virtual worlds

MDPI

Article

Effect of Cognitive Distractors on Neonatal Endotracheal Intubation Performance: Insights from a Dual-Task Simulator

Yan Meng [1,*], Shang Zhao [2], Xiaoke Zhang [3], John Philbeck [4], Prachi Mahableshwarkar [4], Boyuan Feng [1], Lamia Soghier [5] and James Hahn [1]

[1] Department of Computer Science, George Washington University, Washington, DC 20052, USA
[2] Institute for Advanced Research, University of Science and Technology of China, Suzhou 215123, China
[3] Department of Statistics, George Washington University, Washington, DC 20052, USA
[4] Department of Psychological & Brain Sciences, George Washington University, Washington, DC 20052, USA
[5] Department of Neonatology, Children's National Hospital, Washington, DC 20010, USA
* Correspondence: mengy@gwu.edu

Abstract: Neonatal endotracheal intubation (ETI) is a complex medical procedure that demands extensive training before practicing on real patients. Clinical studies indicate that the conventional training approach, typically conducted in idealized conditions with task trainers, has a low skill transferability rate compared to performance in the dynamic environments common in intensive care units (ICUs). According to cognitive load theory, novices encounter difficulties in multitasking scenarios, exhibiting performance declines due to competition among tasks for cognitive resources; experts, having achieved automaticity, have more cognitive resources to handle additional tasks present in high-stress environments and therefore exhibit less performance degradation. Current ETI skill assessment methods do not capture these differences in expertise. To bridge this gap, we develop an innovative dual-task mixed-reality (MR) simulator to evaluate the influence of cognitive distractors on ETI and substantiate effective performance measurement metrics. Results affirm that experts demonstrate superior proficiency in handling extraneous cognitive loads compared to novices. This has important implications for understanding how to measure novice performance in ETI settings. Taken together, the dual-task ETI training simulator and the associated automated skill evaluation metric system hold promise for enhancing training in neonatal ETI practice and ultimately leading to improved patient care outcomes.

Keywords: computer simulation; human–computer interaction; bioinformatics; medical training; cognitive psychology; ANOVA

check for updates

Academic Editor: Anton Nijholt

Received: 29 March 2025
Revised: 12 May 2025
Accepted: 15 May 2025
Published: 20 May 2025

Citation: Meng, Y.; Zhao, S.; Zhang, X.; Philbeck, J.; Mahableshwarkar, P.; Feng, B.; Soghier, L.; Hahn, J. Effect of Cognitive Distractors on Neonatal Endotracheal Intubation Performance: Insights from a Dual-Task Simulator. *Virtual Worlds* **2025**, *4*, 20. https://doi.org/10.3390/virtualworlds4020020

Copyright: © 2025 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (https://creativecommons.org/licenses/by/4.0/).

1. Introduction

Neonatal endotracheal intubation (ETI) is a commonly employed medical intervention to establish an artificial airway for newborns. This procedure becomes imperative when infants face challenges in autonomous respiration or require mechanical ventilation assistance. Given its time-sensitive nature, the successful execution of this procedure mandates the precise placement of the endotracheal tube by the operator within a narrow window of 30 s [1]. However, pediatric residents maintain low intubation success rates (23–25%) throughout three years of residency, with an average of only three supervised clinical ETI opportunities [2]. This is well below the 26–75 procedures needed to achieve a consistent 90% success rate [3,4]. To foster skill proficiency before clinical exposure, conventional training methods have relied upon task trainers coupled with instructor feedback. Complementing these approaches, diverse simulators have been developed for

ETI training [5,6], providing a secure environment for trainees to refine intubation skills free from the constraints of a neonatal intensive care unit (NICU). Nevertheless, the apparent inadequacy in the transfer of skills acquired through task trainer-based simulators to the actual NICU underscores shortcomings in current training paradigm [7]. Our hypothesis is that traditional training frameworks evaluate trainees' procedural skills within idealized conditions, devoid of the environmental stressors inherent in medical procedures involving real patients. Novices can learn to perform manual skills relatively easily when no distractors are competing for cognitive resources. However, when required to perform two tasks simultaneously, performance often drops because the secondary task consumes the attentional resources that are so crucial for good performance at early skill levels [8]. This perspective contends that insufficient cognitive demands during training may result in an overestimation of trainees' proficiency levels. This overestimation, in turn, may lead to the premature advancement of trainees, or their untimely cessation of training, well before they have garnered the requisite proficiency and experience for skill transfer.

Cognitive load theory explains how the human cognitive system processes and learns new information [9]. It points to a crucial difference between experts and novices in how they allocate attention resources. Experts exhibit the ability to adeptly manage multiple tasks simultaneously with minimal performance degradation, a characteristic referred to as "automaticity" [10]. The evaluation of automaticity, indicative of skilled performance, often involves the assessment of secondary tasks that consume attention resources. Secondary tasks, when used for performance assessment, have proven sensitive enough to discern performance variations among skilled individuals compared to traditional metrics [11–14]. The attainment of automaticity, as indicated by secondary task measures, typically demands more prolonged training periods compared to proficiency determined by single-task metrics.

These findings underscore the need for more sophisticated training methods and nuanced performance metrics, which encompass secondary tasks, to address the transfer of skills learned in simulated settings to real-world medical scenarios. To substantiate these assertions, we designed and implemented a mixed-reality ETI simulator equipped with a dual-task training module, followed by an empirical assessment of participants' performance. Our study compares procedural skills using both single-task and dual-task metrics, providing insights into ETI skill automaticity in complex environments. Data analysis results confirm the efficacy of our dual-task simulator in discerning differences between expert and novice ETI performances, pointing to a direction for future training regimen design for ETI and other medical skill development.

The key contributions of this study are outlined as follows:

- Identify the limitations inherent in the current single-task ETI training paradigm, particularly its inadequacy in transferring skill from simulation to clinical settings.
- Develop the first dual-task mixed-reality (MR) simulator for neonatal endotracheal intubation. The simulator effectively captures expert-novice performance differences that traditional training paradigms fail to detect.
- Facilitate seamless transition between single-task and dual-task modes in future works using the MR simulator.
- Develop an automated skill evaluation strategy for trainee performance.
- Ground the efficacy of the multitask ETI training method in theoretical and empirical research from cognitive psychology, thereby paving the way for the evaluation of novel training regimens in future endeavors.

The subsequent sections of this paper are structured as follows: Section 2 provides a comprehensive review of existing literature pertaining to ETI training and the utilization of dual-task training in cognitive psychology study. Section 3 outlines the innovative design of a dual-task MR training ETI simulator, automated scoring system, and data collection process. In Section 4, we propose a validation methodology aiming at assessing the difference between experts and novices. Subsequently, Section 5 discusses data analysis results. Finally, Section 6 concludes our work and introduces future research.

2. Related Work

Medical simulators serve as invaluable tools for medical professionals, spanning from students to seasoned practitioners, facilitating the practice and refinement of techniques within a controlled environment. For instance, simulators for otolaryngology, bone surgery, and vascular reconstruction have been shown to significantly improve surgical outcomes [15–17]. Advanced technologies such as virtual reality (VR) and augmented reality (AR) have been integrated into simulator design to further enhance surgical training and performance [18–21]. These innovations contribute substantially to improving proficiency and confidence, laying a strong foundation for clinical practice and ultimately leading to better patient outcomes.

Infants admitted to the NICU are highly vulnerable to respiratory insufficiency and often require intubation and mechanical ventilation. Proficiency in ETI is a critical resuscitation skill for pediatric trainees and is essential for effective infant care [22]. Figure 1 illustrates the intubation procedure. First, insert the laryngoscope with the left hand in the oral cavity; second, press down the tongue and the epiglottis until the vocal cords are in sight; lastly, hold the left hand steadily and insert the intubation tube with the right hand. To ensure competency, trainees must undergo ETI training prior to clinical exposure [23,24]. Among the various training modalities, physical task trainer simulators are the most widely utilized for ETI instruction. However, current training protocols employing these simulators demonstrate limited skill transferability to real clinical settings [7]. Despite successfully completing task trainer-based training, practitioners frequently encounter challenges in achieving optimal performance during actual patient encounters. This discrepancy can be attributed to several factors. Firstly, the lack of comprehensive understanding of the oral cavity in task trainers limits the efficacy of training guidance and feedback, and the mechanical proficiency gained from blindly memorizing procedural movements within task trainer oral cavities may not translate effectively to real-life patient scenarios. Secondly, the haptic tactile feedback provided by task trainers often lacks fidelity compared to that encountered with real neonates [25]. Lastly, the training environment within a classroom setting differs significantly from the dynamic and distraction-laden ambiance of actual clinical settings in NICU [26].

To improve training outcomes, several computer-aided simulators have been developed to augment the task-trainer-based training experiences [6,27,28]. While VR ETI simulators promise more versatile and realistic training experiences, the low-fidelity versions remain unsuitable for comprehensive educational purposes [27–30]. Conversely, high-fidelity VR simulators, though promising, are relatively new, expensive to develop, and have not been widely adopted within pediatric training programs to investigate their skill transferability [5,31].

Recognizing the need for enhanced training methodologies, our group has developed both AR and VR simulators to elevate training experiences and outcomes. However, existing studies and simulators have yet to comprehensively address the environmental distraction factor impacting skill transferability. In this study, we bridge this gap by introducing a dual-task simulator designed to replicate the complexities of real-world

clinical settings. The primary objective of this study is to validate the discriminative capacity of a dual-task simulator in delineating subtle differences between expert practitioners and novices, where conventional single-task simulators have proven inadequate. Moreover, we aim to establish robust metrics capable of quantifying these differentiating factors. These insights will pave the way for developing more effective training paradigms in medical education.

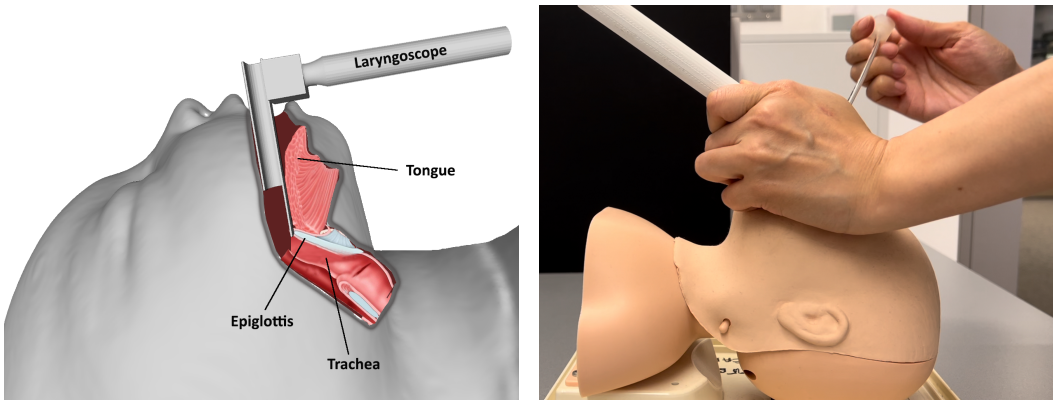


Figure 1. ETI procedure demonstration. (**Left**) Profile anatomical illustration; (**Right**) The completion of an intubation procedure on a task trainer.

Dual-task training is a method used in cognitive psychology and neuroscience to improve cognitive functions by simultaneously engaging in two different tasks. The idea behind dual-task training is based on the concept of cognitive load theory, which suggests that cognitive tasks impose demands on our mental resources. According to this theory, when our mental resources are overloaded, our performance in these tasks may suffer [9,32]. Novices typically experience higher cognitive load when learning a new skill due to the inherent complexity of the task and the need to actively process and integrate new information. As individuals gain expertise and proficiency in a skill, the intrinsic cognitive load decreases as tasks become more automated and require less conscious effort. However, this reduction in intrinsic load may be offset by increases in extraneous load if the task environment becomes more complex or if instructional materials are poorly designed. Nonetheless, with continued practice and skill refinement, individuals can achieve a state of automaticity where tasks are performed with minimal cognitive effort [33].

The dual-task paradigm is the most widely used method to measure automaticity. It typically involves performing a primary task, which is usually the task that the individual wants to improve, along with a secondary task that is designed to place additional demands on cognitive resources. By training under conditions of increased cognitive load, the theory posits that individuals can improve their ability to manage multiple tasks simultaneously, leading to enhancements in cognitive performance and skill automaticity [34]. Therefore, the development of a dual-task simulator not only mirrors real-world clinical scenarios but also serves as a tool for discerning experts and novices, when single-task measurement fails, by measuring the secondary task performance and potentially can be integrated into new medical training regimens.

3. Study Design

We conducted the dual-task neonatal intubation study in collaboration with the Children's National Hospital. The human subject research and data collection were approved by their Institutional Review Board.

3.1. Simulator Design

We designed the ETI dual-task simulator in alignment with best-practice guidelines for healthcare simulation [35], and iteratively refined the design based on user feedback from neonatologists. The ETI dual-task simulator provides automatic skill assessment and dual-task settings. It consists of four distinct modules: a primary task ETI module, a secondary audio-based number reading and response module, a skill analysis and visualization module, and an AR module designed to provide visual guidance during training. This study aims to differentiate expertise levels using a dual-task paradigm; the AR module is not employed during the data collection process as it is specifically utilized in training sessions.

The primary task ETI simulation module includes a Miller 1 blade laryngoscope manufactured by Teleflex in Wayne, PA, USA, an endotracheal tube, and a full-term Laerdal® task trainer from Ellerslie Auckland, New Zealand. The CT-scanned virtual model of the task trainer is visualized by volume rendering. A virtual laryngoscope was modeled by 3D scanning both the Miller 1 blade and handle. Kinematic data for both the task trainer and the laryngoscope are tracked and recorded by an Ascension 3D Guidance trakStar™ 6 degrees of freedom (DOF) tracking system with 3 electromagnetic (EM) sensors. It operates at a default measurement rate of 80 Hz, with a positional accuracy of 1.40 mm and angular accuracy of 0.5°. An EM tracking system comprises a transmitter that emits an electromagnetic field and sensors attached to body segments or instruments to detect the field, in our study, one EM is attached to the laryngoscope handle tip, one to the tube, and one to the task trainer head. These sensors capture field characteristics and relay the data to a central management unit, which processes the information to compute the position and orientation of the sensors, thereby determining the spatial configuration of the tracked body segments in three-dimensional space. The motion data are registered and mapped to their virtual counterparts in the simulator. Figure 2 shows the primary-task ETI module GUI, it can record a new trial or review existing trials. On the left pane, user can select the existing trials to review, the timer is used for time tracking with current frame being shown on the bottom; the main window shows the intubation view, where the user can see the anatomy information of the task trainer (the green and red bars are optional trajectory visual aids of the current and suggested motions); the bottom pane include a view for pitch or force change over time, an ISO threshold for controlling the transparency of the task trainer volume, and a head ratio to set the visible portions of the task trainer.

The AR module is developed using HoloLens™ from Microsoft, which enables users to visualize a superimposed representation of the internal anatomy over a real task trainer. The system employs a distributed framework to facilitate data streaming between the computer and the HoloLens. The visual representation from the user's perspective is illustrated in Figure 3, highlighted with a black circle.

The grading feedback visualization module displays classification results using a polar chart, supplemented by key performance feature values over time. The polar chart is constructed on a five-point scale, with each radial axis representing a distinct performance feature. The enclosed area, determined by the assigned grades for each feature, serves as an indicator of the trainee's overall performance, where a larger enclosed area corresponds to a higher level of proficiency. An overview of this visualization is shown in Figure 4. This visual representation provides trainees with an intuitive feedback mechanism, offering clear insights into their performance and highlighting areas for improvement.

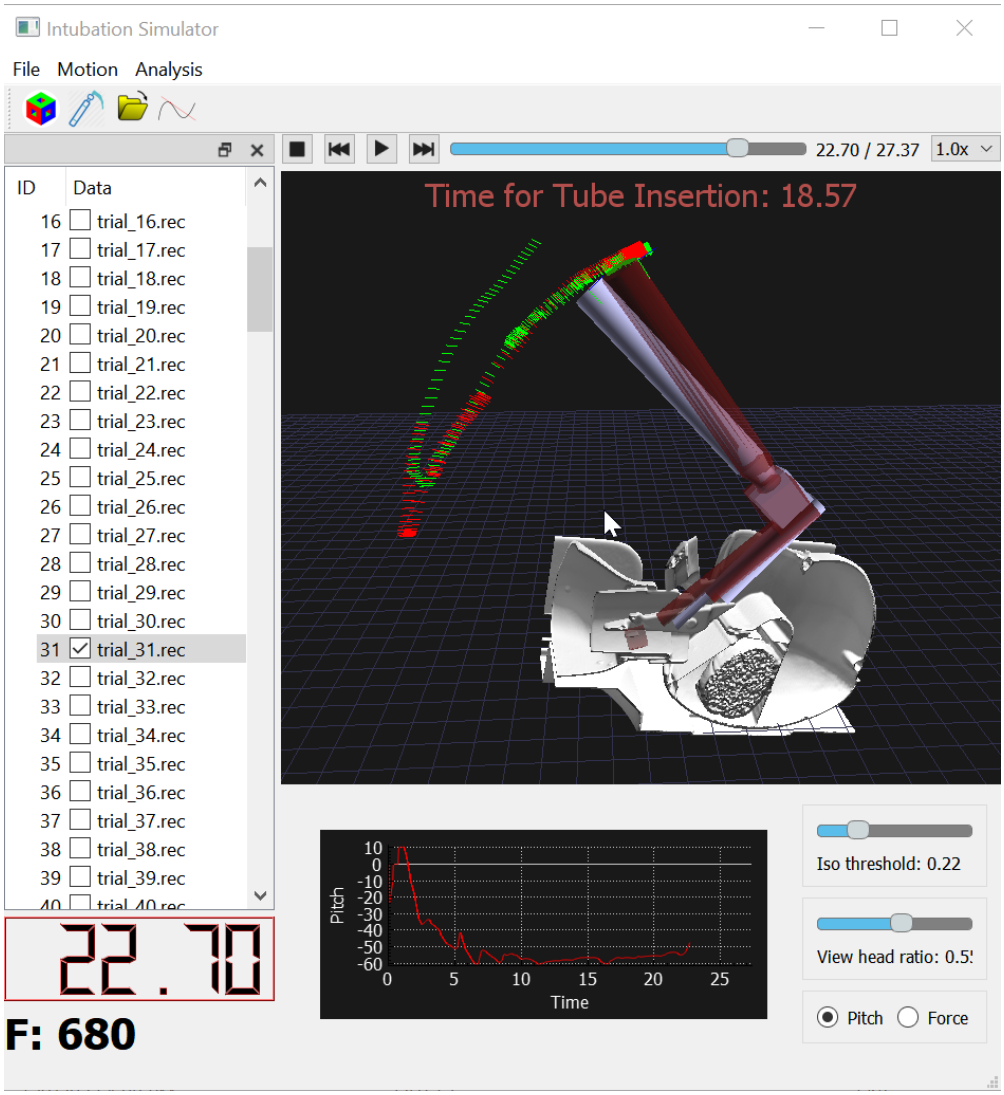


Figure 2. Primary task ETI module: on the left pane, the user can select existing trials to view and check the prcedure timer; the main window shows the task trainer anatomical details with optional visual aids where red trajectory and laryngoscope is visual guidance, and the green one is the user's trajectory. The bottom pane allows user to select motion feature to view and set the visible portion or the head and transparency of the anatomy.

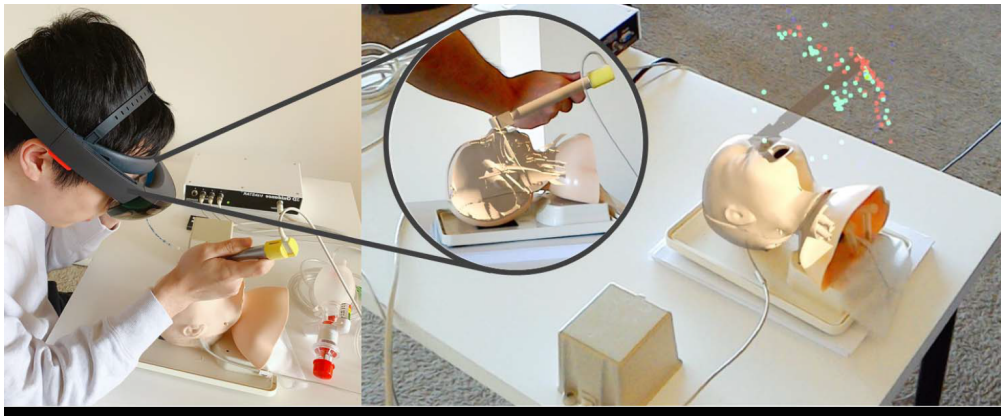


Figure 3. AR module allows trainee to perform intubation with augmented visual guidance. The HMD provides see-through visualization during real-time motion tracking, and the post-trial feedback under playback mode provides color-coded motion analysis with warm colors (in red) indicating the regions that need more attention for improvement.

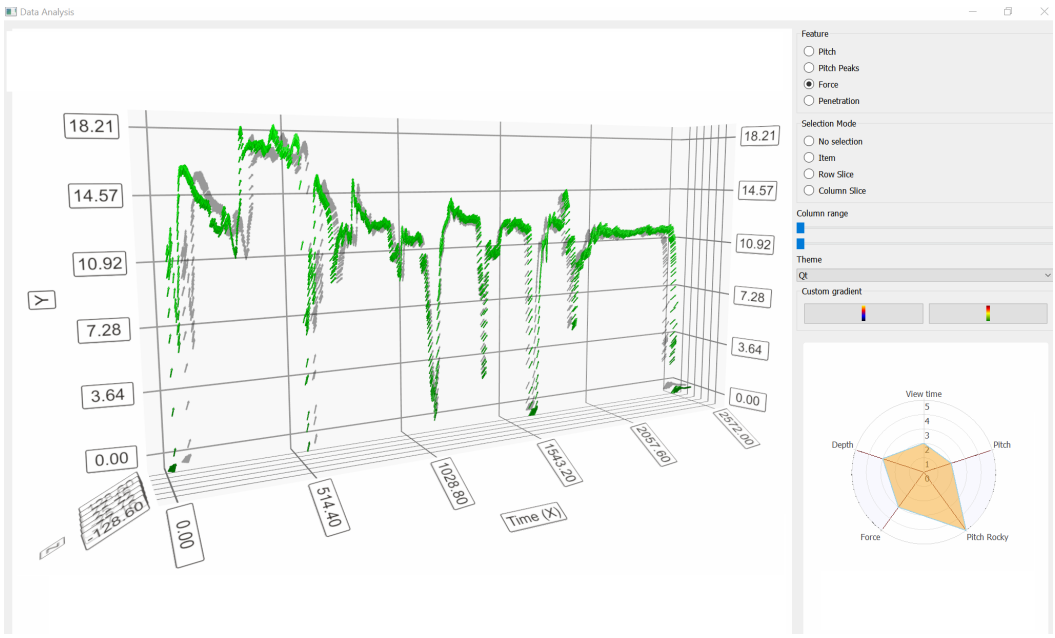


Figure 4. Skill assessment visualization module: user can select the feature to see their change over time; the polar chart shows the details scoring for each feature.

In intubation scenarios, environmental distractions are primarily auditory, as practitioners need process information communicated by other team members. To simulate this, our secondary task module incorporates a speech synthesizer and a footswitch to emulate auditory distractions and capture user responses [36]. Participants are instructed to tap the footswitch during the ETI procedure upon detecting a sequence that meets predefined criteria, referred to as a "hit event". The secondary task incorporates two discernible levels of difficulty:

- **1-back:** Upon detecting a number's immediate repetition, participants are instructed to activate the footswitch, enabling the MR simulator to capture the corresponding timestamp for a hit event recording. For example, a user will tap the footswitch upon hearing the second occurrence of the number 3 within the sequence {...6, 2, 3, 3, 5,...}. The number and order of the sequence are randomized for each trial.
- **2-back:** Upon detecting a number's recurrence after a different one, participants are instructed to activate the footswitch for a hit event recording. For example, a user will hit the footswitch when hearing the second occurrence of the number 8 within the sequence {...2, 7, 8, 5, 8,...}. The number and order of the sequence are randomized for each trial.

We employ the term "0-back" to denote single-task setting, where subjects only perform primary ETI tasks.

Figure 5 shows the ETI dual-task simulator during the data capturing process.

3.2. Dataset Collection

ETI procedures were conducted under both single-task and dual-task conditions, involving 10 experts and 19 novices (average 29.8 ± 6.9 years old). The expert group is drawn from neonatology and pediatric residents and fellows with on average 10.7 years of experience. The novice group is drawn from medical students and pediatric residents with on average 3.3 years of experience.

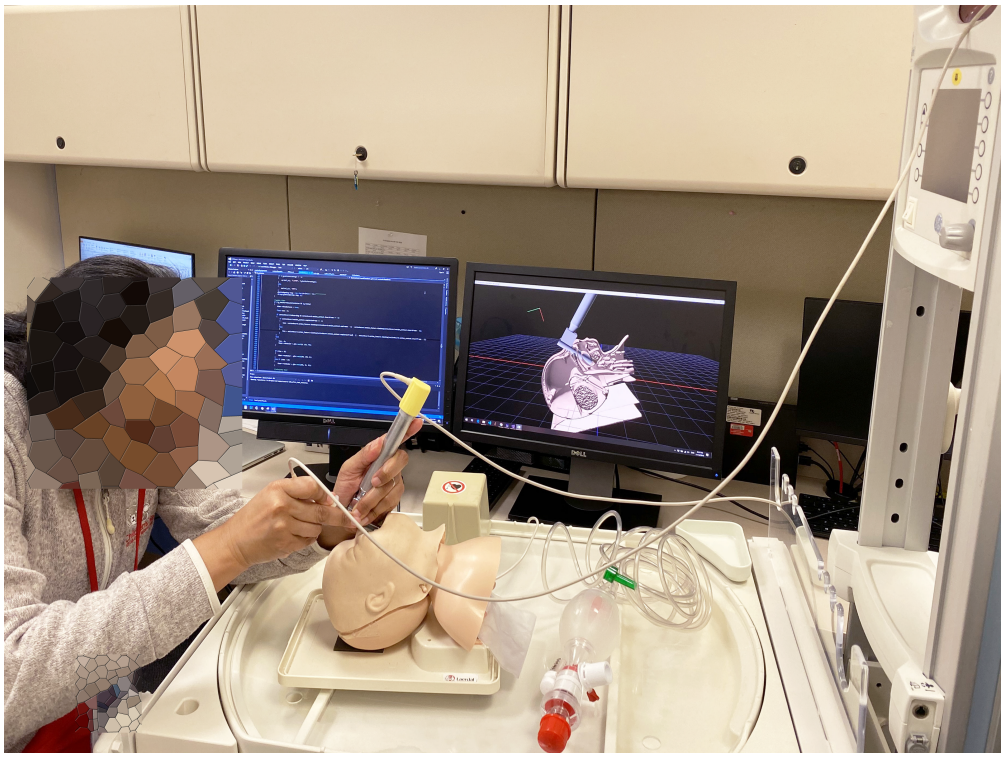


Figure 5. An overview of the dual-task data capture process.

Each participant completed a total of 15 trials. The initial 10 trials alternated between 0-back and 1-back, with a starting trial of 0-back or 1-back (counterbalanced across participants). The remaining 5 trials were 2-backs. To reduce variance associated with participants differently prioritizing performance on the primary versus the secondary task, participants were instructed to emphasize performance on ETI while allocating attention to the secondary task as their cognitive resources permitted. These instructions, coupled with engagement in the secondary task as a nominal motivator, have been shown to be effective in past studies measuring medical simulator proficiency [12].

After purging the data of anomalies in cases when the EM lost track, we curated a dataset of 295 trials. The 3D motions of these trials were then replayed and evaluated by three expert neonatologists. They conducted an unbiased assessment of overall scores through blind review using a 5-point Likert scale, where 1 represents poor, and 5 represents excellent. We used these scoring results to train a deep learning model for automated ETI skill assessment [6].

4. Methods

We conducted a two-way `group` (expert vs. novice) $\times$ `task` (0-back vs. 1-back vs. 2-back) repeated measures analysis of variance (ANOVA) [37]. The dependent measures include (i) primary ETI task performance, and (ii) the task engagement ratio and performance of the secondary task. Presentation of all data is in the form of means with statistical significance established at $p < 0.05$. The F-statistic (for group and task performance variances) and t-statistic (for group and task performance means) were used to investigate if there were statistically significant differences between expertise levels and task complexities.

4.1. Primary ETI Task Metrics

The ETI performance was evaluated with an overall score graded by experts on a 5-point Likert scale, and the scores of the individual features were recommended by expert neonatologists. These individual features include trial time, duration of glottis visibility,

force applied against the gum, penetration depth into the tongue plane, pitch peaks, and yaw peaks as defined in Table 1. Fast Fourier transformation was used to filter out high frequency data in pitch and yaw peaks. They were automatically graded on a 5-point Likert scale with a multi-task convolutional neural network classification algorithm [6].

Table 1. Primary ETI task performance metrics, which are evaluated on a 5-point Likert scale.

Metrics	Descriptions
Overall score	The performance is evaluated by experts based on the entire procedure.
Trial time	Time used to complete an ETI procedure. The shorter the better.
Duration of glottis visibility	The duration of glottis being in the line of sight of the laryngoscope. The longer the better.
Force	The force exerted by the laryngoscope on the upper gum. The smaller the better.
Penetration depth	The penetration depth of the laryngoscope into the tongue plane. It has to be the right amount to view glottis, but not too deep to damage the tongue.
Pitch peaks	The pitch angle peaks of the laryngoscope to measure the smoothness of motion.
Yaw peaks	The yaw angle peaks of the laryngoscope to measure the smoothness of motion.

4.2. Secondary Task Metrics

The performance measurement of the secondary task was conveyed by participant engagement ratio as in Equation (1) and footswitch hit accuracy in (2):

$$Second_task_engage = \frac{Total\ user\ hits}{System\ given\ hits} \tag{1}$$

$$accuracy = \frac{Correct\ user\ hits}{System\ given\ hits} \tag{2}$$

The secondary task engagement ratio measure functions as an indicator of cognitive resource allocation, independent of response accuracy, whereas accuracy serves as a metric for assessing the proficiency of performance in the execution of the secondary task.

4.3. Statistic Metrics

The *t*-test was used to determine whether there exists a statistically significant difference in the means of two groups.

$$t(df) = \frac{\bar{x}_1 - \bar{x}_2}{\sqrt{s^2(\frac{1}{n_1} + \frac{1}{n_2})}} \tag{3}$$

where df is the degree of freedom of the test, which is the number of samples for equal-sized groups, and weighted value of the sample sizes from the two unequal-sized groups. s is the standard error, n_1 and n_2 are the sample size of the two groups. If the corresponding p-value for this t-value is less than 0.05, the null hypothesis that the means of the two groups are the same is rejected.

The ANOVA result was reported as an F-statistic, accompanied by its associated degrees of freedom and p-value. The F-value measures the ratio of between-group variance and within-group variance (4).

$$F_{df_1, df_2} = \frac{Between\ groups\ variance}{Within\ group\ variance} \tag{4}$$

where df_1 and df_2 are the degree of freedom of the numerator and the degree of freedom of the denominator. If the corresponding p-value for this F-value is less than 0.05, then the null hypothesis that there is no significant effect of the factors on the dependent variables is rejected.

We conducted data analysis in the following aspects to see the performance difference between experts and novice, the impact of the dual-task on both groups, and the effect of repetition on performance. The null hypotheses for these tests are as follows:

- There is no performance difference between experts and novices for each task setting.
- There is no performance difference under different task settings for the same group (expert or novice).
- There is no difference between experts and novices on how the dual-task impacts their performance.
- There is no difference in ETI performance under intensive training.

5. Results and Discussion

The results of the data analysis are reported categorically in accordance with the respective hypotheses. Table cells that show statistical significance are colored in purple.

5.1. Expert vs. Novice

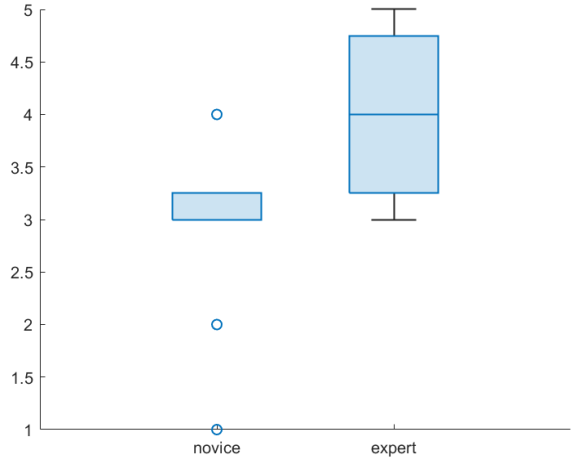
5.1.1. Primary Task Performance Comparison

The overall performance of the expert and novice groups under 0-back, 1-back, and 2-back conditions are plotted in Figure 6, and the corresponding t-test and F-test results are shown in Table 2. The results show that the overall performance is significantly different between experts and novices for 0-back (single task) and 1-back, whereas they are statistically the same for 2-back settings. All conditions considered, experts scored better than novices.

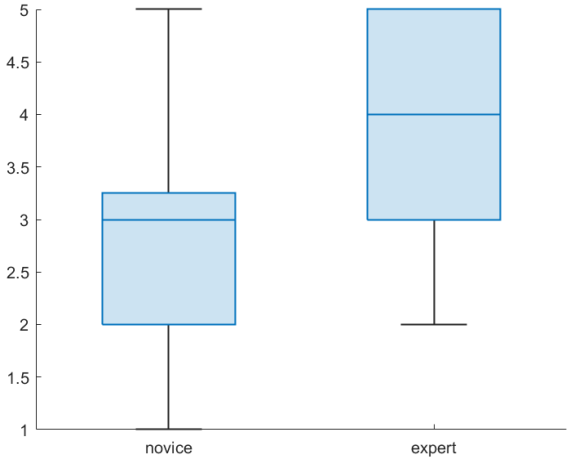
It is worth noting that novice group performance decreased first at 1-back, then increased at 2-back, whereas the expert group was less affected by task conditions. We will further analyze the within-group performance in later sections. For the current between-group analysis result, it is necessary to further examine why 1-back and 2-back affect experts and novices differently: Is it because there are differences in attention allocation? Or did the novice recalibrate and find the optimal path over trials? This can be determined by comparing the secondary task engagement rate and accuracy. The repetition effects will be discussed in later sections.

Table 2. Mean score for expert and novice group, t-value, F-value, and their corresponding p-value.

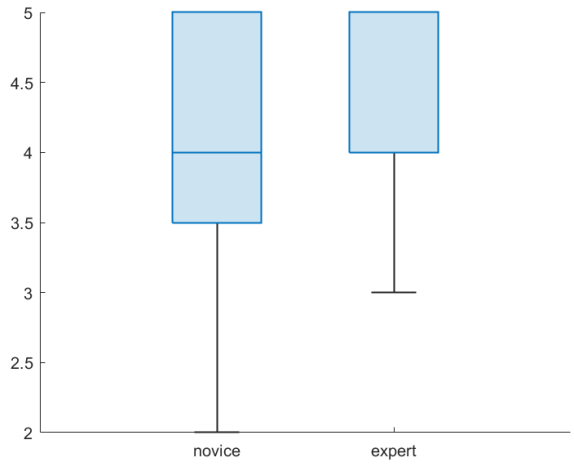
Test	0-Back	1-Back	2-Back	Overall
Expert score	4.00	3.74	4.09	3.94
Novice score	3.03	2.85	4.05	3.29
t-value	$t(98) = -6.35$	$t(98) = -4.02$	$t(93) = -0.2134$	$t(293) = -5.55$
F-value	$F_{1,99}$ = 40.32	$F_{1,99} = 16.13$	$F_{1,94} = 0.05$	$F_{1,294} = 30.79$
p-value	6.74×10^{-9}	1.16×10^{-4}	0.83	6.44×10^{-8}



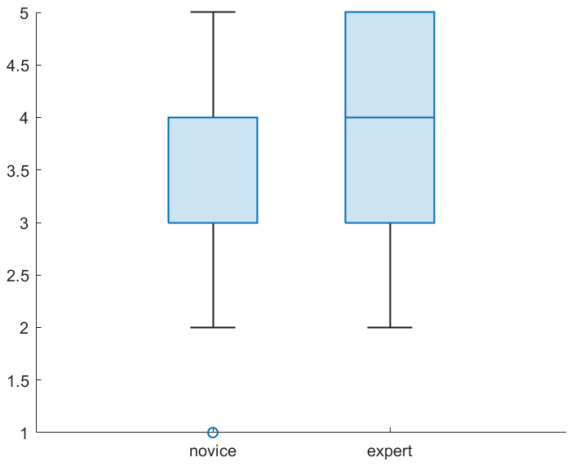
(**a**) Expert vs. novice score under 0-back



(**b**) Expert vs. novice score under 1-back



(**c**) Expert vs. novice score under 2-back



(**d**) Expert vs. novice score overall difference.

Figure 6. Expert vs. novice overall performance scores on 5-point Likert scale boxplot under different n-back task conditions.

5.1.2. Secondary Task Performance Comparison

The comparison of secondary task engagement ratios and accuracy statistics between expert and novice groups is depicted in Figure 7 and summarized in Table 3. The overall secondary task engagement ratio is low for both expert and novice groups in both dual-task conditions. Specifically, in the 1-back setting, both experts and novices allocate a similar level of attention, with novices exhibiting significantly lower accuracy compared to experts; the collective primary task mean score for both groups decreases. We will discuss the statistical significance of this effect in subsequent within-group analyses. In the 2-back setting, both groups experience a further reduction in secondary task engagement ratios and accuracy, with novice performance significantly lagging behind that of experts.

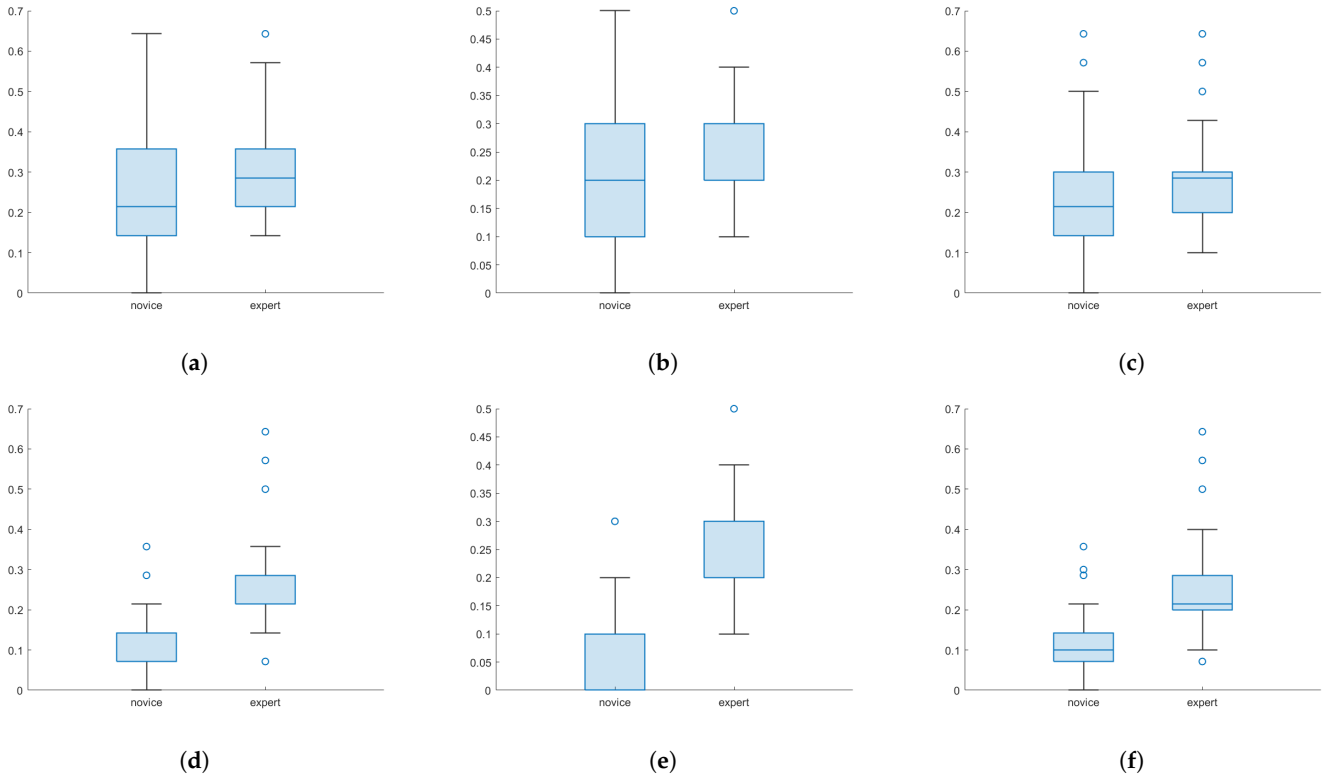


Figure 7. Boxplot comparison of secondary task performance of expert and novice group. (**a**) Expert vs. novice secondary task engagement ratio under 1-back. (**b**) Expert vs. novice secondary task engagement ratio under 2-back. (**c**) Expert vs. novice secondary task engagement ratio under both 1-back and 2-back. (**d**) Expert vs. novice secondary task accuracy under 1-back. (**e**) Expert vs. novice secondary task accuracy under 2-back. (**f**) Expert vs. novice secondary task accuracy under both 1-back and 2-back.

Table 3. Mean secondary task engagement ratio and accuracy for expert and novice group, and between-group difference by *t*-value, *F*-value, and their corresponding *p*-value.

Test	1-Back	2-Back	Overall
Expert engagement	0.30	0.26	0.28
Novice engagement	0.27	0.20	0.24
Engagement *t*-value	$t(98) = -1.24$	$t(93) = -2.78$	$t(193) = 2.50$
Engagement *F*-value	$F_{1,99} = 1.55$	$F_{1,94} = 7.74$	$F_{1,194} = 6.23$
Engagement *p*-value	0.23	0.01	0.01
Expert accuracy	0.27	0.23	0.25
Novice accuracy	0.11	0.07	0.10
Accuracy *t*-value	$t(98) = -8.06$	$t(93) = -9.63$	$t(193) = -12.01$
Accuracy *F*-value	$F_{1,99} = 64.94$	$F_{1,94} = 92.73$	$F_{1,194} = 144.18$
Accuracy *p*-value	1.90×10^{-12}	1.24×10^{-15}	3.61×10^{-25}

5.2. Dual-Task Interference Effects on Expert

One-way repeated measures ANOVA (dual-task setting as factor) is employed to assess the influence of the secondary task on the primary ETI task. A comprehensive comparison of both overall performance and individual features is illustrated in Figure 8 and Table 4. The colored table cells indicate features that yielded significantly different performance metrics between task settings. The *F*-value statistics reveal that the simultaneous execution of two tasks has an insignificant impact on both the overall performance and various features. Notably, the yaw peaks feature demonstrates a statistically signif-

icant difference between the 0-back and 1-back conditions ($t(34) = 1.44$, $p = 0.048$, 1 vs. 0-back t-value is a mirror of the 0- vs. 1-back t in Table 4), indicating an increase in mean yaw peaks during the 1-back dual-task condition. A noteworthy observation arises when examining the performance of experts: a significant increase is found in their overall score ($t(34) = 2.24$, $p = 0.03$) after switching from 1-back to 2-back. Furthermore, trial times exhibit a significant reduction ($t(34) = -2.04$, $p = 0.048$, 1- vs. 2-back), suggesting that, initially, experts may display slight susceptibility to dual-task interference in the experiment. However, as they accumulate more trial experience and progressively allocate less attention to the secondary task, their skill proficiency improves.

Table 4. The impact of secondary task on expert ETI performance measured by t-value, F-value, and their corresponding p-value.

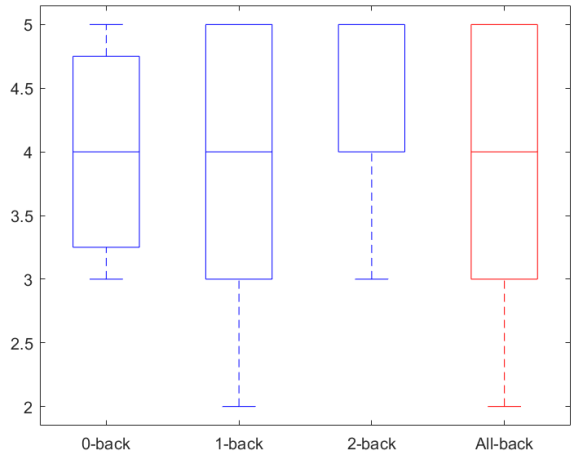
Feature	0 vs. 1-Back $t(34)$ (p)	0 vs. 2-Back $t(34)$ (p)	1 vs. 2-Back $t(34)$ (p)	$F_{2,68}$ (p)
Overall Score	1.27(0.21)	−0.25(0.61)	−2.24(0.03)	2.08(0.13)
Trial time	−0.90(0.38)	1.26(0.21)	2.04(0.048)	2.38(0.10)
Glottis view	0.47(0.64)	0.50(0.62)	0.05(0.96)	0.16(0.85)
Force	−0.63(0.53)	0.25(0.80)	0.76(0.45)	0.37(0.69)
Depth	0.67(0.51)	0.45(0.66)	−0.13(0.90)	0.22(0.80)
Pitch Peaks	−0.88(0.39)	0.31(0.76)	1.03(0.31)	0.81(0.45)
Yaw peaks	−1.44(0.048)	0.69(0.51)	1.68(0.10)	2.21(0.12)

5.3. Dual-Task Interference Effects on Novice

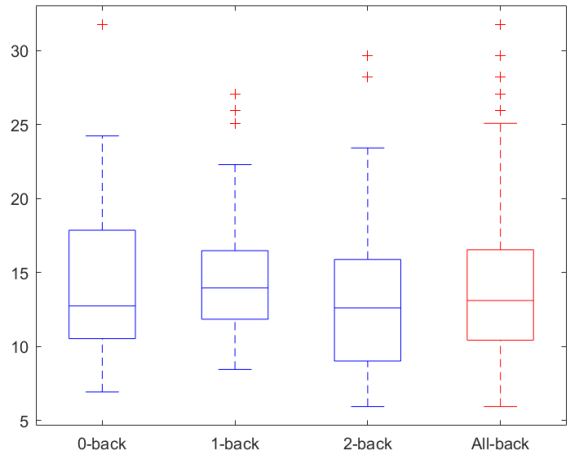
A comprehensive comparison of both overall performance and individual features of novices under different task conditions is illustrated in Figure 9 and Table 5. The tabulated data reveal that, during the initial phase of the experiment, novices exhibited susceptibility to the distractions posed by the alternating execution of 0-back and 1-back task settings. This was reflected in extended execution times ($t(59) = 1.28$, $p = 0.04$) for the ETI task and a reduction in the duration of glottis being visible ($t(59) = -1.66$, $p = 0.03$), alongside an increase in yaw peaks ($t(59) = 1.64$, $p = 0.045$). However, as the subjects transitioned to the 2-back dual-task setting and accumulated practice, novices began to allocate reduced attention to the secondary task. This resulted in improvements in all key performance features, with the exception of yaw peaks, which remain the same level as the beginning of the training.

Table 5. The impact of secondary task on novice ETI performance measured by t-value, F-value, and their corresponding p-value.

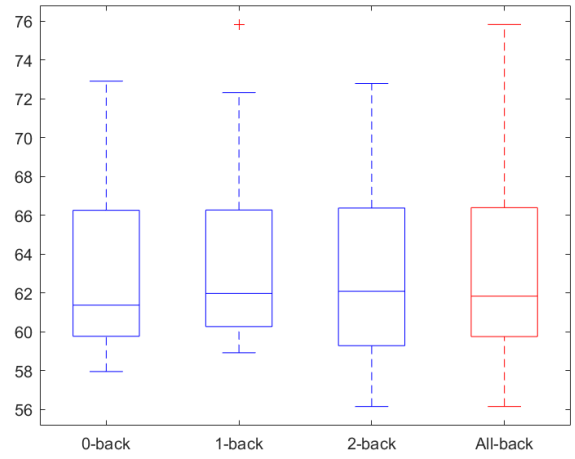
Feature	0 vs. 1-Back $t(59)$ (p)	0 vs. 2-Back $t(59)$ (p)	1 vs. 2-Back $t(59)$ (p)	$F_{2,118}$ (p)
Overall Score	0.98(0.33)	$-8.70(3.62 \times 10^{-12})$	$-7.95(6.48 \times 10^{-11})$	$45.30(2.52 \times 10^{-15})$
Trial time	−1.28(0.04)	2.31(0.02)	$-3.79(3.52 \times 10^{-4})$	6.95(0.001)
Glottis view	1.66(0.03)	2.77(0.008)	−3.45(0.001)	7.01(0.001)
Force	−0.40(0.69)	$3.55(7.69 \times 10^{-4})$	$3.81(3.30 \times 10^{-4})$	$10.28(7.65 \times 10^{-5})$
Depth	0.45(0.66)	2.76(0.01)	2.57(0.01)	5.60(0.005)
Pitch Peaks	−0.82(0.42)	2.41(0.02)	3.28(0.02)	5.59(0.005)
Yaw peaks	−1.64(0.045)	1.67(0.10)	3.30(0.001)	5.49(0.005)



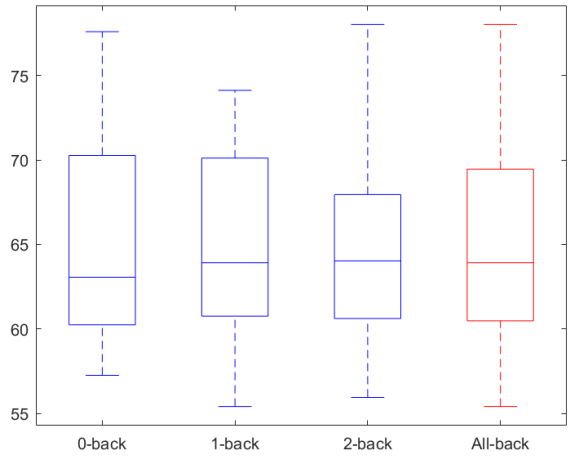
(**a**) Score comparison



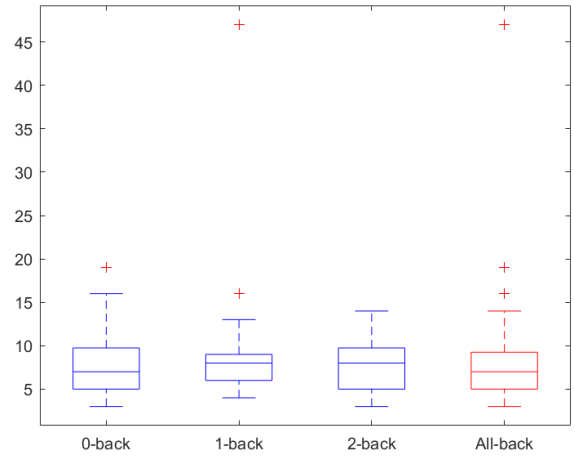
(**b**) Trial time comparison



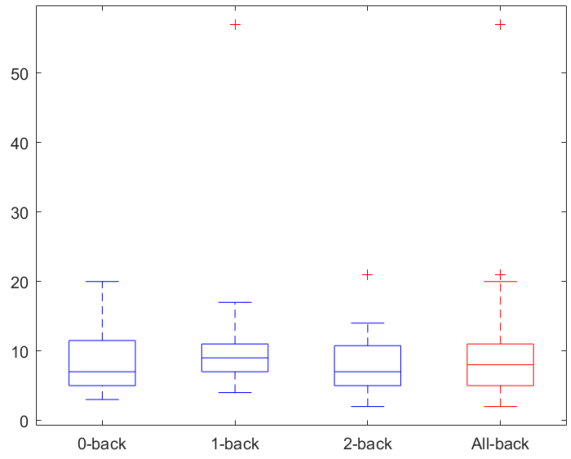
(**c**) Force comparison



(**d**) Depth comparison

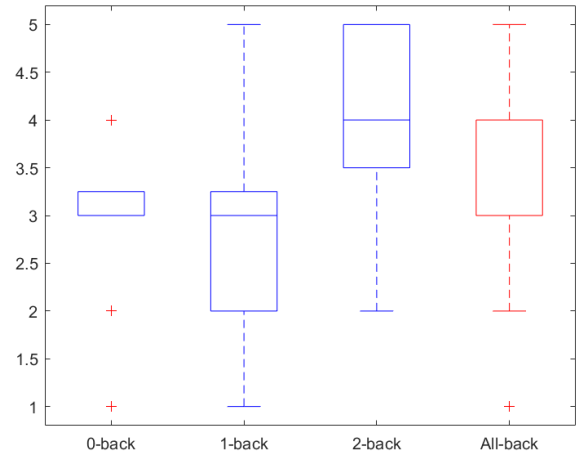


(**e**) Pitch peak comparison

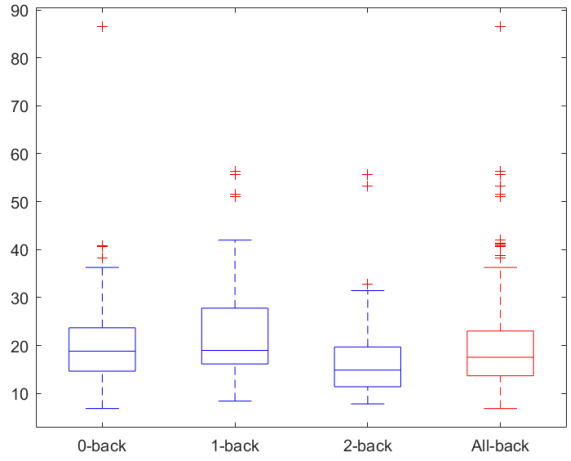


(**f**) Yaw peak comparison

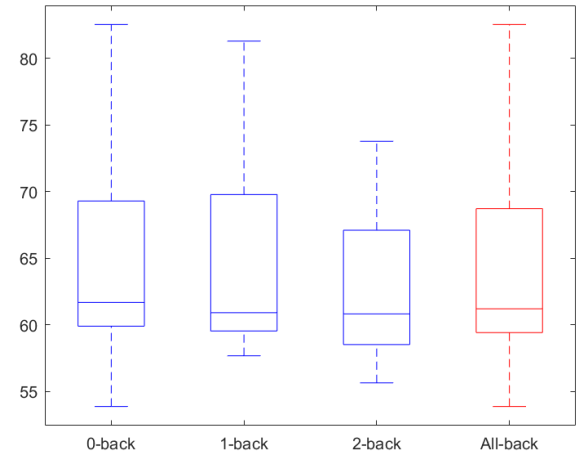
Figure 8. Expert ETI performance boxplot comparison in 0-back, 1-back, and 2-back conditions.



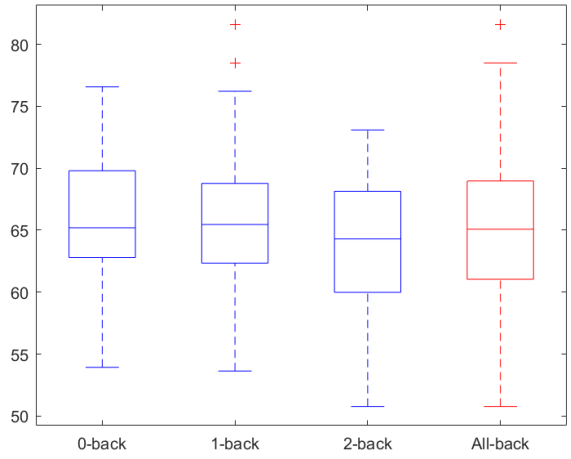
(**a**) Score comparison



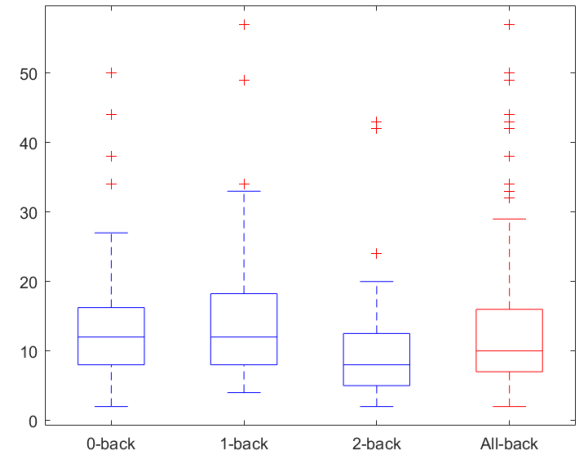
(**b**) Trial time comparison



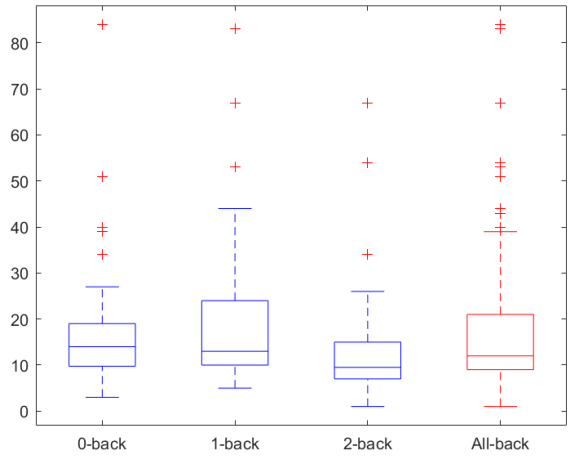
(**c**) Force comparison



(**d**) Depth comparison



(**e**) Pitch peak comparison



(**f**) Yaw peak comparison

Figure 9. Novice ETI performance comparison in 0-back, 1-back, and 2-back conditions.

5.4. ETI Performance Variation over Trials

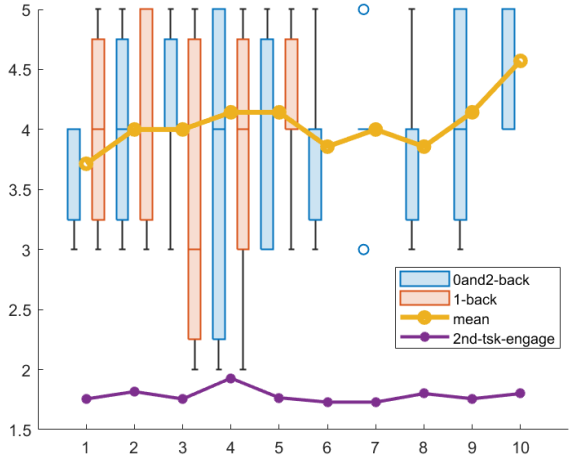
The preceding sections have substantiated the effectiveness of the dual-task simulator in capturing the differences between experts and novices, even when they exhibit equivalent levels of proficiency in the primary task. Meanwhile, over the course of 15 trials, participants demonstrated notable enhancements in their primary task performance, particularly evident among novice participants. In order to quantitatively investigate this, we conducted a one-way repeated measures ANOVA (time as factor). To mitigate the potential interference stemming from the relative heightened secondary task engagement observed in the 1-back dual-task condition, we deliberately omitted the 1-back and selected four performance measures from the trial sequence: 0-back trial 1, 0-back trial 5, 2-back trial 11, and 2-back trial 15.

5.4.1. Expert Performance Change over Trials

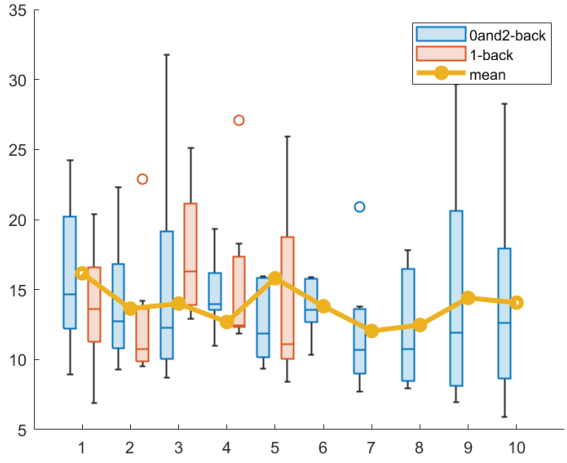
In Section 5.2, our data analysis reveals that experts do not exhibit significant degradation under dual-task conditions. Within this section, we investigate whether these experts' performance benefited in any discernible way over the course of the testing session. While we did not observe statistically significant changes in individual performance metrics, as illustrated in Table 6, there is a cumulative effect wherein these metrics collectively demonstrate improvement over time, resulting in a significant increase in the overall score ($t(6) = 3.29, p = 0.02$) when compared to the initial stage of the training program. Figure 10 illustrates these individual performance feature changes over time, wherein the blue box chart represents the 0-back (single-task) condition from time points 1 to 5, the orange box chart corresponds to the 1-back condition from time points 1 to 5, and the blue box chart from time points 6 to 10 signifies the 2-back condition. The grouping of 0-back and 1-back conditions is necessitated by the distinct alternative dual-task patterns for each subject, which do not align with the time axis. The mean value curve is computed from the means of the 0-back and 2-back conditions. At the lower part of the overall score chart in Figure 10a, we include a plot of the secondary task engagement ratio to provide an estimate of the combined cognitive load. The secondary task engagement ratio has been augmented by 1.5 for clarity in visualization. We observe that the secondary task engagement ratio remains the same level over time.

Table 6. Expert performance change over trials. The significance is measured by *t*-value, *F*-value, and their corresponding *p*-value.

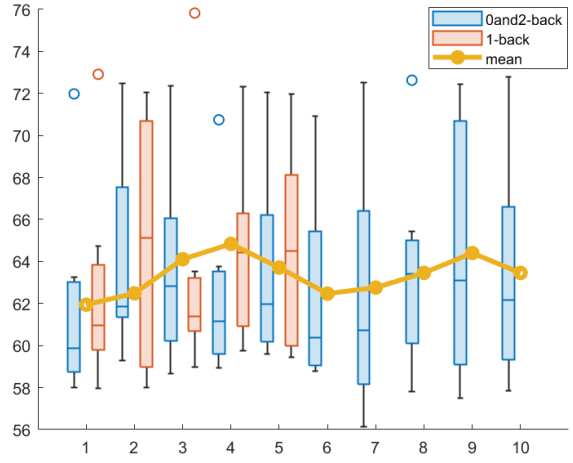
Feature	Trial 5 vs. 1 $t(6)$ (p)	Trial 11 vs. 1 $t(6)$ (p)	Trial 15 vs. 1 $t(6)$ (p)	$F_{3,18}$ (p)
Overall Score	1.16(0.29)	0.55(0.60)	3.29(0.02)	2.57(0.09)
Trial time	−0.10(0.93)	−1.12(0.31)	−0.61(0.56)	0.37(0.77)
Glottis view	−0.75(0.48)	−0.39(0.71)	−1.15(0.29)	0.80(0.51)
Force	1.75(0.13)	0.69(0.51)	1.82(0.12)	1.39(0.28)
Depth	0.35(0.74)	−0.06(0.95)	0.26(0.80)	0.11(0.95)
Pitch Peaks	0.17(0.87)	0.09(0.93)	−0.60(0.57)	0.33(0.80)
Yaw peaks	−0.32(0.76)	−0.81(0.45)	−1.15(0.29)	0.75(0.54)



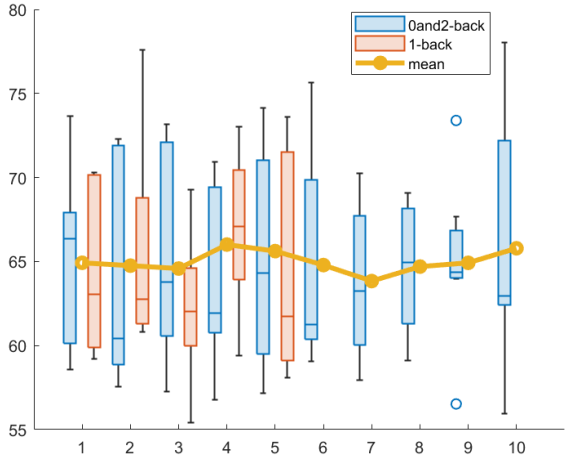
(**a**) Overall score



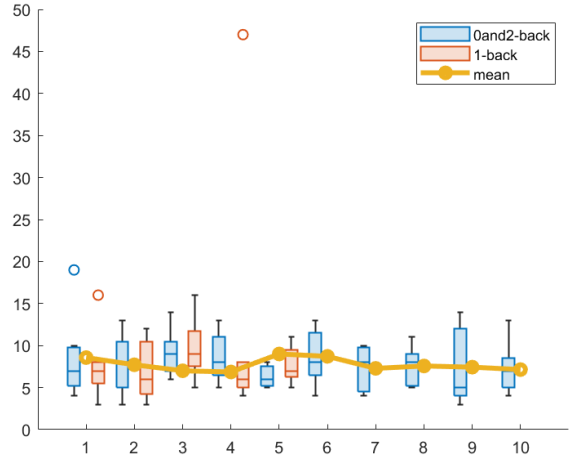
(**b**) Trial time



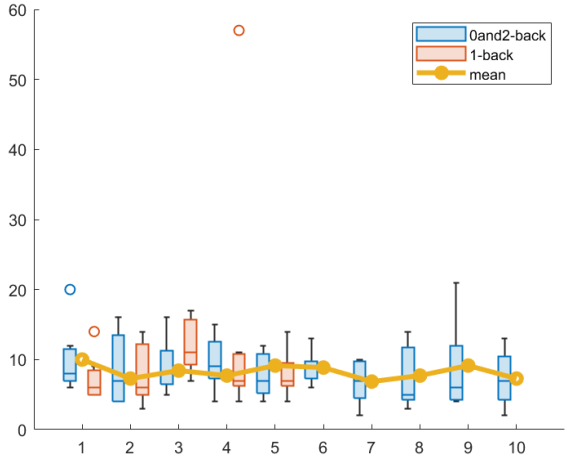
(**c**) Force



(**d**) Depth



(**e**) Pitch peaks



(**f**) Yaw peaks

Figure 10. Expert performance change over trials.

5.4.2. Novice Performance Change over Trials

In Section 5.3, the data analysis result shows that novice performance is substantially influenced by dual-task conditions, with a significant decline at the beginning, which gradually improves in subsequent 2-back trials with significantly reduced secondary task engagement ratio. In this section, we study the impact of repetition on novice performance. As detailed in Table 7, our observations indicate that at the initial stage of their testing session, following five single-task practice trials, a cumulative enhancement across various individual performance features becomes evident, resulting in a significant increase in the overall score. As the number of practice trials progresses, by the 11th trial, a substantial reduction in the time required to complete a given procedure becomes evident, contributing to a significant overall score increase. By the conclusion of the testing session, all performance features exhibit noticeable improvement, except for the depth feature.

The mean value plot for the novice group in Figure 11 provides a clear illustration of their skill development over time. In line with Section 5.4.1, we maintain consistent color-coding and grouping strategies for different task conditions, and the secondary task engagement ratio is adjusted by an increment of 1 to enhance visualization clarity. We observe that the secondary task engagement ratio is lower in the 2-back setting compared to 1-back.

This in-depth data analysis reveals that while novice participants achieved high ETI procedure outcomes, their secondary task performance indicates they have not yet reached automaticity. According to cognitive load theory, this suggests a high cognitive load that limits available working memory resources for concurrent tasks. The findings highlight the limitations of traditional single-task assessment metrics, which may overlook the cognitive demands trainees face in real clinical settings. Incorporating multitask evaluation provides a more comprehensive measure of performance and readiness by capturing the cognitive load experienced during complex procedures.

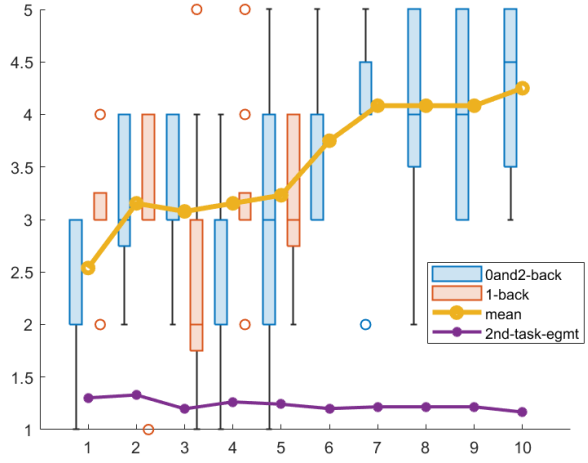
Table 7. Novice performance change over trials. The significance is measured by *t*-value, *F*-value, and their corresponding *p*-value.

Feature	**Trial 5 vs. 1** $t(12)$ (p)	**Trial 11 vs. 1** $t(12)$ (p)	**Trial 15 vs. 1** $t(12)$ (p)	$F_{3,33}$ (p)
Overall Score	2.92(0.01)	7.00(2.37×10^{-5})	8.04(6.21×10^{-6})	18.11(4.00×10^{-7})
Trial time	−1.36(0.20)	−2.52(0.03)	−3.23(0.01)	3.91(0.02)
Glottis view	−2.99(0.10)	−2.87(0.02)	−2.56(0.03)	6.40(0.002)
Force	−0.16(0.87)	−0.66(0.52)	−3.54(0.005)	1.83(0.16)
Depth	−0.19(0.85)	−1.21(0.25)	−0.40(0.70)	0.87(0.47)
Pitch Peaks	−2.07(0.06)	−1.93(0.08)	−3.13(0.01)	4.12(0.01)
Yaw peaks	−1.30(0.22)	−1.22(0.25)	−3.19(0.01)	2.37(0.09)

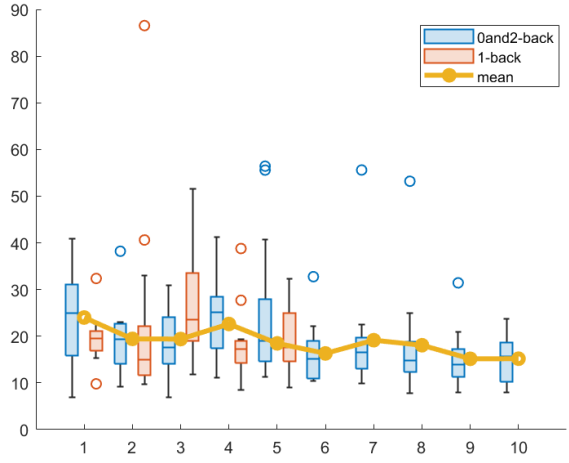
5.5. Discussion

The data analysis results are in line with our initial hypotheses:

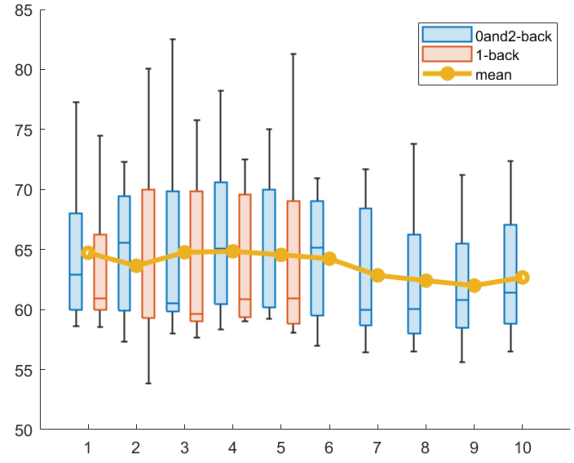
- Single-task performance measurements are insufficient to evaluate practitioner's readiness for clinical practice.
- Individuals with different levels of expertise exhibit distinct responses when subjected to dual-task conditions.
- Experts exhibit greater resilience to the adverse effects of elevated cognitive load, as compared to novices.



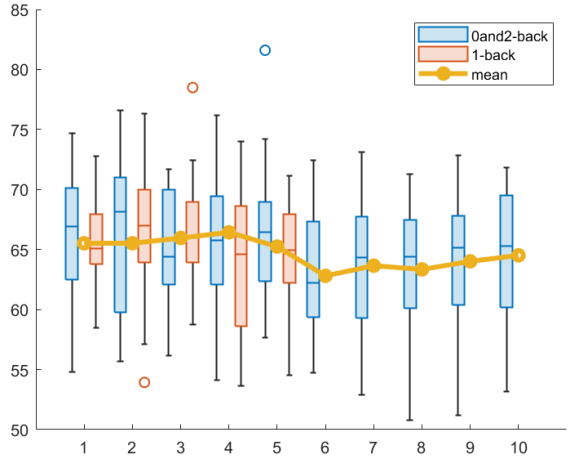
(**a**) Overall score



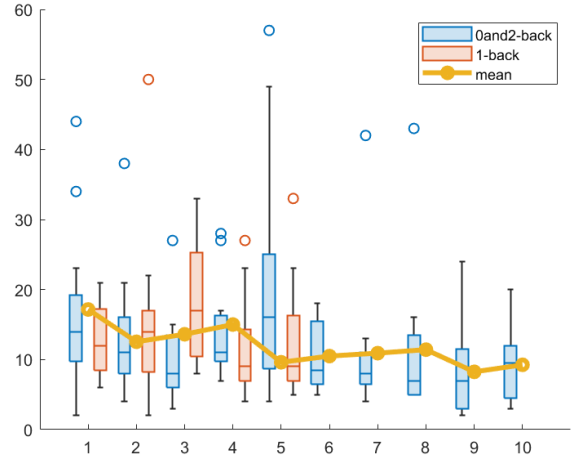
(**b**) Trial time



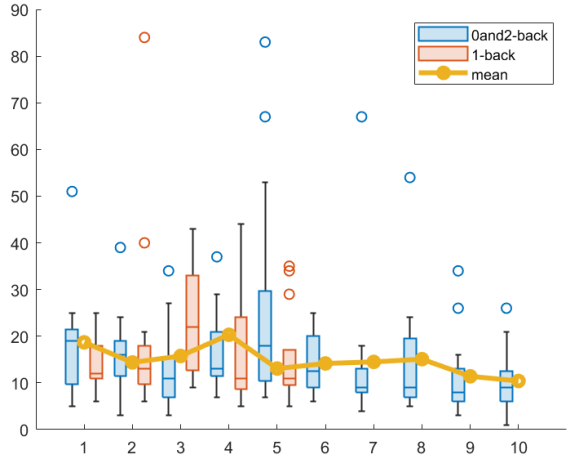
(**c**) Force



(**d**) Depth



(**e**) Pitch peaks



(**f**) Yaw peaks

Figure 11. Novice performance change over trials.

With multiple factors affecting the ETI performance, it is of practical interest to study which factor plays a pivotal role in the overall performance, or if it is the combined effect of all factors. To address this, we employed a stepwise linear regression approach to iteratively

derive an optimal regression model [38]. The resulting model is presented in Figure 12. Our analysis clearly demonstrates that the subject group (`isExpt`), the dual-task mode (`NMode`), the trial order (`orderID`) , and secondary task engagement ratio (`SecondTaskEngage`) are independently significant factors influencing the final performance scores. Furthermore, the interactions of expertise level with dual-task mode and expertise level with secondary task engagement also contribute to the overall performance.

```
Linear regression model:
    Score ~ 1 + orderID + isExpt*NMode + isExpt*SecondTaskEngage

Estimated Coefficients:
```

	Estimate	SE	tStat	pValue
(Intercept)	2.5349	0.12678	19.994	2.2311e-56
isExpt	1.0135	0.16637	6.0921	3.5612e-09
NMode	0.39636	0.12365	3.2056	0.0014994
orderID	0.084712	0.017484	4.8451	2.0706e-06
SecondTaskEngage	-2.5315	0.65742	-3.8507	0.0001452
isExpt:NMode	-0.6818	0.16018	-4.2564	2.8138e-05
isExpt:SecondTaskEngage	2.5441	0.92764	2.7425	0.0064784

Figure 12. ETI performance in response to group, dual-task mode, secondary task engagement ratio, and trial order.

6. Conclusions

In this paper, we present, to the best of our knowledge, the first practical and efficient dual-task ETI simulator that offers an alternative medical skill training paradigm. The incorporation of a high extrinsic load secondary task during training not only mirrors the complexities of the real-world clinical environment but also facilitates the differentiation between learners and expedites the learning process. Statistical analysis results demonstrate the efficacy of dual-task performance metrics in assessing the automaticity of skills. As expertise increases, ETI can be executed more "automatically", resulting in better outcomes even when attention is divided among multiple tasks.

Achieving true automaticity on a task as complex as ETI is understood to necessitate intensive practice over many training intervals. The subsequent phase of our research involves the in-depth examination of skill retention. Our ongoing studies focus on the comprehensive investigation of the long-term retention of acquired skills, facilitated by the systematic assessment of trainee performance at regular intervals throughout the participants' residency. This approach allows us to track the durability and maintenance of these skills over time. We are currently engaged in the data preprocessing stage, and the results of our research will be reported in a future publication.

Author Contributions: Conceptualization, Y.M., S.Z., X.Z., J.P. and J.H.; methodology, Y.M., X.Z. and J.P.; software, S.Z. and Y.M.; validation, Y.M., S.Z. and L.S.; formal analysis, Y.M.; investigation, Y.M., S.Z. and P.M.; writing—original draft preparation, Y.M.; writing—review and editing, J.P., X.Z., L.S. and J.H.; visualization, Y.M. and B.F. All authors have read and agreed to the published version of the manuscript.

Funding: This research was supported by the Eunice Kennedy Shriver National Institute of Child Health & Human Development of the National Institutes of Health under Award Number R01HD091179. The content is solely the responsibility of the authors and does not necessarily represent the official views of the National Institutes of Health.

Institutional Review Board Statement: The study was approved by the Institutional Review Board of Children's National Hospital under the protocol FWA00005945 on 1 November 2019.

Informed Consent Statement: Informed consent was obtained from all subjects involved in the study.

Data Availability Statement: Experimental data are not published along with this study because we have more research topics to be published based on the current dataset.

Acknowledgments: We thank Wei Li for the simulator conception, Rita Daditz, and Sarah Volz for scoring the procedures.

Conflicts of Interest: The authors declare no conflicts of interest.

Abbreviations

The following abbreviations are used in this manuscript:

ETI	Neonatal endotracheal intubation
ICUs	Intensive care units
ANOVA	Analysis of variance
MR	Mixed reality
AR	Augmented reality

References

1. Weiner, G.M.; Zaichkin, J. (Eds.) *Textbook of Neonatal Resuscitation*, 8th ed.; American Academy of Pediatrics and American Heart Association: Washington, DC, USA, 2021.
2. DeMeo, S.D.; Katakam, L.; Goldberg, R.N.; Tanaka, D. Predicting neonatal intubation competency in trainees. *Pediatrics* **2015**, *135*, e1229–e1236. [CrossRef] [PubMed]
3. Evans, P.; Shults, J.; Weinberg, D.D.; Napolitano, N.; Ades, A.; Johnston, L.; Levit, O.; Brei, B.; Krick, J.; Sawyer, T.; et al. Intubation competence during neonatal fellowship training. *Pediatrics* **2021**, *148*, e2020036145. [CrossRef]
4. Konrad, C.; Schupfer, G.; Wietlisbach, M.; Gerber, H. Learning manual skills in anesthesiology: Is there a recommended number of cases for anesthetic procedures? *Anesth. Analg.* **1998**, *86*, 635–639. [CrossRef]
5. Xiao, X.; Zhao, S.; Meng, Y.; Soghier, L.; Zhang, X.; Hahn, J. A physics-based virtual reality simulation framework for neonatal endotracheal intubation. In Proceedings of the 2020 IEEE Conference on Virtual Reality and 3D User Interfaces (VR), Atlanta, GA, USA, 22–26 March 2020; pp. 557–565.
6. Meng, Y.; Hahn, J.K. An Automatic Grading System for Neonatal Endotracheal Intubation with Multi-Task Convolutional Neural Network. In Proceedings of the 2023 IEEE EMBS International Conference on Biomedical and Health Informatics (BHI), Pittsburgh, PA, USA, 15–18 October 2023; pp. 1–4.
7. Haubner, L.Y.; Barry, J.S.; Johnston, L.C.; Soghier, L.; Tatum, P.M.; Kessler, D.; Downes, K.; Auerbach, M. Neonatal intubation performance: Room for improvement in tertiary neonatal intensive care units. *Resuscitation* **2013**, *84*, 1359–1364. [CrossRef]
8. Schneider, W.; Shiffrin, R.M. Controlled and automatic human information processing: I. Detection, search, and attention. *Psychol. Rev.* **1977**, *84*, 1. [CrossRef]
9. Sweller, J. Cognitive load during problem solving: Effects on learning. *Cogn. Sci.* **1988**, *12*, 257–285. [CrossRef]
10. Logan, G.D. Automaticity, resources, and memory: Theoretical controversies and practical implications. *Hum. Factors* **1988**, *30*, 583–598. [CrossRef]
11. Wickens, C.D. Processing resources and attention. In *Multiple Task Performance*; CRC Press: Boca Raton, FL, USA, 2020; pp. 3–34.
12. Stefanidis, D.; Scerbo, M.W.; Korndorffer, J.R., Jr.; Scott, D.J. Redefining simulator proficiency using automaticity theory. *Am. J. Surg.* **2007**, *193*, 502–506. [CrossRef]
13. Szafranski, C.; Kahol, K.; Ghaemmaghami, V.; Smith, M.; Ferrara, J.J. Distractions and surgical proficiency: An educational perspective. *Am. J. Surg.* **2009**, *198*, 804–810. [CrossRef]
14. Haji, F.A.; Khan, R.; Regehr, G.; Drake, J.; de Ribaupierre, S.; Dubrowski, A. Measuring cognitive load during simulation-based psychomotor skills training: Sensitivity of secondary-task performance and subjective ratings. *Adv. Health Sci. Educ.* **2015**, *20*, 1237–1253. [CrossRef]
15. Zirkle, M.; Roberson, D.W.; Leuwer, R.; Dubrowski, A. Using a virtual reality temporal bone simulator to assess otolaryngology trainees. *Laryngoscope* **2007**, *117*, 258–263. [CrossRef] [PubMed]
16. Zhou, Y.; Bailey, J.; Ioannou, I.; Wijewickrema, S.; O'Leary, S.; Kennedy, G. Pattern-based real-time feedback for a temporal bone simulator. In Proceedings of the 19th ACM Symposium on Virtual Reality Software and Technology, Singapore, 6–8 October 2013; pp. 7–16.
17. Pratt, P.; Ives, M.; Lawton, G.; Simmons, J.; Radev, N.; Spyropoulou, L.; Amiras, D. Through the HoloLens™ looking glass: Augmented reality for extremity reconstruction surgery using 3D vascular models with perforating vessels. *Eur. Radiol. Exp.* **2018**, *2*, 2. [CrossRef] [PubMed]

18. Hamza-Lup, F.G.; Rolland, J.P.; Hughes, C. A distributed augmented reality system for medical training and simulation. *arXiv* **2018**, arXiv:1811.12815.
19. Alismail, A.; Thomas, J.; Daher, N.S.; Cohen, A.; Almutairi, W.; Terry, M.H.; Huang, C.; Tan, L.D. Augmented reality glasses improve adherence to evidence-based intubation practice. *Adv. Med. Educ. Pract.* **2019**, *10*, 279–286. [CrossRef]
20. Desselle, M.R.; Brown, R.A.; James, A.R.; Midwinter, M.J.; Powell, S.K.; Woodruff, M.A. Augmented and virtual reality in surgery. *Comput. Sci. Eng.* **2020**, *22*, 18–26. [CrossRef]
21. Dwivedi, S.; Hayes, J.; Pedron, I.; Kang, J.; Brenner, L.J.; Barnes, C.; Bailly, A.; Tanous, K.; Nelson, C.; Moats, J.; et al. Comparing the efficacy of AR-based training with video-based training. In Proceedings of the Human Factors and Ergonomics Society Annual Meeting, Atlanta, GA, USA, 10–14 October 2022; Volume 66, pp. 1862–1866.
22. Foglia, E.E.; Ades, A.; Napolitano, N.; Leffelman, J.; Nadkarni, V.; Nishisaki, A. Factors associated with adverse events during tracheal intubation in the NICU. *Neonatology* **2015**, *108*, 23–29. [CrossRef]
23. Hatch, L.D.; Grubb, P.H.; Lea, A.S.; Walsh, W.F.; Markham, M.H.; Whitney, G.M.; Slaughter, J.C.; Stark, A.R.; Ely, E.W. Endotracheal intubation in neonates: A prospective study of adverse safety events in 162 infants. *J. Pediatr.* **2016**, *168*, 62–66. [CrossRef]
24. Sanders, R.C., Jr.; Giuliano, J.S., Jr.; Sullivan, J.E.; Brown, C.A., III; Walls, R.M.; Nadkarni, V.; Nishisaki, A. National Emergency Airway Registry for Children Investigators and Pediatric Acute Lung Injury and Sepsis Investigators Network. Level of trainee and tracheal intubation outcomes. *Pediatrics* **2013**, *131*, e821–e828. [CrossRef]
25. Russo, S.G.; Bollinger, M.; Strack, M.; Crozier, T.A.; Bauer, M.; Heuer, J. Transfer of airway skills from manikin training to patient: Success of ventilation with facemask or LMA-SupremeTM by medical students. *Anaesthesia* **2013**, *68*, 1124–1131. [CrossRef]
26. Prabhu, A.; Smith, W.; Yurko, Y.; Acker, C.; Stefanidis, D. Increased stress levels may explain the incomplete transfer of simulator-acquired skill to the operating room. *Surgery* **2010**, *147*, 640–645. [CrossRef]
27. Mayrose, J.; Kesavadas, T.; Chugh, K.; Joshi, D.; Ellis, D.G. Utilization of virtual reality for endotracheal intubation training. *Resuscitation* **2003**, *59*, 133–138. [CrossRef] [PubMed]
28. Demirel, D.; Alexander, Y.; Halic, T.; Sankaranarayanan, G.; Ryason, A.; Spindler, D.; Butler, K.L.; Caroline, C.; Petrusa, E.; Molina, M.; et al. Virtual airway skills trainer (VAST) simulator. *Stud. Health Technol. Inform.* **2016**, *220*, 91. [PubMed]
29. Mayrose, J.; Myers, J.W. Endotracheal intubation: Application of virtual reality to emergency medical services education. *Simul. Healthc.* **2007**, *2*, 231–234. [CrossRef]
30. Rodrigues, M.A.F.; Gillies, D.; Charters, P. A biomechanical model of the upper airways for simulating laryngoscopy. *Comput. Methods Biomech. Biomed. Eng.* **2001**, *4*, 127–148. [CrossRef]
31. Meng, Y. VR Enhanced Interactive Intubation Simulation. In Proceedings of the International Conference on Modeling, Simulation and Visualization Methods (MSV). The Steering Committee of The World Congress in Computer Science, Computer Engineering and Applied Computing (WorldComp), Las Vegas, NV, USA, 1–5 August 2018; pp. 81–87.
32. Schumacher, E.H.; Seymour, T.L.; Glass, J.M.; Fencsik, D.E.; Lauber, E.J.; Kieras, D.E.; Meyer, D.E. Virtually perfect time sharing in dual-task performance: Uncorking the central cognitive bottleneck. *Psychol. Sci.* **2001**, *12*, 101–108. [CrossRef]
33. Paas, F.G.; Van Merriënboer, J.J. Instructional control of cognitive load in the training of complex cognitive tasks. *Educ. Psychol. Rev.* **1994**, *6*, 351–371. [CrossRef]
34. Sweller, J. Cognitive load theory. In *Psychology of Learning and Motivation*; Elsevier: Amsterdam, The Netherlands, 2011; Volume 55, pp. 37–76.
35. Cook, D.A.; Hamstra, S.J.; Brydges, R.; Zendejas, B.; Szostek, J.H.; Wang, A.T.; Erwin, P.J.; Hatala, R. Comparative effectiveness of instructional design features in simulation-based education: Systematic review and meta-analysis. *Med. Teach.* **2013**, *35*, e867–e898. [CrossRef]
36. Ward, M.D.; Helton, W.S. Dual-task interference while receiving information on a head mounted display and manual tracking with and without auditory warnings. *Appl. Ergon.* **2022**, *101*, 103713. [CrossRef]
37. Neter, J.; Kutner, M.H.; Nachtsheim, C.J.; Wasserman, W. *Applied Linear Statistical Models*; McGraw-Hill/Irwin: New York, NY, USA, 1996.
38. Ralston, A.; Wilf, H.S. (Eds.) *Mathematical Methods for Digital Computers*; Wiley: New York, NY, USA, 1960.

Disclaimer/Publisher's Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.