ORIGINAL ARTICLE

Real-Time Holographic Feedback via Mixed Reality Sensorised Laryngoscope Training System Enhances Paediatric Intubation Training

Jiaqi Li[1] | Ryman Hashem[1,2] | Ningzhe Hou[1,3] | Perla Maiolino[1,3] | Louis Halamek[4] | Liang He[1,2]

[1]Department of Engineering Science, University of Oxford, Oxford, UK.

[2]The podium institute for sports medicine and technology, University of Oxford, Oxford, UK.

[3]Oxford Robotics Institute, University of Oxford, Oxford, UK.

[4]Department of Pediatrics, Stanford University, USA.

Correspondence
Liang He, Department of Engineering Science, University of Oxford, Oxford OX1 3PJ, UK
Email: liang.he@eng.ox.ac.uk

Funding information
This work was funded by Royal Society Research Grant RGS\R2\242066

Paediatric endotracheal intubation (ETI) is a technically demanding procedure in which novices often apply excessive force and torque. Traditional manikin training and instructor verbal feedback lack quantitative and real-time performance data, limiting their effectiveness. This work introduces a Mixed Reality Sensorised Laryngoscope Training System (MR-SLTS), which integrates a sensorised laryngoscope with head-mounted holographic feedback to visualise forces and torques in real time. Two controlled user studies were conducted. In Study 1, six subjects performed a torque-tracking task using a simplified laryngoscope setup while viewing three holographic visualisations (bar, ring, and line graphs). Bar-graph feedback most effectively reduced tracking error and improved force control ($p < 0.05$). Eye-tracking data showed that ring graphs elicited significantly more saccades than line graphs ($p < 0.05$), indicating higher visual demand, while bar and line graphs produced comparable gaze stability. In Study 2, ten novice subjects, split into mixed-reality and control groups, performed paediatric ETI on a manikin with or without real-time MR-SLTS torque feedback. Trainees receiving mixed-reality feedback kept forces and torques more consistently within expert-defined ranges without increasing intubation time. These results demonstrate that real-time mixed-reality feedback enhances force control during paediatric ETI training while preserving procedural efficiency, offering guidance for future telesimulation-based training systems.

KEYWORDS
Mixed reality, Intubation training, Sensorised laryngoscope, Real-time holographic feedback, Force and torque visualisation, Medical simulation

1 | INTRODUCTION

Endotracheal intubation (ETI) is a critical procedure for airway management in various clinical settings: both emergency and surgical. It involves guiding the endotracheal tube from the oral cavity into the trachea, typically using a laryngoscope to lift the epiglottis and allow passage through the vocal cords. Successful intubation requires precise hand-eye coordination and dexterous manipulation of the laryngoscope; however, novice practitioners often apply excessive force or adopt suboptimal techniques compared to experts [1, 2]. These errors can cause mechanical complications such as oropharyngeal soft-tissue trauma and dental injury, and, in more severe cases, repeated attempts or failed intubations are associated with resultant hypoxaemia, cardiovascular instability, or cardiac arrest [3].

Traditional ETI training methods rely on simulated intubation performed on manikins and verbal feedback from expert instructors. While these approaches are useful for teaching procedural steps, they often fall short in conveying the fine motor control and detailed force application needed for safe and effective intubation [4]. To address this, Hou et al. [5] developed the Sensorised Laryngoscope Training System (SLTS), embedding sensors within a laryngoscope to provide real-time quantitative feedback on applied forces and device motion.

While the SLTS enabled objective and quantitative skill assessment, its monitor-based feedback required users to divide attention between the manikin and an external display, motivating the development of a mixed-reality alternative.

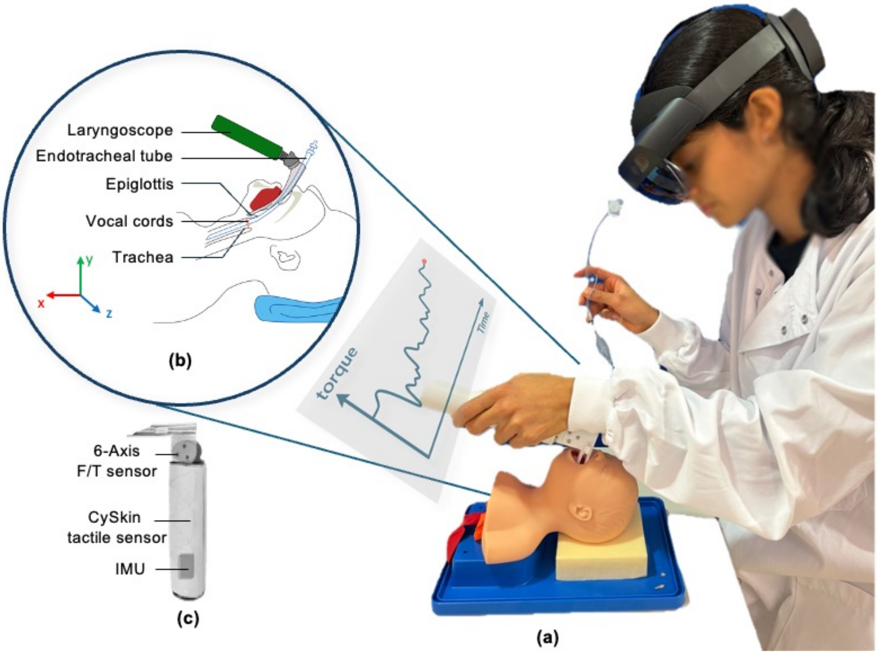

FIGURE 1 **(a)** Illustration of the Mixed Reality Sensorised Laryngoscope Training System (MR-SLTS). Real-time visual feedback of contact force/torque is delivered to the trainee through a head-mounted holographic device, supporting skill acquisition in contact force/torque regulation. **(b)** The paediatric endotracheal intubation procedure. The laryngoscope is used to lift the epiglottis and visualise the vocal cords through coordinated rotational and lifting manoeuvres, allowing the endotracheal tube to pass through the vocal cords into the trachea. **(c)** Schematic of the sensorised laryngoscope

To overcome these limitations, this study introduces the Mixed Reality Sensorised Laryngoscope Training System (MR-SLTS), which delivers live visual feedback directly within the user's field of view using a head-mounted holographic device (Microsoft HoloLens 2). A schematic of the system can be found in Fig. 1. By overlaying information onto the physical environment, the system aims to improve force interpretability, minimise attention shifts, and reduce cognitive burden. Two controlled user studies were conducted to evaluate its effectiveness. Study 1 examined how different formats of holographic feedback design influence force calibration and visual attention during a simplified sensorimotor task, while Study 2 assessed the impact of MR torque feedback on paediatric ETI training outcomes. The results demonstrate the feasibility of real-time MR feedback for skill acquisition and provide design insights for next-generation robotic training simulation.

2 | RELATED WORK

| Sensorised Training Systems for Endotracheal Intubation

Research on endotracheal intubation (ETI) training systems has increasingly focused on using sensor technologies to address the complexity of the procedure. These systems typically employ real-time feedback to help trainees perceive forces and positions that are difficult to judge through conventional means. For instance, Zhou et al. [6] developed a sensor-embedded phantom that provided vibrotactile cues in response to lifting and pushing forces applied to the tongue, which significantly reduced excessive force application.

Other approaches incorporate a broader range of sensing modalities to monitor device positioning and force distribution across the oral cavity. Notable examples include the sensorised video laryngoscope neonatal intubation simulator developed by Covelli et al. [7] and the Difficult Airway Management Simulator Evaluation System by Noh et al. [8].

Expanding upon these efforts, Hou et al. [5] introduced the SLTS, which integrates multiple sensors into a single laryngoscope to provide a more comprehensive assessment of novice performance. The system uses a six-axis force/torque sensor (Nano 17, ATI Industrial Automation) to measure blade forces, a nine-axis inertial measurement unit (IMU) (GY-95T, Guangyun Electronics) for orientation tracking, and distributed CySkin tactile sensors [9] to monitor grip pressure, all without altering the tool's form factor. Real-time sensor signals were displayed on an external monitor to give users immediate feedback during trials. While preliminary studies confirmed the robustness of sensor integration and data synchronisation, subjects reported challenges related to cognitive load due to frequent attention shifts between the manikin and monitor. Additionally, the feedback's numerical presentation made interpretation difficult during the task. These insights motivated the development of the MR-SLTS, which seeks to deliver more intuitive and immersive feedback using a mixed reality interface, as described in the following sections.

| Mixed Reality in Medical Training

MR has become an increasingly valuable tool in simulation-based medical simulation and surgical training, offering holographic overlays that present procedural information directly within the trainee's field of view. MR systems using optical see-through head-mounted displays (OST-HMDs), such as the HoloLens 2, have been applied to surgical guidance [10], preoperative planning [11], and intraoperative navigation [12]. Studies show that presenting information directly in the user's line of sight reduces the cognitive effort associated with attention switching, thereby improving learning outcomes [13, 14].

While MR is increasingly used in surgical contexts, relatively few studies have explored its application in airway management training. Bhavsar et al. [15] demonstrated that the HoloLens 2 could be used effectively to deliver instructional videos, remote coaching, and augmented visual annotations during intubation training. Extending this line of research, our study investigates the impact of using MR to provide real-time quantitative feedback through holographic visualisation and its influence on ETI training performance.

| Eye-Gaze Tracking and Cognitive Load in Medical Training

Medical procedures that require fine motor control and high levels of attention, such as ETI, are particularly susceptible to cognitive overload when trainees must process large volumes of information simultaneously. Eye tracking has been widely used to assess cognitive load in surgical and robotic training environments [16]. Evidence suggests that frequent rapid eye movements (saccades) may fragment visual perception and disrupt the temporal alignment between visual input and motor output, which is critical for efficient sensorimotor coordination [17, 18].

This study investigates how different holographic visualisation formats influence gaze fixation patterns and attentional focus. The resulting insights inform the design of feedback strategies within the MR-SLTS and contribute to a deeper understanding of cognitive load management in immersive training environments.

3 | DESIGN AND METHODS

3.1 | System Architecture

The architecture of the MR-SLTS extends the previously developed SLTS [5], incorporating a mixed reality interface for enhanced visual feedback as illustrated in Fig. 2. The MR-SLTS preserves the original data acquisition and processing

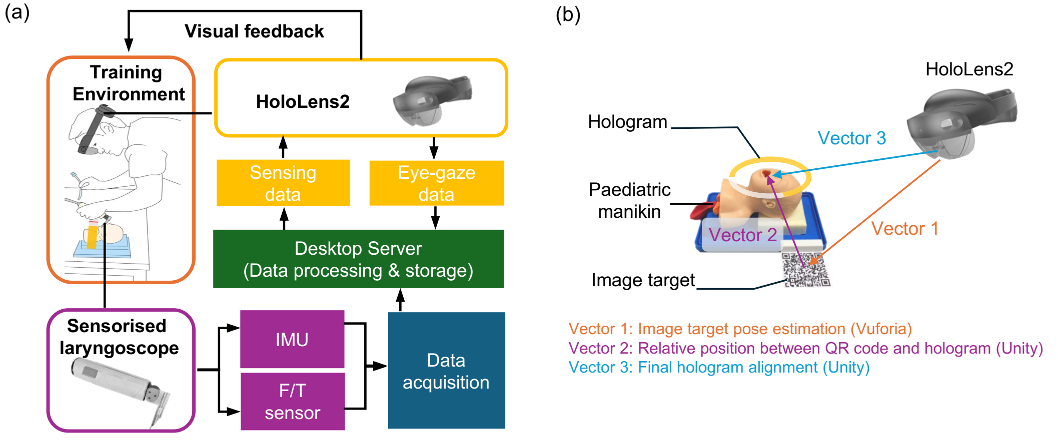

FIGURE 2 **(a)** Overview of the MR-SLTS data pipeline. Sensing signals from the F/T sensor and IMU are acquired via NI DAQ and Arduino, processed on a desktop-based MATLAB server, and streamed to the HoloLens 2 for real-time visual feedback in the training environment, with gaze data sent back for synchronised analysis. **(b)** QR code–based tracking and hologram rendering process. Vuforia detects the QR code and estimates its pose (Vector 1). Unity applies the predefined relative transform (Vector 2). Unity then computes the final hologram position (Vector 3), ensuring consistent alignment above the manikin's face across sessions.

pipeline while introducing an MR visual feedback interface via HoloLens 2. The data acquisition, processing and visualisation pipeline of MR-SLTS is illustrated in Fig. 2a.

Specifically, sensing signals from the embedded F/T sensor and IMU are collected using an NI DAQ device (USB-6363, National Instruments) and an Arduino UNO board, respectively. These signals are then transmitted to a desktop based server running MATLAB, which is responsible for real-time data processing, storage, and bidirectional communication with the HoloLens 2.

The HoloLens 2, wirelessly interfaced with the MATLAB server via the TCP protocol, subscribes to the processed sensing data and renders corresponding holographic visualisations in the user's field of view. In parallel, the HoloLens 2 also collects eye-gaze data throughout the training session, which are transmitted back to the MATLAB server for synchronised storage and later analysis.

3.2 | Development of the MR Feedback Platform

The MR-SLTS software platform was developed using Unity (Unity Technologies, San Francisco, CA, USA), a widely used engine for creating holographic applications tailored to the HoloLens 2 headset. The development process was supported by the Mixed Reality Toolkit (MRTK), an open source framework from Microsoft that streamlines the creation of mixed and augmented reality applications by providing prebuilt components and interaction models optimised for HoloLens devices.

Visual Studio 2022 (Microsoft Corporation, Redmond, WA, USA) was used to build, debug, and deploy the application to the HoloLens 2. The project also leveraged the Universal Windows Platform (UWP) Software Development Kits (SDKs), which offer the runtime libraries and APIs necessary for running Unity based applications on Windows-compatible MR hardware.

These tools collectively enabled the creation of an interactive Unity application capable of rendering real-time numerical feedback within the HoloLens 2's MR environment.

3.2.1 | Visual Feedback Registration

In the MR-SLTS design, the precision requirement for holographic visual feedback registration is relatively modest, with deviations below 5 mm considered acceptable based on informal usability testing. This tolerance reflects the acceptable visual displacement between the intended hologram location and its perceived position in the user's field of view. Since the feedback is not anatomically anchored but is instead presented near the manikin to support user interpretation, such minor deviations do not impair usability.

To ensure robust and repeatable spatial anchoring of holograms, the Vuforia Engine SDK (PTC Inc., Boston, MA, USA) was employed to track a feature-rich 8 × 8 cm QR-code image target. By relying on 2D image target detection, which offers higher precision and lower latency than 3D object detection, the system maintained a stable spatial reference for registration.

The QR-code target was mounted on a 3D-printed tracking guide attached to the lower-left corner of the paediatric manikin. Figure 2b illustrates the tracking and hologram-rendering workflow. Details of the procedure can also be found in the supplementary material. The Vuforia Engine first detects the QR code and estimates its pose (Vector 1) within the HoloLens spatial-mapping coordinate system. Unity then applies the predefined relative transform between the image target and the hologram (Vector 2) to render the overlay in a consistent position and orientation above the manikin's face (Vector 3). This configuration ensured that the hologram remained spatially aligned both within a single session and across repeated training sessions, even when the manikin was physically repositioned.

3.2.2 | Holographic Feedback Design

To minimise cognitive load during ETI, MR-SLTS simplifies how quantitative sensor data are presented. Previous user studies indicated that the error bar format used in the design of SLTS was difficult to interpret in real time [5], often overwhelming trainees during motor-intensive tasks [19, 20].

Comparative analysis between experts and novices has identified torque about the z-axis (coordinate system defined in Figure 1a) as a key indicator of improper wrist rotation in novices [21], and one that improves measurably with training [22]. This torque corresponds to the rotational force applied when lifting the laryngoscope blade, which, if excessive or misaligned, can lead to patient injury or failed intubation. Measurements of linear force are confounded by the relatively stiff jaw, upper airway, and tracheal structures, as well as variable contact conditions, whereas torque more directly reflects the operator's wrist rotation and blade manipulation. Due to its clear interpretability and sensitivity to skill progression, this single parameter was selected as the target for real-time visual feedback. Focusing on one dimension also reduces information overload and helps trainees concentrate on a specific skill.

To deliver continuous and easily interpreted visual feedback, three feedback formats were prototyped: a bar graph, which emphasises simplicity but lacks temporal context; a ring graph, which conveys porportions effectively but similarly omits temporal changes; and a line graph, which illusrates trends over time but is visually more complex.

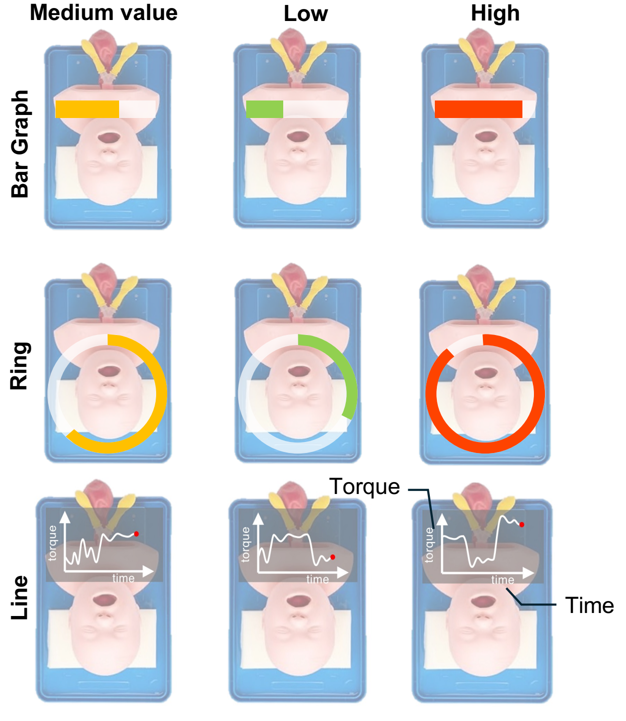

FIGURE 3 Comparison of visual feedback formats. Each row corresponds to one feedback format: bar (top), ring (middle), and line graph (bottom). Each column shows the same sensor reading at the same time point rendered in the three formats. This arrangement illustrates the different ways in which identical data are conveyed, highlighting trade-offs between simplicity, proportional representation, and temporal trend visualisation.

The selected visualisation was spatially calibrated to remain within the user's consistent field of view relative to the manikin, ensuring accessibility and reducing distraction during ETI trials. The holographic feedback designs are shown in Fig. 3.

3.2.3 | Wireless Communication and Latency Considerations

Reliable and low latency communication between the HoloLens 2 and desktop server is essential for effective real-time visual feedback. Several wireless protocols were considered, including TCP, UDP, HTTP, and Bluetooth-based alternatives. TCP over WiFi was ultimately selected for its high reliability, consistent data ordering, and well documented implementation support in both Unity (via MRTK) and MATLAB.

Although UDP may offer lower latency, our preliminary tests showed that TCP communication incurred an average delay of only 4.17 ms per transmission. The HoloLens 2 introduced an additional display latency of 16 ms [23], while the improved sensor data readout rate of 70 Hz corresponded to a delay of approximately 14 ms. Together, these factors kept the total system latency well below the 100 ms threshold commonly cited as acceptable for surgical and motor feedback tasks [24, 25]. Consequently, the holographic feedback appeared in near real time and did not perceptibly disrupt trainee performance.

3.2.4 | Eye-Gaze Tracking Integration

The HoloLens 2 provides built-in eye tracking capabilities, which were utilised through the MRTK in Unity to collect eye gaze direction and head movement data during intubation trials. These data were transmitted in real time via TCP communication to a desktop-based server for storage and post analysis. This setup enabled detailed examination of trainee visual attention patterns, including fixation behaviour and saccade frequency.

4 | EXPERIMENTAL PROTOCOL AND SETUP

This research comprises two sequential studies designed to inform and evaluate the holographic feedback design used in MR-SLTS. Study 1 compares how different visualisation formats (e.g., bar, ring, line graphs) convey real-time quantitative torque data during a simplified intubation-related task in an MR environment. The goal is to identify the most effective and intuitive format based on user performance. The outcome of Study 1 informs the design of Study 2, which evaluates the training efficacy of the selected holographic feedback format. In Study 2, subjects engage with the full MR-SLTS system in a realistic paediatric intubation scenario to assess how the chosen visualisation supports skill acquisition. An overview of the study timeline and design is provided in Supporting Figure S1.

4.1 | Study 1: Holographic Feedback Comparison

Study 1 aimed to evaluate the impact of different holographic feedback formats, including bar, ring, and line graphs, on users' performance in a simplified sensorimotor coordination task using HoloLens 2. The findings would inform the optimal feedback design for the MR-SLTS system, later validated in Study 2.

Six subjects with no prior experience with OST-HMDs were recruited. The sample included both left- and right-handed individuals. Subjects who normally used corrective lenses or contact lenses wore them during the experiments. Although clinical laryngoscopes are conventionally designed for left-handed use, subjects were permitted to use their

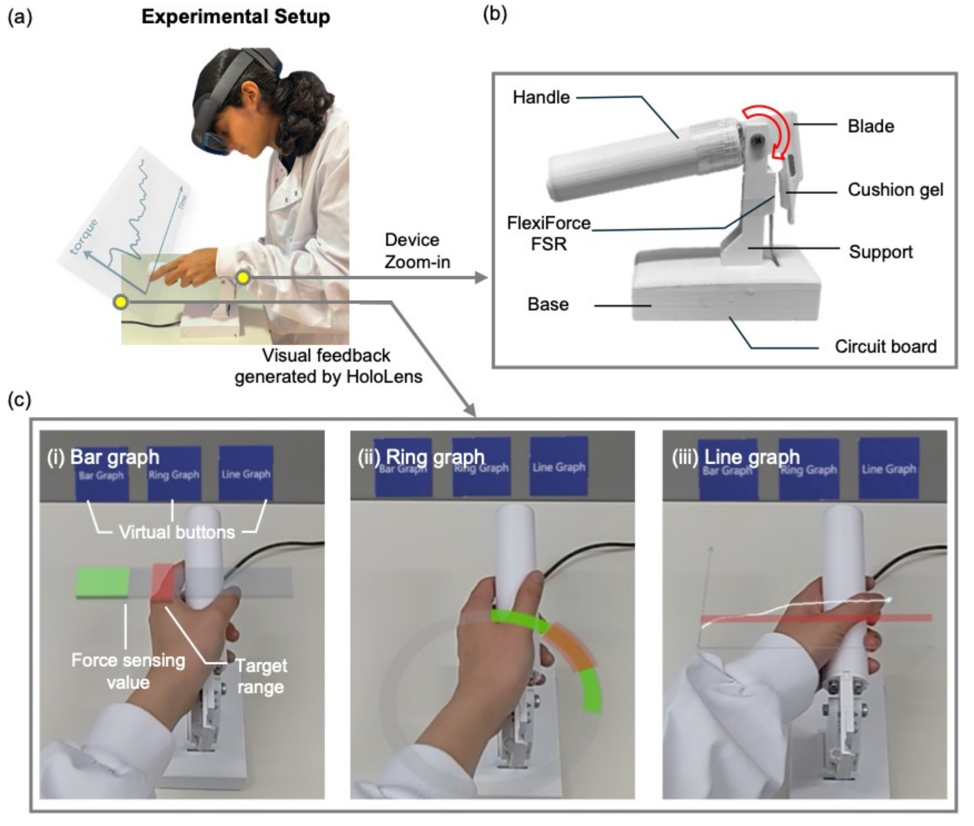

FIGURE 4 **(a)** 3D-printed laryngoscope prototype with compression force measured by a FlexiForce sensor. **(b)** Experimental setup schematic. **(c)** HoloLens screenshots showing the subject's view for (i) bar graph, (ii) ring graph, and (iii) line graph.

preferred hand when conducting the experiments. All subjects provided their informed consent under an approved ethics protocol (CUREC R91525/RE001).

| Experimental Setup

The experimental setup mirrored the MR-SLTS architecture in simplified form. A 3D-printed laryngoscope prototype with a single rotational degree of freedom was mounted on a support stand, as shown in Fig. 4a. Compression forces applied by the blade were captured using a FlexiForce™ sensor and streamed via an Arduino MEGA microcontroller to a laptop-based MATLAB server. Holographic feedback was rendered on HoloLens 2 using Unity with the MRTK framework. Eye-gaze data and force outputs were synchronised via bidirectional TCP communication for storage and analysis. The experimental setup for Study 1 is illustrated in Fig. 4b.

Each graphical format (bar, ring, line) was displayed above the laryngoscope in the user's field of view (HoloLens 2 photoshots shown in Fig. 4c), with position calibrated using a QR code image target. Prior to trials, each subject completed eye-gaze calibration to ensure accurate data collection.

| Protocols

Each subject completed 21 trials, including six initial training trials excluded from analysis. The remaining 15 test trials consisted of five trials per feedback format, presented in randomised order. Each 20-s trial required subjects to adjust the applied force to remain within a time-varying target range $[L_{\min}, L_{\max}]$, defined by step-changing upper and lower limits that mimicked the dynamics of sensorimotor control tasks typical of endotracheal intubation procedures (see Supplementary Fig. S2). Force data were normalised to a 0–100 scale, with targets set between 10 and 45 to reflect delicate procedural demands. Eye-gaze direction and applied force were recorded continuously for post-trial analysis.

| Performance Metrics and Data Processing

The performance of the subjects was quantified using three metrics.

First, the tracking error was defined by comparing the applied force M with the target force range $L_{\min}$ and $L_{\max}$. At each time step, the instantaneous error was calculated as

$$\text{error}(i) = \begin{cases} M(i) - L_{\max}(i) & \text{if } M(i) > L_{\max}(i) \\ L_{\min}(i) - M(i) & \text{if } M(i) < L_{\min}(i) \\ 0 & \text{otherwise} \end{cases} \tag{1}$$

The root-mean-square error (RMSE) was then computed over the valid duration of each trial (excluding the first 2 seconds) and averaged across trials per feedback type.

Second, the eye-gaze shift distance defined the displacement between consecutive eye-gaze direction vectors $\vec{g}_i = (x_i, y_i, z_i)$, and it was calculated as:

$$G_{\text{shift}}(i) = \sqrt{(x_{i+1} - x_i)^2 + (y_{i+1} - y_i)^2 + (z_{i+1} - z_i)^2} \tag{2}$$

A peak detection algorithm was applied on G_{shift} to identify gaze shifts exceeding 0.03 m, which were classified as saccades.

Third, the response time and fluctuation periods were computed using a threshold-based algorithm. The algorithm identifies the force calibration state as stabilised when the applied force remains within the target range continuously for more than 0.6 s after each change of target range, as illustrated in Supporting Fig. S3. Applied to the force-sensing data, it segments periods of force fluctuation (i.e., active adjustment) and stabilisation. From this, the average and median fluctuation durations following target changes were calculated, providing a measure of response speed.

All data were processed with Python. Statistical comparisons between feedback types were performed using paired t-tests or Mann-Whitney U tests, depending on normality.

4.2 | Study 2: MR-SLTS Evaluation

Study 2 aimed to evaluate the effect of real-time holographic visual feedback on z-torque in improving novice performance during endotracheal intubation (ETI) training. It was hypothesised that graphical torque feedback would help reduce excessive force application without significantly increasing procedure time or eye distraction.

A randomised controlled experiment was conducted with ten subjects (aged 22–27) with no prior ETI experience. subjects were assigned to either a feedback group (received real-time Z-torque feedback via a bar graph in the HoloLens 2 display) or a control group (no feedback, but wore HoloLens for gaze data collection). Ethical approval was obtained from the Central University Research Ethics Committee (CUREC, R91525/RE001).

| Experimental Setup

Experiments were conducted using the full MR-SLTS system, illustrated in Fig. 5a. The sensorised laryngoscope transmitted z-torque and angular position data to a MATLAB-based server, which streamed processed values to the HoloLens 2 via TCP. CySkin tactile sensors were disabled to improve data throughput.

Feedback consisted of a bar graph displaying real-time Z-torque values in the HoloLens 2 view. This format was selected based on the optimal results obtained in Study 1. Only negative torque was visualised, as it corresponds to the lifting action of the tongue during laryngoscopy. A red threshold line was included to indicate the maximum allowed z-torque derived from expert performance reported in prior literature (370.65 Nmm), based on measurements from three licensed surgeons [5], beyond which the feedback signaled an excessive force. Fig. 5b shows the in-situ visualisation as seen by subjects during training trials.

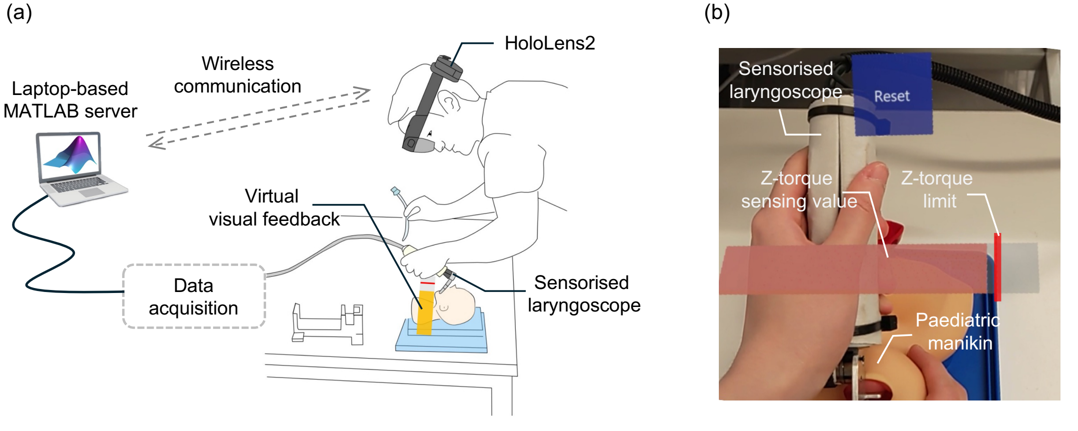

FIGURE 5 Study 2: Mixed-Reality Sensorised Laryngoscope Training System (MR-SLTS). **(a)** Physical MR-SLTS setup and data flow. The sensorised laryngoscope transmits torque and position data to a MATLAB-based server, which streams the processed values to the HoloLens 2. The headset generates virtual visual feedback using real-time torque data and overlays it above the manikin within the user's field of view. **(b)** HoloLens 2 view showing real-time holographic feedback during intubation. The bar-graph feedback indicates the applied torque relative to the defined safe threshold, providing intuitive guidance to the operator.

| Protocols

After an instructional video and three baseline trials with a standard laryngoscope, subjects completed ten training trials using MR-SLTS. Each ETI trial required them to insert an endotracheal tube into the manikin's trachea within a 30-second time limit. A trial was considered successful if, upon blowing air into the endotracheal tube using an resuscitation bag, the lung-representing balloons could be visibly inflated. Between trials, tube placement was verified and the laryngoscope was recalibrated using IMU data. Feedback was enabled only for the experimental group. After a 5-minute rest period, all subjects completed three trials without visual feedback.

Force/torque, eye-gaze, and laryngoscope angle data were recorded for all trials and used to evaluate differences in training outcomes between groups.

| Performance Metrics and Data Processing

The ETI performance in Study 2 was evaluated using multiple quantitative metrics.

First, a Z-torque analysis was conducted to assess real-time feedback during training for the experimental group. Z-torque values were continuously recorded throughout each intubation trial. Trials with unsuccessful tube placement or durations exceeding 30 seconds were excluded from analysis. For each successful trial, z-torque signals were time-normalised to 100% trial completion. A sliding window of 1% width with a 0.5% step was then applied to compute smoothed mean and standard deviation curves.

Second, the percentage errors of force and torque were calculated for each of the six components using the RMSE, following the same formulation as in Study 1 (Equation (1)). In this study, however, L_{max} and L_{min} were redefined as the expert-defined directional limits of force and torque, corresponding to the maximum allowable magnitudes in the positive and negative orientations of each axis ($F_x, F_y, F_z, T_x, T_y, T_z$). These limits were determined by Hou *et al.* [5] based on the mean performance of three licensed surgeons, each performing twenty intubation sessions. The corresponding directional limits for each force and torque component are summarised in Table 1. The percentage error was obtained as

$$\text{Percentage Error} = \left(\frac{\text{RMSE}}{\text{Max Absolute Range Value}} \right) \times 100\% \tag{3}$$

TABLE 1 L_{max} and L_{min} of F/T Sensing

	F_x	F_y	F_z	T_x	T_y	T_z
L_{max}	4.31	1.53	1.16	66.9	79.73	15.53
L_{min}	-3.76	-4.21	-1.38	-28.31	-47.79	-370.65

Third, intubation duration was recorded for every trial and analysed across three training stages: early (trials 1–4), late (trials 5–8), and final testing (trials 9–11). The distribution of trial durations across these three phases were analysed.

Eye-gaze shifts were processed following the same procedure as in Study 1. Gaze shift distances between consecutive frames were computed, and peaks exceeding a 0.03 m threshold were classified as saccades. The total number of saccades per trial was used to compare eye-movement patterns between groups and across time.

5 | RESULTS

| Tracking Error

To assess tracking performance, RMSE values were calculated based on deviations between the applied force and the target force range. The first two seconds of each trial were excluded to remove initial transients. Fig. 6a shows the average RMSE per subject for each type of feedback. The bar graphs produced the lowest average error, with a statistically significant difference observed between the bar and ring conditions ($p < 0.05$, Student's t-test). The difference between bar and line feedback was not statistically significant ($p > 0.05$). The bar graph condition also exhibited the most consistent performance across and within subjects as indicated by the lowest standard deviation among the three feedback methods.

| Eye-Gaze Behaviour

Eye-gaze behaviour was analysed by counting saccadic shifts, based on displacement thresholds applied to consecutive gaze direction vectors. We applied a peak detection algorithm in Python using a single threshold to identify all peaks with a shift distance greater than 0.03 m, classifying them as saccades [26]. An example of the detected saccades from a example experimental trial is illustrated in Supporting Fig. S5. Fig. 6b presents the histogram of the total number of saccades per condition. The first bin (0–1 s) was excluded from analysis due to its notably higher count, reflecting the initial gaze adjustment phase. Trials with ring graphs yielded significantly more saccades than those with line graphs ($p < 0.05$, Mann–Whitney U test), while the difference between bar and line graphs was not significant ($p > 0.3$). Most saccades occurred shortly after changes in the target force values.

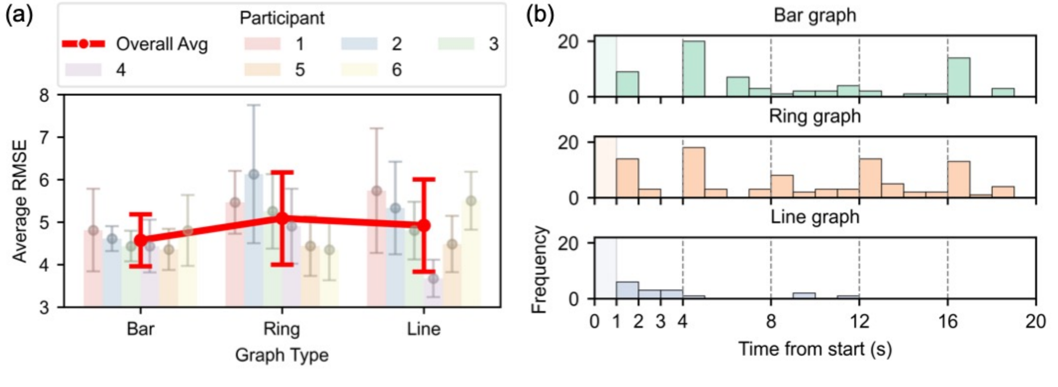

FIGURE 6 Comparison of target-tracking performance and visual demand across different graph visualisation formats in Study 1. **(a)**Average target tracking RMSE for each graph type across individual subjects in Study 1. The plot includes per-subject results, along with the overall average and standard deviation for each visualisation format, highlighting differences in tracking accuracy. **(b)**Histogram showing the total number of saccades recorded for each visualisation type during the target tracking task. This metric reflects the visual demand and eye movement load associated with interpreting each graph format.

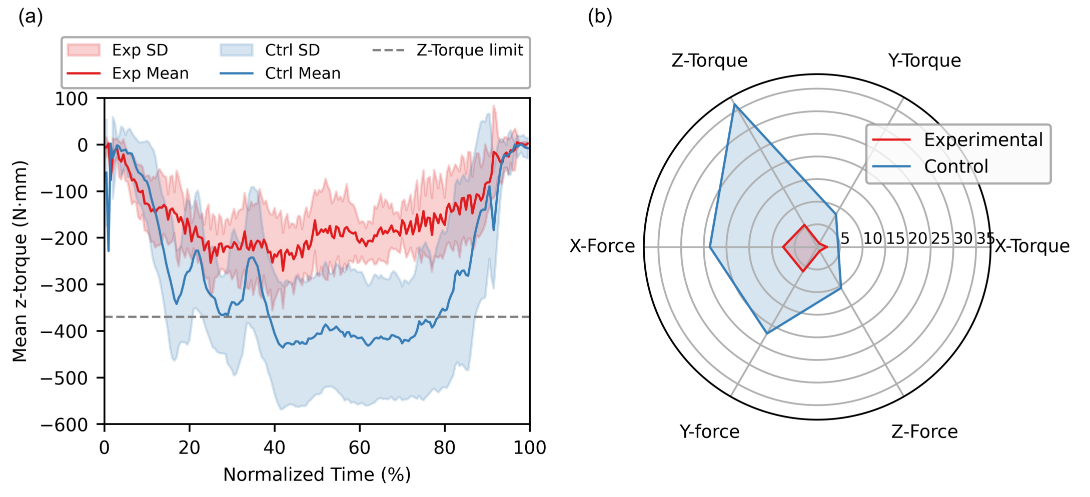

FIGURE 7 **(a)** Average Z-axis torque measurements plotted over normalised time for all successful test trials. The curve represents the typical torque profile applied during endotracheal intubation, providing insight into force patterns over the course of the procedure. **(b)** Comparison of average percentage errors in F/T sensing between the experimental group (with Z-torque feedback) and the control group (without feedback). The results illustrate the impact of real-time visual feedback on F/T control during endotracheal intubation training.

| Response Time

To further evaluate force control of subjects, a threshold-based segmentation algorithm was applied to identify periods of force fluctuation following each target change. An example trial illustrating the segmented fluctuation and stabilised periods of the force sensing signal is shown in Supporting Fig. S4.

Fluctuation durations were comparable across the three feedback formats. The mean (± median) durations were 2.04 s (1.77 s) for the bar graph, 2.09 s (1.86 s) for the ring graph, and 2.34 s (1.94 s) for the line graph. Although the line-graph feedback yielded slightly longer fluctuation periods than the bar and ring formats, the differences were not statistically significant (Mann–Whitney U tests: bar vs. ring, $p = 0.821$; bar vs. line, $p = 0.088$; ring vs. line, $p = 0.105$). These trends may reflect more cautious motor adjustments in response to the line graph's historical trend information, resulting in smoother and more controlled force modulation.

Based on the findings from Study 1, where bar graph feedback resulted in the lowest tracking error, moderate eye distraction, and faster force calibration, this feedback modality was adopted in Study 2 to evaluate training efficacy.

| Force/Torque

In the experiments of Study 2, out of 30 test trials conducted by the 10 subjects, 9 were excluded due to unsuccessful intubation or exceeding the 30-second limit, resulting in 21 valid trials for analysis. Fig. 7a shows the normalised z-torque trajectories for the remaining trials. The experimental group exhibited consistently lower torque values across the procedure compared to the control group.

Percentage errors for all six force and torque dimensions are summarised in Fig. 7b. The experimental group demonstrated lower average percentage errors across all dimensions. The Z-torque error in the control group was the highest among all axes, and an approximate 30% reduction in the experimental group reflects the effectiveness of feedback in reducing excessive rotational compression.

| Duration and Saccades

The duration of intubation was analysed, and no statistically significant differences were observed between the experimental and control groups or throughout the training progression ($p > 0.3$, Mann-Whitney U test). Eye-gaze analysis revealed that the experimental group exhibited a higher average number of saccadic gaze shifts than the control group across all 11 trials ($p < 0.05$, Mann–Whitney U test), a trend that persisted in both the training and testing phases.

6 | DISCUSSION AND CONCLUSION

This study introduced and evaluated the Mixed-Reality Sensorised Laryngoscope Training System (MR-SLTS), a novel training platform that delivers real-time quantitative feedback through HoloLens 2. The system integrates intuitive graphic visualisation, low-latency wireless communication, and eye tracking to support sensorimotor skill development during paediatric ETI training.

In a preliminary sensorimotor task (Study 1), the bar graph feedback produced the lowest tracking error, minimal visual distraction, and the fastest force calibration compared to the ring and line graphs. These findings informed its implementation in the MR-SLTS system. In the subsequent training study (Study 2), subjects using bar graph feedback showed a significantly reduced force and torque application without increasing the procedure time. Their performance also resembled more closely the expert benchmarks. Eye tracking data revealed fewer attentional shifts and supported insights into user's evolving motor strategies. Taken together, the findings underscore the important role of MR-based feedback in advancing technical performance and sustaining contact force/torque regulation in high-stakes clinical training settings.

The current MR-SLTS system focused exclusively on z-torque feedback. Although effective, this single-axis feedback may not capture the full spectrum of trainee performance. Expanding to multidimensional feedback including x- and y-forces and integrating haptic or animated cues could provide more comprehensive guidance. Machine learning algorithms could also enable adaptive and personalised feedback strategies. Future work should explore the combination of electromyography (EMG), motion tracking, and expert modeling to allow richer performance evaluation. Larger clinical trials will be essential for validating the system's readiness for real-world deployment and for benchmarking its educational impact across different learner populations.

In general, MR-SLTS demonstrates the feasibility and benefits of integrating real-time sensor feedback with immersive visualization for next-generation medical training tools.

Acknowledgements

This work was funded by the Royal Society Research Grant RGS/R2/242066. Dr. Louis P. Halamek's time was supported by the Endowment for the Center for Advanced Pediatric and Perinatal Education at Packard Children's Hospital at Stanford.

Supporting Information

Supplementary Video

See attachment

Study Timeline and Structure

An overview of the two-part study design is shown in Figure 8. Subjects were first recruited for Study 1, following the development of a 3D-printed prototype setup used for the simplified sensorimotor task. Study 1 compared different holographic feedback designs to identify the optimal format. The first round of data analysis evaluated the efficiency and intuitiveness of each feedback format based on tracking error, eye-gaze shift frequency, and response speed. The selected feedback method was subsequently implemented in the MR-SLTS setup for Study 2, a full endotracheal intubation (ETI) training simulation. Experimental data collected in Study 2 were examined in the final data analysis round to derive the overall results.

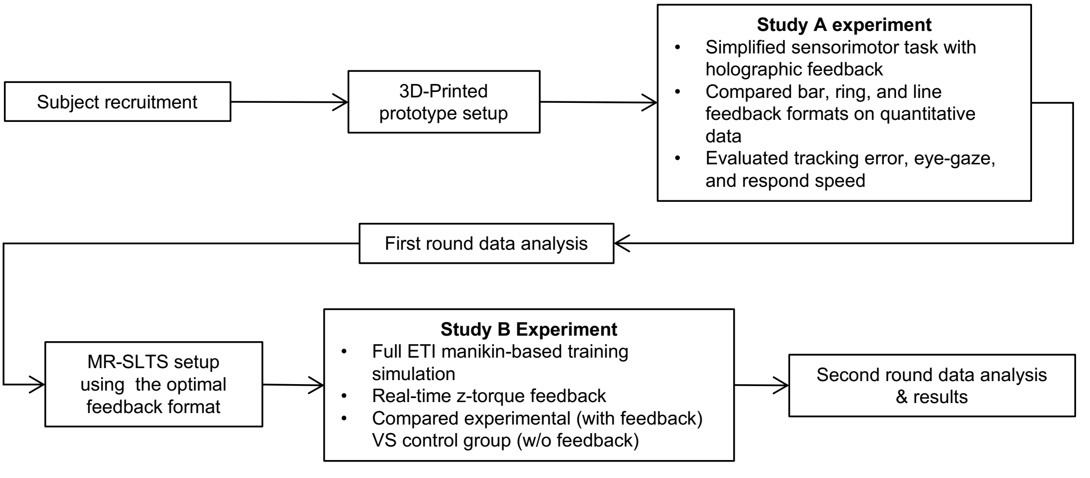

FIGURE 8 Overview of the experimental timeline and structure of the Study 1 and Study 2.

Force Tracking Target

In Study 1, each subject completed a total of 21 target-tracking trials. Each 20-second trial featured a dynamically changing target range, to which subjects continuously adjusted the force output measured by the laryngoscope prototype. Figure 9 illustrates the target values plotted against time, which remained consistent across all trials. The target value was updated every 4 seconds, and a tolerance range of 10 (i.e., the interval between the upper and lower limits within which the error between the applied and target force was considered zero) was defined and normalised to a full-scale value of 100.

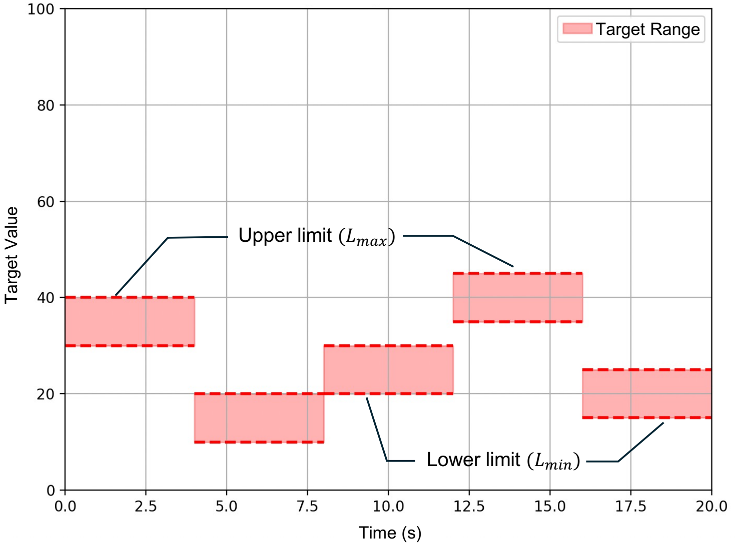

FIGURE 9 Study 1 step-changing force tracking target range against time

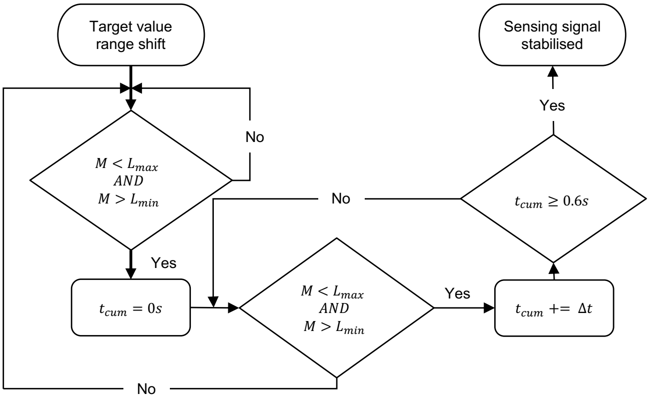

FIGURE 10 Flowchart of the threshold-based segmentation algorithm used to identify periods of force fluctuation.The threshold algorithm identifies the force calibration state as stabilised when the applied force remains within the target range continuously for more than 0.6 s after each change of target range.

Fluctuation Segmentation Algorithm

Figure 10 details the logic of the threshold-based segmentation algorithm used to distinguish periods of force fluctuation and stablisation in force sensing signals of Study 1. An example of its output is shown in Figure 11, with the duration of each fluctuation period annotated.

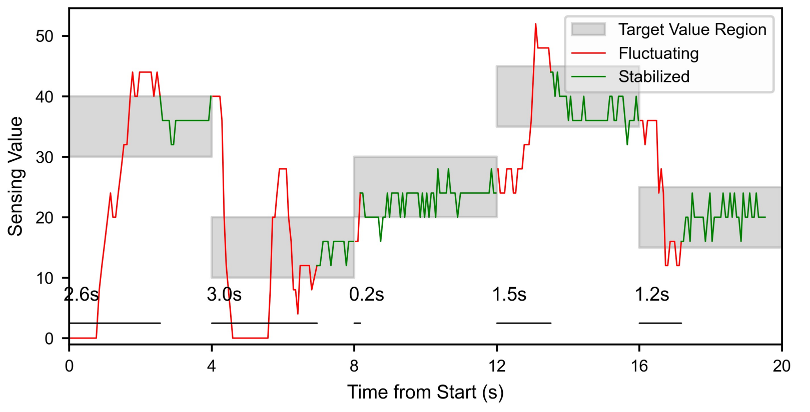

FIGURE 11 Fluctuation and stabilised periods of an example experimental trail.

Saccades Detection

Figure 12 shows an example of the eye-gaze shift distance over time from a representative experimental trial, with distinct sharp peaks indicating saccadic movements. Peaks were detected using the `find_peaks` function in SciPy (Python), applying a fixed amplitude threshold of 0.03 m to identify all saccadic shifts.

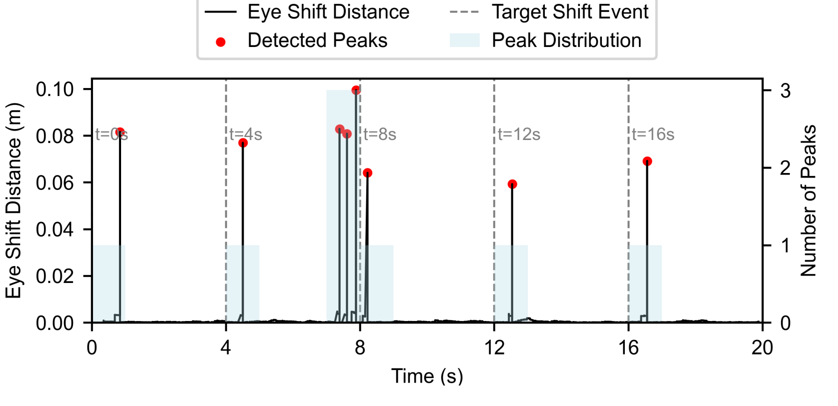

FIGURE 12 Eye-shift distance over time for one example experimental trail.

references

[1] Rahman T, Chandran S, Kluger D, Kersch J, Holmes L, Nishisaki A, et al. Tracking manikin tracheal intubation using motion analysis. Pediatric Emergency Care 2011 8;27:701–705. https://journals.lww.com/pec-online/fulltext/2011/08000/tracking_manikin_tracheal_intubation_using_motion.4.aspx.

[2] Carlson JN, Das S, Torre FDL, Callaway CW, Phrampus PE, Hodgins J. Motion capture measures variability in laryngoscopic movement during endotracheal intubation: A preliminary report. Simulation in healthcare : journal of the

Society for Simulation in Healthcare 2012 8;7:255. /pmc/articles/PMC3412923//pmc/articles/PMC3412923/?report=abstracthttps://www.ncbi.nlm.nih.gov/pmc/articles/PMC3412923/.

[3] Alvarado AC, Panakos P. Endotracheal Tube Intubation Techniques. StatPearls 2023 7;https://www.ncbi.nlm.nih.gov/books/NBK560730/.

[4] Miller AG. Endotracheal Intubation Training and Skill Maintenance for Respiratory Therapists. Respiratory Care 2017 2;62:156–162. https://rc.rcjournal.com/content/62/2/156https://rc.rcjournal.com/content/62/2/156.abstract.

[5] Hou N, He L, Albini A, Halamek L, Maiolino P. The design of a sensorized laryngoscope training system for pediatric intubation. 2024 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS) 2024 10;p. 2414–2420. https://ieeexplore.ieee.org/document/10801567/.

[6] Zhou H, Yang S, Halamek L, Nanayakkara T. A Method to Use Haptic Feedback of Laryngoscope Force Vector for Endotracheal Intubation Training. Proceedings - IEEE International Conference on Robotics and Automation 2023;2023-May:6810–6816.

[7] Covelli A, Bardelli S, Scaramuzzo RT, Sigali E, Ciantelli M, Pistoia MD, et al. Effectiveness of a new sensorized videolaryngoscope for retraining on neonatal intubation in simulation environment. Italian Journal of Pediatrics 2020 2;46:1–11. https://ijponline.biomedcentral.com/articles/10.1186/s13052-020-0774-z.

[8] Noh Y, Wang C, Tokumoto M, Matsuoka Y, Chihara T, Ishii H, et al. Development of the airway Management Training System WKA-5: Improvement of mechanical designs for high-fidelity patient simulation. 2012 IEEE International Conference on Robotics and Biomimetics, ROBIO 2012 - Conference Digest 2012;p. 1224–1229.

[9] Maiolino P, Maggiali M, Cannata G, Metta G, Natale L. A flexible and robust large scale capacitive tactile system for robots. IEEE Sensors Journal 2013;13:3910–3917.

[10] Cleary K, Peters TM. Image-guided interventions: Technology review and clinical applications. Annual Review of Biomedical Engineering 2010 8;12:119–142. https://www.annualreviews.org/content/journals/10.1146/annurev-bioeng-070909-105249.

[11] bo Zou Y, min Chen Y, ke Gao M, Liu Q, yu Jiang S, hui Lu J, et al. Coronary Heart Disease Preoperative Gesture Interactive Diagnostic System Based on Augmented Reality. Journal of Medical Systems 2017 8;41.

[12] Eom S, Kim S, Rahimpour S, Gorlatova M. AR-Assisted Surgical Guidance System for Ventriculostomy 2022;p. 402–405.

[13] Herrlich M, Tavakol P, Black D, Wenig D, Rieder C, Malaka R, et al. Instrument-mounted displays for reducing cognitive load during surgical navigation. International Journal of Computer Assisted Radiology and Surgery 2017 9;12:1599–1605. https://link.springer.com/article/10.1007/s11548-017-1540-6.

[14] Paas F, Tricot A, Ouwehand K, Lespiau F, Mugisha G, Arguel A. Procedural Learning in Mixed Reality: Assessing Cognitive Load and Performance. Education Sciences 2025, Vol 15, Page 339 2025 3;15:339. https://www.mdpi.com/2227-7102/15/3/339/htmhttps://www.mdpi.com/2227-7102/15/3/339.

[15] Bhavsar N, Sriram S, Balasubramanian S, Davidson C, Piechowski W, Zimm J, et al., AIRWAY-XR: Augmented Instruction to Refine Wayfinding and Yielding Skills in Emergency Medicine Residents for Intubation using Mixed Reality Technology; 2025. http://medrxiv.org/lookup/doi/10.1101/2025.01.06.24319788.

[16] Junior FCS, Litchfield D, Sandars J, Cecilio-Fernandes D. Use of eye tracking in medical education. Medical Teacher 2024;https://www.tandfonline.com/doi/abs/10.1080/0142159X.2024.2316863.

[17] Nachmani O, Coutinho J, Khan AZ, Lefèvre P, Blohm G. Predicted Position Error Triggers Catch-Up Saccades during Sustained Smooth Pursuit. eNeuro 2020 1;7. https://www.eneuro.org/content/7/1/ENEURO.0196-18.2019https://www.eneuro.org/content/7/1/ENEURO.0196-18.2019.abstract.

[18] Coudiere A, Danion FR. Eye-hand coordination all the way: from discrete to continuous hand movements. Journal of Neurophysiology 2024 4;131:652–667. /doi/pdf/10.1152/jn.00314.2023.

[19] Eisma YB, Borst C, van Paassen R, de Winter J. Augmented Visual Feedback: Cure or Distraction? Human Factors 2021 11;63:1156–1168. /pmc/articles/PMC8521352//pmc/articles/PMC8521352/?report=abstracthttps://www.ncbi.nlm.nih.gov/pmc/articles/PMC8521352/.

[20] Puttemans V, Vangheluwe S, Wenderoth N, Swinnen SP. Bimanual Directional Interference: The Effect of Normal versus Augmented Visual Information Feedback on Learning and Transfer. Motor Control 2004 1;8:33–50. https://journals.humankinetics.com/view/journals/mcj/8/1/article-p33.xml.

[21] McWilliam P, Halamek LP, King B, Martineau N, Fuerch J, Yamada NK, et al. Sensor-Integrated Laryngoscope. Journal of Clinical Engineering 2020 1;45:39–44. https://journals.lww.com/10.1097/JCE.0000000000000386.

[22] Hou N, He L, Albini A, Halamek L, Maiolino P, The design of a sensorized laryngoscope training system for pediatric intubation; 2023.

[23] von Haxthausen F, Moreta-Martinez R, de la Lastra APD, Pascau J, Ernst F. UltrARsound: in situ visualization of live ultrasound images using HoloLens 2. International Journal of Computer Assisted Radiology and Surgery 2022 11;17:2081–2091.

[24] Khan T, Zhu TS, Downes T, Cheng L, Kass NM, Andrews EG, et al. Understanding Effects of Visual Feedback Delay in AR on Fine Motor Surgical Tasks. IEEE Transactions on Visualization and Computer Graphics 2023 11;29:4697–4707.

[25] Honda T, Hagura N, Yoshioka T, Imamizu H. Imposed visual feedback delay of an action changes mass perception based on the sensory prediction error. Frontiers in Psychology 2013;4. /pmc/articles/PMC3805955//pmc/articles/PMC3805955/?report=abstracthttps://www.ncbi.nlm.nih.gov/pmc/articles/PMC3805955/.

[26] Purves D, Augustine GJ, Fitzpatrick D, Katz LC, LaMantia AS, McNamara JO, et al. Types of Eye Movements and Their Functions 2001;https://www.ncbi.nlm.nih.gov/books/NBK10991/.