Project Interim Report: Rolando Del Maestro September 15/ 2022

Grant Title: A Randomized Controlled Trial to Assess the Efficacy of Expert Benchmark Feedback Systems in Simulated Brain Tumor Resection Training

Summary: The randomized controlled trial to assess the efficacy of expert benchmark feedback systems in simulated brain tumor resection training has been completed. Participants in all feedback groups significantly improved performance compared to baseline performance. Visual and visuospatial feedback protocols resulted in better performance improvement. Providing visual feedback resulted in significantly better improvement in comparison to providing the identical information with no-visual feedback. Simulations with autonomous visual computer assistance may be efficient pedagogical tools in teaching bimanual operative skills while training efficiency is significantly influenced by the methodology of feedback information delivery. Further studies are ongoing to assess the utilization of the automatic visual feedback system proven effective in this study.

Purpose of the research
The goal of this study is to test the efficacy of four training systems, with and without automated personalized feedback, in the teaching of bimanual technical skills during a complex simulated brain tumor surgical procedure.

Our objectives of this grant are to complete a randomized controlled trial:
To determine the effect of providing personalized feedback on trainee learning performance compared to a no feedback approach on simulated brain tumor surgery training.

Power Analysis: Based on a power of 0.8 with effect size of 0.2 required 23 participants per group, 112 participants in total. Our study aimed to include 120 Quebec medical students from 4 universities.

Progress and Results: The randomized controlled trial proposed in this grant has been completed and as seen in Figure 1, 120 medical students were randomly allocated to four groups for the trial. The four groups completed 5 Practice Simulation tasks on the NeuroVR virtual reality platform. The simulation task involved using a simulated bipolar forceps in the non-dominant hand to aid in resection of a simulated brain tumor and a simulated ultrasonic aspirator in the dominant hand to remove the simulated brain tumor (see Figure 2). Participants' performance progress was tracked across five repetitions of the task. Fourteen-performance metrics from four categories (1) safety, (2) quality, (3) efficiency, and (4) bimanual cognitive were calculated for each repetition of the task. Safety category included six metrics: (1) brain volume removed (cc), (2) amount of blood loss (cc), (3) maximum force applied with dominant hand (N), (4) maximum force applied with non-dominant hand (N), (5) sum of forces applied with dominant hand (N), and (6) sum of forces applied with non-dominant hand (N). Quality category included only tumor percentage removed. Efficiency category included 5 metrics: (1) total tip path length dominant hand (mm), (2) total tip path length non-dominant hand (mm), (3) path length index (mm/s), and (4) efficiency index. Bimanual cognitive category included (1)

average instrument tips separation distance (mm), (2) coordination index, and (3) bimanual forces ratio.

Group 1 trainees received no feedback while the three other groups received three different types of feedback with two groups receiving visual and visual spatial feedback after completing each of the five-practice brain tumor simulation tasks. Group 2 trainees received paper-printed feedback outlining their results on 14 different factors used to assess their performance as compared to an expert benchmark range previous developed utilizing the results from neurosurgeons. Group 3 trainees were provided with their results by visual feedback on each of the 14 factors being assessed. If their performance value was within the neurosurgeon benchmark range the trainee visualized the result on a green scale, if very close to the benchmark (within one standard deviation) on a yellow scale and if outside this range on a red scale and if very good on a purple scale. Group 4 trainees received the same visual information as Group 3 but were also provided with visual spatial feedback results of their performance using three dimensional representations.

Figure-1: Flow diagram. One-hundred-twenty students were randomly allocated into four different feedback groups including practice-alone with no-feedback group. No participant/data was excluded from the analysis.

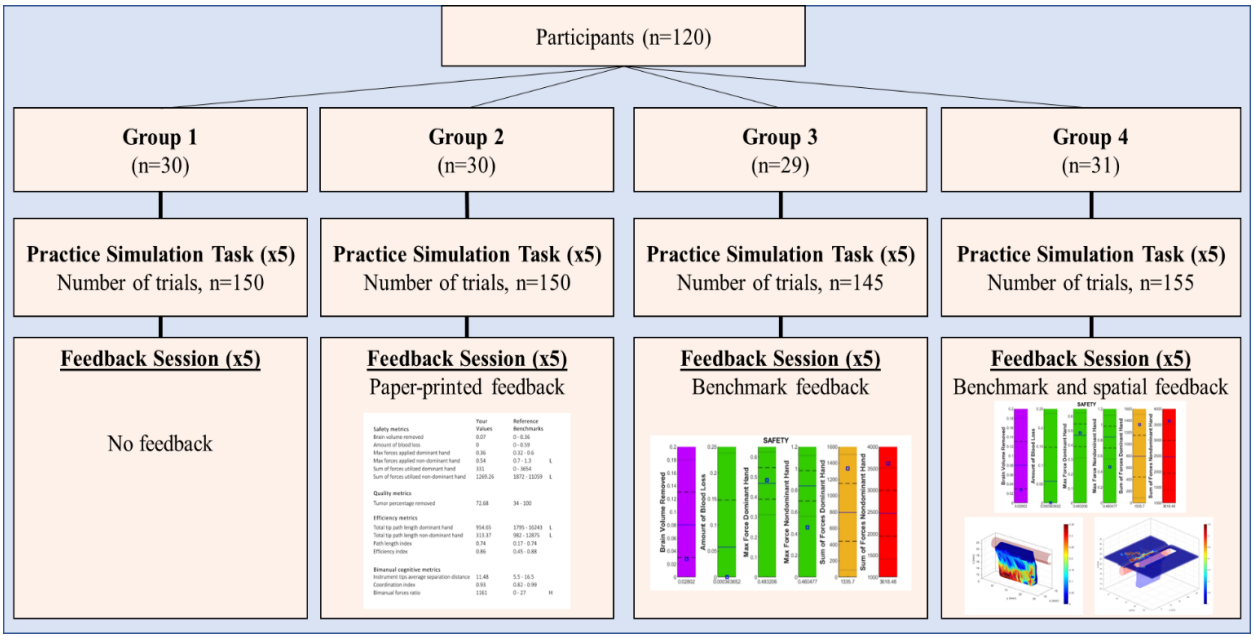
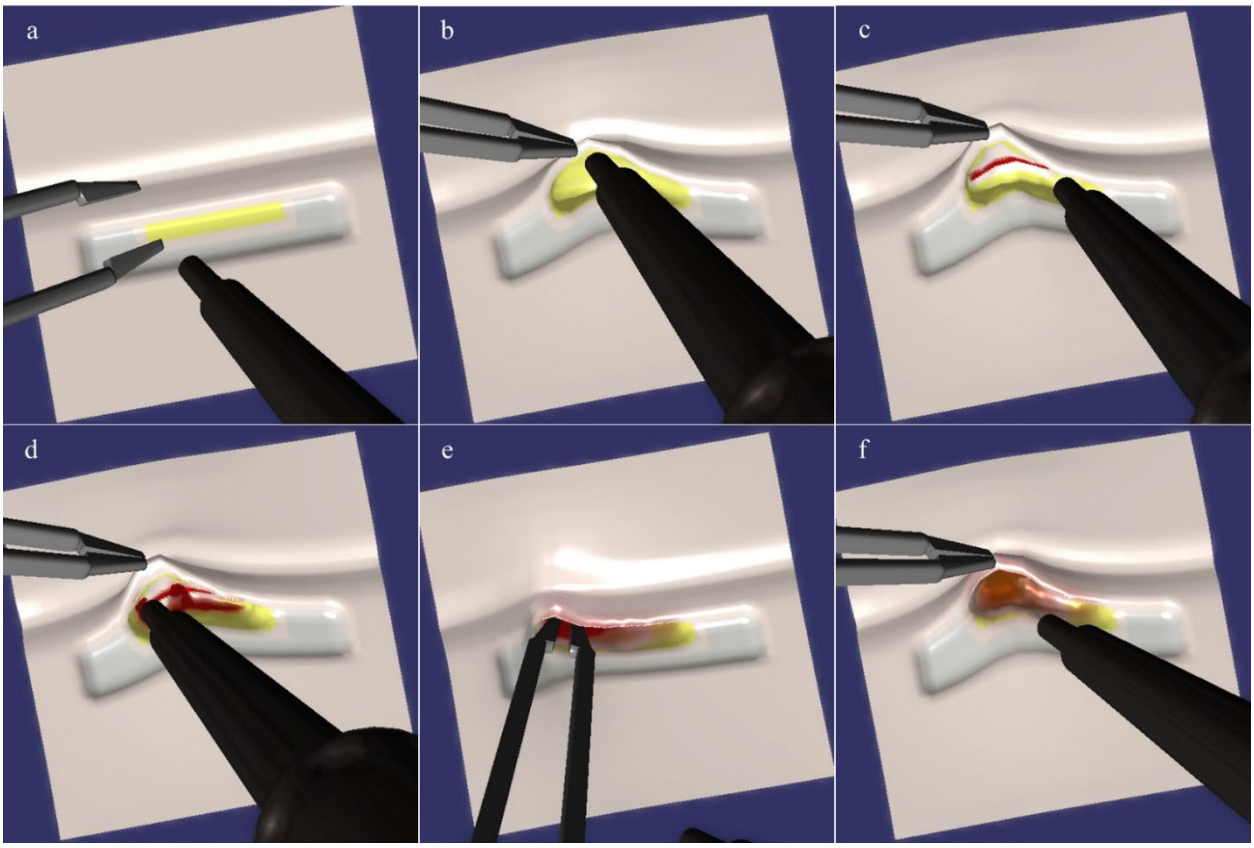


Figure-2: Simulated Scenario. The virtually simulated task involved the subpial resection of a rectangular yellow tumor using an ultrasonic aspirator in the dominant hand and a bipolar forceps in the non-dominant hand (a). The tumor was to be removed completely while minimizing injury to the surrounding tissue (b). There was a blood vessel near the tumor with bleeding capacity (c). Any damage to this blood vessel resulted in bleeding (d). Ultrasonic

aspirator was used to aspirate the blood (d) and bipolar was used to cauterize bleeding tissues (e). (f) demonstrates the appearance of the tissue after successful cauterization.



The Table 1 outlines the demographic characteristics and responses to other questions asked before participation in the trial. Participants were from each of the 4 Quebec universities.

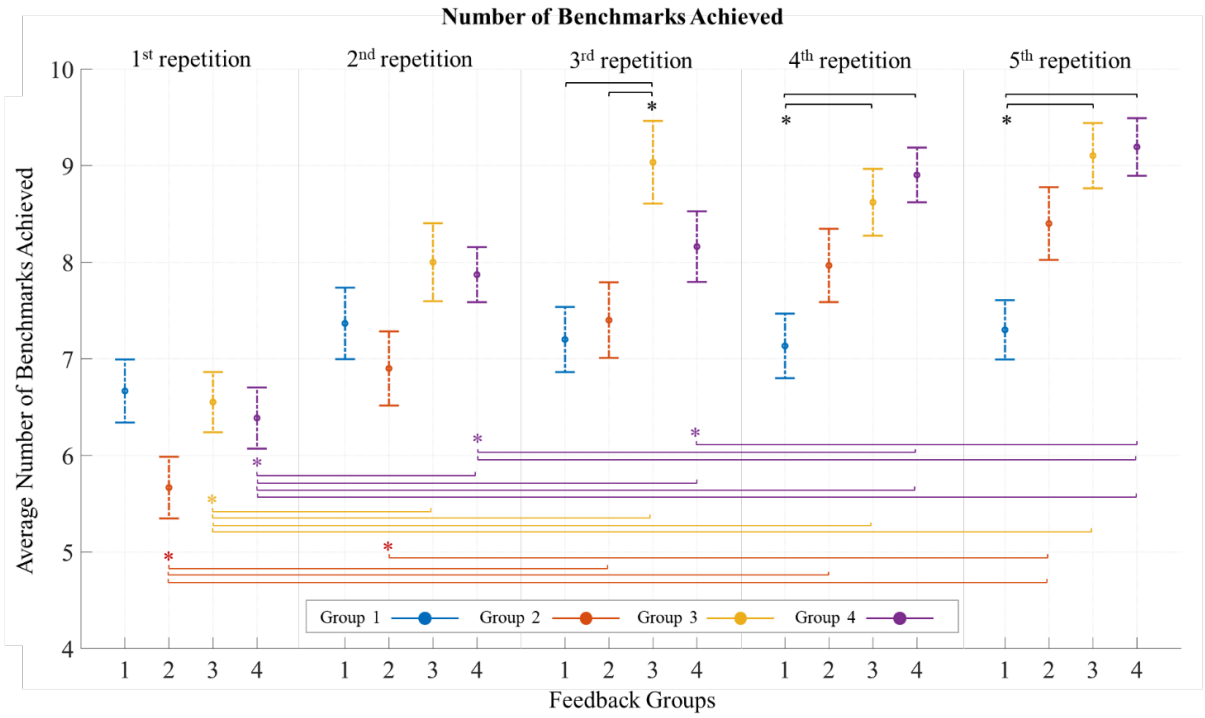
	Group 1 No feedback (n=30)	**Group 2** No-visual feedback (n=30)	**Group 3** Visual feedback (n=29)	**Group 4** Visuospatial feedback (n=31)	**All Participants** (n=120)
Mean age +/- SD (range)	23.6 +/- 4.8 (19-44)	22.8 +/- 3.3 (19-31)	22.4 +/- 2.6 (19-28)	23.6 +/- 3.5 (18-33)	23.1 +/- 3.6 (18-44)
Male/Female	18/12	18/12	18/11	17/14	71/49
Handedness (Right/Left/Ambidextrous)	27/3/0	28/2/0	24/4/1	29/1/1	108/10/2
Year in medical school:					
1st	16	21	18	20	75
2nd	10	6	7	8	31
3rd	3	2	2	1	8
4th	1	1	2	2	6
Level of interest in surgery, median (range)	4 (2-5)	4 (1-5)	4 (1-5)	4 (1-5)	4 (1-5)
Completed surgical rotation (Y/N)	2/28	1/29	2/27	2/29	7/113
Playing video games:					
Not at all	12	13	13	13	51
Occasionally (less than 2 hours per week)	9	9	7	9	34
Often (2- 10 hours per week)	6	8	6	6	26
Very often (more than 10 hours per week)	3	0	3	3	9
Playing musical instruments:					
I don't play any musical instrument	11	14	9	17	51
Yes, I am at beginner level	6	4	6	3	19
Yes, I am at intermediate level	6	7	8	6	27
Yes, I am at advanced level	6	4	4	5	19
Yes, I am at master level	1	0	2	0	3
Previously used virtual reality simulation (Y/N)	1/29	2/28	0/29	2/29	5/115

Our hypotheses were:

1- All three feedback groups provided with quantitative feedback will improve significantly faster than the control no-feedback, Group (1) in global performance score.
2- That a dose-response pattern in subpial simulated tumor resection performance will be observed in groups provided with quantitative feedback, in which Group (4) will improve significantly faster than Group (3), and Group (3) will significantly outperform Group (2).

The initial results can be seen in Figure 3. The values seen in Figure 3 outline the average number of the 14 benchmarks achieved at the end of the 1st repetition compared to after each of the other five repetitions. There was no significant difference in the number of benchmarks achieved by the no- feedback group between the 1st and the 5 repetitions. All three feedback groups outperformed Group 1, the no-feedback group. However, Group 3 participants achieved a significant improvement by the 3rd repetition, and this was only achieved by Group 4 by the 4th repetition.

Figure-3: Number of Benchmarks Achieved. X-axis represents the four feedback groups. Each feedback group is color-coded (see the legend). Y-axis represents the average number of benchmarks achieved by each feedback group. *Horizontal lines represent statistically significant difference ($p<.05$). For within group differences, horizontal lines are represented with the assigned color of the group. Vertical lines represent standard error bars. Group 3 and Group 4 improved significantly compared to the baseline performance by the second repetition. Group 2 improved significantly compared to baseline performance by the third repetition. Group 3 outperformed practice-alone Group 1 by the third repetition. Group 4 outperformed practice-alone Group 1 by the fourth repetition.



Initial conclusions:
Our results demonstrate that our first hypothesis is correct in that all three feedback groups outperformed Group 1, the no-feedback group. Our second hypothesis however is not supported since Group 3 outperformed both Group 2 and Group 4. Simulations with autonomous visual computer assistance may be efficient pedagogical tools in teaching bimanual operative skills while training efficiency is significantly influenced by the methodology of feedback information delivery.

Discussion:
Our present hypothesis is that the reason that Group 4 does not improve as quickly as Group 3 may be related to cognitive overload. Group 4 trainees presented with their results on 14 metrics presented visually and two 3-dimentional representations of their performance may not be able to integrate such a large amount of information effectively dure to the volume of the information presented.

Present investigations:

Our present research involves outlining the response of each of the 4 groups to all 14 metrics assessed in this study to further clarify which of the metrics assessed do not improve, which improve and if any metrics worsen by presenting the metric information.

It is felt that this manuscript will be ready to submit for publication in the next 2 -3 months.

Brain Tumour Foundation of Canada Brain Tumor Research Grant Funding of Research Projects in the Neurosurgical Simulation and Artificial Intelligence Learning Centre.

Our group has developed an Intelligent Continuous Expert Monitoring System (ICEMS) to continuously assess trainee performance during virtual reality brain tumor surgery and this research has been completed with help from this Brain Tumour Foundation of Canada Brain Tumour Research Grant and published in npg Digital Medicine (attached). Since this article is published in this Nature publication it is available to all researchers online free of charge which should increase the impact of this article.

Yilmaz R, Winkler-Schwartz A, Mirchi N, Reich A, Christie S, Tran DH, Ledwos N, **DelMaestro RF**.(2022) Continuous monitoring of surgical bimanual expertise using deep neural networks in virtual reality simulation. *npj Digit Med.* **5,** 54 (2022) https://doi.org/10.1038/s41746-022-00596-8

The ICEMS has been utilized in a randomized controlled trial which demonstrated that virtual Operative Assistant feedback assessed by the ICEMS demonstrated superior performance outcome and skill transfer compared with remote expert instruction, indicating advantages for its use in simulation training (attached). This study was also supported by this Brain Tumour Foundation of Canada Brain Tumour Research Grant. This article is published in the Journal of the American Medical Association (JAMA) and this article is available to all researchers online free of charge which should increase the impact of this article.

Fazlollahi AM, Bakhaidar M, Alsayegh A, Yilmaz R, Winkler-Schwartz A, Mirchi N, Langleben I, Ledwos N, Bajunaid K, Sabbagh A, Harley J, **Del Maestro RF**. (2022) Effect of artificial intelligence tutoring vs expert instruction on learning simulated surgical skills among medical students: a randomized clinical trial. JAMA Netw Open. 2022;5(2):e2149008. doi:10.1001/jamanetworkopen.2021.49008

Submitted manuscripts: This manuscript also supported by the present grant involves the assessment of Electroencephalography (EEG) and artificial intelligence (artificial neural networks) to assess expertise

Natheir S, Christie S, Yilmaz R, Winkler-Schwartz A, Bajunaid K, Sabbagh AJ, Werthner P, Fares J, Azarnoush H, **Del Maestro RF.** Utilizing Artificial Intelligence and Electroencephalography to Assess Expertise on a Simulated Neurosurgical Task. In review: Computers in Biology and Medicine

Our second Objective:

To develop optimal methodology of personalized automated feedback on a neurosurgical platform that most effectively enhances bimanual skills required for safe brain tumor resection.

Our present research involves utilizing the feedback system developed in Group 3 in this study by using a randomized control trial to test non-inferiority of a real-time intelligent coaching system (ICEMS) in technical skills training, during virtually simulated brain tumor resection tasks, by comparing its efficacy in student learning to in-person real-time expert human instruction. This study will explore the efficacy of real-time intelligent instruction to in-person human instructions in virtual reality training. Intelligent systems may facilitate learning in all procedural based medical disciplines without the need for supervision.

Published Presentations (see attached)

Two presentations were given at the Canadian Conference for the Advancement of Surgical Education (C-CASE) 2021: Post-Pandemic and Beyond Virtual Conference and published as abstracts.

Continuous Monitoring and Assessment of Surgical Technical Skills Using Deep Learning Recai Yilmaz, Alexander Winkler-Schwartz, Aiden Reich, Rolando Del Maestro

Machine learning distinguishes between skilled and less-skilled psychological performance in virtual neurosurgical performance. Sharif Natheir, Sommer Christie, Recai Yilmaz, Alexander Winkler-Schwarz, Khalid Bajunaid, Abdulrahman J. Sabbagh, Penny Werthner, Rolando Del Maestro.