Student Evaluation of a Learning Enhancement in Anatomy Program (LEAP) during the COVID-19 Pandemic: A Retrospective Study

Jose V. Tecson, III, MD, DHPEd,[1,2] Rafael C. Bundoc, MD,[1] Abdel Jeffri A. Abdulla, MD, MHA,[1] Ronnie E. Baticulon, MD,[1] Ryner Jose D. Carrillo, MD, MSc,[1] Christopher S. Constantino, MD,[1] Karen June P. Dumlao, MD,[1] Rowena F. Genuino, MD, MSc,[1] Blesile Suzette S. Mantaring, MD,[1] Jose Leonard R. Pascual, V, MD,[1] Donnel Alexis T. Rubio, MD,[1] Kenny S. Seng, MD,[1] Florida F. Taladtad, MD,[1] Sylvia Karina L. Alip, MD,[1] Joannes Luke B. Asis,[3] Pio Renato F. Villacorta, MD[1] and Charlotte M. Chiong, MD, PhD[3]

[1]*Department of Anatomy, College of Medicine, University of the Philippines Manila*
[2]*Department of Ophthalmology and Visual Sciences, College of Medicine and Philippine General Hospital, University of the Philippines Manila*
[3]*College of Medicine, University of the Philippines Manila*

ABSTRACT

Background and Objective. The pandemic acted as an accelerator for the development of online teaching formats in anatomy and histology worldwide. The authors introduce a bridging program that reinforces the knowledge and understanding of gross and correlative anatomy and histology acquired in a virtual environment in preparation for its future clinical application. The study aims to evaluate the Learning Enhancement in Anatomy Program (LEAP) conducted among first-year medical students at the College of Medicine, University of the Philippines Manila.

Methods. This descriptive cross-sectional study aims to determine the initial experience of implementing a learning enhancement program and assess areas for its improvement. An internally validated questionnaire was given to students after the program to gauge students' reactions (Kirkpatrick Level 1 evaluation). Pre- and post-tests were administered to evaluate knowledge acquisition (Kirkpatrick Level 2 evaluation). Short-term behavioral peer evaluation (Kirkpatrick Level 3 evaluation) was also instituted.

Results. One hundred fifty-two (152) students participated in the study. General reactions from students to the LEAP were consistently positive, with a steady majority of the students rating '5' or a 'Strongly Agree' to positive statements regarding the program. Higher ratings for more traditional teaching methods, such as cadavers, formalinized specimens, and bones, compared to virtual systems were apparent. However, inter-student variation in preference for teaching modalities was observed. All stations of the LEAP were evaluated satisfactorily, with most gross anatomy stations rated higher than histology stations. A significant increase was noted in the total post-test scores compared to pre-test scores. This improvement in test scores was observed in the anatomy and histology subcategories and in six of the seven organ system modules. Perceived behavioral outcomes were also generally positive.

Conclusion. The LEAP is a worthwhile endeavor, garnering overwhelmingly positive reactions and a significant improvement in test scores. Future studies are necessary to fine-tune teaching and training in a blended learning environment.

Keywords: anatomy, COVID-19, medical education, program evaluation

eISSN 2094-9278 (Online)
Published: October 26, 2023
https://doi.org/10.47895/amp.v57i10.7300

Corresponding author: Jose V. Tecson, III, MD, DHPEd
Department of Anatomy
College of Medicine
University of the Philippines Manila
547 Pedro Gil St., Ermita, Manila 1000, Philippines
Email: jvtecson3@up.edu.ph
ORCiD: https://orcid.org/0000-0003-1237-7051

INTRODUCTION

All educational institutions, including medical schools, have been placed under lockdown as a result of the COVID-19 pandemic. However, the use of cadaveric material in laboratory dissection classrooms remained the mainstay of anatomy education in the Philippines until the pandemic. This included our institution, which taught anatomy using lectures, full-body cadaveric dissection, and other modalities like models, microscopic slides, bones, specimens, and other materials. To maintain continuity of instruction in response to the viral spread and the requirement for physical distance, medical schools were obligated to transition their teaching from face-to-face classrooms to web-based formats.[1]

Anatomy is the foundation that physicians build their clinical skills.[2] Cadaveric dissection remains the "gold standard" in anatomy instruction.[3] Students lost access to cadavers and various other learning materials because of the lockdown, which could impact their anatomy knowledge. It is widely accepted that the human cadaveric dissection process helps students understand the 3D relationships among anatomical structures and reinforces the contents of textbooks and lectures.[4,5]

In response to the COVID-19 pandemic, the University of the Philippines Manila's College of Medicine introduced a bridging program - Learning Enhancement in Anatomy Program (LEAP). The program, conducted by the Department of Anatomy, was designed for first-year medical students, and aimed to reinforce the knowledge and understanding of gross and correlative anatomy and histology that students acquire in a virtual environment. Furthermore, the program seeks to fulfill the learning outcomes missed in a virtual learning environment. Prior to the pandemic, eleven learning outcomes are addressed in each of the anatomy learning modules. Unfortunately, since there was no face-to-face instruction, a virtual platform was unable to meet the following learning outcomes: effective communication skills; adherence to professional and ethical standards; volunteerism, nationalism, and internationalism; inter-professional practice; and advocacy for social equity and accountability.

Evaluation is defined as "the systematic determination of the quality or value of something."[6] In medical education, evaluation has been defined "as a systematic approach to the collection, analysis, and interpretation of information about any aspect of the conceptualization, design, implementation, and utility of educational programs."[7]

This study used Kirkpatrick's model of Program Evaluation Design.[8] The framework assesses an educational program based on four levels of training program outcomes. Level 1 refers to the participant's description of the overall reaction on the basic program components (Reaction Evaluation). Level 2 measures how much participants have learned in terms of knowledge gained, skills developed or improved, and attitudes changed (Learning Evaluation). Level 3 focuses on how much participants are applying those they learned from the program to their actual workplaces (Behavior Evaluation). Level 4 concentrates on institutional changes and improvements, overall productivity, and operations efficiency practices that can be associated with the program (Results Evaluation).

With the end of the pandemic still nowhere in sight, this method of learning could become the new norm in anatomy education. To improve the program, it is essential to determine the participants' overall satisfaction and experience of it.

OBJECTIVES

This study evaluated the Learning Enhancement in Anatomy Program (LEAP) conducted among first-year medical students at the College of Medicine, University of the Philippines Manila during the COVID-19 pandemic. Specifically, the study:

1. Described the overall reactions of the participants to the LEAP in terms of the program's administration, content, organization, and teaching methodologies (Level 1);
2. Determined how much learning in terms of knowledge the participants learned from the program (Level 2); and
3. Evaluated the behavior changes in the participants (Level 3).

METHODS

Study Design

This is a descriptive cross-sectional retrospective study to evaluate the Learning Enhancement in Anatomy Program (LEAP) of the University of the Philippines College of Medicine using secondary analysis of data collected during the LEAP. The data included student demographics, reaction to the program (Kirkpatrick Level 1), individual performance in the tests (Kirkpatrick Level 2), and perceived behavior assessment (Kirkpatrick Level 3).

Learning Enhancement in Anatomy Program

The LEAP was a face-to-face learning activity organized by the Department of Anatomy from June 13 to 17, 2022. This activity was a bridging program to enhance the learning of the students in Anatomy by presenting prosected cadavers, plastinated cadavers and organs, anatomic models, specimens, bones, microscopic glass slides, and electronic images for further learning. Virtual dissection tables and headsets were also used. The following organ systems were covered by the 14 LEAP stations: OS 201 (Organ System 201: Human Cell Biology), 202 (Integration and Control System), 203 (Skin, Muscles, and Bones), 204 (Head and Neck), 205 (Thorax), 206 (Abdomen and Pelvis), and HD 201 (Human Ontogeny and Parturition). The students were divided into 14 groups of 12-13 members. Each group was allotted 2 hours to complete a station. After each station, the students were provided evaluation forms to solicit their feedback. Pre-test and post-test on gross anatomy (Anatomy) and microscopic

anatomy (Histology) were administered. At the conclusion of the program, the students were requested to answer a Google (Google LLC, Mountainview CA, USA) survey form measuring their reaction to the program's administration, content, and teaching methods.

Prior to the activity, students accomplished an informed consent detailing the need for compliance with the Commission on Higher Education (CHED) guidelines on face-to-face learning, assurance of confidentiality, and that responses would not influence their overall grade for the learning activity.

Questionnaire

The questionnaire was divided into several sections and was sent to all participants online. The first and second sections included a description of the study, the study objectives, and a data privacy statement. The introductory sections stressed the voluntary nature of the survey. The third section was devoted to collecting demographic information from participants. The fourth through seventh sections included questions assessing the students' reactions and reception to the program organization, specific program components, teaching materials and methods, and a general evaluation. All questions were internally validated by study team members and questions deemed ambiguous in wording were revised. Participants provided their assessment of the different program components using a five-point modified Likert scale, with 5 indicating "strongly agree" and 1 indicating "strongly disagree."

Data Management and Statistical Analyses

Data collected during the LEAP were retrieved from the database of the Department of Anatomy for secondary analysis. Only the data collected from students who consented to participate in the study were included. Data encoding was done using Microsoft Excel software (Microsoft Corporation, Albuquerque NM, USA) by a research assistant who is not employed by the Department of Anatomy and did not participate in any process of implementation of the LEAP.

The distribution of student responses in the evaluation of the program were presented using horizontal percentage stacked bar charts. Measures of central tendency and dispersion were also computed. Kolmogorov-Smirnov test was used to test for normality of data. Differences in student ratings among teaching materials and among stations were performed using the Friedman test. Post hoc Dunn's multiple comparison test was performed for pairwise comparison of different methods and stations.

Pre-test and post-test scores were summarized using measures of central tendency, dispersion, skewness, and kurtosis. Moreover, a histogram was used to illustrate the change in distribution of scores before and after the examination. Kolmogorov-Smirnov test was used to test for normality of data. Wilcoxon matched-pairs signed rank test was used to determine significant differences between the pre-test and post-test scores of the students. Analyses of the subcategories of the tests were also performed.

Differences were considered statistically significant when $P < 0.05$. Statistical analyses were performed in statistical software R version 4.2.1. (http://www.r-project.org) and GraphPad Prism Version 9.0.0. All graphs were generated using the ggplot R package.

RESULTS

Profile of Respondents

During the LEAP, 178 students participated out of the 179 first-year medical students enrolled in the University of the Philippines College of Medicine invited to participate. A total of 152 students consented to participate in the study, with 152 digital forms received without duplicates.

Table 1 shows the demographic profile of the LEAP participants. The participants were almost equally divided between males and females. One student identified as non-binary. The median age of the sample was 23, with 88.8% of the participants between 21 and 25 years old. The youngest student was 19 years old, while the oldest student was 33 years old.

Reaction Evaluation (Kirkpatrick Level 1)

Figure 1 shows the distribution of scores for the overall assessment of the administration and organization of the program. Scores of 4 or 5 on the Likert scale were considered satisfactory, a score of 3 was considered neutral, and scores of 1 or 2 were interpreted as unsatisfactory. The program organization and design and facility setup received satisfactory ratings (i.e., 4-5) from more than 95% of the students. Greater than 65% of students gave the highest rating of 5 for these items. These indicate high student satisfaction with these aspects of the LEAP. These were followed by the number of students per subgroup and the number of stations with around 75% positive reviews. Time allocation per station was rated the lowest in the overall assessment, with only 60% of students satisfied.

Figure 2 shows the score distribution and median for the various teaching materials and methodologies employed in the LEAP. Among teaching materials/methods, cadavers received the highest rating, with more than 90% of the students rating it 5 and no ratings below 4. This was followed by bones, formalinized specimens, and skills stations. On the other hand, virtual reality (VR) headset and virtual tables received the lowest ratings, with more than 35% of scores below 4.

Analysis using the Friedman test showed a significant difference among the different teaching materials ($Q_8 = 451.0$, $P < 0.0001$). Multiple comparisons using Dunn's test show that VR headset and virtual tables received significantly lower scores compared to all other modalities ($P < 0.0001$ for most comparisons). Skills stations, formalinized specimens, cadavers, and bones did not differ from each other ($P > 0.05$)

but scored higher than plastinated specimens and models ($P < 0.05$). Thus, teaching materials can be classified into three tiers based on student perceptions of effectiveness: high tier (cadavers, bones, formalinized specimen, and skills stations), middle tier (plastinated specimen and models), and low tier (VR headset and virtual tables).

Students were also asked to rank the teaching materials and methods from their most preferred (1) to their least preferred (8), as seen in Table 2. Cadavers were the most preferred modality in 60% of students, while formalinized specimens ranked second for 36% of participants. VR headset was ranked lowest by more than 40% of the students,

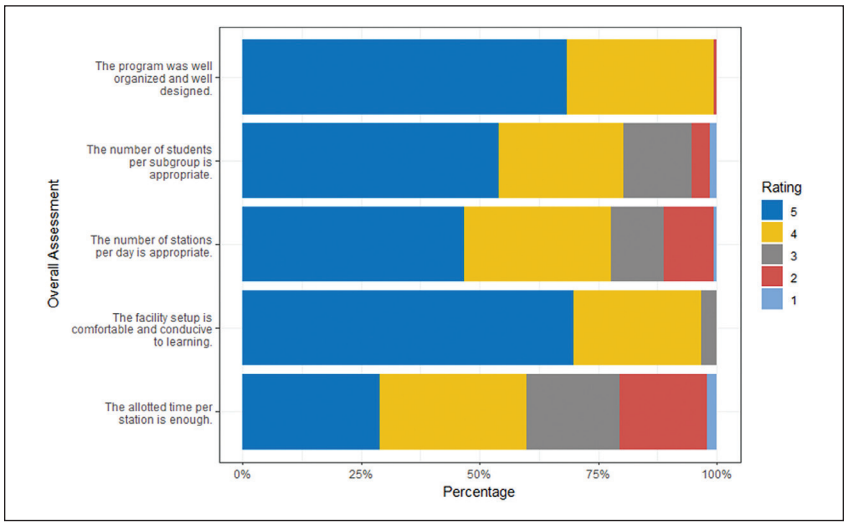

Figure 1. Distribution of scores given by LEAP participants regarding the overall assessment of the administration and organization of the program.

(1-2: Unsatisfactory, 3: Neutral, 4-5: Satisfactory)

Table 1. Demographic Profile of LEAP Participants

Variable	Count	%
Sex		
Male	76	50.0
Female	75	49.3
Non-binary	1	0.7
Age		
Below 21	6	3.9
21-25	135	88.8
26-30	9	5.9
Above 30	2	1.3

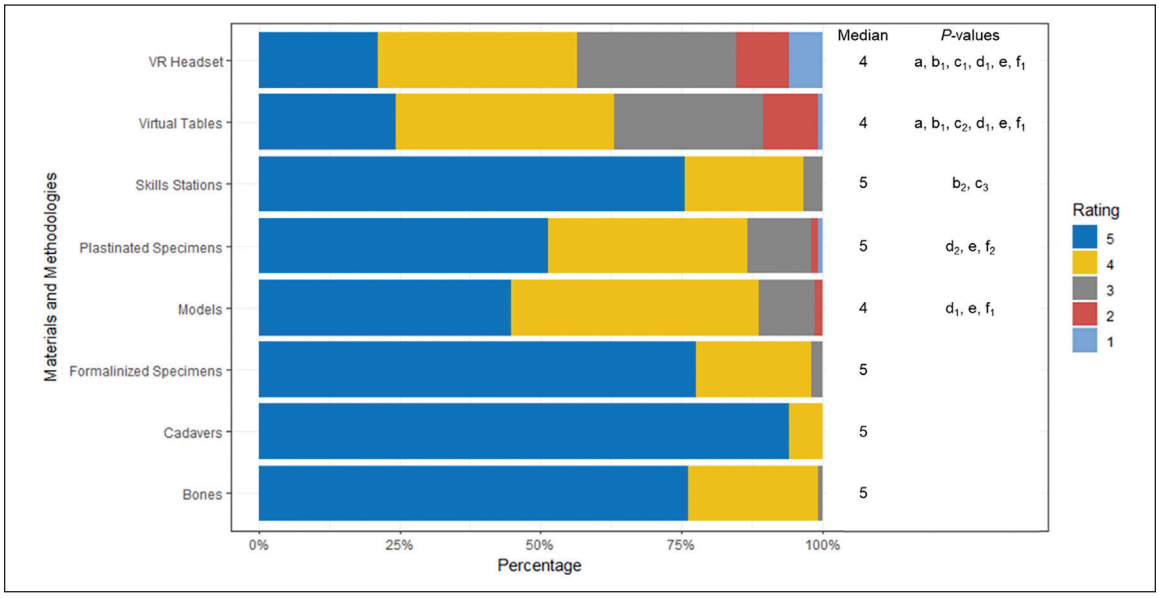

Figure 2. Distribution and median of scores given by LEAP participants for the teaching materials and methods.

(1-2: Unsatisfactory, 3: Neutral, 4-5: Satisfactory)

Compared to Skills Station – a: $P < 0.0001$
Compared to Plastinated Specimens – b: $P < 0.0001$, b2: $P = 0.0100$
Compared to Models – c_1: $P < 0.0001$, c_2: $P = 0.0002$, c_3: $P = 0.0003$
Compared to Formalinized Specimens – d_1: $P < 0.0001$, d_2: $P = 0.0028$
Compared to Cadavers – e: $P < 0.0001$
Compared to Bones – f_1: $P < 0.0001$, f_2: $P = 0.0022$.

similar to the trend observed in the ratings. However, 15% of students ranked VR headsets as their most preferred method. Thus, preference for the teaching methods followed similar trends to perceived effectiveness scores in the sample, with traditional methods such as cadavers favored by students over virtual methods. However, individual preferences vary, with some low-tier methods ranked first by some students.

Figure 3 shows the score distribution and median given by the participants to the gross anatomy and histology stations. All anatomy stations received higher ratings from more than 86% of students and five out of the six anatomy stations had a median score of 5. Among anatomy stations, OS 204 (Head and Neck) and OS 205 (Thorax) stations were given the highest ratings, with about 70% of students giving these stations a score of 5. The OS 202 (Integration and Control) anatomy station, which consisted of the nervous and endocrine systems, was rated the lowest, with a median score of 4. On the other hand, histology stations

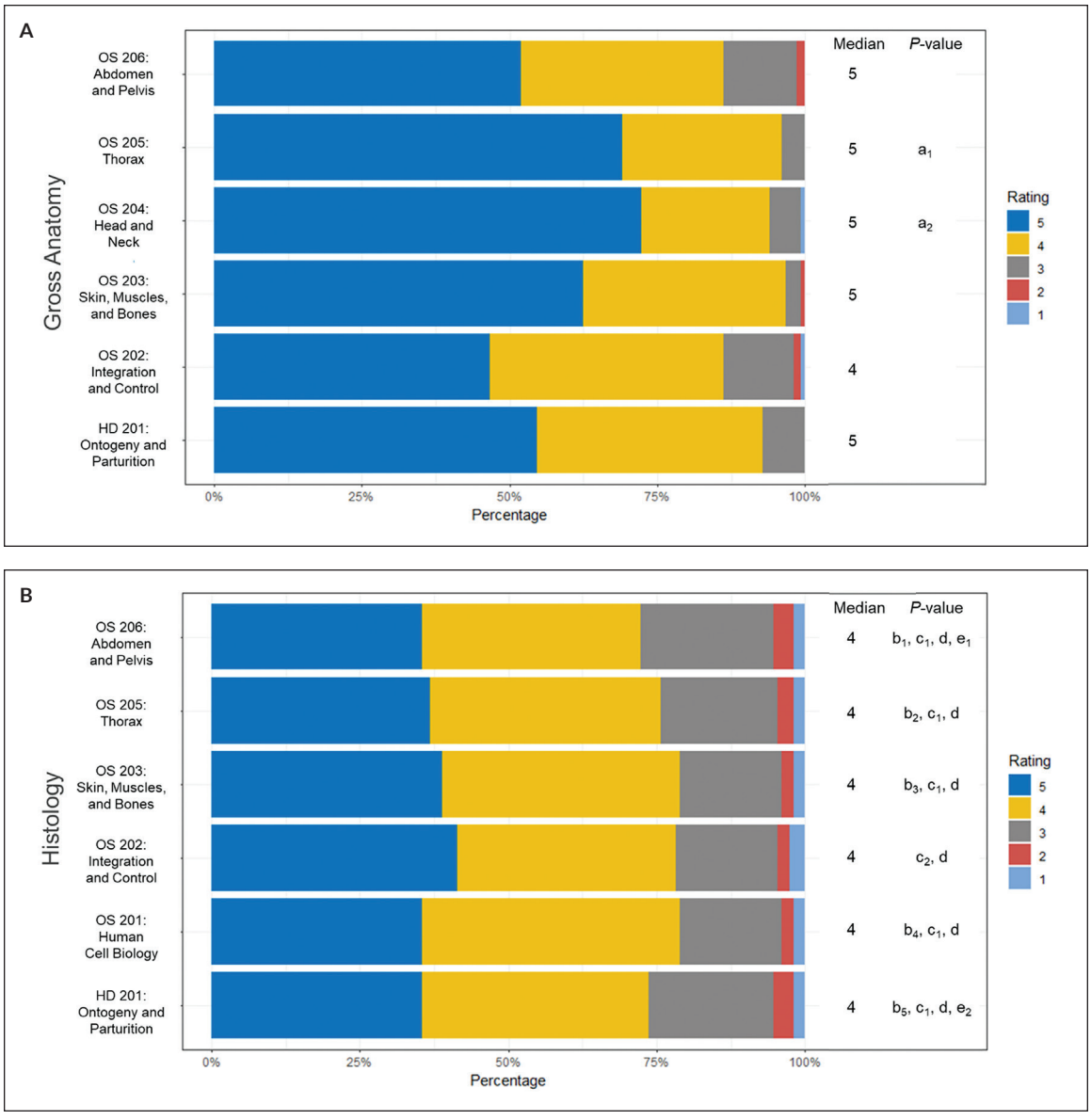

Figure 3. Distribution and median of scores for **(A)** gross anatomy and **(B)** histology stations.

(1-2: Unsatisfactory, 3: Neutral, 4-5: Satisfactory)

Compared to OS 202 anatomy station – a_1: P =0.0119, a_2: P = 0.0184
Compared to HD 201 anatomy station – b_1: P =0.0002, b_2: P =0.0024, b_3: P = 0.0273, b_4: P = 0.0048, b_5: P = 0.0003
Compared to OS 203 anatomy station – c_1: P < 0.0001, c_2: P = 0.0002
Compared to OS 204 or OS 205 anatomy station – d: P < 0.0001
Compared to OS 206 anatomy station – e_1: P = 0.0076, e_2: P = 0.0108.

received positive ratings from at least 72% of participants. However, gross anatomy stations received generally higher scores than microscopic anatomy stations, with median scores of 4 for all histology stations. The OS 202 station received the highest proportion of the highest rating among histology stations, while HD 201 (Human Ontogeny and Parturition) and OS 206 (Abdomen and Pelvis) stations scored lowest.

Statistical analyses of the anatomy and histology station scores using Friedman test showed significant difference among groups ($Q_{12} = 290.7, P < 0.0001$). Pairwise comparisons using Dunn's test among anatomy station ratings show that OS 204 ($P = 0.0184$) and OS 205 ($P = 0.0119$) scores were significantly higher compared to OS 202, the lowest rated anatomy station. In contrast, no difference in the scores of histology stations was observed ($P > 0.05$). However, almost all histology stations were given significantly lower ratings than each of the anatomy stations ($P < 0.05$) except the OS 202 station, suggesting that gross anatomy stations were perceived to be more effective than histology stations.

Table 3 presents the median scores of the general evaluation of the program. All students found the LEAP a worthwhile endeavor for first-year medical students and would strongly recommend the program to their peers. However, less students agreed that LEAP reinforced their knowledge of histology compared to anatomy, with the question on histology receiving the least number of a rating of 5. Participants were generally satisfied with regard to the other aspects of the LEAP, as seen in the high proportions of 5 or 4 ratings, suggesting that there are points of improvement in the program in terms of safety and security/ health protocols, faculty supervision during the program, contents and methodologies of LEAP, and administration and conduct of LEAP. The domain related to pandemic health and safety received the sole score of 1 and a few scores of 2. This may still be related to the general fear of contracting COVID-19 in the school setting.

Learning Evaluation (Kirkpatrick Level 2)

Figure 4 shows the histogram of pre-test and post-test scores of participants taken during the LEAP. The histograms show a rightward shift in the distribution of test scores from the pre-test to the post-test, indicating improvement in scores. This is supported by the increase in the median from 23 in the pre-test to 30 in the post-test and an increase in the mean by 11.17 percentage points. Wilcoxon matched-pairs signed rank test of the scores revealed a significant increase in scores ($W^+ = 12081$, $W^- = -165$, $z = -10.542$, $P < 0.0001$), suggesting that learning occurred in the participants. The

Table 2. Ranking of Teaching Materials and Methods

Teaching Materials and Methods	Rank (%)							
	1	2	3	4	5	6	7	8
Cadavers	60.53	8.55	0.66	2.63	0.00	0.66	3.95	23.03
Formalinized Specimens	2.63	36.84	20.39	7.24	7.24	13.82	11.84	0.00
Skills Stations	5.26	12.50	19.08	15.13	15.79	18.42	8.55	5.26
Bones	0.66	3.95	14.47	30.26	34.21	11.84	3.29	1.32
Models	1.32	1.97	9.87	29.61	24.34	19.74	8.55	4.61
Plastinated Specimens	5.26	13.82	25.00	11.84	11.18	15.79	9.87	7.24
Virtual Tables	9.21	11.18	7.24	1.32	4.61	10.53	37.50	18.42
VR Headset	15.13	11.18	3.29	1.97	2.63	9.21	16.45	40.13

(1: Most effective, 8: Least effective)

Table 3. Median Ratings on General Evaluation of the Program

Questionnaire Item	Median
LEAP reinforced my knowledge of Histology during this pandemic.	4
LEAP reinforced my knowledge of Anatomy during this pandemic.	5
LEAP is a worthwhile activity for first year medical students.	5
LEAP completed my first-year medical education during this pandemic.	5
I will recommend to my peers to participate in future LEAP offerings.	5
I feel safe/ secure and completely satisfied with health and safety protocols before, during, and after the program.	5
I am completely satisfied with the faculty supervision during the program.	5
I am completely satisfied with the contents and methodologies of LEAP.	5
I am completely satisfied with the administration and conduct of LEAP.	5

(1-2: Unsatisfactory, 3: Neutral, 4-5: Satisfactory)

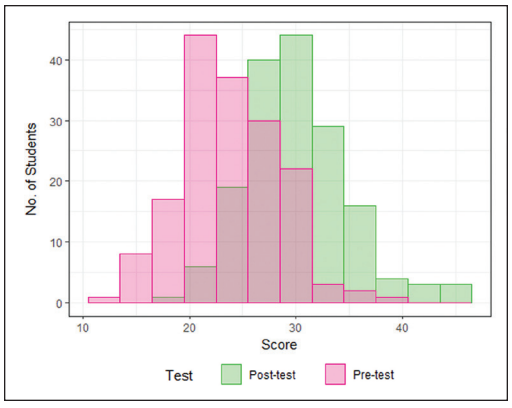

Figure 4. Histograms of Pre-test and Post-test Scores of LEAP participants.

Table 4. Comparison of Pre-test and Post-test Scores by Subcategories

Subcategories	Pre-test Score		Post-test Score		P-value
	Mean (SD)	Median	Mean (SD)	Median	
a. Anatomy	15.50 (3.58)	15	19.49 (3.48)	19	<0.0001****
b. Histology	8.34 (2.43)	8	10.49 (2.31)	11	<0.0001****
c. HD 201: Human Ontogeny and Parturition (Anatomy and Histology)	1.79 (1.05)	2	3.52 (0.95)	4	<0.0001****
d. OS 201: Human Cell Biology (Histology only)	1.78 (0.92)	2	2.01 (0.95)	2	0.0098**
e. OS 202: Integration and Control System (Anatomy and Histology)	4.73 (1.52)	5	5.32 (1.74)	5	0.0007***
f. OS 203: Skin, Muscles, and Bones (Anatomy and Histology)	4.80 (1.60)	5	6.77 (1.54)	7	<0.0001****
g. OS 204: Head and Neck (Anatomy only)	2.28 (1.22)	2	2.91 (1.03)	2	<0.0001****
h. OS 205: Thorax (Anatomy and Histology)	4.77 (1.58)	5	6.17 (1.50)	6	<0.0001****
i. OS 206: Abdomen and Pelvis (Anatomy and Histology)	3.69 (1.58)	4	3.30 (1.73)	3	0.0301*

** P < 0.05, ** P < 0.01, *** P < 0.001, **** P < 0.0001*

variance of the test scores increased from 20.95 to 24.12 which suggests that some students learned more than the others in the same amount of time. This is consistent with an increase in kurtosis from 3.03 to 3.77, indicating the presence of outliers – high scorers in this case.

Table 4 shows the pre-test and post-test scores by subcategories. Significant increases were seen in the items on Anatomy (W^+ = 11467, W^- = -623.5, P < 0.0001) and Histology (W^+ = 9465, W^- = -1267, P < 0.0001). For the individual organ system modules, higher post-test scores were seen in the test questions for HD 201 (Human Ontogeny and Parturition, P < 0.0001), OS 201 (Human Cell Biology, P = 0.0098), OS 202 (Integration and Control System, P = 0.0007), OS 203 (Skin, Muscles, and Bones, P < 0.0001), OS 204 (Head and Neck, P < 0.0001) and OS 205 (Thorax, P < 0.0001). In contrast, students scored lower in the post-test than pre-test scores for OS 206 (Abdomen and Pelvis, P = 0.0301), although only by a margin of 0.39 points.

Behavior Evaluation (Kirkpatrick Level 3)

The distribution of the perceived behavioral outcomes of the LEAP rated by students are shown in Figure 5. The majority of the students agreed that LEAP was able to strengthen the knowledge acquired by the students in a virtual platform and the students were able to demonstrate good interpersonal relationships, appreciate the art and science of anatomy, and demonstrate sensitivity and respect for life and death, and the dignity and rights of persons.

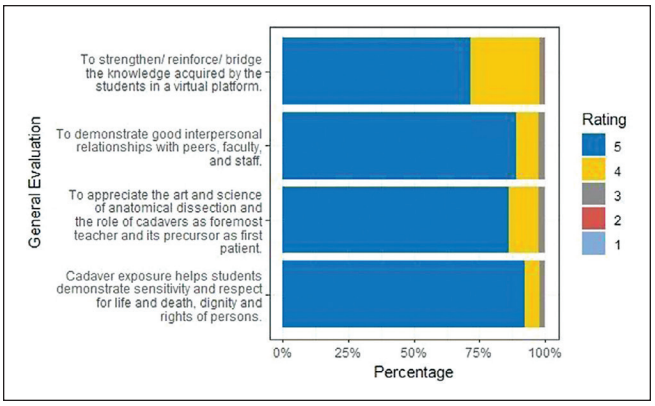

Figure 5. Perceived behavioral assessment of the program.
(1-2: Unsatisfactory, 3: Neutral, 4-5: Satisfactory)

DISCUSSION

The Organ System Integrated (OSI) Curriculum of the University of the Philippines, College of Medicine was approved for full implementation in April of 2003 and was fully implemented for academic year 2004-2005 completely replacing the existing curriculum. The concept of the integrated curriculum is an organization of the horizontal and vertical contents of a traditional medical curriculum into coherent learning units that aim to bring the students beyond the level of mere acquisition of facts to a higher plane of scientific understanding and fluency.[9] Horizontal integration involves the unification of various traditional disciplines, while vertical integration requires the synchronization of basic science concepts with clinical skills acquisition. The OSI curriculum uses organ systems as foci for this integrated learning.

It is through the same lens that the LEAP may be viewed. Shoemaker et al. defined an integrated curriculum as "education that is organized in such a way that it cuts across subject-matter lines, bringing together various aspects of the curriculum into meaningful association to focus upon broad areas of study."[10] It mirrors how teaching and learning are gleaned from the real world: dynamic and not static, interactive and not isolated, an all-senses-at-once affair but within the confines of a structured yet adaptive curriculum. In 2020, with the looming COVID-19 pandemic, all instruction was transferred online. Although the benefits of this were obvious as it served to mitigate disease spread, an education crisis was globally imminent. For anatomy, this meant that education was cadaverless due to the lockdown, with a predicted shortage of human cadavers due to limitations in

body donation (Due to the virus, most cadavers were disposed of immediately, as there was no proven screening to rule out donor infection).[2] Anatomy learning without cadavers is a practice that is generally seen as less favorable, but one which has arguable merits and has been used as standard in many institutions.[7,11] However, when prosections, models, and other learning materials are also removed, learning becomes difficult.[12] To maintain adequate exposure of medical students to cadaveric dissections and other physical learning materials, the Seoul National University College of Medicine (NUCM) in South Korea adopted simultaneous and rotating dissection schedules. To maintain physical distance, the 154 students who participated in the anatomy laboratory were divided into three classes, and each class conducted a cadaver dissection for 3 hours. In a three-hour laboratory dissection session, approximately 50 students were divided into 10 groups, with five students per cadaver.[13] The Lee Kong Chian School of Medicine in Singapore adopted a similar setup,[14] while the Korea University College of Medicine re-opened its dissection laboratory under strict public health protocols, reducing total anatomy laboratory hours to 85 in 2020 and 88 in 2021, from 96 hours in 2019.[15]

The UP College of Medicine pioneered an enhancement program in anatomy and histology in the country. The LEAP is intended to provide a comprehensive laboratory experience where students are provided with anatomy learning materials in their physical form that will stimulate learners to retrieve prior knowledge from the virtual sessions, perform tasks (per station) in groups, plan and discuss application and relationship of each task to the overall understanding of anatomy. This evaluation of the LEAP experience gathered information on its fitness of purpose as an enhancement program and as an essential component of learning anatomy in an integrated medical curriculum (OSI). It can also show the areas for improvement as a curricular intervention.

Reaction Evaluation (Kirkpatrick Level 1)

Keeping in mind the learner's reaction to the learning environment, as in the Kirkpatrick model's Level 1 of evaluation, the dissatisfaction in traditional subject-based curriculum has been well studied[16], as has been low efficiency in terms of critical-thinking and problem-solving.[17] In the present era, the debate is elsewhere: several studies have been conducted comparing online or virtual learning platforms to classroom or laboratory-based education. Pre-pandemic, Mathiowetz et al., reported significantly lower rates of self-perceived learning and satisfaction in students adopting an online anatomy program compared to those in a gross anatomy course.[18] With the pandemic, student concerns on the lack of anatomy laboratory sessions were reported exhaustively, with some looking forward to the time when they will be allowed to return to it.[10]

In this study, it would be wise to keep in mind that these students, having started medical school already in a virtual environment, have no 'pre-pandemic setup' comparison. General evaluation of the LEAP has been positive, with most domains evaluated by more than 50% of the class with a score of 5 on a modified Likert scale (Figure 1 and Table 3).

As regards satisfaction with teaching materials, virtual platforms (such as a virtual dissection table and a virtual-reality headset program) were consistently the least ranked (Figure 2 and Table 2). Conversely, cadavers were consistently the best ranked and rated across all stations. These results are aligned with the findings of most studies on virtual platforms. Attardi et al. noted that students in the online laboratory had difficulty using the 3D models and preferred the unique and hands-on experience with the cadaver.[19] However, differences in individual preferences were observed, suggesting that different learners prefer to learn through different means. Thus, traditional methods can be prioritized in anatomy education but other methods should still be offered to students.

Differences in perceived effectiveness were also seen in the anatomy and histology stations (Figure 3). Students generally gave higher scores to gross anatomy stations than microscopic anatomy stations. The underlying reason for these findings warrants further exploration to improve the overall experience of students for all LEAP stations.

Learning Evaluation (Kirkpatrick Level 2)

In his seminal text, Kirkpatrick writes, "learning has taken place when one or more of the following occurs: attitudes are changed, knowledge is increased, skill is improved."[8] Learning is a compendium of knowledge, skill, and attitude, and one or more of these changes must take place if a change in behavior (Level 3) is to occur.

Students scored higher in the post-test than the pre-test when considering the total score and most subcategories of the test (Figure 4 and Table 4), particularly anatomy, histology, ontogeny and parturition (HD 201), human cell biology (OS 201), nervous and endocrine systems (OS 202), integumentary and musculoskeletal systems (OS 203), head and neck (OS 204), and thorax (OS 205). Test scores for the abdomen and pelvis module (OS 206) showed a marginal but significant decrease from pre-test to post-test. The stations on this anatomical region ranked low among both anatomy and histology stations, although no significant difference was observed. Potential mismatches in Kirkpatrick levels have been expounded in some studies.[20,21] Proposed reasons for mismatch include a lack of validated tools to evaluate knowledge domains more accurately. Future studies can further explore the relationship between Levels 1 and 2 for anatomy and histology in each anatomical region/organ system. This can be achieved by expounding on self-perception of knowledge acquisition by adding multiple questions or by using an externally validated questionnaire.

Behavior Evaluation (Kirkpatrick Level 3)

Kirkpatrick Level 3 can be defined as the extent to which change in behavior has occurred because the participant

attended the training program. Level 3 evaluation aims to answer the question: what happens when trainees leave the program and return to their work? For Levels 1 and 2 (reaction and learning), the evaluation can and should take place immediately. When you evaluate change in behavior, some important decisions must be made when to evaluate, how often to evaluate, and how to evaluate.[8] Although this study focused on the first two levels of evaluation, the authors have set up the questionnaire to tackle Level 3 preliminarily. As Kirkpatrick suggests evaluators should survey and interview one or more of the following to evaluate behavior: trainees, their immediate supervisor, their subordinates. In post-graduate programs such as medicine, a wealth of information regarding behavioral change can be gleaned, not only from immediate supervisors and subordinates, but from colleagues and peers themselves. This critical peer evaluation of behavior is indispensable in the practice of medicine, and is a key point in the Level 3 Evaluation of this study.

Over 90% of students rated their peers '4' or '5' in a positive modified Likert Scale (Figure 5). Assessments mean that most students recognize behavioral changes regarding interpersonal relationships, and empathy and deep respect for the human body. Definitive behavioral changes determined from other sources (supervisors, subordinates) and upon return of the LEAP participants to their previous academic environment can be explored in the future.

CONCLUSION

General reactions from students to the Learning Enhancement in Anatomy Program were consistently positive, with a steady majority of the students rating '5' or a 'Strongly Agree' to positive statements regarding the program. Students gave higher ratings for more traditional anatomic teaching materials, such as cadavers, formalinized specimens, and bones over novel materials, such as virtual dissection tables and virtual-reality headsets. There were no consistent preferences of students across teaching methods, as each method had its strengths and weaknesses evident in the rating. Knowledge and skills acquisition were evident in the significant increase in post-test scores in almost all modules. Short-term behavioral evaluation by peers was also positive. The LEAP is a worthwhile endeavor which garnered overwhelmingly positive reactions and resulted in a significant improvement in test scores. Future studies are necessary to fine-tune teaching and learning in a blended learning environment. Interactions and mismatches between student reactions and student scores may be explored in further studies. An externally validated and rigorous assessment questionnaire may also be employed in the future.

Acknowledgments

We are thankful to our laboratory technicians for the daily preparation of all the materials used in the program. We also wish to express our appreciation to all the members of UPCM Class of 2026 who participated in this study for their time and willingness to share their experiences. Their contributions have been invaluable in helping us to understand the topic and draw meaningful conclusions.

Statement of Authorship

All authors certified fulfillment of ICMJE authorship criteria.

Author Disclosure

All authors declared no conflicts of interest.

Funding Source

None.

REFERENCES

1. Roy SF, Cecchini MJ. Implementing a structured digital-based online pathology curriculum for trainees at the time of COVID-19. J Clin Pathol. 2020 Aug;73(8):444. doi: 10.1136/jclinpath-2020-206682. PMID: 32366598.
2. Singal A, Bansal A, Chaudhary P. Cadaverless anatomy: Darkness in the times of pandemic COVID-19. Morphologie. 2020 Sep;104(346): 147-50. doi: 10.1016/j.morpho.2020.05.003. PMID: 32518047; PMCID: PMC7254017.
3. McLachlan JC. New path for teaching anatomy: living anatomy and medical imaging vs. dissection. Anat Rec B New Anat. 2004 Nov;281(1):4-5. doi: 10.1002/ar.b.20040. PMID: 15558778.
4. Ghosh SK. Cadaveric dissection as an educational tool for anatomical sciences in the 21st century. Anat Sci Educ. 2017 Jun;10(3): 286-99. doi: 10.1002/ase.1649. PMID: 27574911.
5. Moore NA. To dissect or not to dissect? Anat Rec. 1998 Feb;253(1): 8-9. doi: 10.1002/(SICI)1097-0185(199802)253:1<8::AID-AR6>3.0.CO;2-Z. PMID: 9556018.
6. Tyler RW, Gagné RM, Scriven M. Perspectives of Curriculum Evaluation. [Internet].1967 [cited 2022 Nov 15]. Available from: https://books.google.com/books/about/Perspectives_of_Curriculum_Evaluation.html?id=zasjtAEACAAJ
7. Mohanna K, Wall, D, Cottrell, E, Chambers, R. Teaching Made Easy: A Manual for Health Professionals, 3rd ed. London: CRC Press; 2016. p.264.
8. Kirkpatrick D. Evaluating Training Programs The Four Levels [Internet]. 1998 [cited 2022 Nov 15]. Available from: https://www.scirp.org/(S(lz5mqp453edsnp55rrgjct55.))/reference/referencespapers.aspx?referenceid=2233869
9. UPCM Web Portal College of Medicine - University of the Philippines Manila [Internet]. College of Medicine - University of the Philippines Manila. [cited 2022 Nov 22]. Available from: https://cm.upm.edu.ph/p/admissions-primer/
10. Shoemaker BJE. Integrative education: A curriculum for the twenty first century. Eugene, Or: Oregon School Study Council. 1989 Oct;33(2).
11. McLachlan JC, Bligh J, Bradley P, Searle J. Teaching anatomy without cadavers. Med Educ. 2004 Apr;38(4):418-24. doi: 10.1046/j.13652923.2004.01795.x. PMID: 15025643.
12. Franchi T. The impact of the COVID-19 pandemic on current anatomy education and future careers: a student's perspective. Anat Sci Educ. 2020 May;13(3):312-5. doi: 10.1002/ase.1966. PMID: 32301588; PMCID: PMC7262338.
13. Yun YH, Jo DH, Jeon SK, Kwon HY, Jeon YM, Shin DH, et al. The impact of the modified schedules of anatomy education on students' performance and satisfaction: Responding to COVID-19 pandemic in South Korea. PLoS One. 2022 Apr 11;17(4):e0266426. doi: 10.1371/journal.pone.0266426. PMID: 35404971 PMCID: PMC9000102.

14. Mogali SR. Education: Anatomy teaching during COVID-19 pandemic [Internet]. Lee Kong Chian School of Medicine [cited 2022 Nov 15]. Available from: https://www.ntu.edu.sg/medicine/newsevents/magazines-and-newsletters/the-lkcmedicine-october-2020/anatomy-teaching-during-covid-19-pandemic
15. Kim D, Yoo H, Lee YM, Rhyu IJ. Lessons from cadaver dissection during the COVID-19 pandemic. J Korean Med Sci. 2021 Jul;36(26): e188. doi: 10.3346/jkms.2021.36.e188. PMID: 34227264; PMCID: PMC8258239.
16. El-Naggar MM, Ageely H, Salih MA, Dawoud H, Milaat WA. Developing an integrated organ/system curriculum with community-orientation for a new medical college in Jazan, Saudi Arabia. J Family Community Med. 2007 Sep;14(3):127-36. PMID: 23012158; PMCID: PMC3410155.
17. Dubin B. Innovative curriculum prepares medical students for a lifetime of learning and patient care. Mo Med. 2016 May-Jun; 113(3):170-3. PMID: 27443039; PMCID: PMC6140046.
18. Mathiowetz V, Yu CH, Quake-Rapp C. Comparison of a gross anatomy laboratory to online anatomy software for teaching anatomy. Anat Sci Educ. 2016 Jan-Feb;9(1):52-9. doi: 10.1002/ase.1528. PMID: 25903289.
19. Harmon DJ, Attardi SM, Barremkala M, Bentley DC, Brown KM, Dennis JF, et al. An analysis of anatomy education before and during COVID-19: May-August 2020. Anat Sci Educ. 2021 Mar;14(2): 132-47. doi: 10.1002/ase.2051. PMID: 33387389.
20. Brinkman DJ, Keijsers CJPW, Tichelaar J, Richir MC, van Agtmael MA. Evaluating pharmacotherapy education: urgent need for hard outcomes. Br J Clin Pharmacol. 2016 May;81(5):1000-1. doi: 10.1111/bcp.12862. PMID: 26663464; PMCID: PMC4834598.
21. Ramaswamy R, Maio V, Diamond JJ, Talati AR, Hartmann CW, Arenson C, et al. Potentially inappropriate prescribing in elderly: assessing doctor knowledge, confidence and barriers. J Eval Clin Pract. 2011 Dec;17(6):1153-9. doi: 10.1111/j.1365-2753.2010.01494.x. PMID: 20630004.