

33% of the respondents reported that they had previously experienced a situation in the real world similar to the AR experience, e.g., examining a (simulated) patient with shortness of breath. The majority of the students indicated that the virtual lungs sounded natural (**Figure 5A**).

For most students, the AR experience helped them to understand the case *a little bit* or *substantially* better (**Figure 5B**). A qualitative analysis to how the AR experience helped students to better understand the case is shown in **Table 3**. Answers to the open ended question "Please, briefly explain how it helped you to better understand the case" were explored and assigned to one of three categories: "3D models," "auscultation of lungs," and "other." A lot of students reported that the virtual auscultation helped them: they indicated for instance that the lung sounds provided additional information and that is was useful that there were different lung sounds on different spots. In addition, they mentioned that they had better discussions about lung sounds. Students remarked they liked the *"Real physical examination with the case"* and *"The combination of seeing realistic pictures 3D, realistic sounds and the fun and interactive setting."* A few students were critical regarding the auscultation process, e.g., *"There was not much added value compared to showing the exact same case on a regular screen, hovering the mouse over the lung sounds."* Students also reported to be supported by the 3D visuals in multiple ways, e.g., *"You really see it in front of you, and because of that I remember it better," "It helped me to see that the right middle lobe cannot be auscultated from the back."* Only a few students reported aspects that were not related to the first two themes: one student mentioned *"It helps to see it in a real person, and the possibility of repetition was helpful."* More quotations are displayed in **Table 3**.

Besides that the AR experience helped students to understand the case, they also indicated that their insight in the portrayed illness had changed *a little bit* (42.2%) to *substantially* (29.4%) (**Figure 5C**). Answers to the question "What changed in your insight in the portrayed illness?" were qualitatively analyzed. Again, the main themes were the 3D models (visualization of the diagnosis) and the auscultation of the lungs. Students learned for instance from the 3D visuals *"[...] what a collapsed lung looks like in 3D"* or *"That it can be a really small area with a subtle abnormality"* (pneumothorax case). Students also reported insights regarding the auscultation process, for instance *"Also listen apical to the lungs,"* after a learning session with a case of an apical pneumothorax which many students missed at first examination. More quotations are provided in the second half of **Table 3**.

More than half of the students reported that the application better prepared them for clinical practice than regular teaching approaches, such as e-learning modules, small group learning, physical examination of peer students, and practicing with simulated patients (**Figure 5D**). After the first use of this AR application students felt generally somewhat more confident about diagnosing lung cases.

With the questionnaire it was explored what students particularly liked about the AR experience. This resulted in different answers. One student answered: *"Novel way of learning. I love learning by interacting in a more "physical manner."* Another student liked *"To be able to walk around the case, hear all the different parts of the lung and discuss with peers."* Another illustrative quotation is: *"To see anatomy and pathology and how they interact. Also the fact that all observers see the same makes discussing more fruitful."*

44.2% of the studied population reported no physical discomfort at all, and 50% reported *a little* to *some* physical discomfort. Only four participants (3.8%) experienced *a lot* of physical discomfort. The most commonly reported complaints were headache, symptoms that could be attributed to the weight of the head-mounted display (such as *"heaviness on the head," "pressure on the nose," "pain on nose," "tight headset"*) and dizziness. The experience of different types of physical discomfort is displayed in **Figure 6**. While analyzing the general question "What did you dislike?" it was noted that 21.2% of the participants provided an answer related to physical discomfort, mainly due to the head-mounted display, e.g., "*The AR glasses were very heavy, had to lift them with my hand," "Annoying headset"* and *"The headset isn" t so comfortable."*

Students were asked how they would evaluate AR as a tool for learning, based on this specific experience. Most evaluated it is as an *excellent* (16.3%), *good* (58.7%) or *average* (20.2%) tool for learning (**Figure 5E**). Many students (39.4%) indicated that they would like to see more AR applications for medical education on the topic of anatomy. Other topics that were suggested include developmental biology, surgery, cardiology, interactive pharmacology and emergency care. Students were also asked "What did you dislike?" in order to identify areas of improvement. Approximately one-third of the responders to this

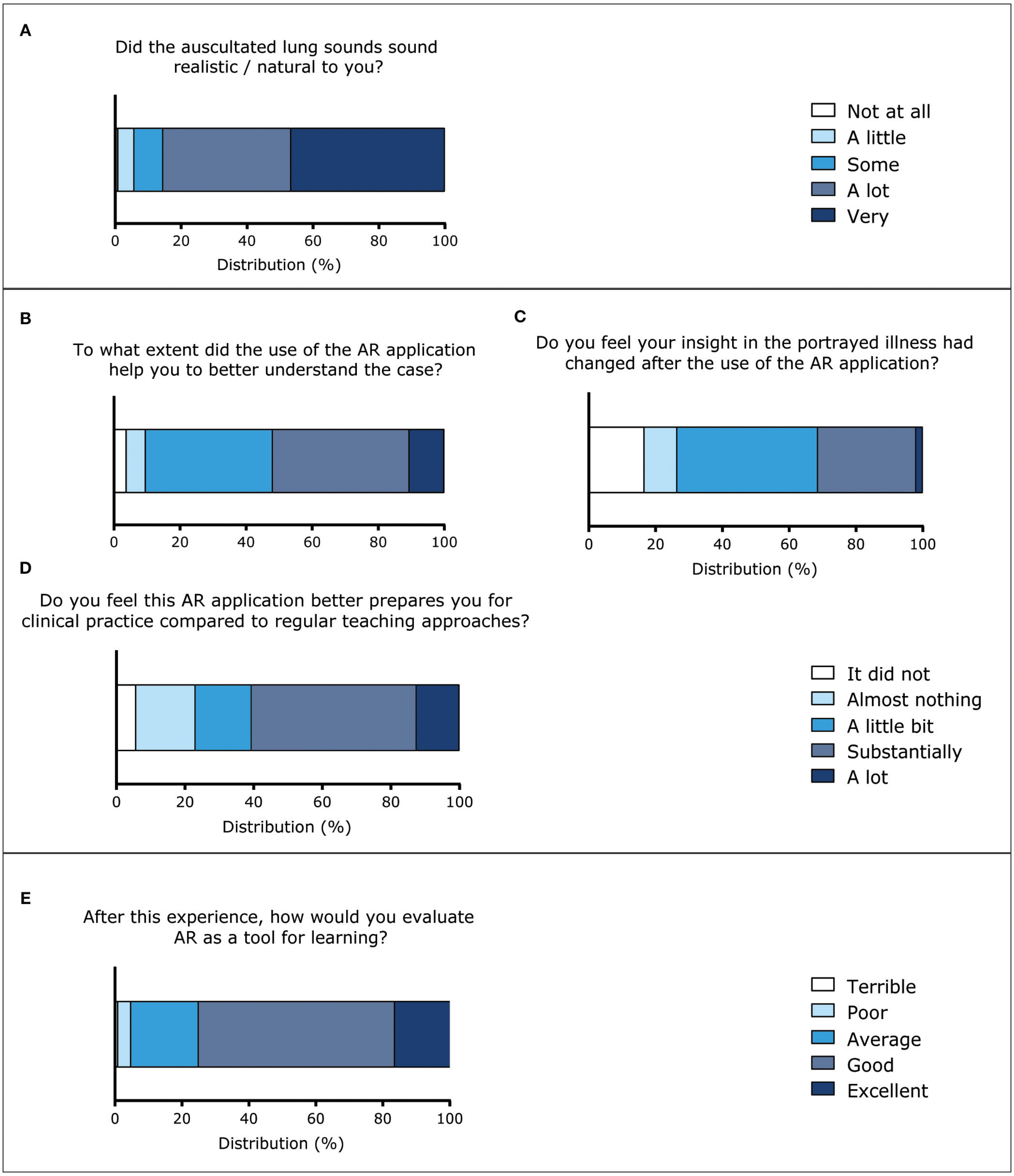

FIGURE 5 | Evaluation of the AR application and AR as a tool for learning (five-point Likert scales). **(A)** Responses to the question "Did the auscultated lung sounds sound realistic / natural to you?" ($n = 103$). **(B)** Responses to the question "To what extend did the use of the AR application help you to better understand the case?" ($n = 104$). **(C)** Responses to question "Do you feel your insight in the portrayed illness had changed after the use of the AR application?" ($n = 102$). **(D)** Responses to the question "Do you feel this AR application better prepares you for clinical practice compared to regular teaching methods?" ($n = 104$). **(E)** Responses to the question "After this experience, how would you evaluate AR as a tool for learning?" ($n = 104$).

TABLE 3 | Illustrative quotations of answers to the questions "Please, briefly explain how it helped you to better understand the case" (divided into three themes) and "What changed in your insight in the portrayed illness?" (divided into two themes).

Answers to the question "Please, briefly explain how it helped you to better understand the case"

Theme	Illustrative quotations
3D model	"You really see it [ed.: the model] in front of you, and because of that I remember it better"
	"A better view of what the problems are and where they are in a 3D body"
	"It helped me to see that the right middle lobe cannot be auscultated from the back"
	"You see what you are hearing. And this makes the understanding of where you should listen a lot better"
	"Made it visual; however, just the case + x-thorax would have been clear as well"
Auscultation of lungs	"It was an essential step in the physical exam, which otherwise would have been played as an audio track or described in words"
	"It helped to understand the difference in lung sounds and also which part of the lungs was affected"
	"It was more clear where to listen on the thorax"
	"Better discussions about lung sounds"
	"Not much added value compared to showing the exact same case on a regular screen, hovering the mouse over the lung sounds"
Other	"It helps to see it in a real person, and the possibility of repetition was helpful"
	"Became more realistic"

Answers to the question "What changed in your insight in the portrayed illness?"

Theme	Illustrative quotations
Visualization of the diagnosis	"The clear image of the infiltration (pneumonia)." (context: viewing a 3D transparent model of a pneumonia in two lung lobes)
	"I now know what a collapsed lung looks like in 3D"
	"The visual presentation [ed.: of the disease] makes it more memorable"
	"That there can also be only a partial pneumothorax."(context: clinical case with a subtle partial pneumothorax that was initially missed by the students)
	"How lungs look like in COPD." (context: viewing a 3D model of bullous emphysema caused by COPD)
Auscultation of lungs	"Also listen apical to the lungs." (context: after missing the diagnosis of an apical pneumothorax, the student learned to carefully auscultate this region)
	"Specifically the "pneumonia sounds" but that could also be done with a X-thorax." (context: we suppose that the student meant that the 3D models were not necessary for learning the lung sounds)
	"I could easily see which part of the lung I was listening to"
	"More clarity how/why/where to auscultate"

question (31.7%) mentioned technical difficulties or problems with the use of the AR application, e.g., *"Problems starting the AR"* and *"We had to reconnect sometimes."* A few students (4.8%) mentioned the limited field of view of the HoloLens® as a disadvantage. Remarks that provided input for future improvements included: *"[...] would" ve been nice if you could actually see the lungs expand with breathing," "sounds etc in the room were a little distractive," "Took quite a while with two groups, we had to wait quite a bit," "Small [learning] space."*

DISCUSSION

In this paper the design and implementation of the AR application, *AugMedicine: Lung cases*, on patient cases with dyspnea is described. Our application was developed in a time period of 22 months by a large team of stakeholders. A small group learning session was designed around the AR application, based on instructional design principles of Merrill (Merrill, 2002) and this session was successfully implemented in our clerkship curriculum. Until now various AR applications for medical education have been developed mainly in the field of surgery and anatomy (Zhu et al., 2014; Sheik-Ali et al., 2019; Gerup et al., 2020; Tang et al., 2020; Uruthiralingam and Rea, 2020). This study contributes to the exploration of the potential of AR in health care education beyond those fields.

Integration of AR technology into the medical curriculum should be carefully designed to avoid focus on the technology itself. In an integrative review it was proposed that appropriate learning theories should be identified in order to guide the integration of AR in medical education (Zhu et al., 2014). In line with this recommendation, the LED has been a prime focus in our design process resulting in successful integration in a small group learning session following the instructional design principles of Merrill (2002). These principles were not only applied for integrating this novel technology into learning sessions, but also in the concept of the AR application itself (first and fourth principle). There was a particular focus on creating a more authentic learning environment for diagnosing dyspnea rather than merely an e-learning or practicing with a simulated patient, which is in line with Merrill's first principle of instruction, engaging learners in relevant and real-world problems.

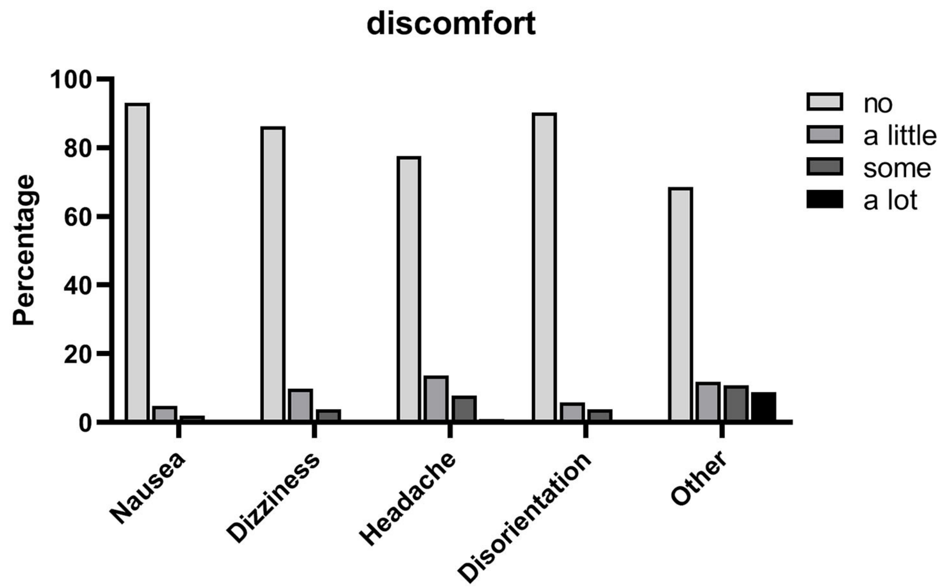

FIGURE 6 | Experience of different types of physical discomfort during the AR experience (four-point Likert scale), $n = 102$.

Kamphuis et al. also suggested that AR technology has the potential to create a situated and authentic learning experience that may better facilitate transfer of learning into the workplace (Kamphuis et al., 2014). In the past, other technologies have been described to increase authenticity while training chest auscultation. Examples are web-based auscultation (Ward and Wattier, 2011), mannequin-based simulators / auscultation torsos (Ward and Wattier, 2011; McKinney et al., 2013; Bernardi et al., 2019), multimedia presentation of acoustic and graphic characteristics of lung sounds (Sestini et al., 1995), "hybrid simulators," a combination of simulated patients with hardware and software of auscultation torsos (Friederichs et al., 2014) and auscultation sound playback devices or stethoscopes (Ward and Wattier, 2011). Although some of these tools improved learning of lung or heart sounds compared to traditional teaching methods, all of these examples have in common that they lack the combination with 3D visuals of the underlying diagnosis. This combination is a major advantage of *Augmedicine: Lung cases*, as students reported that the combination of auscultation and 3D visuals was helpful to them. The results suggest that most students recognize and confirm the authenticity of the learning setting, although some reported that it felt a bit unnatural. This effect may have been caused or aggravated by technical difficulties such as problems with manipulating the virtual stethoscope or drifting of the model. This could be further explored by performing focus groups or interviews to analyze the answers in more depth.

The results of our study indicate that the AR experience helped students to better understand a case of dyspnea and gain more insight into the portrayed illness. An explanation for these effects could be the stimulation of active learning by the application. In the past decades, medical education has shifted from traditional lecture-based teaching to more active and collaborative learning. In science education, active learning has proven to be more effective than lecture-based learning in the study of Freeman et al. (2014). It was defined in this study as: "*Active learning engages students in the process of learning through activities and/or discussion in class, as opposed to passively listening to an expert. It emphasizes higher-order thinking and often involves group work.*" The small group learning session with the AR application has multiple learning activities that relate to active learning: the students walk around a 3D hologram, auscultate the virtual lungs, and discuss the case with their peers and teacher. The better understanding of the case might also be explained by the effect of Embodied Cognition. The concept of Embodied Cognition is described as *"Cognition is embodied when it is deeply dependent upon features of the physical body of an agent, that is, when aspects of the agent's body beyond the brain play a significant causal or physically constitutive role in cognitive processing."* (Website Stanford Encyclopedia of Philosophy, 2020). While learning with the studied AR application, students use their physical body for examining the 3D models, manipulating the virtual stethoscope by a touch-free gestural interface in order to listen to the lung sounds. These are relevant motion patterns for diagnosing patients in the real-life environment and therefore could support students to construct a mental model of the learned concept. Although various clinical cases are available in the AR application, the students explored only one within the context of the formal learning session. This could explain why students generally did not feel substantially more confident about diagnosing lung cases after the AR experience. However, the other cases can be further explored independently at a later moment in time.

A possible disadvantage of introducing AR/VR technologies into education is the perception of simulator sickness or cybersickness, which is a form of motion sickness with adverse symptoms and observable signs that are caused by elements of visual display and visuo-vestibular interaction (Vovk et al., 2018). Although there is extensive literature on this phenomenon in virtual reality (e.g., Gavgani et al., 2017; Gallagher and Ferrè, 2018; Weech et al., 2019; Saredakis et al., 2020), publications in the context of AR are scarce, especially regarding AR with head-mounted displays. However, visuo-vestibular interactions differ considerably between VR and AR environments. In a study by Moro et al., participants were allocated to either use of an VR, AR or tablet-based application for learning skull anatomy (Moro et al., 2017). Adverse physical symptoms were compared, and were significantly higher in the VR group. Vovk et al. investigated simulator sickness symptoms in AR training with the HoloLens® with a validated 16-item simulator sickness questionnaire (SSQ), with three subscales relating to gastrointestinal symptoms, oculomotor symptoms, and disorientation. They concluded that interactive augmented interaction is less likely to cause simulator sickness than interaction with virtual reality. Although some of the major symptoms were encountered as well in our study, we did not explore this in a structured way. Some of the students reported symptoms, such as heaviness on the nose or pain on the nose, that may be caused by the weight of the HoloLens®Perhaps the recently released HoloLens® two could solve some of these physical discomfort issues as it is designed to be more comfortable by shifting most of the weight from the front to the back of the headset. However, this remains to be studied.

Limitations

A limitation of this study is that an isolated questionnaire was used, instead of a questionnaire in combination with e.g., a focus group or structured interviews. Sometimes students provided answers which could be interpreted in multiple ways. Focus groups or interviews with students could have helped, but were not part of this research. Possible confounders in this study were variable learning spaces, variable occurrence of technical incidents, variable teachers (a total of four teachers) and use of two different cases (*pneumothorax* and *COPD plus pneumonia*). We cannot definitively know if the other clinical cases in the application would result in comparable user experiences. Educational spaces that provided the HoloLens®'s positional tracking sensors with more environmental landmarks seemed to result in less technical problems, although we did not formally test this.

Practical Implications

Institutions that consider to use AR in their medical curriculum, should be aware that besides the costs for developing and maintaining an AR application and the purchase of hardware devices, additional resources should be allocated. In our experience, this type of technology requires a dedicated infrastructure, including training and allocating technical support (for maintenance and repair of the head-mounted displays and software updates), arranging a group of teaching assistants for supporting teachers, and training of several dedicated teachers. For the described AR application, the applied group size for the learning session was a maximum of 12 students (of which six simultaneously wearing a head-mounted display). It is not (yet) possible to increase the group size, because of limited availability of head-mounted displays. The scalability of comparable AR applications is also likely to be limited, in contrast to VR applications for which in our experience the group size can more easily be increased.

Future Directions

The applicability and extension of this AR application for other healthcare professionals, such as residents, medical specialists, general practitioners, physician assistants and paramedics, in currently under consideration. Although now applied as a teacher-led small group learning session, it is planned to extend the application to facilitate asynchronous learning. By adapting the AR application to a standalone version it could be used without teacher presence, offering students the option to practice as frequently as needed and experience more cases as part of a more individualized learning process. In terms of researching AR in medical education, a recent systematic review proposed an analytical model to evaluate the potential for an AR application to be integrated into the medical curriculum, in order to result in both higher quality study designs and formal validity assessments (Tang et al., 2020). The model consists of four components: quality, application content, outcome, and feasibility. Although the current study provided relevant insights on the application content and feasibility, the study design is observational and outcomes in terms of efficacy remain to be studied. In alignment with the model, a next step to provide higher grade evidence for the implementation of this AR application could be performing a randomized-controlled trial to compare the effectiveness of this AR experience to traditional learning methods alone.

CONCLUSIONS

Until now various AR applications for medical education have been developed mainly in the field of surgery and anatomy. This study contributes to the exploration of the potential use of AR in health care education beyond those fields. We integrated *AugMedicine: Lung cases* into the curriculum through a specially designed small group learning session. In general, students reported that the AR experience helped them to better understand the clinical case and most students value AR as a good tool for learning. The user evaluation was predominantly positive, although physical discomfort seemed to be an issue for some students.

DATA AVAILABILITY STATEMENT

A limited dataset (without qualitative data) is available as **Supplementary Material**. The complete dataset is available on request. Requests should be directed to: Arianne D. Pieterse, a.d.pieterse@lumc.nl.

ETHICS STATEMENT

The studies involving human participants were reviewed and approved by the Educational Research Review Board of the Leiden University Medical Center. The patients/participants provided their written informed consent to participate in this study. Written informed consent was obtained from the individual(s) for the publication of any potentially identifiable images or data included in this article.

AUTHOR CONTRIBUTIONS

AP, BH, JK, and MR were part of the team that designed and developed the AR application. AP, BH, PJ, and MR designed the study. AP and BH analyzed the data. AP wrote the first draft of the manuscript. BH, PJ, JK, and MR provided further first-hand writing and edited the manuscript. JK and LW contributed in reviewing the manuscript. All authors have approved the final version of the manuscript and agree to be accountable for all aspects of the work.

ACKNOWLEDGMENTS

We thank all members of the *AugMedicine: Lung cases* team for their contribution to the design process. We thank Argho Ray, Leontine Bakker, and Soufian Meziyerh for teaching small group learning sessions with the AR application. We thank Franka Luk and Manon Zuurmond for their contributions to this study and paper.

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/frvir.2020.577534/full#supplementary-material

REFERENCES

Barmaki, R., Yu, K., Pearlman, R., Shingles, R., Bork, F., Osgood, G. M., et al. (2019). Enhancement of anatomical education using augmented reality: an empirical study of body painting. *Anat. Sci. Educ.* 12, 599–609. doi: 10.1002/ase.1858

Bernardi, S., Giudici, F., Leone, M. F., Zuolo, G., Furlotti, S., Carretta, R., et al. (2019). A prospective study on the efficacy of patient simulation in heart and lung auscultation. *BMC Med. Educ.* 19:275. doi: 10.1186/s12909-019-1708-6

Bogomolova, K., Hierck, B. P., Looijen, A. E. M., et al. (2020). Stereoscopic three-dimensional visualization technology in anatomy learning: a meta-analysis. *Med. Educ.* doi: 10.1111/medu.14352. [Epub ahead of print].

Bogomolova, K., van der Ham, I. J. M., Dankbaar, M. E. W., van den Broek, W. W., Hovius, S. E. R., van der Hage, J. A., et al. (2019). The effect of stereoscopic augmented reality visualization on learning anatomy and the modifying effect of visual-spatial abilities: a double-center randomized controlled trial. *Anat. Sci. Educ.* 13, 558–567. doi: 10.1002/ase.1941

Cleland, J. A., Abe, K., and Rethans, J. J. (2009). The use of simulated patients in medical education: AMEE guide No 42. *Med. Teach.* 31, 477–486. doi: 10.1080/01421590903002821

Demonstration of AugMedicine: Lung Cases (2020). Available online at: https://youtu.be/GGiigVMKPTA (accessed September 14, 2020).

Eckert, M., Volmerg, J. S., and Friedrich, C. M. (2019). Augmented reality in medicine: systematic and bibliographic review. *JMIR Mhealth Uhealth* 7:e10967. doi: 10.2196/10967

Edelbring, S., Dastmalchi, M., Hult, H., Lundberg, I. E., and Dahlgren, L. O. (2011). Experiencing virtual patients in clinical learning: a phenomenological study. *Adv. Health Sci. Educ. Theory Pract.* 16, 331–345. doi: 10.1007/s10459-010-9265-0

Freeman, S., Eddy, S. L., McDonough, M., Smith, M. K., Okoroafor, N., Jordt, H., et al. (2014). Active learning increases student performance in science, engineering, and mathematics. *Proc. Natl. Acad. Sci. U.S.A.* 111, 8410–8415. doi: 10.1073/pnas.1319030111

Friederichs, H., Weissenstein, A., Ligges, S., Moller, D., Becker, J. C., and Marschall, B. (2014). Combining simulated patients and simulators: pilot study of hybrid simulation in teaching cardiac auscultation. *Adv. Physiol. Educ.* 38, 343–347. doi: 10.1152/advan.00039.2013

Gallagher, M., and Ferrè, E. (2018). Cybersickness: a multisensory integration perspective. *Multisensory Res.* 31, 645–674. doi: 10.1163/22134808-20181293

Gavgani, A. M., Nesbitt, K. V., Blackmore, K. L., and Nalivaiko, E. (2017). Profiling subjective symptoms and autonomic changes associated with cybersickness. *Auton. Neurosci.* 203, 41–50. doi: 10.1016/j.autneu.2016.12.004

Gerup, J., Soerensen, C. B., and Dieckmann, P. (2020). Augmented reality and mixed reality for healthcare education beyond surgery: an integrative review. *Int. J. Med. Educ.* 11, 1–18. doi: 10.5116/ijme.5e01.eb1a

HoloLens App AugMedicine: Lung Cases (short) (2020). Available online at: https://www.youtube.com/watch?v=feiOz7UAxbs (accessed June 25, 2020).

Kamphuis, C., Barsom, E., Schijven, M., and Christoph, N. (2014). Augmented reality in medical education? *Perspect. Med. Educ.* 3, 300–311. doi: 10.1007/s40037-013-0107-7

Kaufmann, H., and Schmalstieg, D. (2003). Mathematics and geometry education with collaborative augmented reality. *Comput. Graph.* 27, 339–345. doi: 10.1016/S0097-8493(03)00028-1

Kotranza, A., Lind, D. S., and Lok, B. (2012). Real-time evaluation and visualization of learner performance in a mixed-reality environment for clinical breast examination. *IEEE Trans. Vis. Comput. Graph.* 18, 1101–1114. doi: 10.1109/TVCG.2011.132

Lamounier, E., Bucioli, A., Cardoso, A., Andrade, A., and Soares, A. (2010). On the use of augmented reality techniques in learning and interpretation of cardiologic data. *Conf. Proc. IEEE Eng. Med. Biol. Soc.* 2010, 610–613. doi: 10.1109/IEMBS.2010.5628019

McKinney, J., Cook, D. A., Wood, D., and Hatala, R. (2013). Simulation-based training for cardiac auscultation skills: systematic review and meta-analysis. *J. Gen. Intern. Med.* 28, 283–291. doi: 10.1007/s11606-012-2198-y

Merrill, D. (2002). First principles of instruction. *ETRandD. 50,* 43–59 doi: 10.1007/BF02505024

Miles, M. B., Huberman, A. M., and Saldana, J. (2014). *Qualitative Data Analysis: A Methods Sourcebook, 3rd edn*. Thousands Oaks, CA: SAGE Publications, Inc.

Moro, C., Stromberga, Z., Raikos, A., and Stirling, A. (2017). The effectiveness of virtual and augmented reality in health sciences and medical anatomy. *Anat. Sci. Educ.* 10, 549–559. doi: 10.1002/ase.1696

Nilsson, S., and Johansson, B. (2008). "Acceptance of augmented reality instructions in a real work setting," in *Conference on Human Factors in Computing Systems, C. H. I* (Florence). doi: 10.1145/1358628.1358633

Nischelwitzer, A., Lenz, F., Searle, G., and Holzinger, A. (2007). "Some aspects of the development of low-cost augmented reality learning environments as examples for future interfaces in technology enhanced learning," in *Universal Access in HCI, Part III, HCII 2007, LNCS 4556*, ed C. Stephanidis (Berlin; Heidelberg), 728–737. doi: 10.1007/978-3-540-73283-9_79

Pretto, F., Manssour, I., Lopes, M., da Silva, E., and Pinho, M. (2009). "Augmented reality environment for life support training," in *SAC '09: Proceedings of the 2009 ACM symposium on Applied Computing*, eds S. Y. Shin, and S. Ossowski (Honolulu, HI: Association for Computing Machinery), 836–841. doi: 10.1145/1529282.1529460

Rasimah, C., Ahmad, A., and Zaman, H. (2011). Evaluation of user acceptance of mixed reality technology. *AJET 27,* 1369–1387. doi: 10.14742/ajet.899

Rosenbaum, E., Klopfer, E., and Perry, J. (2007). On location learning: authentic applied science with networked augmented realities. *J. Sci. Educ. Technol.* 16, 31–45 doi: 10.1007/s10956-006-9036-0

Sakellariou, S., Ward, B., Charissis, V., Chanock, D., and Anderson, P. (2009). "Design and implementation of augmented reality environment for complex anatomy training: inguinal canal case study", in *Virtual and Mixed Reality, Third International Conference, VMR 2009*, ed R. Shumaker (San Diego, CA), 605–614. doi: 10.1007/978-3-642-02771-0_67

Saredakis, D., Szpak, A., Birckhead, B., Keage, H. A. D., Rizzo, A., and Loetscher, T. (2020). Factors associated with virtual reality sickness in head-mounted displays: a systematic review and meta-analysis. *Front. Hum. Neurosci.* 14:96. doi: 10.3389/fnhum.2020.00096

Schwaber, K., and Beedle, M. (2008). *Agile Software Development With Scrum*. London: Pearson Education (Us).

Sestini, P., Renzoni, E., Rossi, M., Beltrami, V., and Vagliasindi, M. (1995). Multimedia presentation of lung sounds as a learning aid for medical students. *Eur. Respir. J.* 8, 783–788.

Sheik-Ali, S., Edgcombe, H., and Paton, C. (2019). Next-generation virtual and augmented reality in surgical education: a narrative review. *Surg. Technol. Int.* 35, 27–35.

Tang, K. S., Cheng, D. L., Mi, E., and Greenberg, P. B. (2020). Augmented reality in medical education: a systematic review. *Can. Med. Educ. J.* 11, e81–e96. doi: 10.36834/cmej.61705

Uruthiralingam, U., and Rea, P. M. (2020). Augmented and virtual reality in anatomical education - a systematic review. *Adv. Exp. Med. Biol.* 1235, 89–101. doi: 10.1007/978-3-030-37639-0_5

von Jan, U., Noll, C., Behrends, M., and Albrecht, U. (2012). mARble – augmented reality in medical education. *Biomed. Tech.* 57(Suppl. 1), 67–70. doi: 10.1515/bmt-2012-4252

Vovk, A., Wild, F., Guest, W., and Kuula, T. (2018). "Simulator sickness in augmented reality training using the microsoft hololens", in *Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems (CHI '18)*, eds R. Mandryk and M. Hancock (New York, NY; Montréal, QC), 1–9. doi: 10.1145/3173574.3173783

Ward, J. J., and Wattier, B. A. (2011). Technology for enhancing chest auscultation in clinical simulation. *Respir. Care* 56, 834–845. doi: 10.4187/respcare.01072

Website Stanford Encyclopedia of Philosophy (2020). Available online at: https://plato.stanford.edu/entries/embodied-cognition/ (accessed June 25, 2020).

Weech, S., Kenny, S., and Barnett-Cowan, M. (2019). Presence and cybersickness in virtual reality are negatively related: a review. *Front. Psychol.* 10:158. doi: 10.3389/fpsyg.2019.00158

Zhu, E., Hadadgar, A., Masiello, I., and Zary, N. (2014). Augmented reality in healthcare education: an integrative review. *PeerJ* 2:e469. doi: 10.7717/peerj.469

Conflict of Interest: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

frontiers in Digital Health

ORIGINAL RESEARCH
published: 18 February 2021
doi: 10.3389/fdgth.2020.613608

Check for updates

Adaptive Physics-Based Non-Rigid Registration for Immersive Image-Guided Neuronavigation Systems

*Fotis Drakopoulos [1], Christos Tsolakis [1,2], Angelos Angelopoulos [1,2], Yixun Liu [1], Chengjun Yao [3], Kyriaki Rafailia Kavazidi [4], Nikolaos Foroglou [4], Andrey Fedorov [5], Sarah Frisken [5], Ron Kikinis [5], Alexandra Golby [5,6] and Nikos Chrisochoides [1,2]**

[1] Center for Real-Time Computing, Old Dominion University, Norfolk, VA, United States, [2] Department of Computer Science, Old Dominion University, Norfolk, VA, United States, [3] Department of Neurosurgery, Huashan Hospital, Shanghai, China, [4] Department of Neurosurgery, Aristotle University of Thessaloniki, Thessaloniki, Greece, [5] Department of Radiology, Brigham and Women's Hospital and Harvard Medical School, Boston, MA, United States, [6] Department of Neurosurgery, Brigham and Women's Hospital and Harvard Medical School, Boston, MA, United States

OPEN ACCESS

Edited by:
Alkinoos Athanasiou, University General Hospital of Thessaloniki AHEPA, Greece

Reviewed by:
Alexander Astaras, The American College of Thessaloniki, Greece
Kostas Nizamis, University of Twente, Netherlands

***Correspondence:**
Nikos Chrisochoides
npchris@gmail.com

Specialty section:
This article was submitted to Connected Health, a section of the journal Frontiers in Digital Health

Received: *02 October 2020*
Accepted: *23 December 2020*
Published: *18 February 2021*

Citation:
Drakopoulos F, Tsolakis C, Angelopoulos A, Liu Y, Yao C, Kavazidi KR, Foroglou N, Fedorov A, Frisken S, Kikinis R, Golby A and Chrisochoides N (2021) Adaptive Physics-Based Non-Rigid Registration for Immersive Image-Guided Neuronavigation Systems. Front. Digit. Health 2:613608. doi: 10.3389/fdgth.2020.613608

Objective: In image-guided neurosurgery, co-registered preoperative anatomical, functional, and diffusion tensor imaging can be used to facilitate a safe resection of brain tumors in eloquent areas of the brain. However, the brain deforms during surgery, particularly in the presence of tumor resection. Non-Rigid Registration (NRR) of the preoperative image data can be used to create a registered image that captures the deformation in the intraoperative image while maintaining the quality of the preoperative image. Using clinical data, this paper reports the results of a comparison of the accuracy and performance among several non-rigid registration methods for handling brain deformation. A new adaptive method that automatically removes mesh elements in the area of the resected tumor, thereby handling deformation in the presence of resection is presented. To improve the user experience, we also present a new way of using mixed reality with ultrasound, MRI, and CT.

Materials and methods: This study focuses on 30 glioma surgeries performed at two different hospitals, many of which involved the resection of significant tumor volumes. An Adaptive Physics-Based Non-Rigid Registration method (A-PBNRR) registers preoperative and intraoperative MRI for each patient. The results are compared with three other readily available registration methods: a rigid registration implemented in 3D Slicer v4.4.0; a B-Spline non-rigid registration implemented in 3D Slicer v4.4.0; and PBNRR implemented in ITKv4.7.0, upon which A-PBNRR was based. Three measures were employed to facilitate a comprehensive evaluation of the registration accuracy: (i) visual assessment, (ii) a Hausdorff Distance-based metric, and (iii) a landmark-based approach using anatomical points identified by a neurosurgeon.

Results: The A-PBNRR using multi-tissue mesh adaptation improved the accuracy of deformable registration by more than five times compared to rigid and traditional physics based non-rigid registration, and four times compared to B-Spline interpolation methods which are part of ITK and 3D Slicer. Performance analysis showed that A-PBNRR could be applied, on average, in <2 min, achieving desirable speed for use in a clinical setting.

Conclusions: The A-PBNRR method performed significantly better than other readily available registration methods at modeling deformation in the presence of resection. Both the registration accuracy and performance proved sufficient to be of clinical value in the operating room. A-PBNRR, coupled with the mixed reality system, presents a powerful and affordable solution compared to current neuronavigation systems.

Keywords: medical image computing, deformable registration, mesh generation, neurosurgery, machine learning, deep learning, mixed reality, neuronavigation systems

INTRODUCTION

Malignant gliomas are the most common primary and metastatic brain tumors, accounting for ∼70% of the 22,500 new cases of primary brain tumors diagnosed annually in adults in the United States (1, 2). Treatment typically includes surgical removal followed by radiotherapy or chemotherapy. Tumor removal provides a tissue diagnosis, relieves mass effect, and intracranial pressure that may be causing pain or other neurological symptoms, and improves prognosis. However, gross total resection is difficult to achieve because of the infiltrative nature of gliomas and because brain tumors are often embedded in critical functional brain tissue. Oncologic outcomes clearly depend on the extent of tumor resection, yet functional preservation is critical for quality of life and survival, so tumor surgery is a delicate balance between removing as much tumor as possible and preserving important functional areas of the brain (3–5).

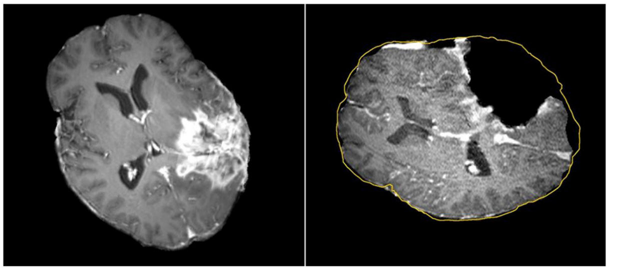

FIGURE 1 | Discrepancies between preoperative and intraoperative MR Imaging before and during neurosurgery. Left: preoperative MRI; Right: intraoperative MRI acquired after a part of the tumor is removed. The yellow outline indicates the preoperative brain outline after a rigid rotation. The large dark cavity is the tumor resection.

During the past two decades, developments in image-guided therapy (6) have allowed surgeons to use preoperative imaging and neuronavigation to facilitate a maximally safe resection of gliomas in eloquent areas of the brain. Preoperative anatomical Magnetic Resonance Imaging (MRI) can be combined with functional MRI (fMRI) to map out areas of the brain near the tumor that are involved with important function such as vision, speech and language, or motor control (7–12). Diffusion Tensor Imaging (DTI) can be used to map out white matter tracts that connect to these important regions and run near or through the tumor (13–19).

Tracking the position of medical tools in patient's brain during surgery is possible with neuronavigation using registration of preoperative image data to patient coordinates. The surgeon can then view the location of tools relative to the preoperative anatomical and functional image data, thereby avoiding damage to eloquent areas during tumor resection (20–28). Commercial neuronavigation systems (e.g., Stealth by Medtronic and VectorVision by BrainLAB) generally use a rigid transformation to map preoperative image data to patient coordinates. Rigid registration is sufficient when mapping between rigid objects (e.g., between the skull in the preoperative image data and the patient's skull in the operating room). During surgery, however, the brain deforms due to several factors such as cerebrospinal fluid leakage, intra-cranial pressure, gravity, the administration of osmotic diuretics, and the procedure itself (e.g., tumor retraction and resection) (17, 29, 30). A rigid transformation will not accurately map preoperative image data to the patient's brain during surgery, particularly as the resection proceeds, with the greatest uncertainty at the most critical portions of the surgery (**Figure 1**).

The adoption of intraoperative MRI (iMRI) has provided a means for monitoring brain deformation during surgery (31). The number of hospitals offering iMRI has grown during the past decade from a handful of research centers to a couple of hundreds clinical sites across the world (32). Currently, access to iMRI is limited by high costs, personnel requirements, and disruption of the operative workflow. Several researchers are investigating methods that combine other imaging modalities, such as 3D ultrasound (33–36) and surface imaging combined with deformation modeling (37–39), to compensate for brain deformation. However, iMRI remains the gold standard for measuring intraoperative brain deformation and for monitoring tumor resection.

Given an intraoperative anatomical MRI image registered to patient coordinates, preoperative fMRI and DTI image data can be mapped to the intraoperative image and then to patient coordinates to provide updated guidance for the surgeon. This mapping is usually performed by first registering the fMRI and DTI images to the preoperative anatomical MRI using a rigid transformation, and then registering the preoperative anatomical MRI to the intraoperative MRI image, which is pre-registered to patient coordinates. Because of brain deformation during surgery, registering from preoperative to intraoperative MRI requires a non-rigid registration. DTI and fMRI are deformed the same way as with PBNRR (20).

There are several approaches for estimating the non-rigid registration between two or more images, as outlined in prior research (40–44). Ranging from control-point registration with spline interpolation to mass-spring models using displacements of anatomical landmarks as force vectors, to physics-based finite element models, many techniques have been applied to non-rigid registration between brain data sets, both for brain mapping (45–47) and for modeling brain shift (20, 21, 24, 25, 37–39, 48–54). However, most of these methods were not designed to model tissue retraction or resection. While Physics-Based Non-Rigid Registration (PBNRR) has been shown to accurately capture brain shift (i.e., volumetric deformations of the brain) during image-guided neurosurgery (21), it fails to accurately fuse iMRI with pre-operative MRI for cases with tumor resection. In this paper, we evaluate the accuracy and performance of the Adaptive Physics-Based Non-Rigid Registration (A-PBNRR) method for modeling brain deformation in the presence of tumor resection. The results of a study on 30 glioma cases from two different hospitals are presented, many of which involved a meaningful tumor volume resection, defined as an Extent Of Resection (EOR) $\geq$70% (55). The majority of the glioma resections also included an EOR $\geq$78–80%, which is considered a significant predictor of Overall Survival (OS) (56, 57). The results of the A-PBNRR are qualitatively and quantitatively compared with the results of three other open-source registration methods that are readily available to researchers and clinicians (58, 59). Acceptance in clinical practice requires that non-rigid registration be completed in the time constraints imposed by neurosurgery (e.g., 2–3 min) and without the cost of high-performance computing clusters (20). Thus, the processing speeds are compared on a readily available 12-core desktop system.

Immersion

The use of NRR within immersive environments together with supporting technologies such as machine learning for reducing the parameter search space can potentially offer a more affordable alternative to highly expensive commercial neuronavigation systems. In the context of immersive environments, an application can be created to enhance the user experience with ultrasound. Ultrasound is a widely available, easy-to-use, and less expensive imaging device than MRI. A main benefit is that it does not use any ionizing radiation and is safe. Moreover, ultrasound can provide real-time imaging, making it a good tool for guidance in minimally invasive procedures. An inherent deficiency of ultrasound, however, is the separation between the image and the transducer probe. The probe is always close to the patient, but the image is shown on a screen several meters away from the patient. Surgeons have to figure out the position of the image relative to the patient anatomy. To deal with this deficiency, HoloLens can be introduced to display the image on the plane defined by the probe. Doing so enhances ultrasound to mixed reality ultrasound. This plane can also be used as a cutting plane to show a pre-operative image slice such as CT or MRI. Users can switch between the preoperative image slice and the ultrasound image to see what happens on this plane before and after surgery.

The HoloLens is an optical see through (OST) device that uses holographic technology to blend the physical world with a virtual space and allow people to interact with virtual holograms. Because HoloLens is an OST device, users can see the real world directly through the display. The framework enabling the OST Image-Guided Neurosurgery System (IGNS) includes two parts: a part similar to traditional IGNS with a data server, and a HoloLens with a data receiver. These parts are shown in **Figure 2**. The data server is used to collect intraoperative data including the intraoperative image and NDI tracking data. These are then sent to the HoloLens *via* a Wi-Fi connection. The second part includes a HoloLens data receiver that receives data from the data server and then relays it to the HoloLens for rendering. The rendering is used to generate left and right eye images of the virtual object, which can be an ultrasound (US) image or DTI fibers etc. The parallax between these two images will be used by the human brain to produce a stereo vision. To correctly render these two images the virtual object must be correctly aligned with the real world, which is usually performed by a spatial registration technique.

Spatial registration is a technique to find the transform between different coordinate spaces. There are multiple coordinate spaces in the operating room (OR) as shown in **Figure 2**. The patient is defined in the coordinate space of the NDI reference tool. During a routine image-guided surgery, a registration procedure is performed to register patient space with the image space such as the pre-operative MRI space. To track the spatial position of a US image an NDI tacking tool is fixed on the NDI probe and a calibration procedure is performed beforehand. Through the NDI tracker, the preoperative MRI scan, intra-operative US scan and the surgical tool can be placed into one reference space. The virtual objects defined in the reference space need to be converted into the HoloLens space to be rendered. A HD color camera equipped with the HoloLens can be used to do the conversion. Since the NDI optical tracker tracks objects based on reflected infrared rays and the HoloLens HD camera tracks objects using computer vision, we would need a device that can track objects using both reflection of infrared rays and computer vision. A simpler way is to use spheres of the NDI reference tool because these spheres can be easily detected using blob detection. Since the spheres can be tracked by both the NDI tracker and the HoloLens camera, the spatial transform between the reference space and the HoloLens space can be found. After the virtual objects are transformed into the HoloLens space, rendering can be performed to get stereo vision.

MATERIALS AND METHODS

Patient Population and Imaging Protocols

This study includes 30 patients from two hospitals:

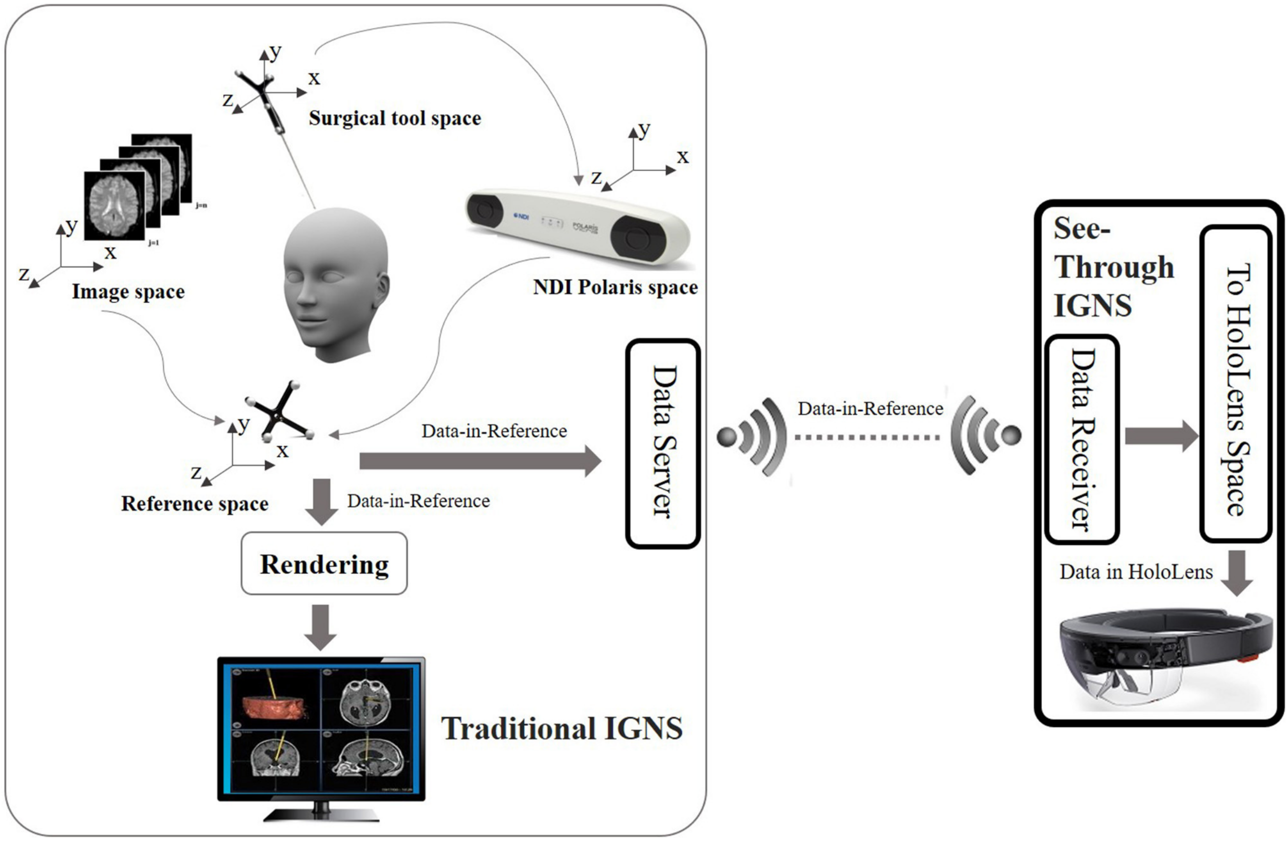

FIGURE 2 | Overview of the OST IGNS framework. The left part is similar to traditional IGNS with the inclusion of a data server. The right part is the OST portion which includes a data receiver, a spatial registration method for conversion of objects to HoloLens space, and HoloLens rendering. The data server and data receiver communicate wirelessly.

Brigham and Women's Hospital

Ten patients (six male, four female) with an age range of 28–62 years and a mean age of 45.2 years, underwent surgery for supratentorial gliomas between April 2005 and January 2006 in Brigham and Women's Magnetic Resonance Therapy (MRT) facility, which was typically used to resect intrinsic brain tumors. In some cases, the lesions were in or adjacent to eloquent brain areas, including the precentral gyrus and corticospinal tract for motor function, as well as Broca's and Wernicke's areas for language function. In general, these were patients undergoing surgery for brain tumors in the intraoperative MRI. Inclusion criteria includes the presence of an intracranial tumor. Exclusion criteria includes contraindication to MRI.

- **Preoperative imaging.** The patients underwent the following imaging protocol on a General Electric (Milwaukee, WI) 3T Signa scanner:
  a. Whole brain sagittal 3D spoiled-gradient-echo (SPGR) imaging (slice thickness, 1.3 mm; time of repetition (TR), 6 ms; time of echo (TE), 35 ms; flip angle (FA), 75°; field of view (FOV), 24 cm; matrix size, 256 × 256).
  b. Axial T2-weighted fast-spin-echo (FSE) imaging (slice thickness, 5 mm; TE, 100 ms; TR, 3,000 ms; FOV, 22 cm; matrix size, 512 × 512).
- **Intraoperative imaging.** The patients underwent the following imaging protocol on a 0.5T iMRI unit (SignaSP; GE Medical Systems, Milwaukee, WI):
  a. Transverse, sagittal, and coronal T1-weighted FSE imaging (TR, 700 ms; TE, 29 ms; FOV, 22 cm; matrix size, 256 × 256; section thickness, 3 mm, intersection gap, 1 mm).
  b. Transverse T2-weighted FSE imaging (TR, 5,000 ms; TE, 99 ms; FOV, 22 cm; matrix size, 256 × 256; section thickness, 3 mm; intersection gap, 1 mm).
  c. Transverse 3D SPGR imaging (TR, 15.5 ms; TE, 5.2 ms; FA, 45°; FOV, 22 cm; matrix size, 256 × 256; section thickness, 2.5 mm; intersection gap, 0 mm).

Huashan Hospital

Twenty patients (eleven male, nine female) with an age range of 19–75 years underwent surgery on single, unilateral, and supratentorial primary gliomas from September 2010 to August 2013. The lesions involved the Pyramidal Tracts (PTs) or were in cortical regions in the motor or somatosensory areas, cortical regions adjacent to the central gyrus, subcortical regions with an infiltrative progression along the PTs, and/or deep temporal or insular regions in relation to the internal capsule.

- **Preoperative imaging.** Preoperative brain images were obtained in the diagnostic room of an iMRI-integrated neurosurgical suite (IMRIS[1], Winnipeg, Manitoba, Canada) using a ceiling-mounted movable 3.0 T MAGNETOM Verio scanner (Siemens AG, Erlangen, Germany) with a 70 cm working aperture:

    a. For suspected high-grade gliomas, which showed obvious enhancement after contrast, contrast magnetization-prepared rapid gradient echo (MP-RAGE) was used as the anatomic base for the anisotropic color map and the tumor anatomic feature analysis. T1 contrast images were acquired with a 3D MP-RAGE sequence (TR, 1,900 ms; TE, 2.93 ms; inversion time, 900 ms; FA, 9°, FOV, 250 × 250 mm^2; matrix size, 256 × 256), after intravenous contrast administration (gadolinium diethylenetriamine penta-acetic acid).
    b. For suspected low-grade gliomas, which showed no obvious enhancement after contrast, a fluid-attenuated inversion-recovery (FLAIR) sequence was used as the anatomic base. T1 contrast images were acquired with FLAIR sequence, with an axial turbo spin echo pulse sequence (TR, 7,600 ms; TE, 96 ms; inversion time, 900 ms; FA, 9°; slices, 60; slice thickness, 2 mm; matrix size, 256 × 180; field of view, 240 × 240 mm^2).

- **Intraoperative imaging.** The same scanner as with preoperative MRI was used:

    a. T1 contrast images were acquired with a 3D MP-RAGE sequence (TR, 1,900 ms; TE, 2.93 ms; FA, 9°, FOV, 250 × 250 mm^2; matrix size, 256 × 215; slice thickness, 1 mm).
    b. T1 contrast images were acquired with FLAIR sequence (TR, 9,000 ms; TE, 96 ms; FA, 150°; slice thickness, 2 mm; FOV, 250 ×2 50 mm^2; matrix size, 256 × 160).

A neurosurgeon estimated the volume of resected tumor for each patient by performing a volumetric analysis on the preoperative and intraoperative MRI. Based on this volumetric analysis, the data were categorized as: (i) brain shift (with no resection), (ii) partial resection, (iii) total resection, and (iv) supra total resection. **Table 1** summarizes the clinical data. The data collections were carried out with Institutional Review Board (IRB) approval from both hospitals. The protocol is detailed in the paper by Yao et al. (60) and Archip et al. (20), the first times the data are used.

Segmentation

Non-rigid registration is performed using a patient-specific brain model derived by segmenting the preoperative anatomical image into brain, tumor, and non-brain regions (20). Segmentation performance is not critical because preoperative imaging is typically performed a couple of days before surgery. Segmentation was performed with a combination of manual and automatic tools. First, the skull and outer tissues were removed using the open-source Brain Extraction Tool (BET) (47). Further segmentation of the brain surface was performed using a combination of automatic operators implemented in 3D Slicer software (i.e., region growing and level-set filters) (61) and a slice-by-slice manual segmentation correction. An evaluation on how segmentation accuracy affects registration accuracy is beyond the scope of this paper but will be included in future work.

Mesh Generation

The segmentation is used to generate a patient-specific finite element mesh for physics-based non-rigid registration methods. The shape of the elements is critical for the accuracy and the convergence of a finite element solution. For example, elements with large dihedral angles tend to increase the discretization error in the solution (62). On the other hand, elements with small dihedral angles are bad for matrix conditioning but not for interpolation or discretization (63, 64).

A parallel Delaunay meshing method is employed to tessellate the segmented brain with high quality tetrahedral elements and to model the brain surface with geometric and topological guarantees (65). Both single-tissue (i.e., brain parenchyma) and multi-tissue (i.e., brain parenchyma and tumor) meshes are generated. **Figure 3** depicts one of the multi-tissue meshes. Parameter δ (**Table 3**) determines the size of the mesh, where a smaller $\delta > 0$ generates a larger mesh.

Rigid Registration

For the purpose of this study, patients first underwent an intraoperative scan after their head was positioned and fixed for the craniotomy but before the skull was opened. As a standard procedure iMRI is performed at the neurosurgeon's request, after dural opening, during or after a significant tumor volume resection or when decided appropriate by the surgeon (66, 67). Assuming minimal brain shift at this point, an initial rigid registration was performed to estimate a rigid transformation from the preoperative to intraoperative image data. This rigid transformation was used to initialize non-rigid registration methods.

The rigid registration was performed using the BRAINSFit module integrated in 3D Slicer v4.4.0 (58). BRAINSFit is a general registration module widely used by the research community. BRAINSFit's rigid registration relies on histogram bins and spatial samples to estimate a Mattes Mutual Information cost metric (**Table 2**). The larger the number of spatial samples, the slower and more precise the fit. The default values for the number of histogram levels and sampling percentage is 50 and 0.2%, respectively. Hundred histogram levels and a 5% sampling percentage were selected to achieve higher accuracy (**Table 2**). The default values were used for rest of the parameters (optimizer type, max number of iterations, min step length, and grid size) to

[1] https://www.imris.com/

TABLE 1 | Clinical MRI data.

Case	Hospital	Genre	Tumor location	Type	Image size (voxels)		Image spacing (mm)	
					Preoperative	Intraoperative	Preoperative	Intraoperative
1	BWH	M	Left perisylvian	BS	256 × 256 × 124	256 × 256 × 60	0.937 × 0.937 × 1.30	0.859 × 0.859 × 2.50
2	BWH	F	Right occipital	BS	256 × 256 × 124	256 × 256 × 60	0.937 × 0.937 × 1.30	0.859 × 0.859 × 2.50
3	BWH	M	Right frontal	BS	256 × 256 × 124	256 × 256 × 60	0.937 × 0.937 × 1.30	0.859 × 0.859 × 2.50
4	BWH	F	Left posterior temporal	BS	256 × 256 × 124	286 × 286 × 90	0.937 × 0.937 × 1.30	0.859 × 0.859 × 2.50
5	BWH	M	Left frontal	BS	512 × 512 × 176	256 × 256 × 60	0.500 × 0.500 × 1.00	0.859 × 0.859 × 2.50
6	BWH	M	Right frontal	BS	256 × 256 × 124	256 × 256 × 60	0.937 × 0.937 × 1.30	0.859 × 0.859 × 2.50
7	BWH	M	Right occipital	BS	512 × 512 × 176	256 × 256 × 60	0.500 × 0.500 × 1.00	0.859 × 0.859 × 2.50
8	BWH	F	Left frontal	PR	512 × 512 × 176	256 × 256 × 60	0.500 × 0.500 × 1.00	0.859 × 0.859 × 2.50
9	HSH	M	Left frontal	PR	448 × 512 × 176	448 × 512 × 176	0.488 × 0.488 × 1.00	0.488 × 0.488 × 1.00
10	HSH	M	Left parietal	PR	448 × 512 × 176	512 × 448 × 176	0.488 × 0.488 × 1.00	0.488 × 0.488 × 1.00
11	HSH	M	Right Frontal	PR	448 × 512 × 80	512 × 456 × 66	0.468 × 0.468 × 2.00	0.468 × 0.468 × 2.00
12	HSH	M	Left parietal occipital (deep)	PR	448 × 512 × 176	512 × 448 × 176	0.488 × 0.488 × 1.00	0.488 × 0.488 × 1.00
13	BWH	M	Fronto temporal	TR	286 × 286 × 90	286 × 286 × 90	0.859 × 0.859 × 2.50	0.859 × 0.859 × 2.50
14	BWH	F	Right Frontal	TR	256 × 256 × 124	256 × 256 × 60	0.937 × 0.937 × 1.30	0.859 × 0.859 × 2.50
15	HSH	M	Right temporal	TR	512 × 448 × 176	512 × 448 × 176	0.488 × 0.488 × 1.00	0.488 × 0.488 × 1.00
16	HSH	F	Left Posterior temporal	TR	484 × 484 × 58	484 × 484 × 58	0.496 × 0.496 × 1.62	0.496 × 0.496 × 1.62
17	HSH	F	Left frontal	TR	448 × 512 × 176	448 × 512 × 176	0.488 × 0.488 × 1.00	0.488 × 0.488 × 1.00
18	HSH	F	Left frontal	TR	448 × 512 × 176	448 × 512 × 176	0.488 × 0.488 × 1.00	0.488 × 0.488 × 1.00
19	HSH	F	Left frontal	TR	448 × 512 × 160	448 × 512 × 160	0.488 × 0.488 × 1.00	0.488 × 0.488 × 1.00
20	HSH	M	Right frontal	TR	448 × 512 × 88	456 × 512 × 66	0.468 × 0.468 × 2.00	0.468 × 0.468 × 2.00
21	HSH	M	Left frontal	TR	384 × 512 × 176	512 × 384 × 144	0.488 × 0.488 × 1.00	0.488 × 0.488 × 1.00
22	HSH	F	Left frontal	TR	448 × 512 × 176	448 × 512 × 176	0.488 × 0.488 × 1.00	0.488 × 0.488 × 1.00
23	HSH	F	Left frontal	TR	384 × 512 × 176	384 × 512 × 176	0.488 × 0.488 × 1.00	0.488 × 0.488 × 1.00
24	HSH	M	Left occipital	TR	448 × 512 × 176	384 × 512 × 176	0.488 × 0.488 × 1.00	0.488 × 0.488 × 1.00
25	HSH	M	Right frontal lobe (deep)	TR	448 × 512 × 176	384 × 512 × 144	0.488 × 0.488 × 1.00	0.488 × 0.488 × 1.00
26	HSH	M	Right frontal	STR	448 × 512 × 144	448 × 512 × 144	0.488 × 0.488 × 1.00	0.488 × 0.488 × 1.00
27	HSH	F	Left frontal	STR	384 × 512 × 176	448 × 512 × 176	0.488 × 0.488 × 1.00	0.488 × 0.488 × 1.00
28	HSH	F	Right frontal	STR	512 × 456 × 66	456 × 512 × 66	0.468 × 0.468 × 2.00	0.468 × 0.468 × 2.00
29	HSH	F	Right parietal	STR	512 × 456 × 66	512 × 456 × 68	0.468 × 0.468 × 2.00	0.468 × 0.468 × 2.00
30	HSH	M	Right temporal insular (deep)	STR	448 × 512 × 176	448 × 512 × 176	0.488 × 0.488 × 1.00	0.488 × 0.488 × 1.00

BWH, Brigham and Women's Hospital; HSH, Huashan Hospital; BS, brain shift; PR, Partial Resection; TR, Total Resection; STR, Supra Total Resection.

ensure the stability of the registration. More information about the parameters of BRAINSFit is available in 3D Slicer.

Non-rigid Registration

As surgery proceeds, the initial rigid preoperative to intraoperative registration becomes increasingly less valid. To make the best use of preoperative image data for surgical guidance (including the use of fMRI and DTI to inform the surgeon of critical structures near the tumor), the preoperative image must be updated accordingly using non-rigid registration.

Recent efforts have aimed to model intraoperative brain deformation due to tissue retraction and tumor resection. Miga et al. (68) introduced a method for modeling retraction and resection using a multi-step procedure, which allows arbitrary placement and movement of a retractor and removal of tissue. Tissue resection is modeled by manually deleting model elements identified as tumor in the preoperative image. Risholm et al. (69) proposed a registration framework based on the bijective Demons algorithm which can handle retraction and resection. Retraction is detected at areas of the deformation field with high internal strain and the estimated retraction boundary is integrated as a diffusion boundary in an anisotropic smoother. Resection is detected by a level set method evolving in the space where image intensities disagree. Ferrant et al. (54) tracked brain deformation due to tumor resection over multiple intraoperative MRI acquisitions. After each scan, the brain surface is segmented, and a surface-matching algorithm is used to drive the deformation of a finite element model of the brain. Vigneron et al. (70) modeled retraction by segmenting the brain surface from two sequential intraoperative MRI image volumes and identifying landmarks on these surfaces. Displacements of the landmarks between the two surfaces are used to drive deformation using a finite element modeling technique that allows discontinuities at the resection boundary. We recently

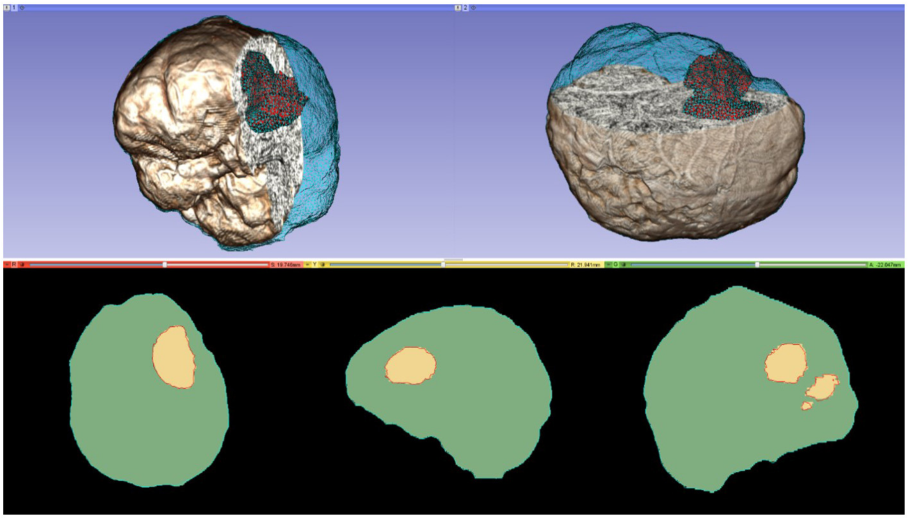

FIGURE 3 | A multi-tissue (brain parenchyma, tumor) finite element mesh used for non-rigid registration (number of tetrahedral elements: 160,179; minimum dihedral angle: 4.41°). Top row: the mesh superimposed on a volume rendering of the MRI data. Cyan and red represent the brain parenchyma and tumor meshes, respectively. Bottom row: mesh fidelity illustrated on an axial, sagittal, and coronal slices. Each slice depicts a 2D cross-section of the mesh surface (cyan and red lines) and the segmented volume (green and yellow regions). The closer the mesh surface is to the segmented boundaries, the higher the mesh fidelity.

TABLE 2 | Parameters used in this study for rigid registration (RR) and B-Spline non-rigid registration methods implemented in 3D Slicer.

Parameter	Value		Description
	RR	**B-spline**	
Cost metric	MMI	MMI	Mattes mutual information
Interpolation mode	Linear	Linear	
Sampling percentage	5%	5%	Percentage of image voxels sampled for MMI
Histogram bins	100	100	Number of histogram levels
Optimizer type	VR3DT	LBFGSB	–
Max number of iterations	1,500	1,500	Maximum number of iterations for optimizer
Grid size	–	15 × 15 × 15	Number of subdivisions of the B-Spline Grid
Min step length	10^{-3}	10^{-3}	Min threshold step for optimizer
Projected gradient tolerance	-	10^{-5}	Used by LBFGSB

MMI, Mattes Mutual Information; VR3DT, Versor Rigid 3D Transform; LBFGSB, Limited memory Broyden Fletcher Goldfarb Shannon minimization with Simple Bounds.

introduced (22, 71, 72) an adaptive form of PBNRR originally described by Clatz et al. (21) and updated and implemented in ITK by Liu et al. (24, 59). This A-PBNRR was specifically developed to model tumor resection in intraoperative images without manual intervention. Unlike other methods (54, 70) it uses image-based registration and thereby does not require segmentation of the intraoperative MRI image, which is time-consuming and may require manual intervention. The algorithm is parallelized to ensure that it is fast enough to be used in a clinical setting.

As a standard for comparison, both the rigid and the non-rigid registration methods were implemented in the open-source BRAINSFit module of a 3D Slicer. The non-rigid registration method is based on a B-spline interpolation scheme, which uses a 3-dimensional cubic control grid to optimize the registration (58). **Table 2** lists the parameters for the B-Spline deformable registration method. To facilitate performance comparisons, all three non-rigid registration methods are parallelized for shared memory multiprocessor architectures.

Adaptive Physics-Based Non-rigid Registration

A-PBNRR (22) augments PBNRR (59) to accommodate soft-tissue deformation caused by tumor resection. PBNRR has been shown to accurately capture brain shift (i.e., volumetric deformations of the brain) during image-guided neurosurgery (21). PBNRR uses the finite element method (FEM) to model deformations and estimates a sparse displacement vector associated with selected features located in the cranial cavity of the preoperative image. The sparse vector is used (as boundary conditions) to drive the deformation of a patient-specific, single-tissue (i.e., brain parenchyma), finite element mesh.

TABLE 3 | Parameters used for PBNRR and A-PBNRR.

Parameter	Value	Description
Initialization transform	Rigid	Rigid transformation to initialize the non-rigid registration
Connectivity pattern	"Face"	Pattern for the selection of blocks
F_s	5%	% selected blocks from total number of blocks
$B_{s,x} \times B_{s,y} \times B_{s,Z}$	3 × 3 × 3	Block size (in voxels)
$W_{s,x} \times W_{s,y} \times W_{s,Z}$	7 × 7 × 3 (BS) 9 × 9 × 3 (PR) 13 × 13 × 3 (TR, STR)	Block matching window size (in voxels)
δ	5	Mesh size
E_b	2.1 KPA	Young's modulus for brain parenchyma
E_t	21 KPA	Young's modulus for tumor (A-PBNRR)
V_b	0.45	Poisson ratio for brain parenchyma
V_t	0.45	Poisson ratio for tumor (A-PBNRR)
F_r	25%	% of rejected outlier blocks
N_{rej}	10	Number of outlier rejection steps
$N_{iter,max}$	10	Max number of adaptive iterations (A-PBNRR)
$N_{b0,min}$	1% × number of selected blocks	Min number of blocks with zero correspondence A-PBNRR)

A-PBNRR adds an iterative method which adaptively modifies a heterogeneous finite element model to optimize non-rigid registration in the presence of tissue resection. Using the segmented tumor and the registration error at each iteration, A-PBNRR gradually excludes the resection volume from the model. During each iteration, registration is performed, the registration error is estimated, the mesh is deformed to a predicted resection volume, and the brain model (minus the predicted resection volume) is re-tessellated. Re-tessellation is required to ensure high quality mesh elements.

The major improvements of A-PBNRR over PBNRR are:

- **Adaptivity.** An adaptive, iterative process allows A-PBNRR to gradually change the preoperative model geometry to accommodate resection.
- **Heterogeneity.** Whereas, PBNRR can only be applied to a homogenous (single-tissue) brain model, A-PBNRR can accommodate a heterogeneous (multi-tissue) model. Two-tissue models (brain parenchyma and tumor) were used in this study, but the method can accommodate any number of tissues. **Figure 3** depicts a heterogeneous brain model, and **Table 3** lists the mechanical tissue properties used in this study.
- **Higher Parallelization.** A-PBNRR uses a parallel framework that can target shared memory multi-core machines. A previous study (22) showed that A-PBNRR exploits additional parallelism over PBNRR with corresponding performance improvements so that, even with multiple iterations, A-PBNRR requires on average <2 min to perform non-rigid registration.

OPTIMIZATIONS

A-PBNRR is a computationally intensive algorithm that must be able to execute during a time-critical IGNS operation. Two ways were explored to improve accuracy and performance: (1) equidistribution of registration points using adaptive refinement for improved accuracy and (2) deep learning for parameter search space reduction for improved accuracy and performance.

Adaptive Refinement for the Optimal Distribution of Registration Points

As noted in section Mesh Generation, the presented pipeline utilizes a Delaunay-based image-to-mesh conversion tool for mesh generation. This approach can generate a mesh that faithfully captures (with geometric guarantees) the surface of the input image and the interface between the two tissues. However, it does not consider any information about the registration points recovered by the Block Matching step. Examples of selected blocks are shown in **Figure 4**. In previous work (73), the distribution of landmarks over the mesh was incorporated into the mesh generation module using custom sizing functions for two different mesh generation methods (Delaunay refinement and Advancing Front). The goal of these modifications is to equidistribute the landmarks among the mesh elements which is expected to improve the registration error. The evaluation presented in Fedorov and Chrisochoides (73) was based on synthetic deformation fields and showed that indeed these modifications reduce the registration error.

In this work, the same sizing function is applied in order to validate the effectiveness of the method. Moreover, preliminary results on applying mesh adaptation methods that originate from the Computational Fluid Dynamics field (74) are presented.

For completeness, a summary of the method employed in Fedorov and Chrisochoides (73) is presented along with the modifications that can turn it into an anisotropic metric-based method. The equidistribution of the registration points can be formulated as assigning the same number of registration points

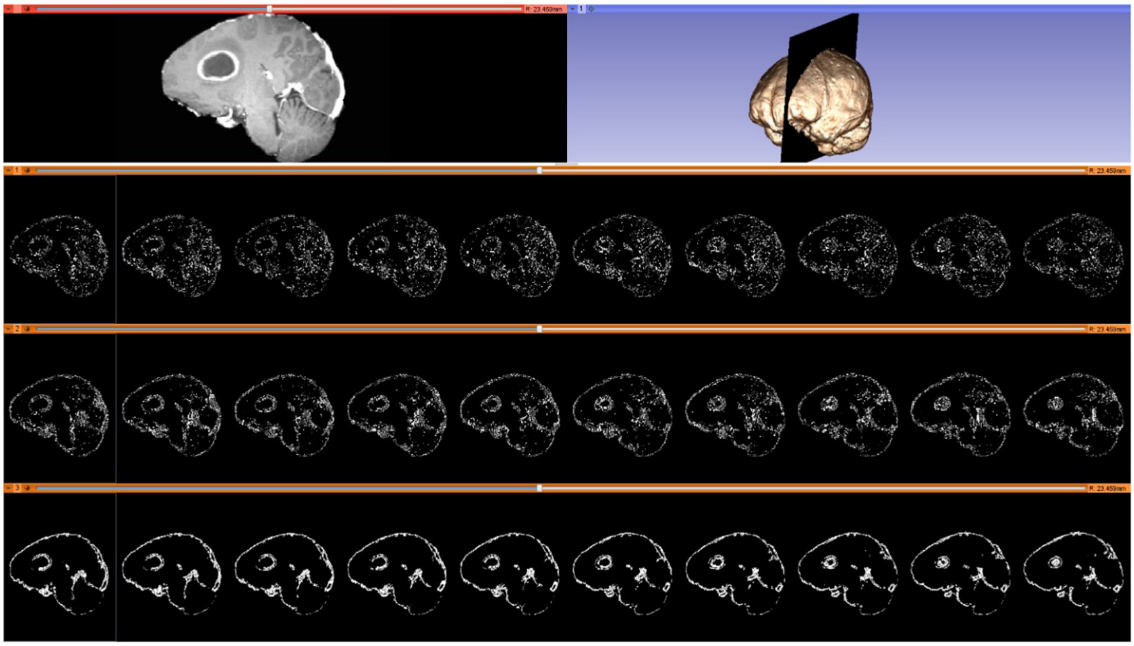

FIGURE 4 | Selected blocks from an MRI volume using various connectivity patterns. Blocks are depicted on 10 consecutive sagittal slices. From top to bottom row: sagittal slice (left) and volumetric MRI rendering (right); selected blocks with a "vertex" pattern; selected blocks with an "edge" pattern; selected blocks with a "face" pattern. Number of selected blocks for all patterns: 322,060.

at each mesh vertex cell complex, where a mesh vertex cell complex is defined as the set of all the elements attached to a vertex. The crux of the method is to set the local spacing at each vertex equal to the distance to the k-th closest registration point. Assuming an ideal spacing the mesh vertex cell complex of each vertex will contain k registration points. An illustration for k = 5 in given in **Figure 5**. Notice that another way to interpret the sizing constraint at each vertex is by a sphere centered at each mesh vertex with a radius equal to the distance to the k-th registration point.

The non-optimized A-PBNRR creates adaptive meshes but it does not capture the local density of the landmarks efficiently due to the fact that only the k-th point is used and the relative locations of the rest k-1 landmarks is ignored. Building upon the observation that the previous method can be seen as placing spheres at each vertex, one can evaluate the smallest bounding ellipsoid that contains the k closest registration points and is centered at the given vertex. Describing the local spacing as an ellipsoid gives the ability to capture the local distribution of the landmarks better thanks to the increased degrees of freedom of an ellipsoid is comparison to a sphere. Creating the minimum volume ellipsoid that encloses a given pointset is a problem well-studied in the optimization literature (75). The constructed ellipsoid has a natural mapping to a 3 × 3 positive definite matrix (76) that can be used as a metric that guides the anisotropic mesh adaptation procedure. In order to give to the mesh adaptation procedure more flexibility an additional "inflation" constant *a* is introduced that is common for all the points and allows to enlarge all ellipsoids by a constant factor. The goal of this parameter is

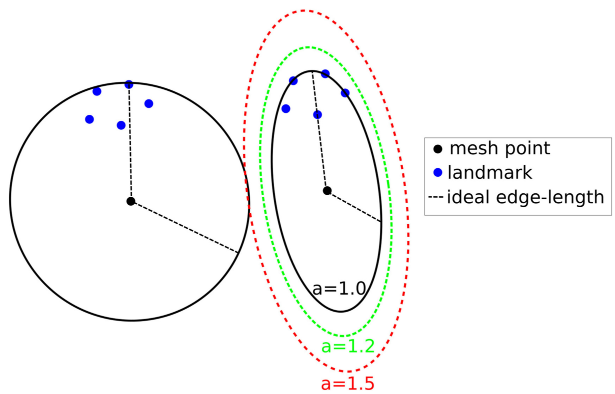

FIGURE 5 | Visualization of the metric construction for mesh adaptation. Left isotropic metric that set the spacing equal to the distance of the 5th closest registration point. Right anisotropic metric based on the five registration points for different values of the inflation parameter *a*.

allow the mesh generation procedure to perform operations that may not conform to the strict size but improve the overall result. See **Figure 5**.

In order to incorporate this approach to A-PBNRR, the mesh generated by the Parallel Optimistic Delaunay Mesher (PODM) method (65) at each iteration, along with the landmarks identified by the Block-Matching step, are used to build a metric field. The metric field is constructed by iterating in parallel the mesh vertices and evaluating the k-closest registration points using a k-nn search from the VTK library (77). The minimum volume bounding ellipsoid is constructed using the Khachiyan algorithm (78). Finally, the mesh is adapted using MMG3D (79).

Deep Learning for Parameter Search Space Reduction

A-PBNRR utilizes many different parameters that drive its results. Every patient's brain is different, and time is a critical resource in IGNS operations. As a result, the issue of determining input parameters to achieve a registration as optimal and as quickly as possible while also accounting for patient-specific details is an open problem. A-PBNRR has many input parameters (see **Table 3**), and the cost for an exhaustive parameter search is prohibitively expensive. For example, for an average case presented in this paper, it takes more than 10 days using a cluster of 400 cores running 24/7 to find sub-optimal parameter values. To address this problem, we have developed a deep feedforward neural network that can predict sets of optimal or suboptimal input parameters that yield a low Hausdorff distance of the registered image from the preoperative image. The deep learning system learns the correlation of the different input parameters, some of which are physical parameters, and how they contribute to a low Hausdorff distance.

The neural network takes as input 14 parameters: 12 A-PBNRR input parameters and two additional patient-specific parameters. The output of the neural network is a single value: the predicted Hausdorff distance of the registered image from the preoperative image if these parameters were to be used as input to A-PBNRR. The two patient-specific parameters are: (1) the location of the tumor in the brain (lobe-wise), each position represented by a numerical value and (2) the degree of brain deformation caused by the tumor, which can be directly inferred from the rigid registration error. These two parameters are necessary to increase the patient-specificity of the model, as a general model does not properly consider differences in the brains of different patients. In our experiments, using the patient-specific parameters yielded significantly better results than simply using the A-PBNRR parameters. As for the architecture, the neural network was implemented using Keras, on a TensorFlow backend. It consists of four hidden fully connected layers, each composed of 128 neurons. We used ReLU as the activation function and stochastic gradient descent (SGD) with Nesterov momentum for optimization. The architecture was determined *via* grid search. The deep learning model was trained on output data from over 2.5 million executions of A-PBNRR, spanning 12 different patient cases, including partial, complete, and extreme tumor resections. Out of the 12 cases, 10 were used for training (∼2.3 million parameter sets), and two for evaluation (∼200,000 parameter sets). The cases used for evaluation are partial tumor resection and complete tumor resection data. We have four classes of data: brain shift, partial, complete, and supra-total tumor resections. The two classes in the middle were used for evaluation since they represent the most-frequently occurring cases. The training and evaluation datasets are mutually exclusive.

The deep learning model is used before the execution of A-PBNRR. The software utilizes as input the parameter sets predicted by the deep learning model to result in the lowest Hausdorff distances. The neural network is given as input each parameter set in a pool consisting of patient-specific parameter sets. This was produced by augmenting a base, general parameter set pool by including the two patient-specific parameters. The neural network iterates through each parameter set and outputs the Hausdorff distance of the registered image that would be produced by A-PBNRR if this parameter set were utilized. The lowest of these predictions are compiled in a file and can be used as input to A-PBNRR.

Choosing good parameters for medical image registration is a difficult task, as there are many possible values and combinations. The deep learning portion of the A-PBNRR framework makes this easier by greatly limiting the set of possible optimal parameters for each individual patient, bringing A-PBNRR one-step closer to being utilized in real-world, time critical IGNS operations.

RESULTS

An evaluation of four registration methods (including the proposed method) is performed using imaging data from thirty patients who underwent partial, total, and supra total glioma resection. For a more comprehensive evaluation, the accuracy was assessed using both qualitative (visual inspection), and quantitative criteria (Hausdorff Distance-based error metric, and a landmark-based error measured by Dr. Chengjun Yao). The four registration methods were:

1. Rigid registration implemented in 3D Slicer v4.4.0 (58).
2. B-Spline non-rigid registration implemented in 3D Slicer v4.4.0 (58).
3. PBNRR implemented in ITKv4.7.0 (59).
4. A modification of PBNRR than handles tumor resections (A-PBNRR) (22).

Tables 2, **3** list the input parameters used for the registration methods.

Visual Assessment

In most applications, careful visual inspection remains the primary validation check due to its simplicity and speed. In this study, a visual inspection of the full registered volumes was performed by a neurosurgeon. The neurosurgeon inspects the brain morphology, relevant landmarks and eloquent areas of the brain, the brain shift, the margins of the tumor and the deformation after the resection. The inspection was performed after subtracting the registered preoperative MRI from the intraoperative MRI. The smaller the differences after the subtraction the more precise the alignment. **Figure 6** presents the registration results for 13 tumor resection cases (three partial, seven total, and three supra total resections). These cases are representative due to the different locations of the tumor resection. For each patient, **Figure 6** shows a 2D section from the

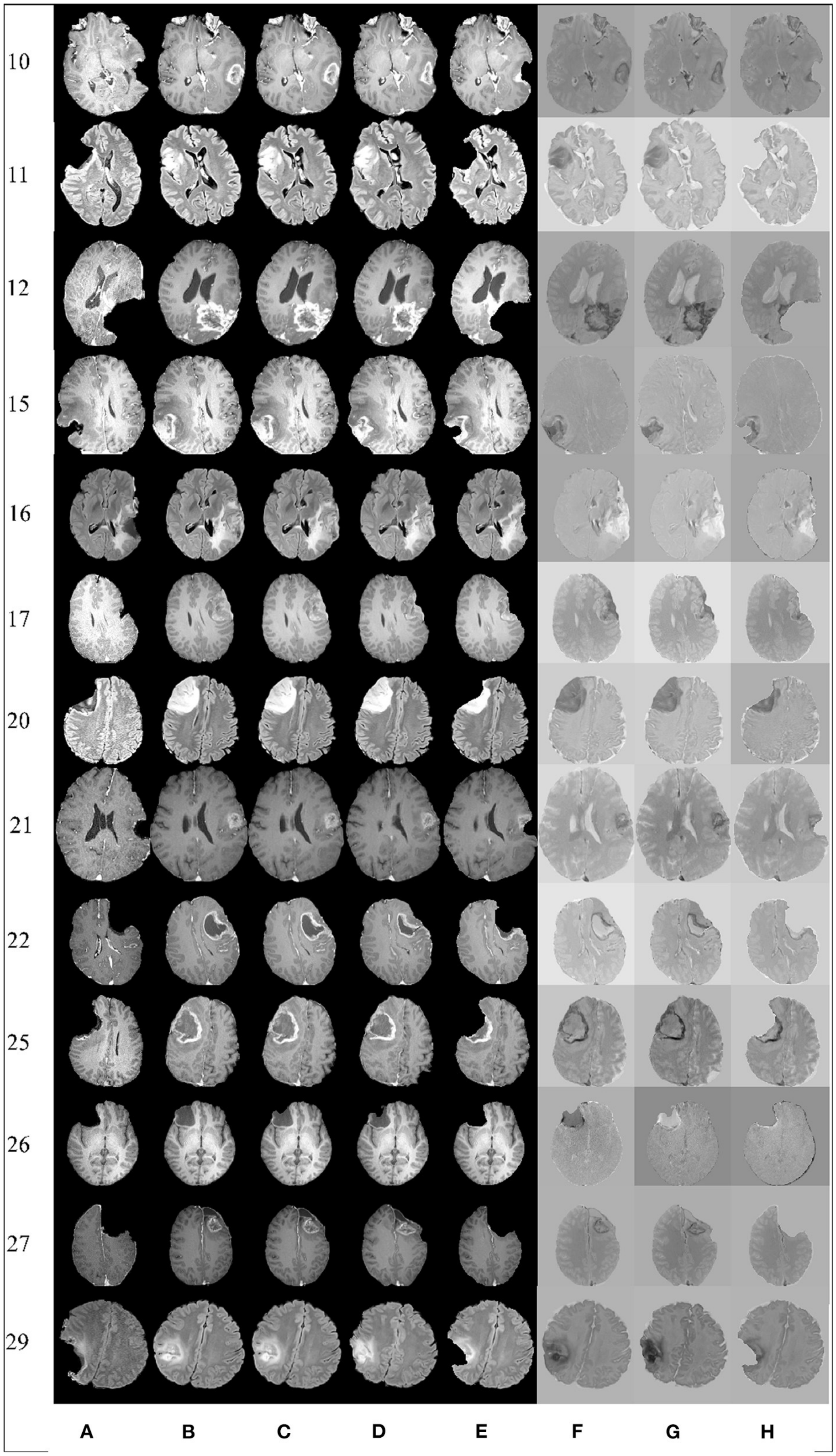

FIGURE 6 | Qualitative results. Each row represents the same slice of a 3D volume for the case numbered on the left. From left to right: intraoperative MRI **(A)**; deformed preop MRI after **(B)** rigid registration, **(C)** B-Spline, **(D)** PBNRR, and **(E)** A-PBNRR; difference between intraoperative MRI and **(F)** B-Spline, **(G)** PBNRR, and **(H)**: A-PBNRR.

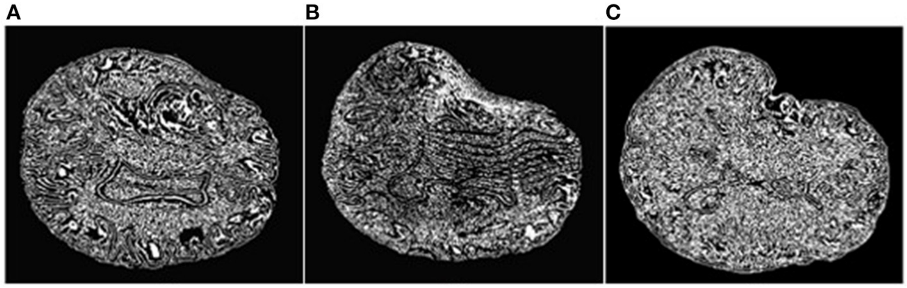

FIGURE 7 | Extracted Canny points in a single slice for quantitative evaluation of registration accuracy using the HD metric. **(A)**: Points extracted from the preoperative MRI; **(B)**: Points of **(A)** after transformation to the intraoperative space; **(C)**: Points extracted from the intraoperative MRI. The HD metric is computed between point sets **(B)** and **(C)**. Note that the Canny points are generally different from feature points used for registration.

intraoperative MRI, the corresponding registered preoperative MRI, and the subtraction of the registered preoperative MRI from the intraoperative MRI. Smaller differences indicate a more precise alignment. As **Figure 6** illustrates, A-PBNRR provides the most accurate alignment and preserves brain morphology in the presence of resection, specifically near tumor margins. In contrast, the other registration methods fail to capture the complex soft-tissue deformation near the tumor resection.

Quantitative Assessment Using Hausdorff Distance (HD)

An objective and automatic method (80) was employed to quantitatively evaluate the registration accuracy. This method was preferred because it is fast and does not require a manual intervention. It relies on Canny edge detection (81) to compute two-point sets. The first point set is computed from the preoperative volume (**Figure 7A**) and then transformed (using the deformation field computed by each registration method) from the preoperative to the intraoperative space. **Figure 7B** depicts a transformed point set. The second point set is computed from the intraoperative volume (**Figure 7C**). A Hausdorff Distance (HD) metric (82) is employed to calculate the degree of displacement between the two-point sets.

Table 4 presents the results of this quantitative evaluation. A smaller *HD* value indicates better registration ($HD \geq 0$), so that perfect registration would have an *HD* of 0. **Table 5** presents the minimum, maximum, and mean errors for each case.

The ratio $= \mathrm{HD_X}/\mathrm{HD_{A-PBNRR}}$ indicates the degree to which the error of the A-PBNRR is lower than the error of the X method, where X ϵ {RR, B-Spline, PBNRR}. A-PBNRR achieved the smallest error in each individual case and the smallest average error (3.63 mm) among all four methods. In 30 test cases, A-PBNRR is 5.47, 4.34, and 5.06 times more accurate in the presence of resection than RR, B-Spline, and PBNRR, respectively. Note that this study utilized a 100% HD metric unlike our previous work (20), which featured a 95% HD metric. **Figure 8A** plots the HD error of data in **Table 4**. **Tables 4**, **5** suggest that independently of the evaluation method the A-PBNRR outperforms all three registration methods used in this evaluation. These quantitative results are consistent with the quality data presented in **Figure 7**.

Quantitative Assessment via Anatomical Landmarks

Registration accuracy was quantitatively evaluated using anatomical landmarks selected by a neurosurgeon, as suggested in Hastreiter et al. (83) (**Figure 9**). The neurosurgeon located six landmarks in each registered preoperative image volume and corresponding intraoperative image volume. Landmarks A and B were selected individually in the cortex near the tumor; C and D were selected at the anterior horn and the triangular part of the lateral ventricle, respectively; E and F were selected at the junction between the pons and mid-brain and at the roof of the fourth ventricle, respectively. Between one and four additional landmarks of functional interest were located on an individual basis by the neurosurgeon. For each case, these additional landmarks were selected depending on the location of the tumor, the surgical approach, and the visibility of the preoperative and intraoperative images. These structures of functional interest include, amongst others, the primary motor cortex, the pyramidal tract, the Sylvian fissure, the lateral border of the thalamus, the basal ganglia, the posterior limb of the internal capsule and major vessels. For each landmark, the error was calculated as the distance between the landmark location in the registered preoperative image and the corresponding intraoperative image. **Table 5** presents the minimum, maximum, and mean errors for each case. The assessment confirms that the A-PBNRR provides the most accurate registration, with an average minimum error of 1.03 mm and average mean error of 3.22 mm.

Performance

All methods were parallelized for shared memory multiprocessor architectures. **Figure 8B** presents the end-to-end execution time for the registration of preoperative to intraoperative images for all 30 cases. Rigid registration, B-Spline, PBNRR, and A-PBNRR required on average 0.84, 8.98, 0.83, and 1.42 min, respectively (including I/O). Note that the B-Spline method is the most computationally intensive, requiring more than 8 min in 17 out of 30 cases. A different set of B-Spline parameters, such as a smaller sampling percentage, a smaller number of histogram bins, or a coarser grid could potentially improve performance at the cost of accuracy. Although A-PBNRR is slower than rigid registration

TABLE 4 | Quantitative registration results using the HD metric.

#	Type	HD_{RR}	$HD_{BSPLINE}$	HD_{PERNRR}	$HD_{A\text{-}PBNRR}$	$\frac{HD_{RR}}{HD_{A\text{-}PBNRR}}$	$\frac{HD_{BSPLINE}}{HD_{A\text{-}PBNRR}}$	$\frac{HD_{PERNRR}}{HD_{A\text{-}PBNRR}}$
1	BS	11.07	9.30	7.63	3.48 *(2)*	3.18	2.67	2.19
2	BS	24.64	24.51	21.39	2.77 *(5)*	8.90	8.85	7.72
3	BS	10.49	7.75	10.53	5.88 *(3)*	1.78	1.32	1.79
4	BS	6.59	6.51	4.97	2.64 *(2)*	2.50	2.47	1.88
5	BS	7.68	5.28	5.73	2.65 *(2)*	2.90	1.99	2.16
6	BS	8.54	8.54	5.55	3.48 *(2)*	2.45	2.45	1.59
7	BS	8.99	8.99	7.36	4.33 *(3)*	2.08	2.08	1.70
8	PR	17.00	17.00	16.49	5.69 *(4)*	2.99	2.99	2.90
9	PR	10.59	5.28	10.76	2.30 *(3)*	4.60	2.30	4.68
10	PR	16.15	13.78	15.12	4.60 *(7)*	3.51	3.00	3.29
11	PR	26.89	15.86	26.89	4.00 *(6)*	6.72	3.97	6.72
12	PR	29.93	21.34	27.76	2.83 *(7)*	10.58	7.54	9.81
13	TR	25.51	25.18	22.50	4.97 *(4)*	5.13	5.07	4.53
14	TR	5.59	5.59	3.43	3.09 *(1)*	1.81	1.81	1.11
15	TR	17.90	16.94	15.56	4.11 *(9)*	4.36	4.12	3.79
16	TR	18.85	17.49	17.38	3.57 *(3)*	5.28	4.90	4.87
17	TR	17.14	7.48	15.41	4.25 *(2)*	4.03	1.76	3.63
18	TR	25.72	25.72	23.90	3.42 *(6)*	7.52	7.52	6.99
19	TR	25.43	17.63	25.22	3.30 *(9)*	7.71	5.34	7.64
20	TR	23.61	21.42	22.89	3.66 *(4)*	6.45	5.85	6.25
21	TR	19.24	14.61	19.89	2.40 *(6)*	8.02	6.09	8.29
22	TR	30.37	21.39	28.96	3.13 *(7)*	9.70	6.83	9.25
23	TR	15.16	11.89	13.96	3.15 *(4)*	4.81	3.77	4.43
24	TR	13.47	8.90	13.66	3.28 *(4)*	4.11	2.71	4.16
25	TR	23.22	14.99	21.44	3.08 *(9)*	7.54	4.87	6.96
26	STR	17.59	17.12	16.63	4.19 *(5)*	4.20	4.09	3.97
27	STR	35.72	27.77	33.57	3.71 *(8)*	9.63	7.49	9.05
28	STR	32.32	29.43	30.13	3.45 *(6)*	9.37	8.53	8.73
29	STR	18.48	13.30	18.15	3.97 *(4)*	4.65	3.35	4.57
30	STR	27.07	15.55	24.91	3.54 *(7)*	7.65	4.39	7.04
Average		19.03	15.22	17.59	3.63	5.47	4.34	5.06

HD_{RR}, $HD_{BSPLINE}$, HD_{PBNRR}, and $HD_{A-PBNRR}$ are alignment errors after Rigid Registration (RR), B-Spline, PBNRR, and A-PBNRR registration, respectively. HD are in mm. The number in the parenthesis denotes the number of adaptive iterations for A-PBNRR. BS, Brain Shift; PR, Partial Resection; TR, Total Resection; STR, Supra Total Resection.

and PBNRR, it has significantly better accuracy in the presence of resection than the other methods, and it is fast enough to satisfy the constraints of image-guided neurosurgery, where registration times of $<$2–3 min are desired.

Adaptive Refinement for the Optimal Distribution of Registration Points

The results of augmenting mesh adaptation to the A-PBNRR method are presented in **Table 6**. The number of registration points per mesh cell vertex are set to k = 500. This value was selected since is produces meshes with a vertex count close to the baseline meshes. The first line of the table corresponds to the base case of using A-PBNRR with no adaptation. "iso" indicates the application of isotropic adaptation as described in (73). Isotropic adaptation reduces the Hausdorff distance by almost 13% for this case but increases the error at the landmarks identified by the neurosurgeon. Moreover, it comes with the price of almost twice the size of the mesh. The rest of the rows correspond to applying anisotropic mesh adaptation as described in section Deep Learning for Parameter

TABLE 5 | Quantitative registration results using six anatomical landmarks (A–F).

Method	Average min error (mm)	Average max error (mm)	Average mean error (mm)
RR	3.19	8.90	5.60
BSPLINE	2.15	8.29	4.40
PBNRR	1.11	6.81	3.47
A-PBNRR	1.03	6.59	3.22

The values are the average minimum, maximum, and mean errors computed over 30 cases for rigid registration (RR), PBNRR, and A-PBNRR.

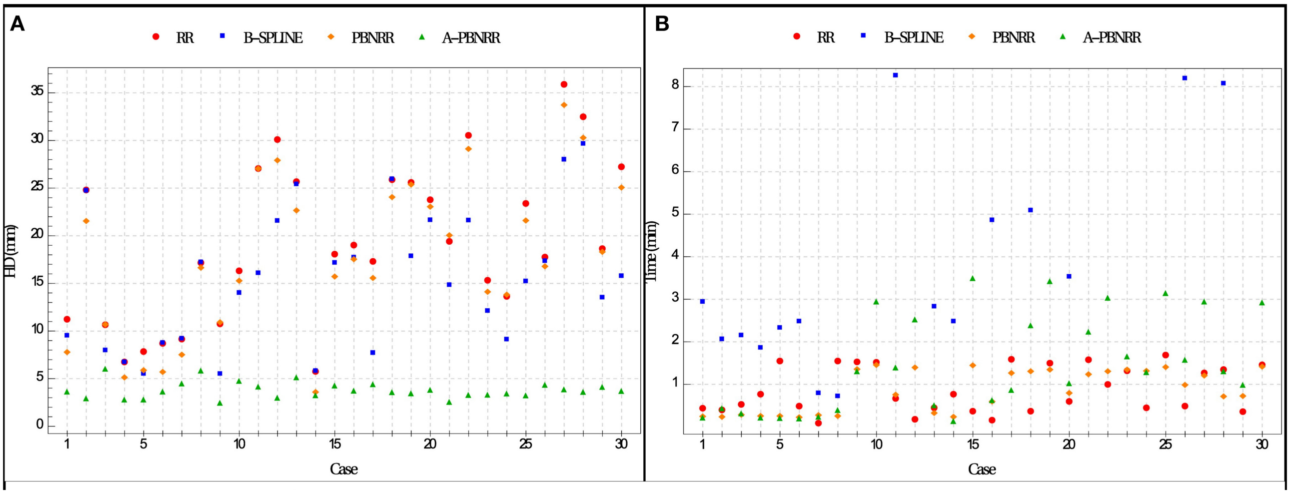

FIGURE 8 | (A) Plot of Hausdorff Distance (HD) errors of **Table 4**. Brain shift: cases 1–7; partial resection: cases 8–12; total resection: cases 13–25; supra total resection: cases 26–30. **(B)** End-to-end execution times for registration from preoperative to intraoperative images (excluding B-Spline times of more than about 8 min). All registration methods were run in parallel on 12 hardware cores on a DELL workstation with 12 Intel Xeon X5690@3.47 GHz CPU cores and 96 GB of RAM. Execution times include I/O. The mesh generation time is excluded from the PBNRR (preoperative step) but is included in the A-PBNRR (intraoperative step).

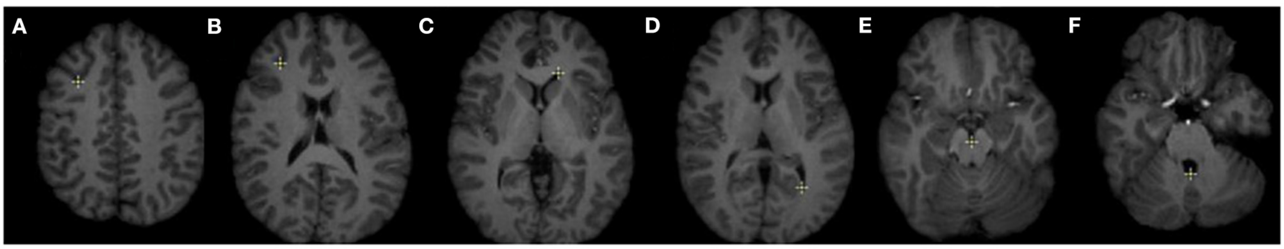

FIGURE 9 | Anatomical landmarks **(A–F)** used for quantitative evaluation of registration accuracy. A neurosurgeon located the landmarks. **(A,B)** cortex near tumor; **(C)** anterior horn of later ventricle; **(D)** triangular part of lateral ventricle; **(E)** junction between pons and mid-brain; **(F)** roof of fourth ventricle.

Search Space Reduction. We also provide a free parameter "a" which corresponds to 'inflating' the generated ellipsoids by a constant amount. The goal of this parameter is to allow the mesh generation procedure to perform operations that may not conform to the strict size but improve the overall result. For a = 1.0 (that is no inflation) the Hausdorff distance is marginally lower and only the minimum landmark-based error is lower. Increasing the inflation parameter to 1.2 and 1.5 one can see an improvement in the minimum and mean landmark-based error and at the same time a reduction in mesh size.

Although these results are preliminary (a complete study will be completed in the future), they indicate that the problem of generating an 'ideal' mesh for image registration purposes includes competing evaluation criteria like the minimum mesh size, Hausdorff distance and the landmark-based error above. Introducing mesh adaptation to A-PBNRR has the potential to improve its effectiveness but further investigation is needed in order to optimize its parameters which makes it a good candidate for the deep learning methods presented in section.

Deep Learning for Parameter Search Space Reduction

Using the deep learning model, we achieved a training root mean squared error (RMSE) of 1.41 and an evaluation RMSE of 1.21 for predicted Hausdorff distances. On average, based on **Table 7**, A-PBNRR with deep learning is ∼8.45 times better than rigid registration, ∼6.71 times better than B-Spline registration, and ∼7.9 times better than PBNRR. It should also be noted that A-PBNRR works very well with deep tumors, which result in great brain deformation, in comparison to the other registration methods, leading to results that are on average ∼16.8 times better. Overall, A-PBNRR with deep learning leads to more accurate results than any of the other registration methods.

DISCUSSION AND FUTURE WORK

Recent advances in neuroimaging such as fMRI and DTI allow neurosurgeons to plan tumor resections that minimize damage to eloquent cortical regions and white matter tracts (11, 13, 84–86). A number of commercial systems (e.g.,

TABLE 6 | Effects of applying mesh adaptation at each iteration of A-PBNRR.

Method	Hausdorff distance (mm)	Min error (mm)	Max error (mm)	Mean error (mm)	# vertices	# tetrahedra
Baseline	2.24	1.07	5.9	3.51	3,264	13,210
Isotropic	1.95	1.22	7.53	3.71	4,177	19,893
Anisotropic (a = 1.0)	2.22	0.55	7.85	3.99	4,520	22,383
Anisotropic (a = 1.2)	2.00	1.01	7.10	3.70	3,629	17,593
Anisotropic (a = 1.5)	2.64	0.93	6.15	3.25	2,838	13,291
Baseline	4.06	2.06	5.37	3.65	2,833	11,040
Isotropic	3.42	2.29	5.76	3.92	4,008	19,466
Anisotropic (a = 1.0)	3.71	2.12	5.50	3.96	4,460	22,342
Anisotropic (a = 1.2)	4.05	2.06	5.05	3.61	3,766	18,077
Anisotropic (a = 1.5)	4.05	1.92	5.17	3.65	2,983	13,812

TABLE 7 | Shows the 12 patient cases that consist the machine learning data set and the results (measured as the Hausdorff distance in mm) achieved with various methods of registration, including with A-PBNRR using deep learning and A-PBNRR using a parameter sweep.

Case	Type	Tumor location	Rigid registration	B-Spline registration	PBNRR	A-PBNRR (default parameters)	A-PBNRR (deep learning)	A-PBNRR (parameter sweep)
8	PR	Left frontal lobe	17.00	17.00	16.49	5.69	2.78	2.78
9	PR	Left frontal lobe	10.59	5.28	10.76	2.30	2.64	1.77
10	PR	Left parietal lobe	16.15	13.78	15.12	4.60	2.40	1.70
12	PR (deep)	Left parietal-occipital lobes	29.93	21.34	27.76	2.83	1.84	1.46
13	TR	Frontal-temporal lobes	25.51	25.18	22.50	4.97	2.64	2.64
15	TR	Right temporal lobe	17.90	16.94	15.56	4.11	2.94	2.00
16	TR	Left posterior-temporal lobe	18.85	17.49	17.38	3.57	3.29	3.24
17	TR	Left frontal lobe	17.14	7.48	15.41	4.25	2.40	1.84
27	STR (deep)	Left frontal lobe	35.72	27.77	33.57	3.71	2.06	1.77
18	TR	Left frontal lobe	25.72	25.72	23.90	3.42	2.64	2.30
11 (evaluation)	PR	Right frontal lobe	26.89	15.86	26.89	4.00	2.85	2.21
21 (evaluation)	TR	Left frontal lobe	19.24	14.61	19.89	2.40	2.00	1.48

Cases with numbers 8, 9, 10, 12, 13, 15, 16, 17, 27, 18 were used for training, and cases 11, 21 were used for evaluation. PR, Partial Resection; TR, Total Resection; STR, Supra Total Resection.

Brainlab Curve Image-Guided Therapy system) can register fMRI and DTI to preoperative anatomical MRI images and then map this data to the patient intraoperatively using rigid registration. It has been shown, however, that there can be significant deformation of the brain during surgery, especially in the presence of tumor resection, making rigid registration insufficient and surgical plans made with preoperative data invalid (87).

Intraoperative MRI can be used to observe the deformed brain during surgery. While it is impractical to acquire fMRI and DTI intraoperatively, the preoperative MRI image can be registered to an intraoperative MRI image using non-rigid registration. The resultant registration can then be applied to the preoperative fMRI, DTI, and the surgical plan, providing more accurate, updated guidance to the neurosurgeon (20).

Non-rigid registration algorithms remain computationally expensive and have proven to be impractical for use in clinical settings in the past. However, this study indicates that parallel/distributed computing and deep learning can provide faster and more effective registration for image-guided neurosurgery (20, 88), which is critical for immersive solutions.

The experimental results confirm that, of the four methods, A-PBNRR provides the most accurate registration, with an average error of 3.2–3.6 mm for landmark-based and Hausdorff Distance-based metrics, respectively, for anisotropic image spacing (e.g., 0.488 × 0.488 × 1.0

mm^3, **Table 1**). The extent of the resection (partial, total, or supra total) does not significantly affect this accuracy (**Figures 7**, **8**; **Table 4**). Performance analysis shows that A-PBNRR is sufficiently fast to be useful in a clinical setting.

As part of future work, improvements on the modeling of major substructures of the brain, such as the ventricles shall be made. As of now, the brain is meshed as one tissue. However, this can sometimes lead to misstructured tissues, such as in the case of the ventricles which might appear twisted. This will be solved by using multi-tissue mesh generation, where the ventricles and the rest of the brain are meshed as independent structures and later combined. Regarding the machine learning, Hausdorff distance results were better than the results using the default parameters. However, the assessment of anatomical landmarks from the neurosurgeons showed little improvement of the machine learning over the default parameters. Therefore, further work is needed.

This paper has shown, on a study of 30 patients, that we can map preoperative image data to the patient with an average error of a few millimeters and with computation times that are acceptable in a clinical environment. Future efforts will continue this focus on improving registration accuracy and decreasing computation times, using Cloud computing and Machine Learning. Decreasing computation times will require exploring ways to improve parallelization of registration methods. Improving accuracy will require an investigation into higher order finite element modeling to study the impact on accuracy and performance. Additionally, the effect of segmentation and model construction on the registration accuracy remains a potential area of future study. Specifically, an investigation into the effect of higher quality segmentation of the brain surface and structures that constrain brain deformation, such as the skull cavity, the falx cerebri, and the tentorium cerebelli, would reveal the impact of incorporating more tissues, including blood vessels and the ventricles, into the brain model. Finally, because intraoperative MRI is not available in many neurosurgical suites, further investigation incorporating the use of intraoperative ultrasound to track brain deformation during surgery would extend these registration methods to map preoperative image data to intraoperative ultrasound.

In summary, although the paper presents promising results, there are two limitations. First, an accuracy of <2 mm needs to be consistently achieved. Second, these results are realized with intra-operative MRIs, which are expensive and thus not widely used. As an alternative, intra-operative ultrasound can be used. However, to achieve high accuracy with ultrasound, a much harder problem, a more computationally friendly modality (i.e., intra-operative MRI) has to first be addressed.

CONCLUSION

This study compared four methods for registering preoperative image data to intraoperative MRI images in the presence of significant brain deformation during glioma resection in 30 patients. The Adaptive Physics-Based Non-Rigid Registration method developed in this study proved to be significantly better than other methods at modeling deformation in the presence of resection. Both the registration accuracy and performance were found to be of clinical value in the operating room, and the combination of A-PBNRR with deep learning and mixed reality can offer a compelling solution for IGNS.

DATA AVAILABILITY STATEMENT

The original contributions presented in the study are included in the article/supplementary material, further inquiries can be directed to the corresponding author/s.

AUTHOR CONTRIBUTIONS

FD and NC developed the Adaptive PBNRR (A-PBNRR) method, which is implemented by FD. AA, CT, and AF (while he was at CRTC) with NC developed the Machine Learning and mesh refinement approaches to improve the accuracy of the approximation method for deformable registration. The immersive aspects of the paper and the use of NRR within mixed reality IGNS environments are done by YL and NC. The collection of data and evaluation of the results and methods are taken care of by CY, KK, SF, NF, AG, and RK. All authors contributed to the article and approved the submitted version.

FUNDING

This work in part was funded by NSF Grant No. CCF-1439079, NASA Grant No. NNX15AU39A, John Simon Guggenheim Foundation, Richard T. Cheng Endowment, and Stavros Niarchos Foundation, SURA CNF'19-04, NIH R01NS049251, NIH 5R01EB027134, and NCIGT P41EB015898. In addition, it is partially supported by Modeling and Simulation Fellowship at Old Dominion University and Dominion Scholarship.

ACKNOWLEDGMENTS

We want to thank the reviewers for their constructive and insightful comments that substantially improved the presentation of this paper.

REFERENCES

1. Louis D, Ohgaki H, Wiestler O, Cavenee W, Burger P, Jouvet A, et al. The 2007 WHO classification of tumors of the central nervous system. *J Acta Neuropathol.* (2007) 114:97–109. doi: 10.1007/s00401-007-0243-4
2. Sathornsumetee S, Rich JN, Reardon DA. Diagnosis and treatment of high-grade astrocytoma. *Neurol Clin.* (2007) 25:1111–39. doi: 10.1016/j.ncl.2007.07.004
3. Evren G, Chang E, Lamborn K, Tihan T, Chang C, Chang S, et al. Volumetric extent of resection and residual contrast enhancement on initial surgery as predictors of outcome in adult patients with hemispheric

anaplastic astrocytoma. *J Neurosurg.* (2006) 105:34–40. doi: 10.3171/jns.2006.105.1.34

4. McGirt M, Mukherjee D, Chaichana K, Than K, Weingart J, Quinones-Hinojosa A. Association of surgically acquired motor and language deficits on overall survival after resection of glioblastoma multiforme. *Neurosurgery*. (2009) 65:463–9; discussion: 469–70. doi: 10.1227/01.NEU.0000349763.42238.E9
5. Shaw E, Berkey B, Coons S, Bullard D, Brachman D, Buckner J, et al. Recurrence following neurosurgeon-determined gross-total resection of adult supratentorial low-grade glioma: results of a prospective clinical trial. *J Neurosurg.* (2008) 109:835–41. doi: 10.3171/JNS/2008/109/11/0835
6. Grimson E, Kikinis R, Jolesz F, Black P. Image-guided surgery. *Sci Am.* (1999) 280:62–9. doi: 10.1038/scientificamerican0699-62
7. Haberg A, Kvistad KA, Unsgard G, Haraldseth O. Preoperative blood oxygen level-dependent functional magnetic resonance imaging in patients with primary brain tumors: clinical application and outcome. *Neurosurgery*. (2004) 54:902–14. doi: 10.1227/01.NEU.0000114510.05922.F8
8. Giussani C, Roux F, Ojemann J, Sganzerla E, Pirillo D, Papagno C. Is preoperative functional magnetic resonance imaging reliable for language areas mapping in brain tumor surgery? Review of language functional magnetic resonance imaging and direct cortical stimulation correlation studies. *Neurosurgery.* (2010) 66:113–20. doi: 10.1227/01.NEU.0000360392.15450.C9
9. Incekara F, Olubiyi O, Ozdemir A, Lee T, Rigolo L, Golby A. The value of pre- and intraoperative adjuncts on the extent of resection of hemispheric low-grade gliomas: a retrospective analysis. *J Neurol Surg A Cent Eur Neurosurg.* (2016) 77:79–87. doi: 10.1055/s-0035-1551830
10. Krishnan R, Raabe A, Hattingen E, Szelenyi A, Yahya H, Hermann E, et al. Functional magnetic resonance imaging-integrated neuronavigation: correlation between lesion-to-motor cortex distance and outcome. *Neurosurgery.* (2004) 55:904–14. doi: 10.1227/01.NEU.0000137331.35014.5C
11. Orringer DA, Vago DR, Golby AJ. Clinical applications and future directions of functional MRI. *Semin Neurol.* (2012) 32:466–75. doi: 10.1055/s-0032-1331816
12. Wiegell MR, Larsson HB, Wedeen VJ. Fiber crossing in human brain depicted with diffusion tensor MR imaging. *Radiology.* (2000) 217:897–903. doi: 10.1148/radiology.217.3.r00nv43897
13. Bello L, Gambini A, Castellano A, Carrabba G, Acerbi F, Fava E, et al. Motor and language DTI fiber tracking combined with intraoperative subcortical mapping for surgical removal of gliomas. *Neuroimage.* (2008) 39:369–82. doi: 10.1016/j.neuroimage.2007.08.031
14. Bello L, Castellano A, Fava E, Casaceli G, Riva M, Scotti G, et al. Intraoperative use of diffusion tensor imaging fiber tractography and subcortical mapping for resection of gliomas: technical considerations. *Neurosurg Focus.* (2010) 28:E6. doi: 10.3171/2009.12.FOCUS09240
15. González-Darder J, González-López P, Talamantes F, Quilis V, Cortés V, García-March G, et al. Multimodal navigation in the functional microsurgical resection of intrinsic brain tumors located in eloquent motor areas: role of tractography. *Neurosurg Focus.* (2010) 28:E5. doi: 10.3171/2009.11.FOCUS09234
16. Kamada K, Todo T, Masutani Y, Aoki S, Ino K, Takano T, et al. Combined use of tractography-integrated functional neuronavigation and direct fiber stimulation. *J Neurosurg.* (2005) 102:664–72. doi: 10.3171/jns.2005.102.4.0664
17. Nimsky C, Ganslandt O, Hastreiter P, Wang R, Benner T, Sorensen A, et al. Preoperative and intraoperative diffusion tensor imaging-based fiber tracking in glioma surgery. *Neurosurgery*. (2005) 56:130–7. doi: 10.1227/01.NEU.0000144842.18771.30
18. Talos IF, O'Donnell L, Westin CF, Warfield SK, Wells WM, Yoo SS, et al. Diffusion tensor and functional MRI fusion with anatomical MRI for image guided neurosurgery. In: *Sixth International Conference on Medical Image Computing and Computer-Assisted Intervention (MICCAI'03)*. Montreal, QC (2003). p. 407–15.
19. Elhawary H, Liu H, Patel P, Norton I, Rigolo L, Papademetris X, et al. Intraoperative real-time querying of white matter tracts during frameless stereotactic neuronavigation. *Neurosurgery*. (2011) 68:506–16. doi: 10.1227/NEU.0b013e3182036282
20. Archip N, Clatz O, Whalen S, Kacher D, Fedorov A, Kot A, et al. Non-rigid alignment of pre-operative MRI, fMRI, and DT-MRI with intra-operative MRI for enhanced visualization and navigation in image-guided neurosurgery. *NeuroImage.* (2007) 35:609–24. doi: 10.1016/j.neuroimage.2006.11.060
21. Clatz O, Delingette H, Talos IF, Golby A, Kikinis R, Jolesz F, et al. Robust non-rigid registration to capture brain shift from intraoperative MRI. *IEEE Trans Medical Imaging.* (2005) 24:1417–27. doi: 10.1109/TMI.2005.856734
22. Drakopoulos F, Chrisochoides N. Accurate and fast deformable medical image registration for brain tumor resection using image-guided neurosurgery. *Comput Methods Biomech Biomed Eng Imaging Visual.* 4:112–26. (2015). doi: 10.1080/21681163.2015.1067869
23. Lang MJ, Kelly JJ, Sutherland GR. A moveable 3-Tesla intraoperative magnetic resonance imaging system. *Neurosurgery*. (2011) 68(1 Suppl.):168–79. doi: 10.1227/NEU.0b013e3182045803
24. Liu Y, Yao C, Drakopoulos F, Wu J, Zhou L, Chrisochoides N. A nonrigid registration method for correcting brain deformation induced by tumor resection. *Med Phys*. (2014) 41:101710. doi: 10.1118/1.4893754
25. Miga M. Computational modeling for enhancing soft tissue image guided surgery: an application in neurosurgery. *Ann Biomed Eng.* (2015) 44:128–38. doi: 10.1007/s10439-015-1433-1
26. Nimsky C, Ganslandt O, Keller B, Romstöck J, Fahlbusch R. Intraoperative high-field-strength MR imaging: implementation and experience in 200 patients. *Radiology.* (2004) 233:67–78. doi: 10.1148/radiol.2331031352
27. Nimsky C, Ganslandt O, Hastreiter P, Fahlbusch R. Intraoperative compensation for brain shift. *Surg Neurol.* (2001) 56:357–64. doi: 10.1016/S0090-3019(01)00628-0
28. Sun G, Chen X, Zhao Y, Wang F, Song Z, Wang Y, et al. Intraoperative MRI with integrated functional neuronavigation-guided resection of supratentorial cavernous malformations in eloquent brain areas. *J Clin Neurosc.* (2011) 18:1350–4. doi: 10.1016/j.jocn.2011.01.025
29. Dorward NL, Olaf A, Velani B, Gerritsen FA, Harkness WFJ, Kitchen ND, et al. Postimaging brain distortion: magnitude, correlates, and impact on neuronavigation. *J Neurosurg.* (1998) 88:656–62. doi: 10.3171/jns.1998.88.4.0656
30. Maurer CR, Hill DLG, Martin AJ, Liu H, McCue M, Rueckert D, et al. Investigation of intra-operative brain deformation using a 1.5-t interventional MR system: preliminary results. *IEEE Trans Med Imaging.*(1998) 17:817–25. doi: 10.1109/42.736050
31. Nimsky C, Ganslandt O, Cerny S, Hastreiter P, Greiner G, Fahlbusch R. Quantification of, visualization of, and compensation for brain shift using intraoperative magnetic resonance imaging. *Neurosurgery.* (2000) 47:1070–80. doi: 10.1097/00006123-200011000-00008
32. Black P. Brains, minds, and the surgical planning laboratory. In: *SPL 25th Anniversary Reception*. Boston, MA: Brigham and Women's Hospital (2016).
33. Moiyadi A, Shetty P. Direct navigated 3D ultrasound for resection of brain tumors: a useful tool for intraoperative image guidance. *Neurosurg Focus.* (2016) 40:E5. doi: 10.3171/2015.12.FOCUS15529
34. Lekht I, Brauner N, Bakhsheshian J, Chang K, Gulati M, Shiroishi MS. Versatile utilization of real-time intraoperative contrast-enhanced ultrasound in cranial neurosurgery: technical note and retrospective case series. *Neurosurg Focus.* (2016) 40:E6. doi: 10.3171/2015.11.FOCUS15570
35. Lunn KE, Hartov A, Kennedy FE, Miga M, Roberts DW, Platenik LA, et al. 3D ultrasound as sparse data for intraoperative brain deformation model. In: *Proceedings of SPIE, Vol. 4325, Medical Imaging 2001: Ultrasonic Imaging and Signal Processing.* (San Diego, CA) (2001). p. 326.
36. Ji S, Fan X, Roberts DW, Hartov A, Schaewe TJ, Simon DA, et al. Chapter 17: Brain shift compensation via intraoperative imaging and data assimilation. In: Neu CP, Genin GM, editors. *Handbook of Imaging in Biological Mechanics*. Boca Raton, FL: CRC Press. (2014). p. 229–39.
37. Audette M, Siddiqi K, Ferrie F, Peters T. An integrated range-sensing, segmentation and registration framework for the characterization of intra-surgical brain deformations in image-guided surgery. *Comput Vis Image Understand.* (2003) 89:226–51. doi: 10.1016/S1077-3142(03)00004-3
38. Miga M, Sinha TK, Cash DM, Galloway RL, Weil RJ. Cortical surface registration for image-guided neurosurgery using laser range scanning. *IEEE Trans Med Imaging.* (2003) 22:973–85. doi: 10.1109/TMI.2003.815868
39. Mostayed A, Garlapati R, Joldes G, Wittek A, Roy A, Kikinis R, et al. Biomechanical model as a registration tool for image-guided neurosurgery: evaluation against B-Spline registration. *Ann Biomed Eng.* (2013) 41:2409–25. doi: 10.1007/s10439-013-0838-y

40. Sotiras A, Davatzikos C, Paragios N. Deformable medical image registration: a survey. *IEEE Trans Med Imaging.* (2013) 32:1153–90. doi: 10.1109/TMI.2013.2265603
41. Goshtasby A, Staib L, Studholme C, Terzopoulos D. Nonrigid image registration: guest editors' introduction. *Comput Vis Image Understand.* (2003) 89:109–13. doi: 10.1016/S1077-3142(03)00016-X
42. Crum WR, Hartkens T, Hill D. Non-rigid registration: theory and practice. *British J Radiol.* (2004) 77:S140–53. doi: 10.1259/bjr/25329214
43. Holden M. A review of geometric transformations for nonrigid body registration. *IEEE Trans Med Imaging.* (2008) 27:111–28. doi: 10.1109/TMI.2007.904691
44. Lester H, Arridge SR. A survey of hierarchical non-linear medical image registration. *Pattern Recogn.* (1999) 32:129–49. doi: 10.1016/S0031-3203(98)00095-8
45. Klein A, Andersson J, Ardekani BA, Ashburner J, Avants B, Chiang MC, et al. Evaluation of 14 nonlinear deformation algorithms applied to human brain MRI registration. *Neuroimage.* (2009) 46:786–802. doi: 10.1016/j.neuroimage.2008.12.037
46. Thompson P, Toga AW. A surface-based technique for warping three-dimensional images of the brain. *IEEE Trans Med Imaging.* (1996) 15:402–17. doi: 10.1109/42.511745
47. Smith SM. Fast robust automated brain extraction. *Hum Brain Mapp.* (2002) 17:143–55. doi: 10.1002/hbm.10062
48. Ferrant M, Nabavi A, Macq B, Jolesz FA, Kikinis R, Warfield SK. Registration of 3-d intraoperative MR images of the brain using a finite-element biomechanical model. *IEEE Trans Medical Imaging.* (2001) 20:1384–97. doi: 10.1109/42.974933
49. Hawkes DJ, Barratt D, Blackall JM, Chan C, Edwards PJ, Rhode K, et al. Tissue deformation and shape models in image-guided interventions: a discussion paper. *Med Image Anal.* (2005) 9:163–75. doi: 10.1016/j.media.2004.11.007
50. Holden M, Schnabel JA, Hill DLG. Quantification of small cerebral ventricular volume changes in treated growth hormone patients using non-rigid registration. *IEEE Trans Med Imaging.* (2002) 21:1292–301. doi: 10.1109/TMI.2002.806281
51. Kay S, Pheiffer TS, Simpson AL, Weis JA, Thompson RC, Miga MI. Near real-time computer assisted surgery for brain shift correction using biomechanical models. *IEEE J Transl Eng Health Med.* (2014) 2:1–13. doi: 10.1109/JTEHM.2014.2327628
52. Markelj P, TomaŽevič D, Likar B, Pernuš F. A review of 3D/2D registration methods for image-guided interventions. *Med Image Anal.* (2012) 16:642–61. doi: 10.1016/j.media.2010.03.005
53. Wittek A, Miller K, Kikinis R, Warfield S. Patient-specific model of brain deformation: Application to medical image registration. *J Biomech.* (2007) 40:919–29. doi: 10.1016/j.jbiomech.2006.02.021
54. Ferrant M, Nabavi A, Macq B, Black PM, Jolesz FA, Kikinis R, et al. Serial registration of intraoperative MR images of the brain. *Med Image Anal.* (2002) 6:337–59. doi: 10.1016/S1361-8415(02)00060-9
55. Simon M, Neuloh G, Lehe M, Meyer B, Schramm J. Insular gliomas: the case for surgical management. *J Neurosurg.* (2009) 110:685–95. doi: 10.3171/2008.7.JNS17639
56. Smith JS, Chang EF, Lamborn KR, Chang SM, Prados MD, Cha S, et al. Role of extent of resection in the long-term outcome of low-grade hemispheric gliomas. *J Clin Oncol.* (2008) 26:1338–45. doi: 10.1200/JCO.2007.13.9337
57. Sanai N, Polley MY, McDermott MW, Parsa AT, Berger MS. An extent of resection threshold for newly diagnosed glioblastomas. *J Neurosurg.* (2011) 115:3–8. doi: 10.3171/2011.2.JNS10998
58. Johnson H, Harris G, Williams K. Brainsfit: mutual information registrations of whole-brain 3d images, using the insight toolkit. *Insight J.* (2007).
59. Liu Y, Kot A, Drakopoulos F, Yao C, Fedorov A, Enquobahrie A, et al. An ITK implementation of a physics-based non-rigid registration method for brain deformation in image-guided neurosurgery. *Front Neuroinformatics.* (2014) 8:33. doi: 10.3389/fninf.2014.00033
60. Yao C, Lv S, Chen H, Tang W, Guo J, Zhuang S, et al. The clinical utility of multimodal MR image-guided needle biopsy in cerebral gliomas. *Int J Neurosci.* (2016) 126:53–61. doi: 10.3109/00207454.2014.992429
61. Antiga L, Piccinelli M, Botti L, Ene-Iordache B, Remuzzi A, Steinman D. An image-based modeling framework for patient-specific computational hemodynamics. *Med Biol Eng Comput.* (2008) 46:1097–112. doi: 10.1007/s11517-008-0420-1
62. Babuska AAK. On the angle condition in the finite element method. *SIAM J Numeric Anal.* (1976) 13:214–26. doi: 10.1137/0713021
63. Fried I. Condition of finite element matrices generated from nonuniform meshes. *AIAA J.* (1972) 10:219–21. doi: 10.2514/3.6561
64. Shewchuk J. *What Is a Good Linear Finite Element? Interpolation, Conditioning, Anisotropy, and Quality Measures (Preprint).* University of California, Berkeley, CA 73:12.
65. Foteinos P, Chrisochoides N. High quality real-time image-to-mesh conversion for finite element simulations. *J Parallel Distributed Comput.* (2014) 74:2123–40. doi: 10.1016/j.jpdc.2013.11.002
66. Wirtz C, Knauth M, Staubert A, Bonsanto M, Sartor K, Kunze S, et al. Clinical evaluation and follow-up results for intraoperative magnetic resonance imaging in neurosurgery. *Neurosurgery.* (2000) 46:1112–22. doi: 10.1097/00006123-200005000-00017
67. Kuhnt D, Becker A, Ganslandt O, Bauer M, Buchfelder M, Nimsky C. Correlation of the extent of tumor volume resection and patient survival in surgery of glioblastoma multiforme with high-field intraoperative MRI guidance. *Neuro-Oncology.* (2011) 13:1339–48. doi: 10.1093/neuonc/nor133
68. Miga M, Roberts DW, Kennedy FE, Platenik LA, Hartov A, Lunn KE, et al. Modeling of retraction and resection for intraoperative updating of images. *Neurosurgery.* (2001) 49:75–84. doi: 10.1227/00006123-200107000-00012
69. Risholm P, Samset E, Talos IF, Wells WM. A non-rigid registration framework that accommodates resection and retraction. *Inf Process Med Imaging.* (2009) 21:447–58. doi: 10.1007/978-3-642-02498-6_37
70. Vigneron LM, Warfield SK, Robe PA, Verly JG. 3D XFEM-based modeling of retraction for preoperative image update. *Comput Aided Surg.* (2011) 16:121–34. doi: 10.3109/10929088.2011.570090
71. Drakopoulos F, Liu Y, Foteinos P, Chrisochoides N. Towards a real time multi-tissue adaptive physics based non-rigid registration framework for brain tumor resection. *Front Neuroinformatics.* (2014) 8:11. doi: 10.3389/fninf.2014.00011
72. Drakopoulos F, Yao C, Liu Y, Chrisochoides N. An evaluation of adaptive biomechanical non-rigid registration for brain glioma resection using image-guided neurosurgery. In: *Computational Biomechanics for Medicine Vol. XI.* Athens: MICCAI (2016).
73. Fedorov A, Chrisochoides N. Tetrahedral mesh generation for non-rigid registration of brain MRI: analysis of the requirements and evaluation of solutions. In: Garimella RV, editor. Proceedings of the 17th International Meshing Roundtable. Pittsburgh, PA: Springer (2008). p. 55–72.
74. Loseille A, Alauzet F. Continuous mesh framework part i: well-posed continuous interpolation error. *SIAM J. Numer. Anal.* (2011) 49:38–60. doi: 10.1137/090754078
75. Todd MJ. *Minimum-Volume Ellipsoids.* Philadelphia, PA: Society for Industrial and Applied Mathematics (2016).
76. Dompierre J, Mokwinski Y, Vallet M-G, Guibault F. On ellipse intersection and union with application to anisotropic mesh adaptation. *Eng Comput.* (2017) 33:745–66. doi: 10.1007/s00366-017-0533-y
77. Schroeder W, Martin K, Lorensen B. *The Visualization Toolkit.* 4th Edn. Clifton Park, NY: Kitware (2006).
78. Khachiyan LG. Rounding of polytopes in the real number model of computation. *Math Oper Res.* (1996) 21:307–20. doi: 10.1287/moor.21.2.307
79. Dapogny C, Dobrzynski C, Frey P. Three-dimensional adaptive domain remeshing, implicit domain meshing, and applications to free and moving boundary problems. *J Comput Phys.* (2014) 262:358–78. doi: 10.1016/j.jcp.2014.01.005
80. Garlapati R, Joldes G, Wittek A, Lam J, Weisenfeld N, Hans A, et al. Objective evaluation of accuracy of intra-operative neuroimage registration. In: Wittek A, Miller K, Nielsen PMF, editors. *Computational Biomechanics for Medicine: Models, Algorithms and Implementation.* New York, NY: Springer (2013). p. 87–99.
81. Canny J. A computational approach to edge detection. *IEEE Trans Pattern Anal Mach Intell.* (1986) 8:679–98. doi: 10.1109/TPAMI.1986.4767851
82. Commandeur F, Velut J, Acosta O. A vtk algorithm for the computation of the hausdorff distance. *VTK J.* (2011). Available online at: http://hdl.handle.net/10380/3322

83. Hastreiter P, Rezk-Salama C, Soza G, Bauer M, Greiner G, Fahlbusch R, et al. Strategies for brain shift evaluation. *Med Image Anal.* (2004) 8:447–64. doi: 10.1016/j.media.2004.02.001
84. Wu J, Zhou L, Tang W, Mao Y, Hu J, Song Y, et al. Clinical evaluation and follow-up outcome of diffusion tensor imaging-based functional neuronavigation: a prospective, controlled study in patients with gliomas involving pyramidal tracts. *Neurosurgery*. (2007) 61:935–48; discussion: 948–9. doi: 10.1227/01.neu.0000303189.80049.ab
85. Kekhia H, Rigolo L, Norton I, Golby A. Special surgical considerations for functional brain mapping. *Neurosurg Clin N Am.* (2011) 22:111–32; vii. doi: 10.1016/j.nec.2011.01.004
86. Tempany C, Jayender J, Kapur T, Bueno R, Golby A, Agar N, et al. Multimodal imaging for improved diagnosis and treatment of cancers. *Cancer*. (2015) 121:817–27. doi: 10.1002/cncr.29012
87. Nimsky C, Ganslandt O, Hastreiter P, Wang R, Benner T, Sorensen AG, et al. Intraoperative diffusion-tensor MR imaging: shifting of white matter tracts during neurosurgical procedures–initial experience. *Radiology*. (2005) 234:218–25. doi: 10.1148/radiol.2341031984
88. Chrisochoides N, Fedorov A, Kot A, Archip N, Black P, Clatz O, et al. Toward real-time image guided neurosurgery using distributed and grid computing. In: *SC 2006 Conference, Proceedings of the ACM/IEEE.* (Tampa, FL) (2006). p. 37.

Disclaimer: The content is solely the responsibility of the authors and does not necessarily represent the official views of the government agencies that support this work.

Conflict of Interest: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

frontiers in Virtual Reality

PERSPECTIVE
published: 19 March 2021
doi: 10.3389/frvir.2021.611740

Check for updates

A Process-Model for Minimizing Adverse Effects when Using Head Mounted Display-Based Virtual Reality for Individuals with Autism

Matthew Schmidt[1]*, *Nigel Newbutt*[1], *Carla Schmidt*[1] *and Noah Glaser*[2]

[1]Educational Technology Program, University of Florida, Gainesville, FL, United States, [2]Educational Psychology, University of Connecticut, Storrs, CT, United States

Interest in the use of virtual reality technologies for individuals with autism spectrum disorders has been increasing for over two decades. Recently, research interest has been growing in the area of head mounted display-based virtual reality technologies, thanks to increased availability and affordability. Affordances and theorized benefits of headset-based virtual reality for individuals with autism spectrum disorders are quite promising. However, very little attention has been given in the literature to implementation safety and ethics. This is a particular concern in light of documented adverse effects associated with headset-based virtual reality. To approach this gap, this article details how the authors approached the issue of minimizing adverse effects with related and overlapping methods, but from two separate, independent research sites—one in the United States and one in the United Kingdom. A structured within- and across-case analysis of the two independent studies was conducted to identify central implementation processes and procedures. Analysis resulted in development of a model for minimizing potential adverse effects of headset-based virtual reality for this population. We assert that our model could provide clarity in terms of design and implementation of headset-based virtual reality for individuals with autism spectrum disorders, guide implementations of future researchers and practitioners, and contribute to minimizing and controlling for potential adverse effects.

Keywords: autism, virtual reality, head-mounted displays, implementation, adverse effects, cybersickness

OPEN ACCESS

Edited by:
Daphne Economou,
University of Westminster,
United Kingdom

Reviewed by:
Darryl Charles,
Ulster University, United Kingdom
Christopher M. Laine,
University of Southern California,
United States

***Correspondence:**
Matthew Schmidt
matthew.schmidt@coe.ufl.edu

Specialty section:
This article was submitted to Virtual Reality in Medicine, a section of the journal Frontiers in Virtual Reality

Received: *29 September 2020*
Accepted: *28 January 2021*
Published: *19 March 2021*

Citation:
Schmidt M, Newbutt N, Schmidt C and Glaser N (2021) A Process-Model for Minimizing Adverse Effects when Using Head Mounted Display-Based Virtual Reality for Individuals with Autism. Front. Virtual Real. 2:611740. doi: 10.3389/frvir.2021.611740

INTRODUCTION

Researchers have been exploring the use of head-mounted displays (HMDs) and virtual reality (VR) in the area of autism spectrum disorders (ASD) since Strickland (1996) seminal work over two decades ago. Recently, interest and research activity have been increasing substantially due to the availability of affordable, consumer-grade HMD-based VR systems. ASD is a lifelong condition characterized by challenges with social communication/interaction and restricted, repetitive patterns of behavior (American Psychiatric Association, 2013; DSM-5 American Psychiatric Association, 2013). Manifestations of ASD present substantial heterogeneity across affected individuals (Masi et al., 2017). Early intervention has been shown to lead to improvements in core autism symptoms (Estes et al., 2015). However, access to and consistency of services in early intervention are known barriers, potentially leading to challenges in adulthood such as chronic underemployment, social isolation, and inability to live independently (Eaves and Ho, 2008; Taylor and Seltzer, 2011; Taylor and Mailick, 2014; Hedley et al., 2017). Prevalence of autism is increasing, with current estimates suggesting one in 54 children receiving an ASD diagnosis in the United States alone (Russell et al., 2014). As such, viable and effective therapeutic interventions are in high demand. Virtual reality has garnered significant attention as a potentially effective solution for delivering such interventions (Bellani et al., 2011; Parsons, 2016).

Virtual reality is considered to hold tremendous promise for individuals with ASD due to: (1) the suitability and natural affinity to computers and technology for this population; (2) the ability of HMD-based VR to simulate real-world situations and contexts in safe and predictable ways; and (3) the ability to control, shape, and tailor VR interventions to participant needs and local contexts (Parsons, 2016; Bozgeyikli et al., 2018). Therefore, VR appears to be intrinsically reinforcing for individuals with ASD, has technological affordances which align closely with instructional needs, and can broaden access to a range of services (Parsons et al., 2020). VR provides individuals with ASD the ability to encounter and practice skills in highly realistic and customizable contexts, for input stimuli to be intentionally manipulated, and to have real-world consequences mitigated or removed. Within this context, a patchwork of research studies have investigated VR-based interventions across multiple domains such as social/emotional skills (Moore et al., 2005; Ke and Im, 2013), safety skills (Self et al., 2007), street-crossing (Josman et al., 2011), daily living (Parsons et al., 2004; Mitchell et al., 2007; Jarrold et al., 2013), and communication (Kandalaft et al., 2013) to name but a few (Wang and Anagnostou, 2014; Mesa-Gresa et al., 2018). While reported research outcomes in this area suggest VR has promise for individuals with ASD, as a whole, research supporting connections between theorized benefits of VR HMDs and empirical evidence supporting effectiveness has been described as limited, piecemeal, and largely inconclusive (Parsons, 2016).

To date, the majority of research has focused on the use of desktop-based VR and virtual worlds. More recently, research interest in HMD-based VR for individuals with ASD has increased, in part due to increased consumer availability of HMDs. Researchers are particularly interested in implementing HMD-based VR for individuals with ASD, as specific affordances of the technology such as heightened sense of presence, immersion and ecological validity all lend themselves to a meaningful environment that potentially can enable opportunities for acquisition, maintenance, and generalization of skills (Freina and Ott, 2015). However, a widely-acknowledged concern related to using HMD-based VR is the potential that users will experience some degree of adverse effects (Palmisano et al., 2017). By adverse effects, we refer to negative effects including cybersickness (e.g., nausea, eye-strain, etc.), safety concerns (e.g., stumbling and falling), increased anxiety, sensory disturbances, etc. (cf. Kellmeyer, 2018). Although an abundance of calls for future research to investigate adverse effects can be found in the literature (e.g., Irish, 2013; Fletcher-Watson, 2014; Wong et al., 2015; Mesa-Gresa et al., 2018), actual research on adverse effects is largely absent. Research in the field has yet to meaningfully and systematically approach questions related to adverse effects for individuals with ASD when applying VR in general, and HMD-based VR specifically (Bradley and Newbutt, 2018; Malihi et al., 2020).

The majority of research on VR interventions for individuals with ASD has been principally concerned with establishing intervention effects. However, researchers' desire to establish intervention effects problematically positions VR technology as the primary driver of intervention outcomes. To-date, no design heuristics or guidelines have been published in this area. Although a variety of heuristics and design principles exist for developing 2D computer interfaces for individuals with ASD (Benton et al., 2011; Khowaja and Salim, 2013), far less is known about how to design 3D interfaces and environments: "[. . .] there is no well-established literature on the best practices in designing VR user interface attributes for individuals with ASD yet" (Bozgeyikli et al., 2018, p. 22:5). With a dearth of guidance in terms of VR design best practices for individuals with ASD, researchers are broadly left to their own devices when confronting how to design effective interventions. The lack of practical, detailed, and considered processes to help minimize possible adverse effects in published research suggests a gap in the literature that could limit the full potential of HMD-based VR for individuals with ASD.

Alongside the lack of guidance for designing 3D interfaces and environments for individuals with ASD in general, there is also a lack of understanding regarding how these interfaces and environments might lead to potential adverse effects when experienced in HMDs specifically. That adverse effects have not been a particular focus within the field poses a potential issue around researchers adopting and using emergent technologies like HMD-based VR before their implications of use are understood or before evidence-based recommendations for implementation are established. Borrowing from Kellmeyer (2018), "very little systematic discussion of the neurophilosophical and ethical challenges from the clinical use of these new VR systems is available" (p. 2). Given the limited research in the field, a significant gap exists related to exploring and reporting potential adverse effects (Bradley and Newbutt, 2018). We argue that with rapid technological advancements in this area, researchers, practitioners, and other professionals working with people with ASD have an urgent responsibility to confront these issues (Parsons, 2016).

Thus far we have established the three following research gaps: (1) consideration of adverse effects of VR for individuals with ASD in published research is limited; (2) published design heuristics and principles related to VR for individuals with ASD are largely absent; and (3) there is a paucity of research that explicitly explores and reports potential adverse effects for this vulnerable population. The purpose of this article, therefore, is to approach these gaps by offering a set of preliminary implementation guidelines for HMD-based VR to minimize potential adverse effects for individuals with ASD. Our work is intended to guide researchers and practitioners alike, and therefore serves to bridge a research-to-practice gap in this area.

We turn to our own practice to assert our design expertise in this area, establish our positionality, and frame the issue from the perspective of expert recommendations and best practices, as informed by ethical guidelines. We, the authors, are among the first to have pursued formal research studies with HMD-based VR for individuals with ASD. Our research was performed in two separate, independent research groups. We have published and presented widely on VR for ASD and iteratively have refined our implementations (e.g., Schmidt et al., 2012; Newbutt et al., 2016; Newbutt et al., 2017; Glaser and Schmidt, 2018; Newbutt, 2019; Schmidt et al., 2019; Newbutt et al., 2020). Over the course of multiple studies and through working with stakeholders, we have developed principles and expert knowledge that inform our professional practice This has been achieved through a process of reflection-on-action that considers a range of critical issues and factors directly affecting the acceptance, perceived utility, implementation quality, and participant well-being of VR experiences for this population.

MATERIALS AND METHODOLOGY

This article seeks to address a research-to-practice gap of insufficient implementation guidelines for minimizing potential adverse effects of using HMD-based VR for individuals with ASD. Our specific aims are (a) to explore how two independent research teams sought to minimize adverse effects for research participants in two separate studies and (b) to distill the lessons learned from case comparison into actionable guidelines that might be used by researchers and practitioners in their local contexts. On this basis, and alongside the gaps in knowledge presented previously, the following research questions guided our inquiry:

1. What are the common elements and/or key points of differentiation across studies?
2. What themes emerge from comparison of the individual studies performed at Study 1 and Study 2?
3. What lessons can be drawn from within- and across-case comparison related to minimizing adverse effects of HMD-based VR for individuals with ASD?

To approach these questions, we analyzed two recently published research studies (Newbutt et al., 2020), associated artifacts, and documentation so as to extract and synthesize salient details of how researchers implemented HMD-based VR in their studies. The first and second authors were the lead researchers. Case study methodology was used to perform a structured within- and across-case analysis of implementation procedures with the goal of promoting analytic generalizations (Yin, 2017). A "case" was defined as one of the two independent research studies led by the first and second author.

The USA-based study led by the first author (Schmidt et al., 2019) utilized a design-based research (DBR) methodological approach that considered ease-of-use, the nature of participants' user experience, feasibility, and relevance of a spherical, video-based virtual reality (SVVR) app as well as a fully immersive, collaborative 3D virtual environment, both of which were delivered using HMDs for a group of adults with ASD in a day program. The UK-based study led by the second author (Newbutt et al., 2020) sought to place children on the autism spectrum at the center of a study examining the potential of VR HMDs used in primary and secondary school classrooms; explored a range of VR experiences, how autistic users reported physical experiences, enjoyment, and potential of VR HMDs in their classrooms, while systematically exploring potential adverse effects.

Our work is presented as an instrumental, collective case study (Thomas, 2015). Our analytic process involved describing each case in detail and then presenting themes within the case, followed by thematic analysis across cases. As articulated in **Figure 1**, we adopted a three-phase approach. Phase 1 consisted of structured critical discussions in which we reviewed and summarized research details, methods, and findings, and established a basis for across-case comparison. In phase 2, artifacts were systematically discussed and reflected upon through formal meetings in which authors interrogated one another's work to identify themes. Finally we synthesized lessons learned (Yin, 2017) into a framework that captured best practices.

Trustworthiness

We implemented a range of techniques to ensure the methodological rigor of our research from Lincoln and Guba (1985) recommendations for promoting trustworthiness. Most important to trustworthiness is establishing credibility (Shenton, 2004), which we sought to accomplish in the current case study in three ways. We aimed to promote transferability, which according to Tobin and Begley (2004), refers to the generalizability of research from one case to another. Thick description of our separate contexts is provided in the results of phase 1 below. By presenting similar findings from two studies, but in two separate settings, we provide a more inclusive overall picture. We also aimed to establish dependability and confirmability by maintaining an audit trail. We provide in the results section below a clear and logical description of our research process. We report our process in detail so that it can be repeated in future work. In this way, readers could adopt the research design we report here as a type of prototype model (Shenton, 2004). To allow the traceability of our work in a stage-wise process, we used software that logged all changes to our work over time. Our audit trail consists of versioned files that include step-by-step histories of the methods and findings presented in this paper.

RESULTS

In the following sections, we present the outcomes of our within- and across-case analysis according to three separate phases of analysis. We present these phases in alignment with our research questions, beginning with detailed case descriptions, followed by thematic analysis and synthesis of lessons learned.

Results of Phase 1: Common Elements and Key Points of Differentiation Across Studies

In this section, we articulate each case's details, background, and context related to: (a) equipment, participants, and virtual

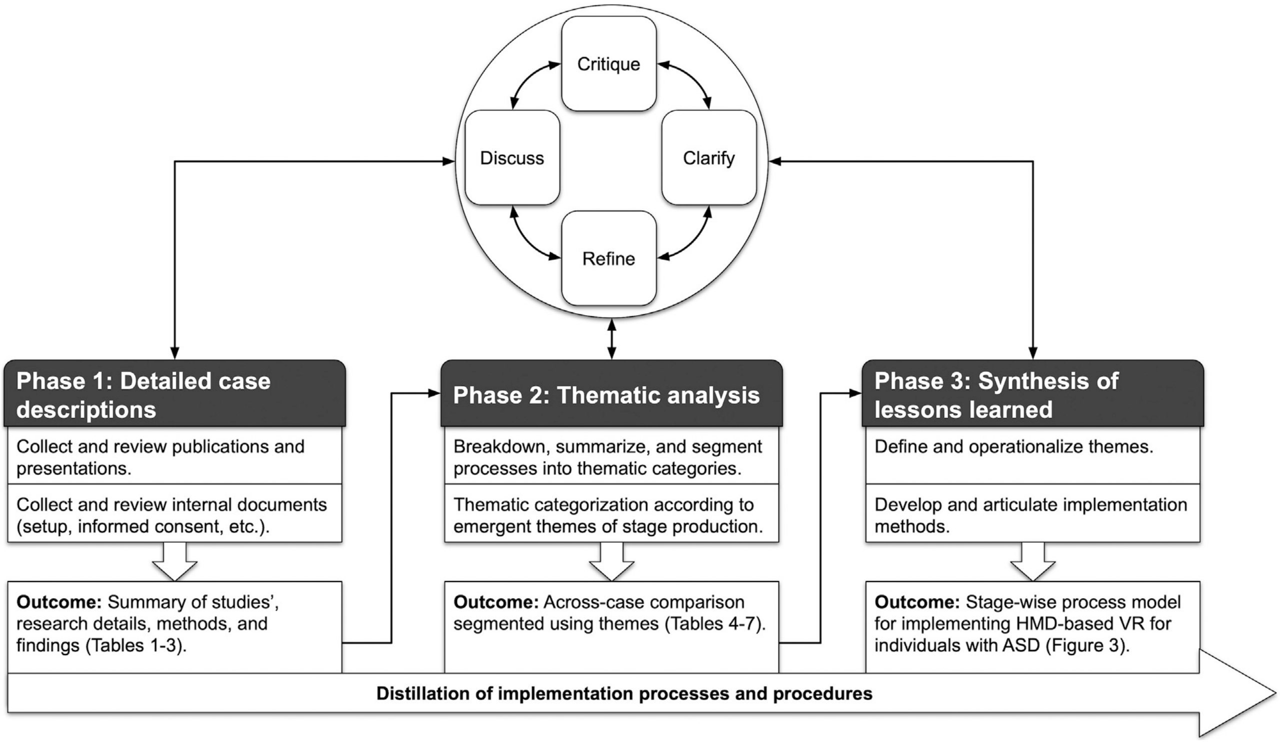

FIGURE 1 | Three-phase structured-iterative-comparative methodology used in this research.

TABLE 1 | Outcomes of Phase 1: Results from analysis of studies (equipment, participants, and virtual environments).

	Study 1	Study 2
Participants	● Expert Testers (neurotypical) ○ *N* = 4 (male = 1; female = 3) ○ 3 PhD-level experts; one staff member ○ Purposive sample ● ASD Participant Testers ○ *N* = 5 (all male, all verbal) ○ Confirmed ASD diagnosis ○ Age range = 23 to 34 Convenience sample	● *N* = 31 (male = 25; female = 6) ● Confirmed ASD diagnosis ● Verbal = 28; non-verbal = 3 ● Age range = 6 to 16 Convenience sample
Apparatus	● Lower tech ○ Google Cardboard ○ Google Daydream View ● Higher tech ○ HTC Vive	● Lower tech ○ Google cardboard ○ ClassVR[a] ● Higher tech ● HTC Vive
Virtual environment	● Bespoke, custom-developed software ● For lower tech: Single-user Android 360° video-based public transportation training simulation ○ For higher tech: Multi-user, avatar-based VR public transportation simulation modeled on adult day program university setting	● Off-the-shelf software ● Historically-focused environments and landscapes (i.e., Egypt) ● Socially-focused environments (i.e., Fun Fair) ● Simulation experience (i.e., making a coffee in a coffee shop) ● All experiences single-user

[a]*ClassVR is a low-cost, stand-alone VR headset designed specifically for fleet deployment in educational contexts.*

environments (**Table 1**); (b) research aims, data collection protocols (**Table 2**); and (c) key findings and limitations (**Table 3**). Study 1 study utilized a design-based research (DBR) methodological approach that considered ease-of-use, the nature of participants' user experience, feasibility, and relevance of a spherical, video-based virtual reality (SVVR) app as well as a fully immersive, collaborative 3D virtual environment, both of which were delivered using HMDs for a

TABLE 2 | Outcomes of Phase 1: Results from analysis of studies (research aims, data collection protocols).

	United States study	United Kingdom study
Ethical review and informed consent	● Research approved by the Universities' Ethics Committee/Institutional Review Board ● Parent and/or participant consent obtained ● Participant assent obtained as/where required	
Purpose and research questions	● To what extent do project prototypes meet design goals of being acceptable, feasible, easy to use, and relevant to the unique needs of participants? ● What is the nature of participants' user experience relative to the lower-tech and higher-tech simulations?	● What type of VR HMD device and experiences therein are preferred by children on the autism spectrum? ● How do children on the autism spectrum report the physical experience, enjoyment, and potential of VR HMDs in their classrooms? ● What would children on the autism spectrum like to use VR for in schools?
Data collection procedures	● Expert Testers (Spring, 2018) ○ Experienced public transportation training simulation in two low tech HMDs for a period of up to 15 min (each) across two sessions at their workplace ○ Sessions facilitated by trained graduate student ○ Experts completed post-experience survey(s) and structured interviews ● Participant Testers (Summer, 2018) ○ Experienced public transportation training simulation with one low tech device and one high tech device for a period of up to 15 min (each) at the offices of the adult day program ○ Sessions facilitated by lead researcher and trained graduate student ○ Participants completed post-experience survey(s) and unstructured interviews ○ An adult day program staff member who knew the participants well helped facilitate this	● Participants tested all three HMDs for a period of up to 20 min (each) across two sessions in their school (Spring 2018) ● Sessions facilitated by lead researcher and teacher who knew the pupil well ● Participants exposed to virtual environments via Google Cardboard, Class VR, and HTC Vive (in that order) ● Participants and teachers completed post-experience survey(s) ● Teachers helped to facilitate this for younger pupils and pupils with learning difficulties
Data collected	● Expert testers ○ Semi-structured interview ○ Self-report questionnaire (designed by research team) ● Participant testers ○ Screen, webcam, and audio recordings ○ Unstructured, post-usage testing interviews ○ Qualitative field notes ● Both ○ System usability scale (SUS) ○ Adjectival user-friendliness scale	● Self-report questionnaires (designed by research team); completed by both ASD and teacher cohorts ● Qualitative field notes ● Focus group protocol

TABLE 3 | Outcomes of Phase 1: Results from analysis of studies (key findings and limitations).

	Study 1	Study 2
Key findings	● Above-average usability and good user-friendliness ● All users were able to complete sessions (but not without help); one returned for additional session ● All participants indicated desire to use again ● Sense of enjoyment (i.e., joy, fun, excitement) ● Some evidence of adverse effects (i.e., nausea, dizziness)	● A preference for using higher tech. A clear preference towards using/deploying VR HMDs for preparing for real life activities and visiting new places (vs. socialising and making friends) ● Preference for using VR to help calm and relax was a key opportunity for participants and teachers ● All participants indicated desire to use again ● Sense of enjoyment and ease of use ● No adverse effects
Limitations	● Small sample ● Exploratory nature of the research ● Virtual environments were functional prototypes, not finished products ● Findings contextualized within adult day program and associated participants	● Small sample ● Exploratory nature of the research ● Exploration of VR software constrained by time ● Findings contextualized within UK schools included in the study and children in those schools

group of adults with ASD in a day program. Study 2 sought to place children on the autism spectrum at the center of a study examining the potential of VR HMDs used in primary and secondary school classrooms; explored a range of VR experiences, how autistic users reported physical experiences, enjoyment, and potential of VR HMDs in their classrooms, while systematically exploring potential adverse effects. Detailed case descriptions are presented in **Table 1**. Cases are presented in a side-by-side tabular format to align each study's research details, methods, findings, etc. for thematic analysis in Phase 2.

TABLE 4 | Results from thematic analysis: Stage 1 setting the stage.

Stage 1—Setting the stage			
Processes		**Common elements**	**Key points of differentiation**
United States Needs assessment, stakeholder input/ meetings, design & development of intervention, IRB approval, expert testing (autism experts, usability experts).	United Kingdom Ethics approval, stakeholder meetings, formation of RQs, school visits, show and tell with parents/ carers, practical issues related to tech in classrooms explored, research mentor meetings with lead researcher.	IRB/ethics approval; needs assessment, stakeholder input/meetings; expert testing (teachers, program managers, collaborators, etc…); designing/revising RQs with stakeholders; design/development of intervention and/or technology integration procedures.	Context and setting.

TABLE 5 | Results from thematic analysis: Stage 2 Dress Rehearsal.

Stage 2—Dress rehearsal			
Processes		**Common elements**	**Key points of differentiation**
United States Internal usage testing; hardware/software stability testing, locating ideal room; rehearsing procedures; developing setup/breakdown checklist; iteratively refining intervention protocol.	United Kingdom Testing equipment in situ; locating ideal room; hardware/software stability testing; demo with teachers and parents; develop set-up checklist; final consultation; finalize protocol.	Testing equipment (hardware/software) in contexts/conditions of actual usage; including facilitators (teachers, staff, etc.); testing and refining protocols.	Age of participants.

TABLE 6 | Results from thematic analysis: Stage 3 First Preview.

Stage 3—First preview			
Processes		**Common elements**	**Key points of differentiation**
United States Present equipment to participants (see and touch devices), explore if participants are interested in using equipment, address any questions or concerns, preview the content, ask if participants want to participate, gain consent	United Kingdom Present equipment to participants (a chance to see and touch devices), discuss any worries/ issues, watch video materials related to others using VR HMDs, preview and consider the content to be presented to participants, discuss any further/final worries concerns, gain consent from parents and pupils	Presentation of equipment and stimuli (in non-VR, screen-based format), continually address questions and queries from participants, complete consent process	Adult vs. child consent/ assent procedures

Results of Phase 2: Emergent Themes from Comparison of Studies

On the basis of the detailed case descriptions from Phase 1 (**Tables 1**–**3**), Phase 2 sought to present themes within and across cases. Results are framed within themes that emerged as lead researchers compared and analyzed cases. Over the process of thematic analysis, the researchers identified common segmentation points shared between cases. Seeking to sufficiently capture the complexity and highly contextual nature of our cases, the metaphor of stage production emerged. Stage production is particularly apt for characterizing our approach, as it seeks to gradually lead to a refined and reliably consistent performance that involves a range of stakeholders working together. As such, we characterize our implementations' segmentation points as "stages", and articulate these as: (a) setting the stage (**Table 4**); (b) dress rehearsal (**Table 5**); (c) first preview (**Table 6**); and (d) opening night (acts 1 and 2; **Table 7**). Within each stage, we briefly detail the processes underwent by each research team and highlight common elements and key points of differentiation. In line with our instrumental, collective case study methodology, the approach of structuring processes across cases in a stage-wise manner (see **Tables 4**–**7**) established the foundation for further interrogation and analytic generalization in Phase 3.

Results of Phase 3: Lessons from Within- and Across-Case Comparison Related to Minimizing Adverse Effects of HMD-Based VR for Individuals with ASD

Results presented in this section represent the final outcome of the 3-phase distillation procedure. Having identified points of

TABLE 7 | Results from thematic analysis: Stage 4 Opening Night (Acts 1 and 2).

Stage 4—Opening Night: Act 1			
Processes		**Common elements**	**Key points of differentiation**
United States Ensure participant is sitting in swivel chair, explain the task structure while showing how to navigate the app, ask if participant is comfortable and ready to begin, fit the headset (Cardboard or Daydream), ask if participant is comfortable and willing to continue, ask participant to complete first task (2 min), provide assistance if needed, ask how participant is feeling and if comfortable and ready to begin next task, repeat after each task (*n* = 4) until activity is completed	United Kingdom Explain how the session will run, sit participant down, introduce first HMD experience using low tech device (Cardboard), check they still want to proceed, fit the HMD, run application (2–3 min) then remove HMD, check for negative effects (ask participant and observe), run further HMD experience, complete activity (<5 min), check for negative effects, follow up 30 min later for signs of negative effects	Deliberate orientation and training process; ensure participant is sitting; initial trial; continually check participants willingness to use at set intervals; continually check for adverse effects; short lengths of exposure	Follow up on negative effects beyond the formalities of the study
Stage 4—Opening Night: Act 2			
Processes		**Common elements**	**Key points of differentiation**
United States Ensure participant is sitting in swivel chair, show HTC Vive controllers, show how to hold controllers, ask participant to press different buttons on controllers, ask participant if they are comfortable and ready to try on the HTC Vive HMD, fit the HMD, ask participant if it is comfortable and they are willing to continue, have participant perform simple navigation task (find their cubicle in the office; 1-2 min), provide assistance if needed, ask participant if task was easy or difficult, ask how participant is feeling, ask if participant is comfortable and willing to continue, engage in full activity (~9 min), check for negative effects	United Kingdom Explain how the session will run, researcher uses HTC Vive to demonstrate activity (and controls), sit/stand participant (depending on their age), check they still want to proceed, fit the HMD and hand controls, ensure comfort, run application (2–3 min) then remove HMD, follow the participant around the space/room (if standing/walking) to ensure stable balance, check for negative effects (ask participant and observe), run further HMD experience, complete activity (<5 min), check for negative effects, follow up 30 min later for signs of negative effects	Deliberate orientation and training process; initial trial; continually check participants willingness to use at set intervals; continually check for adverse effects; provide assistance	Length of exposure; standing vs. sitting; multi-user vs. single-user virtual environment

convergence and divergence across cases in phase 2, our focus transitioned in phase 3 to integrating the processes from both studies (**Tables 4–7**). Considerations of how to convey the lessons learned from prior analysis presented an opportunity to merge the processes and common elements from both cases into a representation of best practices. This required a translation of the specific details of each case into a set of generalizations that described each of the implementation stages. Attempts to develop textual descriptions of phases led to tensions, as text-based descriptions did not sufficiently represent the linear, step-wise nature of the implementation processes uncovered in our analysis. This led us to develop an initial linear visualization of our processes, which we subsequently refined using the approach shown in **Figure 1** above. As a result, we articulated the stage-wise implementation framework as a linear process with descriptions and intent of each stage. Resonant in this process was the concept of gradual acclimation, which we discuss as follows.

Gradual Acclimation for Minimizing Adverse Effects

To minimize potential adverse effects *gradual acclimation* is needed (**Figure 2**). This is the process through which VR is continually optimized within and across implementation stages. We articulate gradual acclimation in two different areas of the process model: (1) *contextual acclimation* (Stages 1 and 2) and (2) *technology acclimation* (Stages 3 and 4). *Contextual acclimation* is concerned with researchers identifying the local variables that potentially could impact successful implementation. Contextual acclimation is primarily informed via input from stakeholders and facilitators (e.g., teachers, staff, program managers, collaborators, directors). Contextual acclimation is followed by *technology acclimation*, which is informed principally by participants when they begin by handling and experiencing the technology in a gradual manner in stages 3 and 4.

Gradual acclimation is, in part, predicated on ethical concerns. When working with a vulnerable population, researchers have a special obligation and greater responsibility to actively take precautions to help minimize real or potential risks (American Educational Research Association, 2011; Behavior Analyst Certification Board, 2014). This extends to possible adverse effects associated with the use of HMD-based VR for individuals ASD. We therefore assert that the potential for adverse effects could be minimized through intentional procedures that promote gradual acclimation within and

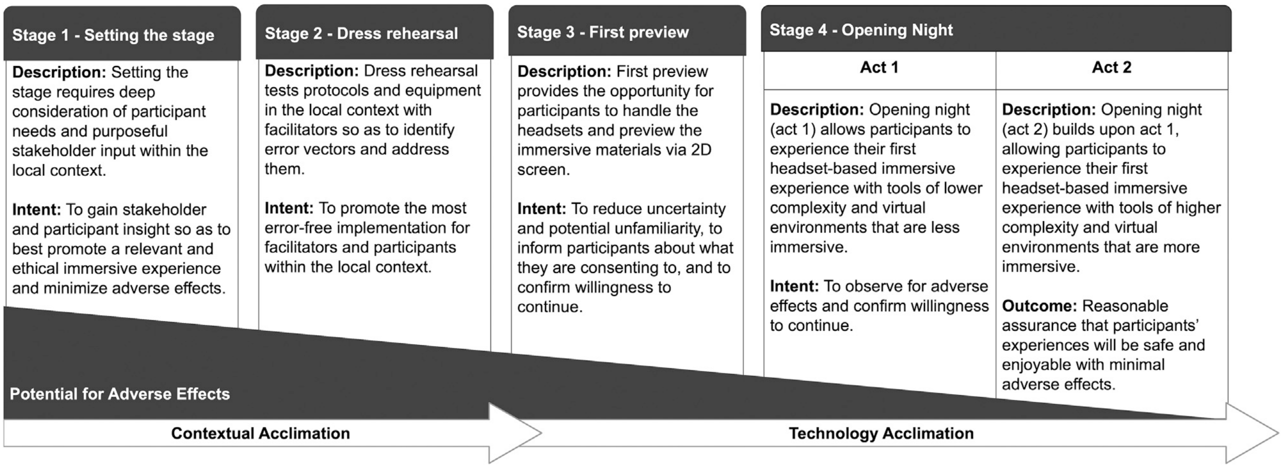

FIGURE 2 | Final stage-wise process model promoting gradual acclimation to minimize potential adverse effects of HMD-based VR for individuals with ASD.

across all stages. This is captured in **Figure 2** as a triangle labeled *Potential for adverse effects*, which illustrates how the potential for adverse effects could be reduced by utilizing the intentional and systematic processes through *contextual acclimation* and *technology acclimation*. **Figure 2** represents our final stage-wise process model of gradual contextual and technology acclimation to minimize potential adverse effects of HMD-based VR for individuals with ASD.

DISCUSSION

Auspicious claims championing the potential of VR for individuals with ASD are abundant in the literature of the past 25 years—from the earliest reports (e.g., Strickland, 1996; 1997) to the most recent (Miller et al., 2019; Ke et al., 2020). However, "While it is apparent that VR has great conceptual potential, research and development are early on the road to making it a reality" (Wang and Anagnostou, 2014, p. 2137). Contributing to this are the many gaps in this research area, including a lack of guidelines for implementing VR interventions with this population, a general lack of consideration of adverse effects of VR in the published research, and very little specific research that reports findings related to adverse effects. Given the known characteristics of people with ASD (Christensen et al., 2018), and compounded by common comorbidities (Deprey and Ozonoff, 2018) such as anxiety (van Steensel and Heeman, 2017) and sensory processing differences (Thye et al., 2018), the prospect of adverse effects related to VR usage for individuals with ASD presents salient risks. Approaching calls and concerns for the safe and ethical applied use of technology for individuals with autism (Parsons, 2016; Bradley and Newbutt, 2018; Schmidt et al., 2019), the current article approaches the urgent need to consider potential adverse effects of HMD-based VR use for individuals with ASD.

To confront this need, we present a stage-wise process model derived from our own research that illustrates how researchers and practitioners alike can seek to reduce potential adverse effects in the design and execution of their HMD-based VR implementations for individuals with ASD. To contextualize the embedded processes and promote recall, we adopt the metaphor of a stage production to describe the different stages of the process. Like stage production, VR implementation requires considerable attention to details, many of which are unknown until they are encountered. Successful VR implementation also requires intentional, continuous coordination with key stakeholders such as teachers, administrators, IT staff, etc. (Newbutt, 2019), similar to a stage production requiring coordination between theater and production staff (e.g., stage manager, stagehands). Further, and perhaps most importantly, planning opportunities for participants to encounter VR technology and practice using it before engaging in an actual intervention is as important to VR implementation as is memorization of lines and table-reads to actors' performance in a stage production. Finally, in stage production, preparation and rehearsal are critically important to bringing the production to a high standard of quality. We posit that application of the stage-wise model presented here could likewise influence the quality of the VR experience for participants by controlling for potential adverse effects.

Stage-Wise Process Model

The stage-wise process model represents a distillation of how two separate research teams approached the challenge of controlling for potential adverse effects. Our intent here is to provide guidance in an approachable, actionable framework and in doing so to portray our model in a minimalist manner so as to be useful to a range of audiences. However, this introduces a risk that it could be interpreted as an oversimplification of what in many cases can be exceptionally complex processes. The reader is therefore cautioned not to interpret the model at face value, but instead to consult the specific information presented in **Tables 4–7** for detailed examples of the specific procedures and activities that were applied by using each stage of our work. We provide further discussion of the stages in the following sections.

Stage 1: Setting the Stage

Stage 1 includes front-end analysis and subsequent design and development of the intervention. These are activities that are well documented in the learning and instructional design literature (Morrison et al., 2012; Dick et al., 2014). What differentiates our approach from such established instructional design processes, however, is intentional solicitation of collaborative and meaningful input from all stakeholders, beginning in Stage 1, and continuing across all stages of the process. Continual input from stakeholders ensures the voices of those who matter the most are heard (Parsons et al., 2020). We highlight that this goes beyond requirements of institutional review boards or ethics committees for human subjects research (Protection of Human Subjects, 2009). Catalyzing this were critiques rebuking research studies that: (1) do not incorporate the voices and perspectives of individuals with ASD, their support communities, and/or relevant stakeholders in the design, implementation, and/or execution of studies; and (2) do not pursue and incorporate the priorities and cultural values of autistic communities (Parsons, 2016; Parsons et al., 2020). Importantly, teams of researchers and stakeholders should explicitly acknowledge in this stage that VR use and/or intervention has inherent potential for adverse effects. As such, they should actively seek to identify factors that could promote reduction of adverse effects as early as this stage. For example, a team might discover that some participants find the fit of one headset to be disagreeable, which could prompt exploration of more adjustable straps or different headsets altogether.

Stage 2: Dress Rehearsal

In Stage 2, the VR technology used to deliver the intervention is tested *in-situ*. This testing allows teams to collectively examine the space where the technology will be used. When working *in-situ*, researchers could find themselves provided with a space that is not ideal for HMD-based VR application. For example, they could be assigned a room that is too small, a cluttered area, or spaces containing tables and chairs. Such a space can present collision and trip hazards, which could lead to adverse health effects. Further, space needs to be amenable to a positive sensory experience for participants. For example, fluorescent lights or loud noises could be severely distracting for individuals with ASD (Grandin, 2002). In addition to this, testing allows the team to set up, configure and test the technology outside of a controlled lab setting. This allows for discovery of a variety of potential error points, for example, connectivity challenges such as restricted networks and prohibitive firewall rules, logistical challenges such as setting up for multi user experiences, or computer problems, such as misconfigurations and unforeseen glitches. Importantly, dress rehearsal is not technology testing; it is preparing for social-behavioral human subjects research (American Psychiatric Association, 2013). As such, it provides opportunities for stakeholders to see and experience the application of the technology and provide feedback, with the ultimate goal of providing an error-free, enjoyable implementation by identifying and addressing vectors that potentially could lead to adverse effects such as confusion, frustration, etc.

Stage 3: First Preview

In Stage 3, focus shifts to technology usage by participants. In line with the principle of gradual acclimation, this stage provides participants the opportunity to examine the technology and learn about the intervention. Many individuals with ASD respond positively to order, structure, and predictability, with deviations potentially leading to increased stress and anxiety (Gotham et al., 2013; Lidstone et al., 2014; Factor et al., 2016). First Preview seeks to address this by providing structured opportunities to communicate the project to participants. Although research studies are often communicated to research participants through approved informed consent procedures (e.g., American Educational Research Association, 2011; British Educational Research Association, 2018), we have found in our work the language used to describe VR projects (as approved by institutional review boards or ethical committees) can be abstract and contextually dependent. Given that many individuals with ASD are quite literal in their thinking (Volden et al., 2009; Deliens et al., 2018), this can frustrate communication efforts. The purpose of providing the opportunity for examining the technology and previewing the VR materials on a monitor is to provide concrete and real examples of what they will be experiencing, thereby enhancing communication with our participants. In this manner, two key aspects are achieved: (1) we are ensuring our ethical obligation to our participants (i.e., clear and concise communication); and (2) vulnerable participants are able to make better-informed decisions about whether they choose to participate in a given research study and what to expect if they do agree to be involved. This again illustrates the importance of privileging the voices and perspectives of participants throughout the implementation process (Parsons and Cobb, 2013; Politis et al., 2019).

Stage 4: Opening Night Act 1

In Act 1 of Opening Night the equipment and experience is delivered in a limited manner so as to promote gradual acclimation. Examples include: limiting the amount of time using the equipment; using lower tech versions of the equipment; and/or using simplified versions of the equipment. First, limiting the amount of time that participants engage in the VR experience could reduce the probability of serious adverse effects and allows researchers to check frequently with participants regarding how they are feeling, as well as with their guardians, caregivers, educators, etc. For example, researchers might design a series of 2-min tutorials or short 360 videos and take breaks in between to evaluate participants' well-being and willingness to proceed. Second, using lower tech versions of the technology could be less complicated and therefore provide superior usability (Parish-Morris et al., 2018; Schmidt et al., 2019). For example, Google Cardboard devices have only a single button and no straps, allowing for the HMD to be removed easily in the case that a participant is experiencing negative effects. Third and finally, using simplified versions of the equipment, software, etc. could provide authentic experiences while reducing potential frustration or confusion (Rojo et al., 2019). For example, a series of 360 videos might auto-play instead of requiring participants to operate controls, or short tutorials might be presented using an "on rails" design in which participants are limited to moving their avatar, making selections, or changing their view, but not controlling the path their avatar takes. Given that the local context necessarily circumscribes how this stage of implementation is designed, these examples are not meant to be prescriptive but instead to illustrate how gradual acclimation is realized in this stage. Importantly, researchers must take care to intentionally and often question participants and other advocates who are involved regarding their well-being. Given recognized communication differences in many individuals with ASD (Masi et al., 2017), soliciting responses is recommended over providing opportunities for self-initiated reporting. This requires careful discussion and agreement to be sure that, for example, non-verbal participants can express their willingness to continue or desire to end the experience and/or study.

Stage 4: Opening Night Act 2

Act 2 of Opening Night is when participants engage comprehensively in the VR experience using fully immersive headsets (i.e., Oculus Quest, HTC Vive). During this stage, researchers should remain vigilant in their observation of participants while they engage in the experience(s), scrutinizing for any signs of adverse effects. Assuming that the prior stages have been enacted in a considered manner, participants should have gradually acclimatized to the technology, and the likelihood that adverse effects would unfavorably impact participants in this stage should be diminished.

Efforts to minimize risk do not end at this point. Even though a VR intervention might be very feasible for certain applications initially, using the VR intervention might become frustrating over time or introduce new risks. Evidence of this is scant given the lack of longitudinal research in this area. However, emerging research suggests longitudinal exposure effects of VR with individuals with ASD (Glaser et al., 2020), with participants rating their perceptions of cybersickness higher in later sessions and as the VR environment increased in complexity. Therefore, after participants have begun a training program or intervention, vigilance in observation should be ongoing and continual adjustments should be made as needed.

Through the stages presented above and the synthesised data from a US and a UK-based study, we have addressed an established lack of explicit guidance (Newbutt, 2019). On the one hand, research has noted the potential for VR HMDs to induce some negative effects (Sharples et al., 2008; Chessa et al., 2019; Weech et al., 2019), while on the other, very limited advice is offered on how to practically overcome these challenges. This is especially true in the field of autism, in which we are working with individuals who can present with sensory concerns (American Psychiatric Association, 2013). Our results provide initial insights to both fill a gap in the literature and also promote safe HMD-based VR practices for individuals with ASD and their support communities.

Implications

In the current article, we have presented a stage-wise process model promoting gradual acclimation to minimize potential adverse effects of HMD-based VR for individuals with ASD. We do not assert that these procedures are novel of-themselves. Instead, the innovation of our approach rests in how we distilled specific processes from successful research studies which could be used by others potentially to guide their own implementations. Further, our process model intentionally embraces and embodies social disability models, foregrounding themes of dignity, agency, and empowerment (Braddock et al., 2013). By locating the perspectives and voices of individuals with ASD centrally, we seek to avoid the mistake of insufficiently considering neurophilosophical issues such as hyper- or hyporeactivity to sensory inputs that individuals with ASD may experience (Tavassoli et al., 2016; Uljarević et al., 2016). While the impetus for developing the stage-wise process model was to provide guidance for reducing potential adverse effects, benefits of using the process model could represent a generalizable set of guidelines that others could use to inform their VR implementations. Inclusion of stakeholders and participants across all stages of implementation could increase their confidence in potential benefits of using the technology, positively influence potential anxiety on the part of participants, and increase stakeholder confidence that the technology will not cause harm. Designers could map existing methods and processes to the various stages in the model and use this to inform and guide their designs. Ultimately, the process model we provide in **Figure 2** and discuss above could provide some assurance that stakeholders are being supportive of their ethical responsibilities when working with a vulnerable population (Bell, 2008; Pittaway et al., 2010).

Limitations

The research presented in the current article represents an important first step in reporting how researchers approached the problem of potential adverse effects in their implementation procedures. However, researchers seeking to utilize and further develop our approach should be aware of its limitations. The purpose of the research presented here was not to validate the stage-wise process model. While we believe it is likely that following a stage-wise implementation process helped to minimize adverse effects, we do not currently have sufficient data to support a causal relationship between the stage-wise process model and reduction in adverse effects. Because this work drew from completed studies, it was not possible to establish whether our stage-wise implementation processes were specifically related to the general lack of adverse effects observed in both studies. In addition, despite our collaborative approach working with individuals with ASD, we have not yet returned to these communities with our final process model to solicit feedback and input—a direction for future research. Furthermore, to extend the work presented here and approach limitations identified above, future research should explore the empirical relationship between utilization of our process model and influence on adverse effects. Moreover, given the heterogenous nature of ASD, future research should address how to tailor the process model in consideration of factors such as necessary level of support, individual differences (i.e., reactivity to sensory inputs, levels of cognition, etc.), and extended exposure to HMD technologies.

CONCLUSION

The focus in the current article on adverse effects of HMD-based VR for individuals with ASD rests on the basis of ethical guidelines published in our respective research contexts, including but not limited to those of BERA (British Educational Research Association, 2018), AERA (American Educational Research Association, 2011), and BACB (Behavior Analyst Certification Board, 2014). We situate the respect and dignity of our participants at the forefront of our work and incorporate practices that promote inclusion and involvement of participants and stakeholders. Paramount to our work is an acute sensitivity to the vulnerability of our target population (Parsons, 2015). In light of this, we endeavor to honor our special obligation and the greater responsibility we have as researchers to actively take precautions to help minimize real or potential risks associated with the use of HMD-based VR for ASD (e.g., emotional and physical harm or intervention side-effects). These potential risks could represent environmental constraints that we as researchers have an ethical obligation to eliminate. To this end, we have developed methods and processes for minimizing adverse effects of HMD-based VR for individuals with ASD. Our sincere hope is that the work presented here can serve as a signpost for how researchers can ethically approach considerations of adverse effects in future HMD-based VR implementations, and we urge researchers to consider applying and empirically interrogating the model proposed here.

DATA AVAILABILITY STATEMENT

The original contributions presented in the study are included in the article, further inquiries can be directed to the corresponding author.

AUTHOR CONTRIBUTIONS

All authors listed have made a substantial, direct and intellectual contribution to the work, and approved it for publication.

REFERENCES

American Educational Research Association. (2011). *Code of ethics*. Available at: https://www.aera.net/Portals/38/docs/About_AERA/CodeOfEthics(1).pdf

American Psychiatric Association. (2013). *Diagnostic and statistical manual of mental disorders*. Washington, DC: American Psychiatric Association.

Bord, K. (2007). Wicked ID: conceptual framework for considering instructional design as a wicked problem. *Cjlt/Rcat* 33 (1), n1. doi:10.21432/t2cg6h

Behavior Analyst Certification Board (2014). *Professional and ethical compliance code for behavior analysts*. Available at: https://www.bacb.com/wp-content/uploads/2020/05/BACB-Compliance-Code-english_190318.pdf

Bell, N. (2008). Ethics in child research: rights, reason and responsibilities. *Child. Geogr.* 6 (1), 7–20. doi:10.1080/14733280701791827

Bellani, M., Fornasari, L., Chittaro, L., and Brambilla, P. (2011). Virtual reality in autism: state of the art. *Epidemiol. Psychiatr. Sci.* 20, 235–238. doi:10.1017/s2045796011000448

Benton, L., Ashwin, E., Johnson, H., Grawemeyer, B., and Brosnan, M. (2011). *IDEAS: An interface design experience for the autistic spectrum, 1759–1764*. Berlin: Springer.

Bozgeyikli, L., Raij, A., Katkoori, S., and Alqasemi, R. (2018). A survey on virtual reality for individuals with autism spectrum disorder: design considerations. *IEEE Trans. Learn. Technol.* 11 (2), 133–151. doi:10.1109/tlt.2017.2739747

Braddock, D., Hoehl, J., Tanis, S., Ablowitz, E., and Haffer, L. (2013). The rights of people with cognitive disabilities to technology and information access. *Inclusion* 1 (2), 95–102. doi:10.1352/2326-6988-01.02.95

Bradley, R., and Newbutt, N. (2018). Autism and virtual reality head-mounted displays: a state of the art systematic review. *J. Enabling Tech.* 12, 101–113. doi:10.1108/jet-01-2018-0004

British Educational Research Association. (2018). *Ethical guidelines for educational research*. Available at: https://www.bera.ac.uk/researchers-resources/publications/ethicalguidelines-for-educational-research-2018

Chessa, M., Maiello, G., Borsari, A., and Bex, P. J. (2019). The perceptual quality of the oculus rift for immersive virtual reality. *Human-Comput. Interact.* 34 (1), 51–82. doi:10.1080/07370024.2016.1243478

Christensen, D. L., Braun, K. V. N., Baio, J., Bilder, D., Charles, J., Constantino, J. N., et al. (2018). Prevalence and characteristics of autism spectrum disorder among children aged 8 Years - autism and developmental disabilities monitoring network, 11 sites, United States, 2012. *MMWR Surveill. Summ.* 65 (13), 1. doi:10.15585/mmwr.ss6513a1

Deliens, G., Papastamou, F., Ruytenbeek, N., Geelhand, P., and Kissine, M. (2018). Selective pragmatic impairment in autism spectrum disorder: indirect requests versus irony. *J. Autism Dev. Disord.* 48 (9), 2938–2952. doi:10.1007/s10803-018-3561-6

Deprey, L., and Ozonoff, S. (2018). *Assessment of comorbid psychiatric conditions in autism spectrum disorder*. Berlin: Springer.

Dick, W., Carey, L., and Carey, J. (2014). *The systematic design of instruction*. 8th Edn. London: Pearson.

DSM-5 American Psychiatric Association. (2013). *Diagnostic and statistical manual of mental disorders*. New York, NY: American Psychiatric Publishing.

Eaves, L. C., and Ho, H. H. (2008). Young adult outcome of autism spectrum disorders. *J. Autism Dev. Disord.* 38, 739–747. doi:10.1007/s10803-007-0441-x

Estes, A., Munson, J., Rogers, S. J., Greenson, J., Winter, J., and Dawson, G. (2015). Long-term outcomes of early intervention in 6-year-Old children with autism spectrum disorder. *J. Am. Acad. Child Adolesc. Psychiatry* 54 (7), 580–587. doi:10.1016/j.jaac.2015.04.005

Factor, R. S., Condy, E. E., Farley, J. P., and Scarpa, A. (2016). Brief report: insistence on sameness, anxiety, and social motivation in children with autism spectrum disorder. *J. Autism Dev. Disord.* 46 (7), 2548–2554. doi:10.1007/s10803-016-2781-x

Fletcher-Watson, S. (2014). A targeted review of computer-assisted learning for people with autism spectrum disorder: towards a consistent methodology. *Rev. J. Autism Dev. Disord.* 1 (2), 87–100. doi:10.1007/s40489-013-0003-4

Freina, L., and Ott, M. (2015). A literature review on immersive virtual reality in education: state of the art and perspectives. *Virt. Real.* 1, 133. doi:10.1007/s10055-020-00489-9

Glaser, N. J., and Schmidt, M. (2018). Usage considerations of 3D collaborative virtual learning environments to promote development and transfer of knowledge and skills for individuals with autism. *Tech. Know. Learn.* 25 (25), 315–322. doi:10.1007/s10758-018-9369-9

Glaser, N., Schmidt, M., and Schmidt, C. (2020). *Fear and loathing in VR: cybersickness evidence in headset-based VR training for adults with Autism Presented at the 2020 international convention of the association for educational communications and technology*. Online.

Gotham, K., Bishop, S. L., Hus, V., Huerta, M., Lund, S., Buja, A., et al. (2013). Exploring the relationship between anxiety and insistence on sameness in autism spectrum disorders. *Autism Res.* 6 (1), 33–41. doi:10.1002/aur.1263

Grandin, T. (2002). *Teaching tips for children and adults with autism, 2.* Fort Collins: Colorado/EUA, 5.

Grynszpan, O., Weiss, P. L., Perez-Diaz, F., and Gal, E. (2014). Innovative technology-based interventions for autism spectrum disorders: a meta-analysis. *Autism* 18 (4), 346–361. doi:10.1177/1362361313476767

Hedley, D., Uljarević, M., Cameron, L., Halder, S., Richdale, A., and Dissanayake, C. (2017). Employment programmes and interventions targeting adults with autism spectrum disorder: a systematic review of the literature. *Autism* 21 (8), 929–941. doi:10.1177/1362361316661855

Irish, J. E. N. (2013). Can I sit here? A review of the literature supporting the use of single-user virtual environments to help adolescents with autism learn appropriate social communication skills. *Comput. Hum. Behav.* 29 (5), A17–A24. doi:10.1016/j.chb.2012.12.031

Jarrold, W., Mundy, P., Gwaltney, M., Bailenson, J., Hatt, N., McIntyre, N., et al. (2013). Attention in a virtual public speaking task in higher functioning children with autism. *Autism Res.* 6, 393–410. doi:10.1002/aur.1302

Josman, N., Ben-Chaim, H. M., and Friedrich, S. (2011). Effectiveness of virtual reality for teaching street-crossing skills to children and adolescents with autism. *Int. J. Disabil. Hum. Dev. IJDHD* 7 (1), 49–56. doi:10.1515/ijdhd.2008.7.1.49

Kandalaft, M. R., Didehbani, N., Krawczyk, D. C., Allen, T. T., and Chapman, S. B. (2013). Virtual reality social cognition training for young adults with high-functioning autism. *J. Autism Dev. Disord.* 43 (1), 34–44. doi:10.1007/s10803-012-1544-6

Ke, F., and Im, T. (2013). Virtual-reality-based social interaction training for children with high-functioning autism. *J. Educ. Res.* 106 (6), 441–461. doi:10.1080/00220671.2013.832999

Ke, F., Moon, J., and Sokolikj, Z. (2020). Virtual reality–based social skills training for children with autism spectrum disorder. *J. Spec. Education Technol.* 14, 016264342094560. doi:10.1109/fie44824.2020.9273818

Kellmeyer, P. (2018). Neurophilosophical and ethical aspects of virtual reality Therapy in Neurology and Psychiatry. *Camb Q. Healthc. Ethics* 27 (4), 610–627. doi:10.1017/s0963180118000129

Khowaja, K., and Salim, S. S. (2013). A systematic review of strategies and computer-based intervention (CBI) for reading comprehension of children with autism. *Res. Autism Spectr. Disord.* 7 (9), 1111–1121. doi:10.1016/j.rasd.2013.05.009

Lidstone, J., Uljarević, M., Sullivan, J., Rodgers, J., McConachie, H., Freeston, M., et al. (2014). Relations among restricted and repetitive behaviors, anxiety and sensory features in children with autism spectrum disorders. *Res. Autism Spectr. Disord.* 8 (2), 82–92. doi:10.1016/j.rasd.2013.10.001

Lincon, Y. S., and Guba, E. G. (1985). Naturalistic inquiry. Sage: Beverly Hills.

Malihi, M., Nguyen, J., Cardy, R. E., Eldon, S., Petta, C., and Kushki, A. (2020). Short report: evaluating the safety and usability of head-mounted virtual reality compared to monitor-displayed video for children with autism spectrum disorder. *Autism* 24, 1924–1929. doi:10.1177/1362361320934214

Masi, A., DeMayo, M. M., Glozier, N., and Guastella, A. J. (2017). An overview of autism spectrum disorder, heterogeneity and treatment Options. *Neurosci. Bull.* 33 (2), 183–193. doi:10.1007/s12264-017-0100-y

Mesa-Gresa, P., Gil-Gómez, H., Lozano-Quilis, J.-A., and Gil-Gómez, J.-A. (2018). Effectiveness of virtual reality for children and adolescents with autism spectrum disorder: an evidence-based systematic review. *Sensors* 18 (8), 114. doi:10.3390/s18082486

Miller, I. T., Wiederhold, B. K., Miller, C. S., and Wiederhold, M. D. (2019). Virtual reality air travel training with children on the autism spectrum: a preliminary report. *Cyberpsychol. Behav. Soc. Netw.* 23, 10–15. doi:10.1089/cyber.2019.0093

Mitchell, P., Parsons, S., and Leonard, A. (2007). Using virtual environments for teaching social understanding to 6 adolescents with autistic spectrum disorders. *J. Autism Dev. Disord.* 37 (3), 589–600. doi:10.1007/s10803-006-0189-8

Moore, D., Yufang Cheng, P., and McGrath, P. (2005). Collaborative virtual environment technology for people with autism. *Focus Autism Other Dev. Disabl.* 20, 231–243. doi:10.1177/10883576050200040501

Morrison, G. R., Ross, S., Kalman, H., and Kemp, J. (2012). *Designing effective instruction*. 7th Edn. Hoboken: John Wiley.

Newbutt, N., Sung, C., Kuo, H.-J., Leahy, M. J., Lin, C.-C., and Tong, B. (2016). Brief report: a Pilot study of the use of a virtual reality headset in autism populations. *J. Autism Dev. Disord.* 46 (9), 3166–3176. doi:10.1007/s10803-016-2830-5

Newbutt, N., Sung, C., Kuo, H-J., and Leahy, M. J. (2017). "The potential of wearable technology (head-mounted displays) and virtual reality to support people with autism: acceptance, challenges, and future applications," in *Recent advances in technologies for inclusive well-being: from worn to off-body sensing, virtual worlds, and games for serious applications*. Editors A. Brooks, S. Brahnam, and L. Jain (Berlin: Springer), 221–241. Chapter 11.

Newbutt, N., Bradley, R., and Conley, I. (2020). Using virtual reality head-mounted display in schools for pupils on the autism spectrum; views, experiences and future directions. *Cyberpsychol. Behav. Soc. Netw.* 23 (1), 22–33. doi:10.1089/cyber.2019.0206

Newbutt, N. (2019). "Assisting people with autism spectrum disorder through technology," in *Encyclopedia of education and information technologies*. Editor A. Tatnall (Switzerland: Springer).

Palmisano, S., Mursic, R., and Kim, J. (2017). Vection and cybersickness generated by head-and-display motion in the Oculus Rift. *Displays* 46, 1–8. doi:10.1016/j.displa.2016.11.001

Parish-Morris, J., Solórzano, R., Ravindran, V., Sazawal, V., Turnacioglu, S., Zitter, A., et al. (2018). Immersive virtual reality to improve police interaction skills in adolescents and adults with autism spectrum disorder: preliminary results of a phase i feasibility and safety trial. *Annu. Rev. CyberTherapy Telemed.* 14 (16), 50–56. doi:10.1002/aur.2352

Parsons, S., and Cobb, S. (2013). *EPSRC observatory for responsible innovation in ICT. Who chooses what I need? Child voice and user-involvement in the development of learning technologies for children with autism*. Berlin: Springer.

Parsons, S., Mitchell, P., and Leonard, A. (2004). The use and understanding of virtual environments by adolescents with autistic spectrum disorders. *J. Autism Dev. Disord.* 34 (4), 449–466. doi:10.1023/b:jadd.0000037421.98517.8d

Parsons, S., Yuill, N., Good, J., and Brosnan, M. (2020). 'Whose agenda? Who knows best? Whose voice?' Co-creating a technology research roadmap with autism stakeholders. *Disabil. Soc.* 35 (2), 201–234. doi:10.1080/09687599.2019.1624152

Parsons, S. (2015). Learning to work together: designing a multi-user virtual reality game for social collaboration and perspective-taking for children with autism. *Int. J. Child-Comput. Interact.* 6, 28–38. doi:10.1016/j.ijcci.2015.12.002

Parsons, S. (2016). Authenticity in virtual reality for assessment and intervention in autism: a conceptual review. *Educ. Res. Rev.* 19, 138–157. doi:10.1016/j.edurev.2016.08.001

Pittaway, E., Bartolomei, L., and Hugman, R. (2010). 'Stop stealing our stories': the ethics of research with vulnerable groups. *J. Hum. Rights Pract.* 2 (2), 229–251. doi:10.1093/jhuman/huq004

Politis, Y., Olivia, L., and Olivia, T. (2019). Empowering autistic adults through their involvement in the development of a virtual world. *Aia* 5 (4), 303–317. doi:10.1108/aia-01-2019-0001

Protection of Human Subjects (2009). *CFR* 45, 46.

Rojo, D., Mayor, J., Rueda, J. J. G., and Raya, L. (2019). A virtual reality training application for adults with Asperger's syndrome. *IEEE Comput. Grap. Appl.* 39 (2), 104–111. doi:10.1109/mcg.2018.2884272

Potel, G., Rodgers, L. R., Ukoumunne, O. C., and Ford, T. (2014). Prevalence of parent-reported ASD and ADHD in the UK: findings from the Millennium Cohort study. *J. Autism Dev. Disord.* 44 (1), 31–40. doi:10.1007/s10803-013-1849-0

Schmidt, M., Laffey, J. M., Schmidt, C. T., Wang, X., and Stichter, J. (2012). Developing methods for understanding social behavior in a 3D virtual learning environment. *Comput. Hum. Behav.* 28 (2), 405–413. doi:10.1016/j.chb.2011.10.011

Schmidt, M., Schmidt, C., Glaser, N., Beck, D., Lim, M., and Palmer, H. (2019). Evaluation of a spherical video-based virtual reality intervention designed to teach adaptive skills for adults with autism: a preliminary report. *Interact. Learn. Environ.* 11, 33. doi:10.1080/10494820.2019.1579236

Self, T., Scudder, R. R., Weheba, G., and Crumrine, D. (2007). A virtual approach to teaching safety skills to children with autism spectrum disorder. *Top. Lang. Disord.* 27, 242–253. doi:10.1097/01.TLD.0000285358.33545.79

Sharples, S., Cobb, S., Moody, A., and Wilson, J. R. (2008). Virtual reality induced symptoms and effects (VRISE): comparison of head mounted display (HMD), desktop and projection display systems. *Displays* 29 (2), 58–69. doi:10.1016/j.displa.2007.09.005

Shenton, A. K. (2004). Strategies for ensuring trustworthiness in qualitative research projects. *Efi* 22 (2), 63–75. doi:10.3233/efi-2004-22201

Strickland, D. (1996). A virtual reality application with autistic children. *Presence Teleop. Virt. Environ.* 5 (3), 319–329. doi:10.1162/pres.1996.5.3.319

Strickland, D. (1997). Virtual reality for the treatment of autism. *Stud. Health Technol. Inform.* 44, 81–86.

Tavassoli, T., Bellesheim, K., Siper, P. M., Wang, A. T., Halpern, D., Gorenstein, M., et al. (2016). Measuring sensory reactivity in autism spectrum disorder: application and simplification of a clinician-administered sensory observation scale. *J. Autism Dev. Disord.* 46 (1), 287–293. doi:10.1007/s10803-015-2578-3

Taylor, J. L., and Mailick, M. R. (2014). A longitudinal examination of 10-year change in vocational and educational activities for adults with autism spectrum disorders. *Dev. Psychol.* 50 (3), 699–708. doi:10.1037/a0034297

Taylor, J. L., and Seltzer, M. M. (2011). Employment and post-secondary educational activities for young adults with autism spectrum disorders during the transition to adulthood. *J. Autism Dev. Disord.* 41 (5), 566–574. https://doi.org/10/czqgq4. doi:10.1007/s10803-010-1070-3

Thomas, G. (2015). *How to do your case study*. New York City, NY:Sage.

Thye, M. D., Bednarz, H. M., Herringshaw, A. J., Sartin, E. B., and Kana, R. K. (2018). The impact of atypical sensory processing on social impairments in autism spectrum disorder. *Dev. Cogn. Neurosci.* 29, 151–167. doi:10.1016/j.dcn.2017.04.010

Tobin, G. A., and Begley, C. M. (2004). Methodological rigour within a qualitative framework. *J. Adv. Nurs.* 48 (4), 388–396. doi:10.1111/j.1365-2648.2004.03207.x

Uljarević, M., Lane, A., Kelly, A., and Leekam, S. (2016). Sensory subtypes and anxiety in older children and adolescents with autism spectrum disorder. *Autism Res.* 9 (10), 1073–1078. doi:10.1002/aur.1602

van Steensel, F. J. A., and Heeman, E. J. (2017). Anxiety levels in children with autism spectrum disorder: a meta-analysis. *J. Child Fam. Stud.* 26 (7), 1753–1767. doi:10.1007/s10826-017-0687-7

Volden, J., Coolican, J., Garon, N., White, J., and Bryson, S. (2009). Brief report: pragmatic language in autism spectrum disorder: relationships to measures of ability and disability. *J. Autism Dev. Disord.* 39 (2), 388. doi:10.1007/s10803-008-0618-y

Wang, M., and Anagnostou, E. (2014). "Virtual reality as treatment Tool for children with autism," in *Comprehensive guide to autism*. Editors V. B. Patel, V. R. Preedy, and C. R. Martin (Berlin: Springer), 2125–2141.

Weech, S., Kenny, S., and Barnett-Cowan, M. (2019). Presence and cybersickness in virtual reality are negatively related: a review. *Front. Psychol.* 10, 158. doi:10.3389/fpsyg.2019.00158

Wong, C., Odom, S. L., Hume, K. A., Cox, A. W., Fettig, A., Kucharczyk, S., et al. (2015). Evidence-based practices for children, Youth, and young adults with autism spectrum disorder: a comprehensive review. *J. Autism Dev. Disord.* 45 (7), 1951–1966. doi:10.1007/s10803-014-2351-z

Yin, R. (2017). *Case study research and applications: design and methods*. 6th Edn. New York City, NY: SAGE.

Conflict of Interest: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

frontiers in Digital Health

ORIGINAL RESEARCH
published: 20 April 2021
doi: 10.3389/fdgth.2021.611813

The Development of a Virtual World Problem-Based Learning Tutorial and Comparison With Interactive Text-Based Tutorials

Trupti Jivram*, Sheetal Kavia, Ella Poulton, Aurora Sésé Hernandez, Luke A. Woodham and Terry Poulton

e-Learning Unit, Centre for Technology in Education, St George's University of London, London, United Kingdom

Edited by:
Panagiotis Evaggelos Antoniou, Aristotle University of Thessaloniki, Greece

Reviewed by:
Daphne Economou, University of Westminster, United Kingdom
Stathis Konstantinidis, University of Nottingham, United Kingdom

****Correspondence:***
Trupti Jivram
tjivram@sgul.ac.uk

Specialty section:
This article was submitted to Connected Health, a section of the journal Frontiers in Digital Health

Received: *29 September 2020*
Accepted: *17 March 2021*
Published: *20 April 2021*

Citation:
Jivram T, Kavia S, Poulton E, Hernandez AS, Woodham LA and Poulton T (2021) The Development of a Virtual World Problem-Based Learning Tutorial and Comparison With Interactive Text-Based Tutorials. Front. Digit. Health 3:611813. doi: 10.3389/fdgth.2021.611813

Collaborative learning through case-based or problem-based learning (PBL) scenarios is an excellent way to acquire and develop workplace knowledge associated with specific competencies. At St George's, University of London we developed an interactive online form of decision-based PBL (D-PBL) for our undergraduate medical course using web-based virtual patients (VPs). This method of delivery allowed students to consider options for clinical management, to take decisions and to explore the consequences of their chosen actions. Students had identified this as a more engaging type of learning activity compared to conventional paper-based/linear PBL and demonstrated improved exam performance in controlled trials. We explored the use of Second Life (SL), a virtual world and immersive 3D environment, as a tool to provide greater realism than our interactive image and text-based D-PBL patient cases. Eighteen separate tutorial groups were provided with their own experience of the same patient scenario in separate locations within the virtual world. The study found that whilst a minority of students reported that the Second Life experience felt more realistic, most did not. Students favored the simpler interaction of the web-based VPs, which already provided them with the essential learning needed for practice. This was in part due to the time proximity to exams and the extra effort required to learn the virtual world interface. Nevertheless, this study points the way towards a scalable process for running separate PBL sessions in 3D environments.

Keywords: interactive tutorials, D-PBL medicine, virtual patients, problem-based learning, virtual worlds

INTRODUCTION

For students in workplace or competency-led courses, collaborative learning through case-based or problem-based learning (PBL) scenarios is an excellent way to acquire and develop workplace knowledge associated with specific competencies, and diagnostic accuracy (1, 2).

The current pandemic has resulted in a swift and dramatic shift away from face-to-face small group learning, and as groups begin to engage more remotely, educators will search for tools which will help compensate for this loss of the more engaging physical presence. An obvious addition would be to add tools which challenge groups to solve tasks collaboratively online, rather than passively follow a fixed narrative. Our work in virtual worlds did not attempt to establish online collaboration, as that was not the primary need at that time. The collaborative space was still the face-to-face interaction based in a single room that is associated with traditional PBL, along with an in-person facilitator. Our goal was that the PBL was run as a blended activity mediated by the group collectively working with a single avatar in the virtual world to provide both the cues for discussion, and the available interactions and decisions at any given time.

Although PBL has proved effective and reasonably popular for more than 60 years, this traditional approach using "paper" cases has been limited by this medium of delivery, in two main ways.

First, traditional paper-based cases were linear, – which could only proceed in a single direction and did not allow students to take own decisions and explore the consequences of their actions. Consequently, at points of acute management, learners can only follow the path set out on the page. Such cases have limited use in developing clinical understanding, competency or reasoning, and are unrealistic for emulating real life, where there are frequently several ways to tackle a problem and mistakes made may not be immediately obvious. The unfolding "paper" case is never able to offer the learner the opportunity to take charge of the scenario or control the management of the patient, thus limiting the role of students as being solely observers.

Secondly, in the modern era, PBL development needs to explore the opportunity for more immersive and engaging methods of delivering the cases that utilize interactive visually oriented technologies. The need for this has been particularly reinforced during the COVID-19 pandemic in 2020, when much of medical education has moved online in an extremely short time frame.

The PBL cases at St George's differ from the normal linear narrative for a written PBL case in that each of the online narrative will contain a number of branching points which allow students to choose between a range of different actions or interpretations. Different branches have different consequences leading to different case outcomes; the students experience the consequences of good and bad decision, and must proceed from that point, just as would have to do in real life and underpinned by cognitive learning theories that focus on linearity problem-solving and the importance of attaining the right outcome (3). A PBL case delivered in a virtual world offers students safe practice, procedural experience, exposure to unseen conditions or diseases and above all, immersive decision-making opportunities (4, 5).

Several studies have considered the use of computer-aided instruction for variants of student-activated learning (6–8). In light of the experience of moving PBL online at St George's, University of London due to the pandemic, we were minded to reflect on our range of previous work toward developing PBL as a blended learning activity, in which we attempted to address the issues described by replacing linear paper cases with online interactive scenarios (9). We believed that it was critical for us to reflect upon this "journey" of iterative changes over more than a decade, such that we would be better able to identify positive improvements that can be made.

St George's has a Transitional year (T-year) within the medical course where the learning is based on PBL. This year is the cross-over period between campus-based learning and clinical apprenticeship, which brings together school-leaver, widening access and graduate-entry students with a total yearly intake of ∼260 students. These students are organized into groups of 7–8, each in its individual location or "baseroom" to work through their PBL.

We had approached the first of the PBL limitations described by developing a more interactive online form of decision-based PBL (D-PBL) (10), which allows students to consider options for clinical management as the cases unfold, to take decisions and to explore the consequences of their chosen actions. Students had identified this as a more engaging learning activity than their conventional paper PBL, and in randomized trials, learners who experienced D-PBL exhibited improved exam performance (11).

Facilitator and student feedback identified that authentic decision making is a significant factor in student engagement (9). Based upon this idea, and building upon established principles for effective case design which establish the importance of providing multiple cues to stimulate discussion (12) we decided in 2010 to explore the use of a virtual world in our PBL tutorials to attempt to provide a more engaging, immersive environment and increase the range of available interactions to students at any given time, beyond the more limited range of menu-based choices available in other online formats. Previous studies had suggested a wide range of roles for virtual worlds in medical education: patient safety (13) skills training (8) social interaction for inter-professional training (14) and for postgraduate Continuing Medical Education (15).

Virtual worlds are three-dimensional online environments which users can adapt to their own needs. By graphically mimicking real-world situations, they can be used to tell a story, a story which the user can interact with and to some extent control its outcome. In this regard, it has similarities with our own D-PBL. The facilitation of teaching and learning through the use of technologies such as virtual worlds has been widely explored in higher education (16, 17). There have been many discussions about the uses and advantages of using virtual worlds (18, 19) but it is more relevant here to consider whether it could bring value to PBL in medical education.

Within an associated project, a trial was conducted for the replacement of Paramedic workplace practice teaching within an immersive 3D virtual world environment to explore whether it could provide greater realism and encourage active decision-making. Virtual patient (VP) scenarios were designed within the virtual world Second Life (SL) (20) by Linden Labs (21) to be used by learners working in groups of four remotely to each other. Feedback established that, despite some technology barriers, SL had the potential to provide a more authentic learner environment than paper or web-based PBL (22). The case study suggested that virtual worlds could offer greater realism, whilst retaining the active collaborative decision-making element of D-PBL. We therefore elected to develop a PBL tutorial that used virtual worlds as an alternative to our interactive online D-PBL, while still retaining the blended setting with students collaborating in a physical room.

Our first research question for a PBL in a virtual world setting was, could an immersive 3D environment provide improved realism over interactive image and text based PBL patient cases, and would that immersion lead to greater student engagement, making the decisions taken more memorable?

Our second research question for PBL was a practical one; was it possible to give each group the opportunity to manage the VP in the scenario individually, while multiple independent instances of the same scenario took place in the virtual world? To avoid excessive workload in the longer term, these scenarios would need to be cloned, and yet remain independent in operation.

In constructing this scenario for the T-Year students, great consideration was given to future proof the scenario, including the inevitable technological progression for more realistic 3D environments. We overcame the technical challenge to provide each PBL group with their own separate area on the virtual island to play the scenario, as well as ensuring a staff member had the ability to monitor multiple groups and restrict students from moving between groups. A major intention of this study was to illuminate these multiple group challenges and solutions for more advanced 3D worlds in future.

MATERIALS AND METHODS

Selection of the PBL Scenario for Adaptation to Second Life

A single PBL case covering chronic renal failure was selected from a 5-week module which covers the gastrointestinal system, liver and kidneys. This module is delivered as part of the Transitional (T) year of the St. George's PBL medical curriculum where students are introduced to clinical placement for half the year and teaching within campus for the other. This whole cohort was separated into two separate groups of 130 students, one group taking this module in May and the other in July, each containing 18 PBL groups of 7–8 students, each PBL group accompanied with a facilitator in each PBL group (18 facilitators). PBL being a student-centered, collaborative and inquiry based learning activity, based on constructivism. In addition to the pedagogies of conventional PBL, our existing online interactive PBL (D-PBL) also offers the student group some degree of choice of direction of the narrative, as the scenario unfolds.

This case was used to teach the basic functions of the kidney, its role in homeostasis, complications of renal failure, transplantation, and public health issues relevant to chronic renal failure. The case was selected because (i) it included more than one location to assist in exploring the value of interactive 3D dimensional representation, (ii) it would have a reasonably high interaction with "tools" e.g., stethoscope, ultrasound machine, and computer health records.

Second Life and St. George's Island

Warburton (18) has suggested that the immersive nature of the virtual world can provide a compelling educational experience, particularly in relation to simulation and role-playing activities. The SL scenarios and SL environment were modeled to follow as closely as possible all the pedagogical approaches and to complete the same narrative, as that covered in the D-PBL case. The additional pedagogical additions were attempting to simulate greater realism and a more immersive environment.

Second Life (SL) is a virtual world which offers users an impressionistic visual representation of a virtual world built by users or "residents." A range of easy-to-use construction tools and a scripting language are provided by the publicly available software for the creation and editing of content in an environment that provides users with a wide range of interactions. Low costs and accessibility (via desktop/laptop device with no additional equipment required) had encouraged educators to explore the potential benefits of virtual worlds for learning. St George's had previously established an engaging and realistic environment on its island, including an orientation and training area, and a simulation of the local environment with street scenes etc. to give context to the teaching of paramedics in accidents and incidents which is key for their practice. All the environments for the scenarios and training were designed specifically for their purpose. SL offers text-chat and voice communication tools, and this enhanced the interaction with the scenario for the students in the paramedic sessions. For the purposes of this project, St George's expanded their island with the purchase of another island attached to the first, to allow more space to build the necessary multiple simulations for the different baserooms.

Creation of the Clinical Environment in Second Life

The first part of the PBL case was to take place in the General Practitioner (GP) surgery (first scene), where a number of tests were required before referring the patient onwards to the renal unit within a hospital setting (second scene in the tutorial). These outline plans were created first as 2-dimensional images (**Figures 1**, **2**), reviewed by a medical professional, who also approved the two environments (scenes) constructed within the SL world (**Figures 3**, **4**). Complexity of the equipment within the environments was optimized to ensure the environment loaded with acceptable speed.

We did not attempt to explore or position the tutorial as a virtual reality (VR) experience for several reasons. Chiefly, we were striving to use the virtual world to foster authenticity in the decision-making process, rather than an experiential authenticity in which students might believe they are present in the environment, and to maintain the collaboration as being face-to-face in the physical room. Authenticity in decision-making is aligned with the principles outlined by Shaffer and Resnick that activities should be aligned with those in real-life practice (23). Additionally, at the time the technology was not sufficiently advanced to provide an adequately realistic environment to make VR a suitably immersive tool for students, and nor was technology such as headsets available or affordable to us.

Development of Multiple Baserooms in Second Life

To allow 18 PBL groups to use the scenarios concurrently, 18 isolated environments were created on the island. Each environment encompassed the GP surgery and then the renal unit.

To achieve this, we used a "holodeck" tool that allowed us to build both scenarios once only but replicate this in multiple locations. The holodeck would load the next part of the scenario

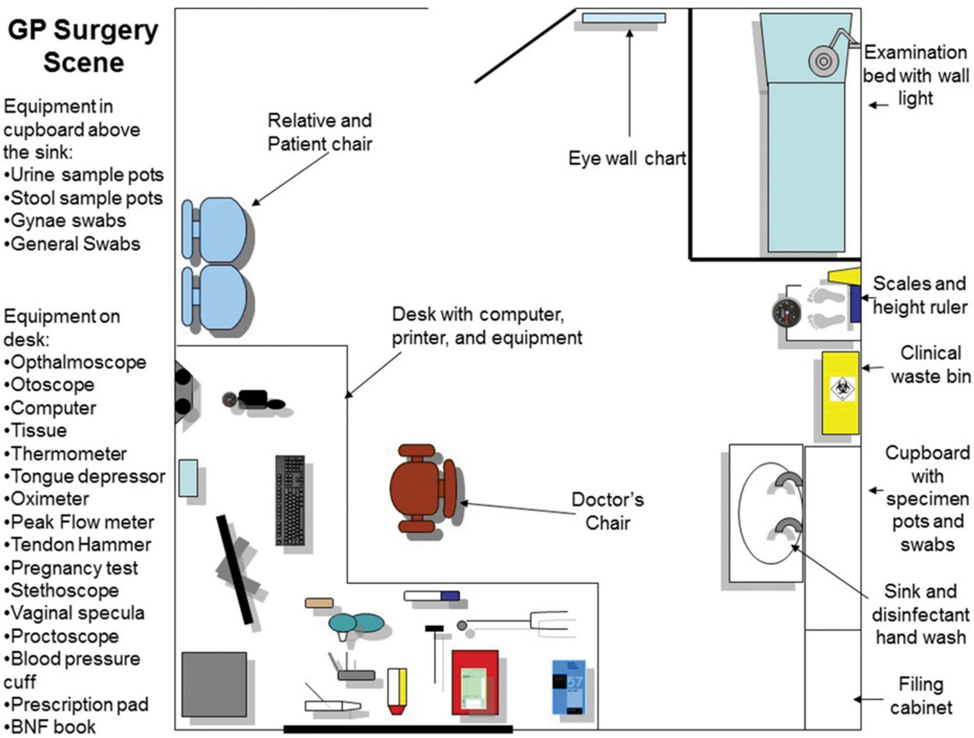

FIGURE 1 | The plan of the GP surgery room. This map was used both as a design for the 3D world, and as a guide to inform students and facilitator which equipment can be found within the scenario, and its location.

in response to a pre-scripted trigger that would only appear once students had completed necessary actions within the GP practice.

Each PBL group was represented in the virtual world by an avatar, which one person in the group could control and use to navigate and interact with the scenario. The avatar's role was initially the doctor in the GP surgery, and subsequently as the clinician in the hospital renal unit. The holodeck which holds the scenario environment was then duplicated, to create 18 virtual "baserooms" distributed across the island (**Figures 5**, **6**). The holodecks were carefully positioned at certain distances from each other to prevent any crossover of chat, or interaction between the baserooms.

Each PBL group was unaware of the progress of other groups due to their avatars inability to fly around the island. However, staff moderators could fly to monitor activity and help with technical issues such as progressing to the next scene within the scenario and observe the progress of each group within their baserooms. These moderators are different from the facilitators who were present in the physical room to guide the students through the scenario as they would in a normal PBL setting.

Creation of the Virtual Patient Avatar in Second Life

The interface of the scenario was kept as simple as possible, since complicated functionality would be frustrating for students new to the environment and functionality of SL. Most of the case interactions took place using a non-player character which represented a virtual patient (VP) (from now on the non-player character will be referred to as a VP in this paper) in each scenario. This VP was physically inanimate in the scenario but provided several interfaces for learners to interact with and advance the narrative: through typed chat using "chatbot" functionality, "touchpoints" to initiate examinations, and equipment within the environment.

Chat

The primary interaction with the chatbot was a SL text-chat interface that allowed learners to question the VP, which would respond with replies based upon the detection of keywords in the question. For example, asking a question containing the keyword "history" would result in a response corresponding to that data, so in this scenario the VP would respond with information on its medical history. The chatbot would only respond to text-chat that was prefixed with the trigger word "ask:" This prevented the chatbot from responding to other in-world communication.

The expected chat responses were assembled initially from the "paper" case, with the addition of alternative wordings identified from prior PBL experiences, before discussion with members of the Communication Skills team and an existing PBL tutor. The chat questions were believed to be acceptable in isolation, but the Communication Team found it difficult to suggest alternative triggers and questions themselves as they teach students to react to the patient depending on patient responses which includes verbal and non-verbal communication.

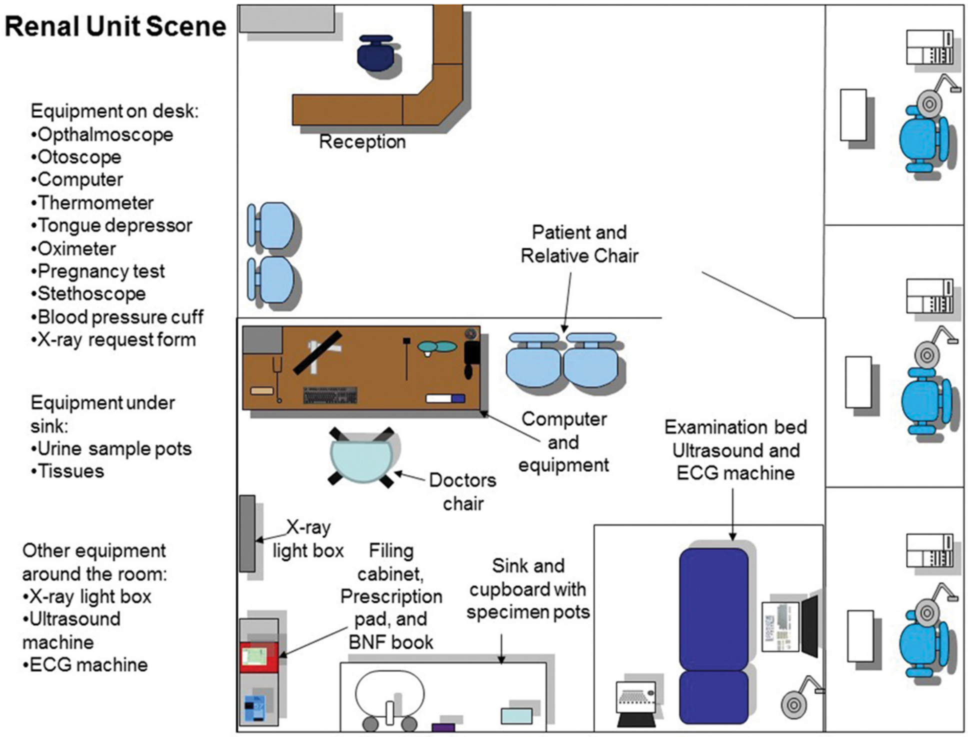

FIGURE 2 | The plan and layout of the renal unit ward where the second part of the tutorial takes place. All interactive equipment is labeled and shown on the map which is provided to the students and facilitator for the session.

FIGURE 3 | The GP surgery, showing the HUD and chat information displayed.

The mapping of keywords to text responses was stored outside the SL scenario. Standard Boolean operators (AND, OR, NOT, ELSE) provided a more sophisticated level of conversational response. This level of interactivity in the chat was appropriate for the scenario given the number of other steps the students needed to take within the given time for the session to manage the VP. A list of trigger words and keywords were provided to the PBL tutor before the session to help them to facilitate the discussions,

FIGURE 4 | Once the case was launched, the 3-dimensional room, in this case the Renal Unit, would appear around the avatar.

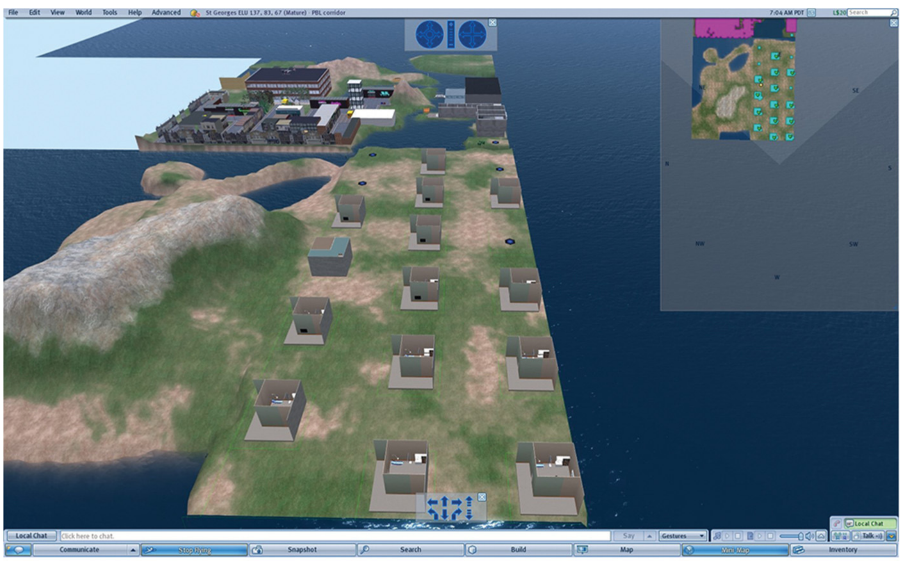

FIGURE 5 | Image from within Second Life showing the virtual base-rooms on the St George's Island with surrounding buildings and landscape. In the foreground on the right, the 18 virtual baserooms are shown by either (i) blue circles, which means the PBL group has not yet activated its virtual GP surgery, (ii) the rectangular boxes of the PBL surgery. Most groups have reached the first phase, the GP surgery, but the group 4 up on the left-hand row, has reached the renal clinic which has a different external size. A by-product of SL/PBL was that differential progress of the PBL groups could be followed.

ensuring key information from the scenario was not missed by the student group.

Virtual Patient Examination

To carry out an examination on the VP the student groups clicked on the specific body part to interact with it. Dependent upon the body part this would trigger the physical examination information, on relevant body part e.g., limbs examination, chest sounds, abdomen examination etc. The results of the examination would be revealed in the chat and visual display (HUD) top left-hand corner (**Figures 3**, **4**).

Use of Equipment

Learners had access to an inventory of equipment that would routinely be used in the relevant scenario in real-world situations. Medical items such as blood pressure cuffs, cannula, stethoscopes

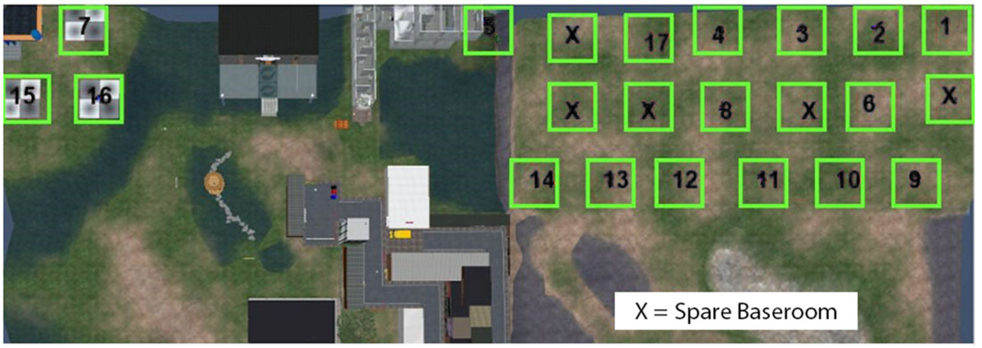

FIGURE 6 | Second life Island plan of which real life baserooms will be located in which virtual baseroom. This plan was used by the development team to keep track of all rooms and to ensure there was no crossover between students.

etc. were placed around the scene within the GP practice setting, student avatars were able to touch these, and then presented with a pop-up offering a choice of possible actions. For example, by touching the octameter, the player is asked "which ear you would like to attach it to." The object would attach itself to the VP and provide a relevant output or reading, which could change dynamically depending upon previous interactions and treatments provided. Furniture and other equipment could be used, such as computers, sinks, cupboards etc. to perform actions such as checking test results via the computer, washing hands or revealing further available equipment, like urine test strips.

Controlling the Scenario

The scenario is driven using a screen controller that can be easily accessed through the use of the HUD, top left (**Figures 3**, **4**), and displays media content associated with the VP, including text, images, audio, and video from an external web application created by Daden Limited (24), who were contracted as part of the previous project to develop the HUD which allowed external web content to be visible in SL from specific tiggers within the scenario. This application records the history of all actions taken by the students' avatar, and provides an interface for authoring the scenario, assembling media content and mapping keywords to text responses.

By taking this scenario content out of SL into a conventional web application, we ensured that the scenario could be duplicated and used by multiple groups concurrently. Also ensuring the content was available to be used via other platforms not only within SL. The steps involved in the development of the scenario are outlined in **Figure 7**.

Testing Procedure

The scenarios were evaluated at St George's during two planned testing days using volunteer students with no prior experience of SL, which accurately represented the target group. At the end of the session, they were asked to give feedback through a focus group on their experience of using the platform and completing PBL via the virtual world.

Training of Students and Tutors in Second Life PBL

Before the PBL went "live," two volunteer students from each baseroom were invited to a training session, in which they were taught how to navigate the St George's orientation maze using a shortened scenario similar to the PBL case. The training session ensured that two students representing each PBL group understood how to move around in SL, control their viewpoint within the scenario, use the equipment and HUD, chat with the VP, and understood how to move onto the next scene of the scenario.

Facilitators were given a demonstration of the SL scenario and provided with tutor notes during this training session to understand how the guides in the notes would help them during the session. A meeting before the PBL session was also organized to ensure all facilitators were comfortable with the SL session. Facilitators were provided with additional guides which included a map of both scenes in the scenario, the trigger words for questions to use during the chat, the list of equipment to interact with and instructions on how to use SL and move on in the scenario.

Delivery of PBL

The PBL scenario was structured in a similar manner to traditional PBL (25), with the important distinction that the scenario now took place within a 3D environment rather than as text, in this one-off session. The students in each group, along with their facilitator were all co-located in the same physical space. The role of the facilitator within each PBL group during the SL scenarios remained the same, to guide student discussion where required. The facilitator was given a SL case-specific tutor guide, similar to the tutor guide of a conventional D-PBL case, which informed the facilitator of both the overall direction of the case, the learning objectives the students needed to cover in the case, and additional information needed for understanding and supporting the SL delivery as mentioned above.

For the first tutorial each base room was set up with their own unique avatar log in details. The representative students in each group who had been trained for this task, logged into SL

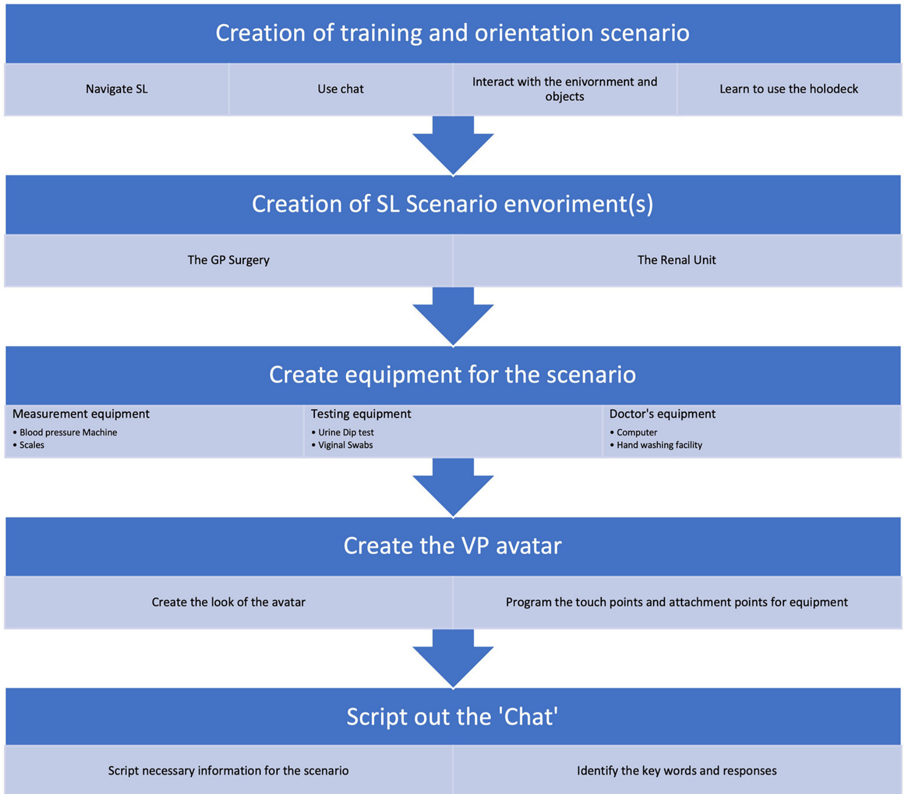

FIGURE 7 | A simple flow of the order in which the elements of the scenario were developed.

and the case was projected onto the smart whiteboard for all the students in the group. The baseroom of each group then appeared as a separate identity on the island (**Figure 8**), on their assigned holodeck pad (as seen as blue circles on the island).

When the group had completed their tasks in the first location (GP surgery scene), they moved to the next scene. If the group failed to complete any of the required steps, they would be prompted to carry out further investigations before moving on. The second location (scene) of the scenario was set in the hospital renal unit. As before, the group discussed how they should proceed, and which investigations should be carried out within the second location.

The session lasted 3 h, the same as the D-PBL session, and the tutorial adhered as closely as possible to the traditional process for either a paper case or a D-PBL case. The students generated learning objectives whilst exploring the scenario, and "interacting with the VP." After the tutorial, students carried out their self-directed learning as usual from the learning objectives generated, for reporting back and discussion at the next tutorial.

Only the first tutorial was delivered in SL, the same students went on to complete the second tutorial which was delivered as St George's standard interactive D- PBL case for the normal allocated 3 h, but this second tutorial was not evaluated as part of this study.

FIGURE 8 | Each PBL group was set up with their own unique avatar. Either of the two students trained to use SL would log in, and their avatar would appear on their assigned holodeck, projected onto the whiteboard. Once the group was ready, they launched the case.

Data Collection and Analysis

Data collection was principally captured in the form of structured student and tutor surveys and a student focus group after delivery

TABLE 1 | Previous experience with virtual world.

Answer options	response Percent ($n = 244$)	Response count
No	73%	177
Yes, for fun	21%	52
Yes, for education	4%	10
Yes, for another purpose	2%	5

of the SL PBL session. The survey was created with the aim of evaluating the virtual worlds resource and had not been validated.

Descriptive statistics were calculated for survey items. Focus group comments and open-ended survey responses were analyzed by an external member of the project team using a directed content analysis approach (26) and the tool Atlas.ti (26), to provide further context and background to the themes identified from the survey.

RESULTS

Student Survey

Experience and Perception of Virtual Worlds Prior to the Intervention

Two hundred forty-four students out of a possible 260 completed the survey, representing 34 out of the possible 36 PBL groups. As shown in **Table 1**, 177 (73%) students had not used virtual worlds before. Fifty-two (21%) had used them for entertainment purposes, though in providing further details many qualified this by indicating that their experience was in computer games such as the Sims (27) which are very different in style, scope and user experience to SL. 10 (4%) students had previous experience of using virtual worlds for education purposes.

One hundred seventy-seven students had not previously used virtual worlds and were asked to provide details from a multiple-choice list. One hundred four students cited a lack of interest as their reason for never having used them, with 82 pointing to not having a prior need to do so. Thirty-seven students saw no value in virtual worlds, while 20 students were unaware of their existence. Twelve students identified that they did not have access to a computer that can support virtual worlds, while 9 were not comfortable with using the software.

Nevertheless, despite this general lack of experience in virtual worlds 152 (62%) of the students thought that virtual worlds could provide useful learning resources for medical education, with 88 (36%) answering that it would not be useful (4 students did not respond). One hundred thirty (53%) students reported that they did not require any more guidance than was provided to participate in the PBL, with 25 (10%) indicating that more guidance was needed, and 89 (36%) electing not to answer the question. The students were also asked "*if the time taken to learn how to use the platform was off-putting*" and the response was mixed, with 131 answering "yes" (54%), 109 (45%) answering "no," and 4 students skipping the question.

Assessment of the Intervention

Quantitative Analysis

The students were asked to rate a range of Likert items using a 4-point scale, based upon "how easy was it to use the [SL] features" (**Table 2**). Mean responses were calculated by assigning a numeric scale of 1.

The easiest feature was using the medical equipment with a mean rating for ease of use of 2.98 followed by conducting a physical assessment with a mean rating of 2.83. By contrast, the most difficult features to use were: (i) history-taking (2.33) and (ii) the scenario as whole (2.46). One hundred forty-five of the students reported that they received the expected responses from the virtual patient chatbot most of the time, although 84 felt that they rarely received the expected response, and 15 skipped the question. One hundred fifty-four respondents felt that the scenario was not realistic, compared with 85 that did. One hundred thirty-five respondents stated that the information in the scenario was presented clearly, while 104 respondents felt that the information presented was unclear.

Qualitative Analysis

The survey included some open-ended questions (**Table 3**) which provided further insight into what the students expected from the scenario and yielded some key themes. Approximately half of the students (78) who responded considered that SL promoted decision making, having to think about different options rather than choosing from a given list. For example, "*allowed us to also think outside the box rather than to just be given the scenario.*", "*..Think of ways to investigate patient rather than pick from list.*"

In contrast, although the ability to examine the patient was praised to some extent, this was an area that the students found limiting. Learners reported being aware that they were not having a conversation with the patient, but that the patient was simply responding to a pattern-matched phrase or word. With this knowledge they then altered their behavior to use minimal phrases and words, one student stating "*you could easily pick up that you're just recognizing key words, and you just put in keywords.*" Student opinion on history-taking was mixed; though some found it engaging, the majority reported that the process question/answer during the history taking did not work well; "*the patient didn't understand even simple questions during the history taking process*" and that *"the responses from the patient were not appropriate."* Responses to the examination elements were more positive and the "key elements" (tests and equipment) were present, realistic and easy to use.

One area in which opinions amongst the students differed greatly was around the realism of the scenario. Some students praised the realism of the approach to addressing the patient: *"I think it is realistic in a way, because clinically, the processes are there that you have to go through systematically; history, examination, investigation, and referral."* However, other students expressed frustration with the boundaries of the realism, commenting that "*if you're a first year and you've never been into a hospital before, it's quite fun to see what a room might look like, but you know, there's no patients, and no people. There's no kids running around. It's not realistic.*" Nevertheless, some

TABLE 2 | Ease of use of the features during the use of virtual world.

Answer options	n	Response count				Mean rating
		Very difficult	Fairly difficult	Fairly easy	Very easy	
History taking	239	36	100	92	11	2.33
Using the medical equipment	241	12	31	149	49	2.98
Conducting a physical assessment	237	13	57	125	42	2.83
Using the HUD (Heads-Up Display Screen)	229	18	63	125	23	2.67
Viewing media e.g., images	238	17	54	138	29	2.75
Moving between the GP surgery/renal unit	233	24	52	126	31	2.70
The scenario as a whole	235	33	71	120	11	2.46

241 students answered this question (3 students skipped).

students considered that the SL PBL was "fun" and "interesting" and encouraged discussion.

Students disliked the lack of guidance and ordering of events, which is a feature of virtual worlds and found the SL platform difficult to use, especially controlling the VP and moving from one scenario to the other.

Preferences on the PBL Teaching Methodology

The students were asked to rank the different forms of PBL: paper PBL, online D-PBL and virtual PBL. One hundred eighty-five (77.08%) of the 240 respondents ranked D-PBL in 1st place; 36 ranked paper PBL in 1st, and only 19 ranked virtual worlds 1st. In all rankings D-PBL was the favored methodology, SL was consistently lowest.

The students were almost equally divided between those who believed "*there was value/potential for value in this method of PBL*" (51.1%), and those did not. However, 77.2% of the students reported that SL "*PBL was hindered by the style of scenario*." A smaller percentage (53.3%) though it "*hindered effective collaboration,*" and 13% thought it *"aided collaboration."* Only 3% of students thought that the "*PBL was improved by this style of scenario*." Less than half (47.4%) of students believed that they covered the appropriate Learning Objectives' while using the SL PBL.

TABLE 3 | Categorization of responses to two open-ended questions ("what worked well during SL PBL," from 155 students) and ("what didn't work well during SL PBL"; from 199 students).

Worked well	Didn't work well
1. Enhanced decision making (78)	1. Slow/time consuming (103)
2. Enjoyable/interesting learning experience (40)	2. Question/answer process during the history taking (86)
3. Design of the scenario during the "examination" (36)	3. Technical design: glitches and appearance of the scenario (58)
4. Possibility of interacting with the patient during history taking (30)	4. Sequence of facts was confusing (45)
5. Capacity of enhancing teamwork and discussion (27)	5. LOBs not achieved (20)
6. Realistic (22)	6. Team work was hindered (19)
7. Interactivity (6)	7. Not realistic (19)
8. Feedback (5)	8. Lack of some necessary examinations (5)

Facilitator Survey

Sixteen out of the 18 facilitators completed the short facilitator survey, mainly composed of open-ended questions exploring the effectiveness of SL PBL and response of the students.

Most facilitators considered SL PBL is less effective than both the interactive D-PBL and the paper PBL. Some commented that it was more difficult to identify learning objectives compared to the other methods. Only 25% reported that the SL PBL was effective. In general, facilitators agreed with student preferences for the different forms of PBL, with D- PBL first (60%), but SL was second with 27%, and paper last.

Most facilitators (60%) reported that it was more difficult to facilitate the virtual world PBL sessions, though a proportion believed that increased training and practice would improve the facilitation process. Furthermore, ∼30% of the facilitators believed that the SL PBL could be useful if further developed.

Facilitators believed that during the facilitation of the SL PBL, the students were actively engaged but at the same time the dynamic was more complicated than usual and 80% of the facilitators believed that technology, at least occasionally, "got in the way of the learning."

Some facilitators found SL case easier tutoring the second time around, with one facilitator stating that it "*was smoother but I think that was due to the fact I have already done this case before it was easier to make use of the virtual map to find appropriate equipment for examination and investigations.*"

DISCUSSION

This study investigated the feasibility of running multiple baserooms simultaneously, but also independently, each with its own group-controlled avatar. We had expected that the virtual world may be ultimately lacking in realism, but nevertheless would provide a pointer to the feasibility of creating separate simulations for each PBL group. This would then address an overarching objective of the PBL curriculum, to provide a unique experience for each group, with the students in charge of the management.

In practice, the theoretical process for establishing the individual baserooms worked very well. The two students who had volunteered to be trained from each PBL group were capable of managing the necessary functionality of the avatar, HUD and chatbot features.

The survey results indicate that students considered that SL PBL provided a decision-making opportunity which encouraged them to think and discuss the different options. Some of the students that had previously used virtual worlds, enjoyed the experience for its entertainment value and believed it to be more realistic, despite the need for design improvements. However, whilst a minority reported that the SL experience felt more realistic, most did not, and the students clearly preferred their simpler text-based online interaction, the D-PBL experience. This finding is also in agreement with Loke's review of proposed mechanisms of learning in virtual world experiences (28) which found that, contrary to a popular conception, students do not appear to experience a physical presence of real-world phenomenon through their virtual world actions; they do not perceive virtual worlds as "learning by doing."

A repeated common theme in student/facilitator surveys and focus groups was it would be necessary to heavily review and improve the design of this tool to make it viable. It was clear that the software had not reached state-of-the-art which would allow students to feel comfortable with either technology or the level of realism. Nevertheless, it is instructive to consider whether, even if realism improved, students would gain significantly from the experience. Students made the point that the online D-PBL already contains many of the characteristics that we are seeking with the use of SL PBL, and therefore further attempts at realism were not necessarily productive or cost-effective. As technology evolves in areas of VR and AR lessons learnt from this study some of those listed in **Table 3** could be applied to future studies. It is possible that the assessment tool developed by Fu et al. (29) to assess student enjoyment of games, could be adapted for future assessment of student satisfaction with 3D educational tools.

This study would appear to be at odds with other findings which have explored the value to students of greater realisms in PBL settings (22), but in these studies, students did not have a comparison with interactive scenarios in a non-3D World. Consequently, this may be the first time for 3D environments that a fair comparison has been made of one interactive tool vs. another i.e., one text-based and 2D, the other a more complex 3D world. More normally the 3D world comparison is made of one interactive tool vs. a static linear, non-decision-making tool, and thereby negates any advantage that these interactive environments may possess over the more linear learning and teaching opportunities. Since this work, there has been a number of other studies and developments in technology for use in education where (30) found increased student learning performance from the use of collaborative learning environments.

A non-controlled variable may have influenced the subjective opinion of students concerning the merit of virtual PBL. The students considered that it was not appropriate to organize this session 3 weeks before the exams, and this may have influenced their perception of virtual world PBL.

What has emerged from this study is a clear view of what students would require before a 3D world would have any advantage (in terms of student engagement) over their existing interactive, text-based system, namely: improving the technical design and organization of the content; increasing the technical sophistication of the virtual world, and reducing the time needed to understand the tool.

An attempt to improve the technical design, would encounter the issue that the open-endedness of a virtual world provides few clues to students for how to progress further in the scenario. The participants' role is not intuitive. By contrast, structured gaming environments present defined possibilities for action. Games are more intuitive, they can guide the participants to know what to do next, and when the activity is finished. Likewise, text also offers a similar sense of structure.

Improvements to a virtual world would require technologists, physicians, and educators to work together to consider all the possible options that would allow the students to cover the scenario satisfactorily, and it is questionable whether this effort would be worthwhile, with so many challenges to overcome.

An obvious example is the chat function. In terms of content design, the chat does not work well enough to be useful. Improved mapping might help, but it must be borne in mind that students and practitioners ask questions in their own style and that mapping may require extensive semantic analysis and a very large vocabulary. With advances in AI the approach to "chat" within the scenario could be explored with tools outside SL which have developed for more realistic interaction and responses. Similar challenges are presented in an open-ended environment with a number of available tests when assessing the patient; those offered are likely choices, rather than unlimited choices.

Aside from the technical limitations of the platform, there are some limitations to the study which must be considered. Primarily, the generalisability of our results are limited by the fact that participants were from a single institution and setting, and received one intervention in the virtual world. There were also a number of non-controllable factors that may have influenced our results, such as the tutorial taking place prior to an exam period which was the students' primary focus. Although participation in the tutorial was part of the regular curriculum, completion of the survey was not mandatory, and a more detailed analysis of responses per group would be needed to properly assess and identify patterns in the responses that might indicate the presence of attrition bias. Similarly, the design of our survey was not validated in this context prior to the study and our analysis does not allow us to account for intra-group correlation, which has the potential to introduce a further bias into our results. These limitations necessarily mean that more evidence would be required to support our conclusions and establish if our findings can be generalized beyond the context of this study.

One potential objective had been to consider running PBL groups for students in different geographic locations as Melus-Parazon et al. (31) had achieved in primary health care settings, but the technical difficulties removed this possibility. A summary of what worked well for this study from the students perspective and what could have been better is provided in **Table 3** of the Results section. From this we can see conflicting views however it is hard to draw a conclusion from this due to students only doing one tutorial in SL.

It is increasingly recognized that many e-learning interventions may be far more costly than they need to be "*Educators often seem to use cutting edge (read that as expensive) technologies because they are available, rather than because they add value commensurate with the higher cost*" (32). Though this study points the way toward a process for running PBL sessions in 3D environments, and at the same time shows the ability to monitor those session more completely than is possible in conventional non-virtual world PBL, it is clear that it would seem to require a very considerable change in the technology before this process can be useful to both learners as well as educators.

Nevertheless, the pursuit of more engaging learning activities, including virtual realities, may well be worth the effort. Norman et al. (33) has noted that a learning tool does not need to mimic real life, it is the learning experience that needs to reflect real-life. Recent events have forced radical changes in face-to-face modalities for PBL which will carry the risk of a more remote and less engaging learning experience. Students may not have seen "the point" of virtual realities beforehand, in their face-to-face tutorial groups, but the switch to online PBL may promote a search for a more engaging experience, by students and tutors alike.

Immediately before the COVID-19 pandemic, Savin-Baden (34) noted that ideally PBL should include practices such as gaming, emotional learning, playful learning as well as student-led cooperation in mentorship, technology support, and even co-production. The issues that this study faced are similar to those found in a recent study by Sancar-Tokmak and Dogusoy (35) who explored using SL to solve the high dropout rate in a Distance Learning Center. Although the learners could recognize the opportunities that SL could provide for PBL, they preferred alternative learning methodologies because of access and usability issues.

It is to be hoped that in accepting the pedagogic losses we experience from the reduction in face-to-face teaching, we can provide compensations in other directions, online. Perhaps cooperative online learning that includes virtual realities may help to contribute, as Castelo-Branco et al. (25) puts it, "a pedagogy of imagination and surprise."

DATA AVAILABILITY STATEMENT

An anonymised version of the data supporting the conclusions of this article will be made available by the authors upon direct request.

ETHICS STATEMENT

Ethical review and approval was not required for the study on human participants in accordance with the local legislation and institutional requirements. Written informed consent for participation was not required for this study in accordance with the national legislation and the institutional requirements.

AUTHOR CONTRIBUTIONS

All authors listed have made a substantial, direct and intellectual contribution to the work, and approved it for publication.

FUNDING

This work was supported by JISC (Joint information systems committee, United Kingdom) funded under its transforming curriculum delivery through technology grant (capital programme), project name; Generation 4 (G4). https://www.webarchive.org.uk/wayback/archive/20140613220103/http://www.jisc.ac.uk/media/documents/programmes/curriculumdelivery/G4_Final_Reportv2.pdf.

ACKNOWLEDGMENTS

The authors would like to acknowledge contributions from Dr. Chris Beaumont for data evaluation as part of the G4 project.

REFERENCES

1. Hoffman K, Hosokawa M, Blake R, Headrick L, Johnson G. Problem-based learnin outcomes: ten years of experience at the university of missouri, columbia school of medicine. *Acad Med.* (2006) 81:617–25. doi: 10.1097/01.ACM.0000232411.97399.c6
2. Schmidt HG, Machiels-Bongaerts M, Hermans H, ten Cate TJ, Venekamp R, Boshuizen HP. The development of diagnostic competence: comparison of a problem-based, an integrated, and a conventional medical curriculum. *Acad Med.* (1996) 71:658–64. doi: 10.1097/00001888-199606000-00021
3. Savin-Baden M, Tombs C, Poulton T, Conradi E, Kavia S, Burden D, et al. An evaluation of implementing problem - based learning scenarios in an immersive virtual world. *Int J Med Educ.* (2011) 2:116–24. doi: 10.5116/ijme.4e92.b22f
4. Vozenilek J, Huff JS, Reznek M, Gordon JA. See one, do one, teach one: advanced technology in medical education. *Acad Emerg Med.* (2004) 11:1149–54. doi: 10.1197/j.aem.2004.08.003
5. Reznek M, Harter P, Krummel T. Virtual reality and simulation: training the future emergency physician. *Acad Emerg Med.* (2002) 9:78–87. doi: 10.1111/j.1553-2712.2002.tb01172.x
6. Bergin RA, Fors UGH. Interactive simulated patient - an advanced tool for student-activated learning in medicine and healthcare. *Comput Educ.* (2003) 40:361–76. doi: 10.1016/S0360-1315(02)00167-7
7. de Jong T, Njoo M. Learning and instruction with computer simulations: learning processes involved. In: De Corte E, Linn M, Mandl H, Verschaffel L, editors. *Computer-Based Learning Environments and Problem Solving.* Springer (1992). p. 411–29.
8. Scalese RJ, Obeso VT, Issenberg SB. Simulation technology for skills training and competency assessment in medical education. *J General Internal Med.* (2008) 23:46–9. doi: 10.1007/s11606-007-0283-4
9. Poulton T, Conradi E, Kavia S, Round J, Hilton S. The replacement of "paper" cases by interactive online virtual patients in problem-based learning. *Med Teacher.* (2009) 31:752–8. doi: 10.1080/01421590903141082
10. Ellaway RH, Poulton T, Jivram T. Decision PBL: a 4-year retrospective case study of the use of virtual patients in problem-based learning. *Med Teacher.* (2015) 37:926–34. doi: 10.3109/0142159X.2014.970627
11. Poulton T, Ellaway RH, Round J, Jivram T, Kavia S, & Hilton S. Exploring the efficacy of replacing linear paper-based patient cases in problem-based learning with dynamic Web-based virtual patients: randomized controlled trial. *J Med Internet Res.* (2014) 16:e240. doi: 10.2196/jmir.3748

12. Dolmans DHJM, Snellen-Balendong H, van der Vleuten CPM. Seven principles of effective case design for a problem-based curriculum. *Med Teacher*. (1997) 19:185–9. doi: 10.3109/01421599709019379
13. Lee A, Berge ZL. Second life in healthcare education: virtual environment potential to improve patient safety. *Knowledge Manage.* (1988) 3:17–23.
14. Boulos MNK, Hetherington L, Wheeler S. Second life: an overview of the potential of 3-D virtual worlds in medical and health education. *Health Inform Libraries J*. (2007) 24:233–45. doi: 10.1111/j.1471-1842.2007.00733.x
15. Wiecha J, Heyden R, Sternthal E, Merialdi M. Learning in a virtual world: experience with using Second Life for medical education. *J Med Internet Res.* (2010) 12:e1. doi: 10.2196/jmir.1337
16. Wang F, Burton JK. Second life in education: a review of publications from its launch to 2011. *Br J Educ Technol.* (2013) 44:357–71. doi: 10.1111/j.1467-8535.2012.01334.x
17. Savin-Baden M, Poulton T, Beaumont C, Conradi E. What Is real? using problem-based learning in virtual worlds. In: Bridges S, Chan L, Hmelo-Silver C, editors. *Educational Technologies in Medical and Health Sciences Education. Advances in Medical Education*. Vol 5. Cham: Springer (2016). doi: 10.1007/978-3-319-08275-2_5
18. Warburton S. Second Life in higher education: assessing the potential for and the barriers to deploying virtual worlds in learning and teaching. *Br J Educ Technol.* (2009) 40:414–26. doi: 10.1111/j.1467-8535.2009.00952.x
19. Savin-Baden M, Tombs G (editors.). *Threshold Concepts in Problem-based Learning*. Leiden: Brill | Sense (2019) doi: 10.1163/9789004375123
20. Second Life. Available online at: https://secondlife.com/ (accessed September 16, 2020).
21. Linden Lab. Available online at: https://www.lindenlab.com/ (accessed September 16, 2020).
22. Beaumont C, Savin-Baden M, Conradi E, Poulton T. Evaluating a Second Life Problem-Based Learning (PBL) demonstrator project: what can we learn? *Interactive Learn Environ.* (2014) 22:125–41. doi: 10.1080/10494820.2011.641681
23. Shaffer DW, Resnick M. "Thick" authenticity: new media and authentic learning. *J Interactive Learn Res.* (1999) 10:195–215.
24. Daden Limited. Available online at: https://www.daden.co.uk/ (accessed September 16, 2020).
25. Castelo-Branco L, Finucane P, Marvão P, McCrorie P, Ponte J, Worley P. Global sharing, local innovation: Four schools, four countries, one curriculum. *Med Teacher.* (2016) 38:1204–8. doi: 10.1080/0142159X.2016.1181731
26. Hsieh H-F, Shannon SE. Three approaches to qualitative content analysis. *Qual Health Res*. (2005) 15:1277–88. doi: 10.1177/1049732305276687
27. Sims. Available online at: https://www.ea.com/games/the-sims (accessed September 16, 2020).
28. Loke S. How do virtual world experiences bring about learning? A critical review of theories. *Austr J Educ.* (2015) 31:112–22. doi: 10.14742/ajet.2532
29. Fu FL, Su RC, Yu SC. EGameFlow: a scale to measure learners' enjoyment of e-learning games. *Comput Educ.* (2009) 52:101–12. doi: 10.1016/j.compedu.2008.07.004
30. Doumanis I, Economou D, Sim GR, Porter S. The impact of multimodal collaborative virtual environments on learning: a gamified online debate. *Comput Educ.* (2019) 130:121–38. doi: 10.1016/j.compedu.2018.09.017
31. Melus-Parazon E, Bartolomé-Moreno C, Palacín-Arbués JC, Lafuente-Lafuente A, García I, Guillen A, et al. Experience with using second life for medical education in a family and community medicine education unit. *BMC Med Educ.* (2012) 12:30. doi: 10.1186/1472-6920-12-30
32. Cook DA, Triola MM. What is the role of e-learning? Looking past the hype. *Med Educ.* (2014) 48:930–7. doi: 10.1111/medu.12484
33. Norman G, Dore K, Grierson L. The minimal relationship between simulation fidelity and transfer of learning. *Med Educ.* (2012) 46:636–47. doi: 10.1111/j.1365-2923.2012.04243.x
34. Savin-Baden M. What are problem-based pedagogies? *J Prob Based Learn.* (2020) 7:3–10. doi: 10.24313/jpbl.2020.00199
35. Sancar-Tokmak H, Dogusoy B. Novices' instructional design problem-solving processes: second life as a problem-based learning environment. *Inter Learn Environ*. (2020). doi: 10.1080/10494820.2020.1799025. [Epub ahead of print].

Conflict of Interest: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Advantages of publishing in Frontiers

OPEN ACCESS
Articles are free to read for greatest visibility and readership

FAST PUBLICATION
Around 90 days from submission to decision

HIGH QUALITY PEER-REVIEW
Rigorous, collaborative, and constructive peer-review

TRANSPARENT PEER-REVIEW
Editors and reviewers acknowledged by name on published articles

Frontiers
Avenue du Tribunal-Fédéral 34
1005 Lausanne | Switzerland

Visit us: www.frontiersin.org
Contact us: frontiersin.org/about/contact

REPRODUCIBILITY OF RESEARCH
Support open data and methods to enhance research reproducibility

DIGITAL PUBLISHING
Articles designed for optimal readership across devices

FOLLOW US
@frontiersin

IMPACT METRICS
Advanced article metrics track visibility across digital media

EXTENSIVE PROMOTION
Marketing and promotion of impactful research

LOOP RESEARCH NETWORK
Our network increases your article's readership