BioLab: Exploring the Benefits of Gamifying Anatomy Education

Eric Baldwin

4th Year Project Report
Artificial Intelligence and Computer Science
School of Informatics
University of Edinburgh

2024

Abstract

Education has been falling behind the constant technological developments, staying put with traditional teaching materials in a classroom setting. Not fully utilising the advantage of the constantly innovative approaches in the technology sector. This has made it hard for education to avoid these traditional methods and look towards new approaches. Gamification has been shown to increase student motivation, engagement, and student knowledge retention. This study is motivated by the potential for gamification to make education more enjoyable for the user and increase the user's knowledge retention in studying anatomy.

The study was designed as a quasi-experimental cross-over setup to compare the efficacy of a developed gamified learning tool called 'BioLab' with the traditional classroom learning materials students would see. Being able to fully understand the impact gamification has on users in understanding the topic, the study assessed the results using Cohen's d, t-tests, and K-means clusters to analyse the improvement rates of the participants. In addition, the study received quantitative and qualitative feedback from the participants to measure their motivation and engagement.

The findings from this study indicated that a gamified learning tool can play a huge role in improving knowledge retention in students. Some students increased their scores more than others by using a gamified tool. This could indicate that students who struggle with traditional teaching methods could gain a lot from using gamified learning tools. The study provided evidence that gamification can significantly increase the study of anatomy.

Research Ethics Approval

This project obtained approval from the Informatics Research Ethics committee.
Ethics application number: 219854
Date when approval was obtained: 2024-01-30

The participant's information sheet and a consent form are in the appendix.

Declaration

I declare that this thesis was composed by myself, that the work contained herein is my own except where explicitly stated otherwise in the text, and that this work has not been submitted for any other degree or professional qualification except as specified.

(*Eric Baldwin*)

Acknowledgements

I would like to thank my supervisor, Professor James Garforth, for his mentorship throughout the past two years. His feedback on my drafts, discussions we had on concepts, and encouragement when challenges arose helped me complete this dissertation.

I would also like to thank my mum and my boyfriend, who have supported my mental health and wellbeing through this process and throughout my whole university experience. I wouldn't have been able to do this without them.

Table of Contents

Chapter 1

Introduction

1.1 Motivation

In modern education, we are at the crossroads of technology and traditional learning methodologies, giving rise to new and technologically advanced approaches that redefine how knowledge is acquired and engaged. Among these, gamification in education is a valuable tool in combating a lack of motivation and engagement, merging the fun of gaming with academic learning. This research was prompted by my interest in how gamification can make learning more enjoyable and effective, especially for a generation that has been brought up with technology.

The digital era has introduced us to learning that doesn't have to be inside a physical classroom, becoming more accessible, engaging, and personalised. This project is born out of a desire to explore the intersection of information technology and education, focusing on gamification to enhance educational outcomes.

Gamification represents a shift in education, promising to make learning more inclusive and adaptive to students' diverse needs. It can combine technology and educational pedagogy by using different tools to aid various aspects of education. It explores how game elements can transform education, making learning more than just a task that the student has to do but something enjoyable.

1.2 Project Goals and Research Questions

The overall goal of this project is to increase learning through gamification. I am doing this by utilising technology and gamification to aid the learning retention of anatomy-related topics. Students who have grown up with technology are more motivated and engaged by using new digital materials to help their education. [38] This indicates increasing knowledge retention in students studying medicine and other related topics to gain helpful knowledge and skills quicker than with traditional tools.

- **RQ1:** How does using a gamified learning app impact people's short-term retention of anatomical knowledge compared with traditional learning methods?

  - **RQ1.1:** To what extent does a gamified learning tool increase knowledge retention over traditional learning materials?
  - **RQ1.2:** Is there a difference in knowledge retention using gamification, which can be measured based on gender?
- **RQ2:** How does student engagement with anatomical content differ between gamified learning environments and traditional paper study?
- **RQ3:** Does the integration of gamification into anatomy education improve student motivation?

1.3 Theoretical Framework

Self-Determination Theory (SDT) Self-Determination Theory [31] suggests that motivation comes from the need for ”autonomy, competence and relatedness”. This can be used with gamification to create an environment in which motivation comes from users feeling in control. Gamification can allow users to control every aspect of the game at their own pace; they can also master the game, fulfilling the need for ”competence”. Gamified learning platforms allow users to pick their path and complete tasks they want at their own pace, allowing them to complete challenges that match their skill level. Using the theoretical framework of SDT highlights how gamification can engage and motivate users.

ARCS Model of Motivational Design The ARCS Model of Motivation Design stands for attention, relevance, confidence, and satisfaction. Gamification can satisfy all of these feelings and be used as a framework to create learning materials and tools to improve users’ motivation.

1.4 Current Problem

Biology education has not been able to keep up with technology changing [19]. Instead, it has relied on traditional teaching methods. Traditional teaching materials are less inclusive and less accessible than digital teaching materials and methods. [46, 17]

In this project, I will increase the efficacy of learning anatomy by increasing students’ motivation and engagement and making the learning tools more accessible to users. Traditional materials are less accessible. For example, you must be in the classroom to hear and see what the teacher is teaching. In general, gamification and digitalisation of education have been used to combat this.

However, digital education does not come without problems. The digital divide and gender inclusivity are essential aspects that must be considered when producing learning materials. Gamification must address these issues to become a transformative educational tool.

This project aims to address these challenges and clarify the framework for effective implementation. It seeks to provide a comprehensive understanding of how gamification can be tailored to the individual and how this can improve the student's motivation and engagement.

1.5 Structure of Dissertation

This dissertation is split into four different chapters:

Chapter One: Explains the motivation for the learning tool being created, the research questions which are trying to be answered, and the current problems with digital education and gamification.

Chapter Two: I look at the background of Biology education and how technology has been integrated with education. I will also present the educational pathology I aim to use in my project. This section also includes sections on gender in gamification and different methods in which IT tools are effective in education.

Chapter Three: In this chapter, I set out the research design, explaining what type of research will occur. Who will participate, and what they will be doing? In addition, I will lay out the framework of what 'BioLab' is and how different aspects of the gamification tool can be used to aid the retention of knowledge from the user. I will then explain the data I'll collect in the study and how it will be used in the analysis.

Chapter Four: In this chapter, I will set out the results collected from the study participants. These results will then be used to prove why traditional learning methods are less effective than gamification in education.

Chapter Five: Discuss the results and how they answer the research questions. I will then also speak about the study's limitations and any potential improvements that can be made.

Chapter 2

Background

2.1 IT in education

As technology has changed, education has changed alongside it. The internet and the downsizing of computers have enabled a whole new way of teaching. New ways of teaching through video format and gamification have taken centre stage in the learning of 'Gen Z' [25]. A whole new generation is being brought up in the digital age. Due to this educators have had to take a different approach to teaching. Not only has technology increased learning efficiency [3] using tools such as digital whiteboards and computers, but it has also increased the capacity of education and enabled people previously unable to access education to access educational resources, resulting in a more inclusive and equal learning environment. [35] In summary, incorporating technology into education is not just a trend but a fundamental shift in how teaching occurs and knowledge is acquired.

Gamification is a relatively new way in which learning institutions can teach students. It harnesses game design to engage and motivate students and deliver on the core curriculum. The syllabus comprehensively and simultaneously employs strategies that support a deeper and more nuanced understanding of a wide range of subjects. Focusing on different methodologies and learning outcomes is essential to understand the benefits of gamification. As with teaching, additional gamification attributes lead towards different set of results. A key gamification element is its flexibility in accommodating diverse teaching methodologies and learning outcomes. Other attributes, such as competition, teamwork, exploration, and problem-solving, can be added into the learning experience. This allows educators to tailor the teaching approach to align learners with specific learning techniques, making it a useful tool for addressing different learning preferences and outcomes. Some of these include:

1. Experience points, Levels, Leader- boards, Challenges, Badges (For Engagement and participation)
2. Points and prizes (For Motivation)

Integrating these features into a gamified learning experience led to higher engagement in the course. [4] One key aspect evident from this research paper was the involvement increase of over “511%” in replies to posts and an “817%” increase in new thread questions created about the course. [4] Students also noted how they found the course more engaging and motivating. Having these attributes would be of high importance at all levels of education. Another example of gamification in the classroom can be seen in secondary school science classrooms, where gamification leads to increased motivation and engagement. [23] A recent study focused on gamification techniques alongside augmented reality found increased student achievement and surpassing learning outcomes in courses it applied to. This study was conducted during the COVID-19 pandemic when virtual learning became the norm, and thus, student engagement became a high priority. The use of gamification played a pivotal role in enabling the students to continue learning while engaging with other students simultaneously. [24]

Gamification has since increased in popularity and is being used worldwide, not just in the classroom. Using technology to further gamification in education has meant a more level playing field regarding who has access to what. It is no longer the case where only a select few students can access all the resources. Technology is bringing down the wall in terms of a more equitable learning landscape, and no matter where you are from or how much money you have, technology has been able to ‘level the playing field’ [35]. Furthermore, gamification has enabled students to tailor their learning experience to their learning styles, allowing people to get more out of using gamification to learn. Technology has strived to make education more and more equitable but hasn’t always succeeded. This is evident from the ‘tech divide’ in education. The schools with the highest budget can also afford the best equipment. Gamification can help tackle this by offering a web experience that anyone can access without needing the most up-to-date technology. Instead, the servers are running most of the background work, which means that the technology used to access the website doesn’t matter.

2.2 Gender roles in gamification

Historically, educational methods have not been looked at through the lens of gender inclusivity. However, analysing gender in gamification is essential to help us understand possible links between gender and how each individual learns. Various techniques and methods for teaching have proven to be more effective for different people.[14] For this reason, it is essential to remember this for gamification. One highlight of gamification is its ability to cater its learning environment to the learner. Hence, gamification is better equipped to teach people with vastly different learning styles than traditional teaching.

Gender roles are a specific topic that seems to come up a lot in the role of gamification. Gamification techniques found disparities in students’ learning outcomes based on gender. [14], [23], [16], [7] While Denden and Khan found differences towards the effectiveness of gamification based on gender and other personality differences, Chapman and Bressler did not. In the Denden study, 70% of the participants were male. This is not an accurate representation of the whole population, and as a result, this study has low external validity. While Bressler’s study focused on the gamification of science in secondary school pupils, it found that gender didn’t play a part in the learning experience and the outcome of student’s enjoyment and engagement of the course. This study focused on a small group of students, only 67. However, there was a relatively equitable split between genders. It again found that gender didn't play a role, as seen in other papers. [15] As with these papers, the younger the students are, gender is seen to have a lower impact on students' engagement. This leads us to the hypothesis that the younger the student is introduced to gamification, the less gender plays a role in the efficacy of such a tool.

We can see the potential for bias in these studies in their participant selection. The gender imbalance, paired with the lack of participants, makes it difficult to convey meaning without bias accurately. In addition, many factors, such as socioeconomic factors, previous gaming experience, and confidence in technology, were not included, which could mean the results have low external validity.

One significant conclusion that we can draw from these studies is that age is a vital variable. We can see that the younger the student is, the lower the gender's impact on their engagement levels. This highlights the early stages as a critical factor in bringing gamification into education.

2.3 IT and biology in education

As the digital age has progressed, so has the education of biology. This integration has resulted in significant advancements, from incorporating Virtual Reality (VR) and Augmented Reality (AR) to using instructional videos and gamified learning experiences. Recent advancements in VR have made it possible for medical students to visualise their learning differently. Visualising different aspects of human anatomy without having a subject in front of them leads to a more interactive and accessible education.[42]

One particular use case for this is when medical professionals use VR for surgical training and planning. This technology has enabled a much cheaper and more reliable way of planning for surgery. Previously, doctors would have to have silicone moulds of bodies made if they wanted to test a set of skills but can now use VR repeatedly. Studies conducted with medical professionals using VR to train found they performed up to "43% faster" [28] than doctors who didn't train using VR. This highlights how using technologies such as VR has enabled faster medical knowledge uptake.[42]

As well as VR, many other aspects of technology have been used to teach students in secondary schools. From collaborative laboratories [36] to multimedia resources. [21] All increase students' motivation to take biology courses. Increasing students' motivation would impact their achievement levels. [20] Using multimedia tools such as YouTube has enabled students to visualise biology on a cellular level thanks to animations, video and photos taken from microscopes, which have been put onto multimedia platforms. All of this has enabled students from around the world to experience and learn together.

Biology has recently embraced gamification towards its teaching. By allowing for games to be created that enable students to learn more about biology, these students will be more motivated and engaged with this subject. By including gamification in biology, we can see a clear trend of increased motivation and engagement about the subject and improved academic performance, which can be gained from these learning tools. [19]

2.4 Assessing the effectiveness of IT in education

There are many different ways to measure the effectiveness of IT in education, as well as the main options such as qualitative assessments and attendance. Such metrics include:

Metrics	Outcome	Cite
Attendance and Punctuality	Higher attendance and punctuality throughout the academic year lead to higher student achievement. This is due to the student being able to absorb more information and engage more in the learning environment.	[18, 13, 44]
Participation in Class	Class and learning materials such as a gamified learning experience has shown that student motivation increases.	[5], [2]
Homework and Assignment Completion	Students' engagement with material throughout the course show how they understand the topics.	[41], [40], [30]
Using Learning Resources	Materials such as videos and games shows heightened learning motivation from students.	[38]
Peer Collaboration	The collaboration between students highlights a greater understanding of the content being taught. It emphasises engagement and teamwork. "When students cooperate in small groups of between four and eight people, they are more creative and have better outcomes than when they work alone."	[48], [43], [1], [31]
Self-Assessment	Observing students' engagement with their assessments and subsequent actions regarding different course components provides valuable insights into their learning needs and strategies for achieving course objectives.	[47], [49], [39]
Student Surveys	Surveys that specifically target student engagement, motivation, and satisfaction with the learning experience.	
Peer Evaluation	Peer Evaluation, teamwork, and contributions to group activities. "A student learns something and thinks she understands it. She is tested on that learning and finds some missing or faulty knowledge, then she is helped to correct it."	[31], [26], [34]

Table 2.1: Metrics for academic success

2.5 Overview of interactive educational tools

Learning has evolved rapidly alongside technology. One of these improvements is the ability to learn interactively. Technology allows us to visualise and test different strategies to reach our final learning goal. Using interactivity builds on the learner's ambition and feeling of achievement. Once they have learnt something new, this feeling motivates them to learn and achieve more. [31]

Various interactive tools are used to aid education. These vary in how they are meant to be used. The use case can affect how people learn; different subjects use different tools to aid learning.

Web based tools One of the leading and most common tools is web-based, which means that it is easily accessible and you don't need expensive technology. All schools in the UK are equipped with computers, which means that the school wouldn't have to go out and purchase iPads or costly technology to use web-based tools. Web-based tools are also easily updated and can be changed remotely rather than directly on the device. This flexibility means the data can be kept updated with the current curriculum. Skull game [27] is a gamified web-based tool that teaches students about anatomy. It uses a treasure hunt style of game to keep people engaged while learning. Other tools such as 'Mathletics' [32] are web-based tools to teach mathematics. They are used in a similar environment to the skull game, where students can use them at home and in the classroom to learn about these different subjects. Web-based tools also have the highest global outreach. Many countries cannot access tablets and virtual reality technology in the classroom. While many countries don't have computer access, it is much more widely used than other technology. [11]

Application-based learning tools These include apps like Duolingo, which aims to help people worldwide learn languages. Duolingo uses gamification techniques to aid in this. Adaptive learning technology is used extensively in this case [33], personalising the learning experience and questions to better cater to the individual's needs. Some apps have a web-based tool, such as Khan Academy [32], which caters to kids learning to code and offers games, books, and other activities that allow for an interactive learning tool.

The educational impact that these interactive tools have is transformative. Allowing for personalised self-paced learning for each student has not been possible with traditional teacher-student classrooms. However, with the help of technology and gamification, interactive tools have become widespread, allowing students and others who want to learn more about subjects to do so at their own pace. Improving the understanding and motivation towards a healthy learning attitude has improved knowledge retention. [22] This, in turn, increases student achievements in the classroom and allows for a better, more well-rounded education.

Chapter 3

Methodologies

3.1 Introduction

In this section, I will explain the strategy that will enable us to understand how well the gamified learning environment ‘BioLab’ performs against more established learning tools currently available. The primary aim of this section is to detail the rationale behind the BioLab’s design and workings. I have compared BioLab to ‘Traditional learning materials’ in today’s classroom environment.

I will also lay out the foundations for ’BioLab’, a tool that will be used to answer the research questions previously posed. I have also designed a study using participants to research the project’s efficacy. I have developed a quasi-experimental study because it can negate the current knowledge on the topic. This is because some of the chosen participants have a background in biology, and to avoid skewing the results. This experiment will enable these participants to act as their control. The research design focuses on comparing the use of both gamified experiences against current education tools to see the effect on gained knowledge from using each learning material. I have also taken ethical considerations to ensure participant consent and privacy, and a section will follow in this chapter dedicated to all the moral implications of this study.

3.2 Research design and rationale

Research design In this study, I used a Quasi-experimental within-subjects, cross-over design. This study allows participants to use both the gamified learning tools and traditional textbook-style learning guides. I split the participants into two groups and gave Group ’A’ the textbook-style learning guides, while Group ’B’ got the gamified learning tool first. Then, once they have spent precisely twenty minutes learning, they will take a small test to gauge their current knowledge of human anatomy. Once they have finished this test, they will swap learning materials and spend another twenty minutes using these materials to learn about the human anatomy. Once this time has finished, I will again test them to gauge their knowledge in the area to see if it has increased and by how much.

I chose a quasi-experimental design because it can conclude whether the gamified app can cause an improvement in learning outcomes. It does this by controlling individual learning differences from prior knowledge and isolating the potential causal effect of the gamified learning tool. Within subjects reduces the variability caused by people's prior knowledge in the area they are being tested on. As the subjects are exposed to both the gamified learning experience and the textbook style, their current knowledge wouldn't hinder any research outcomes. Cross-over helps to mitigate the effect of potential confounding variables like gender, age, and discipline.

Rationale To ensure that the comparison of both of the learning tools is the same, we make sure that the conditions of the study are constant. This is done through the use of identical test questions for both groups. This means we can attribute any observed improvement in scores to the learning materials used. The random assignment of the groups ensures that there is no selection bias in the groups, enhancing the validity of the study's results. Once the results have been received, statistical analysis, such as t-tests and Choen's d, will be performed. These are used to measure and evaluate the significance and importance of the differences in the learning outcomes.

There will be some limitations to this study that will not be able to be fixed due to the time scale of the test. As each participant will have to spend an hour using both sets of materials, there is a potential for fatigue among the participants. This could hinder their results in the second test as they get bored from the repetitive nature of the study. There is also the chance that there are potential learning effects from repeated exposure to materials, meaning that the participant's knowledge will hypothetically increase over time just by being exposed to any set of materials on the subject. However, this is catered for as we measure the knowledge increase from tests one to two. We can still measure the increased performance from BioLab versus traditional learning materials.

The study can go beyond statistical significance to measure the efficacy of the learning materials. It will also ask participants for feedback on both sets of learning materials to gain robust evidence of the effectiveness of each tool to increase motivation and engagement.

3.3 Participants

3.3.1 Selection Criteria

The study's participants were selected to align with the overall objective of evaluating the efficacy and motivational promise of the gamified learning environment, 'BioLab'. Given the focus on education at the university level, it was essential to the study that the participants primarily consisted of university-level students who are used to learning at such a level. A select group of educators and professionals supported these participants. Some educators specialise in biology to ensure a broad perspective on the learning experience across different user groups.

Primary Participant Group The main group of participants in the study were university students. These included students currently studying an undergraduate or master's course from various degrees. The choice of this group as the primary participant group is because of the assumption that these individuals are the most representative of the end users of 'BioLab'.

Age Range: The mean age of the student participants was around 21 years, accurately reflecting the age group of undergraduate students at university. As these university students learned during the coronavirus pandemic, it was assumed that they were familiar with both traditional and digital teaching methods. They provided the study with a balanced perspective on the gamified learning experience.

Additional Participant Group The study needed educators and professionals to compare and contrast the two methods to ensure that the goal of increasing academic achievement was met, and including educators in the evaluation meant that the feedback received from this group of participants would be able to identify better differences between the benefits of gamified learning and traditional learning. It also meant that we could study the appeal of gamified learning from participants less versed in digital learning techniques.

Background Knowledge Some educators have a biology background, meaning they have a critical understanding of the content in BioLab. The professional insights educators from all backgrounds will bring to assess the content's accuracy, engagement level, and effectiveness are valuable.

Exclusion Criteria Participants who have previously learned anatomy using gamified learning environments were not included in this study, which ensured that it was the participants' first time using this learning environment.

I spoke with a biology educator during the curriculum development for 'BioLab'. They helped guide the curriculum that the users would learn. Using this professional knowledge in the area meant that the curriculum of BioLab closely matched what students would learn as they studied in university. Ensuring the efficacy of the curriculum was very important throughout the study as this gave the users a well-rounded education, starting with 'easier' to learn bones and moving on to gradually harder and more complicated bones. The educator who aided in the curriculum was not included in the study as they already had a reference to what they were expected to know.

3.3.2 Demographics

The study also ensured that demographic information, such as age, gender, field of study for student participants, and area of teaching for teacher participants, was collected. This data was secured in an online digital format before the start of the study. All of the data was secure and anonymous for the privacy of each participant.

An essential aspect of our study was to focus on participants from various demographics. This is to ensure the efficacy of the gamified learning materials. It also allows for a more in-depth analysis of the study's results. For example, comparing the difference between the sciences and a humanities subject could help us uncover a learning retention difference in a subject discipline. It enables us to identify any differences in the effectiveness of 'BioLab' for each group of students. Another possible correlation would be age and gamified learning experiences' efficacy.

3.4 Materials

3.4.1 Traditional Teaching Materials

This study will use traditional teaching materials and compare them with BioLab. Traditional teaching materials consist of materials used in a classroom teaching setting. This includes annotated diagrams that the teacher could run through and give handouts to students to learn in their own time.

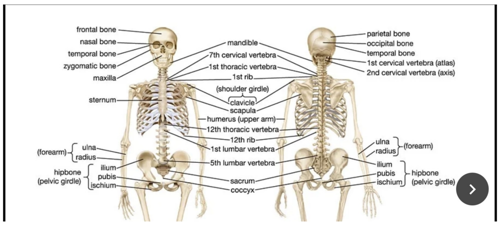


Figure 3.1: Annotated skeleton taken from Encyclopedia Britannica

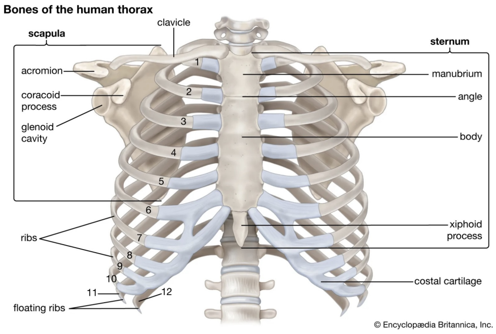


Figure 3.2: Annotated skeleton chest taken from Encyclopedia Britannica

The traditional teaching materials used in this study came from the Britannica website. [8, 9, 10] This was to ensure the accuracy of the control information being used for learning. It was also crucial as Britannica, one of the world's most famous scientific journals, is a 'standard' in today's textbooks. Therefore, it was the perfect candidate to be the traditional teaching materials.

The content from the Britannica website aligns with the overall objective of medical students' requirement to learn and memorise the bones in the human body. The content had annotations around each bone, allowing for complete chest and spine coverage. These clear and precise annotations meant that the diagrams printed three pages worth of diagrams.

Specific sections of the Britannica were taken out and printed for participants to learn; they were free to write on and highlight anything they required. The sections we focused on were the chest and spine; this is in to coordinate with the curriculum in 'BioLab'. The participants were guided through how the study expected them to use the annotated diagrams. They were given materials such as highlights, pens, pencils, and some spare paper to allow them to learn just as they would with traditional textbook materials.

By leveraging the comprehensive and authoritative content available through Britannica, we could establish a solid foundation for discussing and comparing the efficacy of the two learning tools.

3.4.2 BioLab

BioLab integrates anatomy topics into a gamified learning experience, increasing engagement and motivation for learning anatomy. The game aims to use different types of pathology to stimulate the users into learning more than they would with traditional learning materials.

3.4.2.1 3D Models

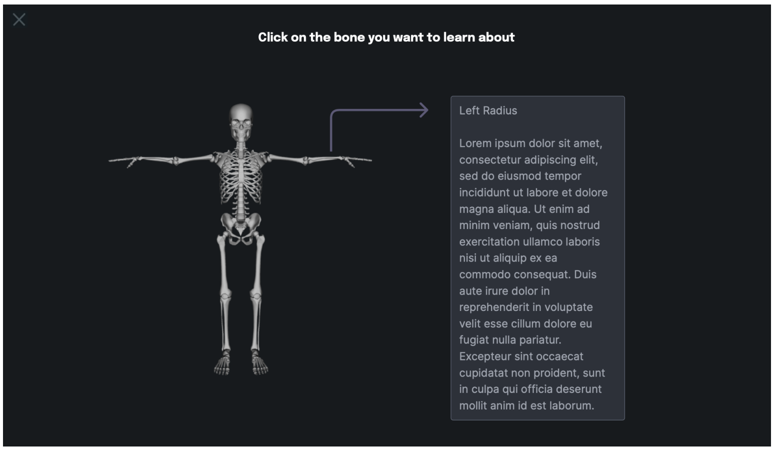


Figure 3.3: Mockup design of BioLab 3D Model

One of the main features of BioLab is the Skeletal 3D Models. These models will be used everywhere in the game, from the quizzes to the AR and exploring the skeleton section.

By providing full, anatomically correct models of the skeletal system, BioLab will be able to provide a comprehensive learning tool. Allowing users to explore the skeleton at their own pace. The educational benefit of enabling the users to manipulate and examine the skeletal structures themselves allows for a greater understanding. [29, 45] It makes the students feel like they managed to find this out by themselves. This reinforces the learning pathology about self-guided learning [12], strengthening their understanding and supporting their motivation to learn more.

The 3D models were also made to have interactive capabilities, such as zooming in and rotating. They allow the user to have a complete and total view of every human body bone from every angle. Clicking on each bone highlights it and enables the user to find more information about it.

3.4.2.2 Augmented Reality

Augmented Reality will also be added to BioLab Mobile. This will allow the user's phone to view the 3D model of the skeleton in the same room as them. They can move around it and inspect the skeleton from all angles. They are visualising the skeleton in the real-world context, which enhances the spatial understanding of the skeleton and increases the user's engagement due to the trivial nature of using AR to view a skeleton in your room.

The final result will allow the user to interact with the skeleton easily. The user can place the skeleton anywhere on a flat surface. This means you can have a life-sized skeleton right before you with just a click of a button. In addition, the user can rotate the skeleton without having to move around the room; this increases the accessibility of the app as it means that users with decreased mobility can still use this aspect of the game. Finally, users will be encouraged to click on different parts of the skeleton, allowing them to find out the name of that bone and highlight the bone so they know which bone corresponds to which. While the bone is highlighted, it can still interact with the skeleton in all previous ways.

This AR feature will only be available on the iOS version of the game, but the web version will have an interactive 3D model where you can use your mouse to interact in the same way as you would in AR.

3.4.2.3 Quiz's

BioLab will allow users to test their knowledge using its specific body part quizzes. These quizzes are tailored to each player and start with much more straightforward questions, designed to get more challenging over time. Helping the user get acquainted with easier bones and build a foundation. As previously mentioned, this was coordinated in tandem with a Biology teacher, enabling the matching of the current university curriculum. The quizzes constantly return to previous questions to reinforce the user's learning. The user can increase the difficulty by going to the next level, which brings out new sets of questions and new ways of answering these questions.

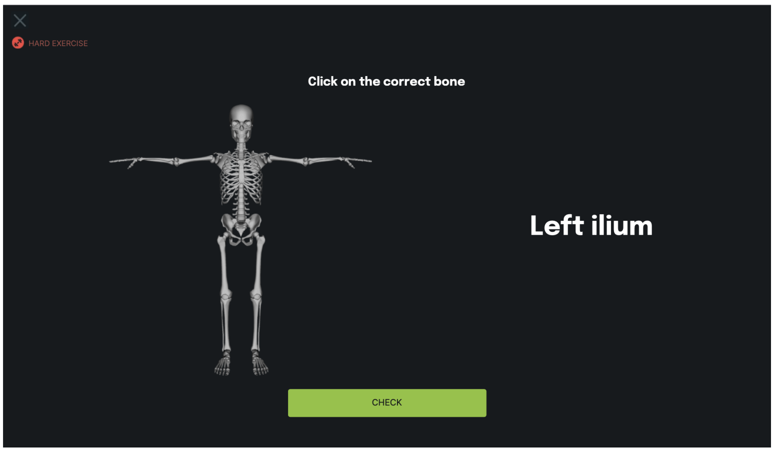


Figure 3.4: Mockup design of BioLab Questions - Using 3D Model

Sets of Questions:

- Multiple Choice allows the user to see possible answers to the question. For example, in level one, the quiz shows a 3D model of the skeleton with a bone highlighted. The users must correctly select the answer out of four possible options. This will help the user form an understanding of the concept that is being shown in the gamified learning experience. [31] These questions are beneficial as they allow the users to have immediate feedback on their answers.

- Fill in the blank - This set of questions allows the user to type in the missing word. The game will highlight the bone and give the words, if any, beside the disappeared word. Allowing the user to make one mistake, the game will notify them that they have made a mistake and encourage them to try again. The algorithm will also check if they have mistyped the word and give them a prompt saying they are close to the correct answer.
- What's this bone - Much like in 'Fill in the blank', this question type will prompt the user for input. However, there will be no hints, and the user must type in the whole bone name. This reinforces the learning that has taken place through the quiz, as it is the last question in the quiz.
- Select the bone - One of the more challenging levels is selecting the bone. This is when the user is given the name of the bone, and they have to use the 3D model to determine the bone they think it is and submit it. Users have lives that they lose if they offer the wrong bone to make it more competitive.

All levels in the quiz will provide immediate feedback. These include incorrect answers, correct answers, and close answers. Feedback is also given at the end of each quiz, summarising how the user did on that level, including their score and how many stars they got. Stars are used as a mechanism to drive users to try and beat their previous score. They will obtain three stars for full marks, two stars for 75% and one star for 50%. These 'Badges' increase engagement and participation in the game.

3.4.2.4 Database integration

Users' progression and scores are all collected inside a Heroku database. This will enable the keeping of a record of all users and their progression statistics. Keeping all this information means we can display usage information for the users and allow them to see their current high score for each level and how they can improve. Seeing previous scores helps motivate the users and increases their engagement in the game. [31] BioLab leverages the desire for achievement and recognition to help build the user's knowledge. The game allows for elements such as 'Stars' to be given to users when they score highly. This will enable them to see their current star total and try to increase their number of stars. Keeping these statistics can allow for future features such as tracking from teachers to see students in their class and track their progress as they take quizzes and obtain stars.

The server's data is encrypted and cannot be accessed without two-factor authentication from Heroku and MySQL. Even if you could access this, the data would be encrypted to be unreadable. This keeps the user's data secure without worrying about any third party seeing sensitive information.

3.4.2.5 Multiplayer

The game will also make use of a multiplayer feature. This feature promotes collaborative learning [48, 43, 1, 31], as the users play against each other as they race to answer as many questions correctly in a given amount of time. The two players will be aware of the other players' scores in real-time as they are updated because they are integrated with Pusher. Pusher notifies the user when the other player has joined, their score, and when they have finished. Pusher is integrated with the Heroku database using PHP calls back and forth. The questions are randomly created when both users join the game.

The randomised questions are generated when the players join to ensure that each multiplayer game is unique and a fair challenge to both users. This prevents rote learning and encourages players to fully understand the material rather than just memorising the answers to specific questions. The social aspect of a multiplayer game benefits the users by promoting both a competitive and collaborative learning experience for them. Social learning environments have many benefits created by the multiplayer feature in BioLab. Learning alongside and against peers increases users' engagement, and their motivation to win increases the retention of anatomical knowledge.

3.5 Learning sessions

The learning sessions are designed to evaluate BioLab's efficacy while negating any possible effect of participants' prior knowledge of anatomy. The session is structured to ensure that the participants must use both traditional learning materials and BioLab and then are assessed on their knowledge retention and understanding.

Introduction and Setup: The session starts with a brief introduction where the participants are asked to fill out the forms required to participate. Then, the procedure is explained to them, such as what they will be learning and what materials are provided. In addition, a short demo of the app is given to the user so they understand how to use it. This included a tour of the app's features and how to access the quizzes. Ensuring the participants were comfortable navigating the app was very important when comparing the two materials. I didn't want the participants to be hindered by not understanding how the application worked.

The instructions included the learning objectives that the participants were expected to learn. In this study, it was the names of the chest and spine bones.

First Learning Phase (20 Minutes): Half of the participants will be learning using traditional paper materials in this phase. They are encouraged to take notes and explore the content thoroughly. The paper materials come from the Britannica website and are printed out for the participants to learn and retain as much knowledge as possible.

While this is taking place, the other half of the participants will enter into BioLab, where they can explore the features and quizzes to learn about anatomy. They will use BioLab to try to retain as many bone names as possible.

First Test: After the 20 minutes, the participants all take a test. To ensure the study is fair, the test must be the same between the two groups. The test covers aspects of the human chest and spine, representing what the participants have been learning for the past twenty minutes. The tests last 10 minutes, but they can finish earlier if they wish. The test is designed to assess their retention of knowledge around anatomy.

Second Learning Phase (20 minutes): After completing the test, participants switch the learning materials they were on previously. This cross-over design ensures that the participants can experience all learning materials to measure each method's efficacy.

Second Test: Once the twenty minutes are up, the participants again take another test assessing their retention of anatomy knowledge. Again, we have to ensure that both of the tests remain the same between the two groups to ensure that the study is fair between the two groups.

Environment Setup

BioLab is available through the web, meaning users must access the internet through their computer. While the participant takes part in the study, support is offered at all times if they have any technical questions relating to BioLab or how the test works. The participants are also welcome to take breaks between the learning and testing phases to ensure they are not exhausted by the questions and the topic.

3.6 Data Analysis

I initiated the analysis with t-tests, which served as the primary tool for detecting immediate differences in test scores for the two learning methods, 'BioLab' and traditional learning materials. This choice was due to the paired nature of the data, with each participant experiencing both materials. The paired t-test is well suited to this study, and it enables us to see the direct effect of each learning method on the same individuals, thereby controlling for inter-participant variability.

In addition to using t-tests, I used 'Cohen's d' for each participant's data. This involved calculating the difference between the mean of the two tests (one test for BioLab and one for the traditional learning materials) and then dividing the difference by the standard deviation of the scores across both conditions. This means I could understand the significance of the difference between the two learning materials and understand and quantify the effect of any gamified learning experience.

I also analysed the users' demographics and compared this to the efficacy of the learning materials using the t-tests and Cohen's d. This will give me the interactions between variables such as age, participant background subject, and previous learning experience. This analysis helped explore whether the learning materials' effectiveness varied over different groups of people and what group saw the most significant increase in test results due to 'BioLab'.

Using all these statistical methods meant that I could assess the results comprehensively. As discussed in the next section of this thesis, I evaluated the efficacy of gamified learning compared to traditional methods. I painted a detailed picture of the landscape of the learning tools, aided by grouping participants into sub-groups of their chosen field or speciality. Due to all of this, I was able to have a deep understanding of how and when using a gamified learning experience surpasses traditional methods of learning.

3.7 Control for confounding variables

In the study, I wanted to ensure that the findings were not biased in any way. To this end, I implemented strategies for controlling confounding variables. Firstly, randomly assigning participants to either the BioLab or the traditional materials group mitigated their prior abilities around anatomy. In addition, the cross-over type of study indicated that each participant was their own control. They would act as their baseline from the first test, and we could see the improvements made from the study materials used.

As the study wants to measure many aspects of the gamified learning materials, the participants take a questionnaire about what they currently study/specialise in at the start of the participation. And their prior exposure to digital learning. This data was used to create subgroups within the study, which could be used to examine the effects of these variables on knowledge retention. To ensure that everything was as unbiased as possible, all the learning materials were standardised across the two groups, and all participants had access to all the same features.

While the participants were studying and taking the test, they were encouraged to mitigate environmental factors. Noise and distractions were urged to be minimal throughout the study for all participants.

3.8 Ethical considerations

When the participants agreed to participate in the study, they received a comprehensive consent form detailing what the study was for and how it would be used. The form also described their rights as participants regarding their data and privacy.

The participants were informed that it was a voluntary study and they could withdraw themselves at any point. If they chose to do this, all data relating to themselves would be removed from the study.

To protect the privacy of all participants, the data collected for the study was anonymised and not linked in any way to any of the participants. The data was stored in an encrypted manner that was compliant with GDPR.

In addition to these considerations, special care was taken to ensure the well-being of all participants using the digital gamified learning experience. To ensure that there was no distress from the usage of 'BioLab,' there was a feedback form to submit any concerns from any participants.

3.9 Limitations

While this study provides valuable insights into the efficacy of gamified learning through BioLab compared to traditional methods, it has limitations. Although the sample size is sufficient for detecting significant differences within our study population, it may limit the generalisability of our findings to broader populations or different educational contexts.

The short-term nature of the study restricts our understanding of long-term retention of anatomical knowledge post-intervention. Future research should consider longitudinal designs to assess the durability of learning gains.

Additionally, the study's reliance on self-reported measures of engagement and motivation may introduce bias, as participants' perceptions may not fully reflect their actual learning experiences. Objective measures of engagement, such as time spent on tasks or interaction rates with the learning materials, could provide a more nuanced understanding.

Finally, the study's setting, predominantly virtual due to the constraints of the current educational landscape, may influence the outcomes. The physical learning environment and its interaction with digital tools warrant further exploration."

3.10 Summary

In this chapter, I explored my methodologies during the study. These methodologies were used to check the effectiveness, engagement, and motivation of the gamified learning environment 'BioLab'. I then compare these results with the results from traditional learning methods when learning about anatomical knowledge. The study used a quasi-experimental within-subjects, cross-over design to ensure that bias and previous knowledge of any participant related to anatomy did not cause any bias. Participants included university students, as these students have had the most exposure to online materials and teachers with and without a background in biology. This ensured a broad range of perspectives on the gamified learning tool.

The study's idea and design were to answer the question about the efficacy of retaining anatomical knowledge. Participants were randomly assigned to one of the two groups to either start with 'BioLab' or use traditional learning materials. These traditional learning materials came from the Encyclopaedia Britannica to enforce strict correctness and ascertain the actual effectiveness of these tools. The groups would then swap over to the other learning materials, with testing taking place between the swap to measure knowledge acquisition and retention.

'BioLab' has unique features such as 3D models, augmented reality (AR), quizzes, and a multiplayer component. These elements are strategically designed to enhance engagement and educational outcomes, making it a promising tool for anatomical learning. Learning sessions were structured around these materials, providing equal opportunity and exposure to both methods before testing.

I also acknowledge some study limitations and how they could be overcome. These included a small sample size because the study took participants over an hour to partake in. In addition, the short-term nature of the study meant that I couldn't measure the effect each learning material had on the long-term retention of information. Finally, the fact that the measures of engagement and motivation were self-reported meant that I was unable to back up these claims as well as could have been done without using, for example, eye tracking software.

Chapter 4

Results

4.1 BioLab

In this section, I plan to discuss the final result of BioLab's development. I will review the challenges I faced and the features that I implemented. I'll also discuss the features that could not be created from the original concept. Finally, I'll delve into the technical considerations made during BioLab's evolution.

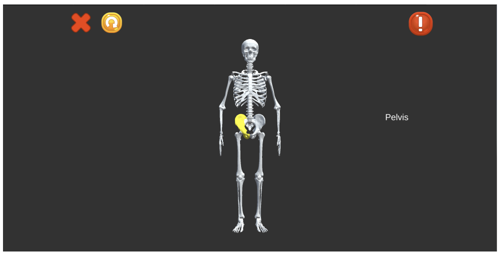


Figure 4.1: BioLab 3D human skeleton annotations

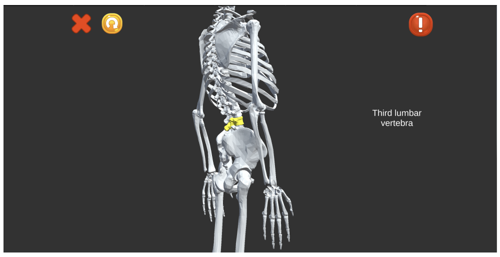


Figure 4.2: BioLab 3D human skeleton annotations - Highlighting the zoom and rotation capabilities

3D Model BioLab's interactive 3D model is central to its approach to creating a unique and interactive experience for users. This is why the 3D model was developed first. The ability to interact with the model by zooming in and rotating the skeleton around is trying to resemble a hands-on learning experience as best as possible.

The 3D model came from a high-fidelity object file [6]; this served as the base of the 3D skeleton. This file was then inputted into Blender, where I separated each bone. Then, the object file was added to the Unity game file, where I used C++ to label every bone in the human body as a separate object. This was a time-consuming process as there were over 200 bones to label individually. This only had to be done once to ensure each bone had an anatomically correct label. Features such as zooming in and rotation required intricate coding through C++ using the prefab of the skeleton to ensure that the positions of all of the bones would stay relative to the skeleton as a whole. To ensure that the user would be able to click on each of the bones, I had to create a raycasting script to detect where the user was clicking down on the skeleton; the raycasting was very important as the script wouldn't know the exact position of the skeleton due to previous rotations and zoom effects. This feature was particularly difficult to implement as the code would have to scan through the JSON file to find to correct name of the bone each time the user clicked on a new one. This information mapping to the 3D model results in a highly interactive model encouraging the user to explore and discover the human anatomy.

Quizzes The quiz feature added to BioLab would help build an educational framework for the users. I wanted to include the interactive 3D model in the questions so that the user could interact with the skeleton to find the correct bone and explore its positioning with respect to the whole skeleton.

The quizzes had differing difficulty levels to enable the user to progress at the rate that they felt comfortable. This builds foundational knowledge before moving on to harder and more intricate sections of the human anatomy. All aspects laid out previously in the methodologies came to fruition and allowed for a seamless integration of the model and all questions.

There were some technical challenges with the questions, such as accessing the JSON file to generate a random question for the user and then mapping the answer to the correct bone to highlight said bone successfully. The score had to be kept by the code, and then at the end of the quiz, the user could assess their final score. This score would be pushed onto the database using PHP and integrated into the C++ script.

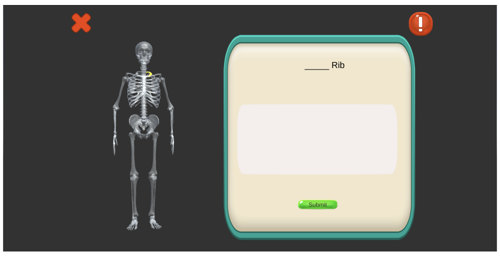


Figure 4.3: BioLab Quiz - Fill in the blank

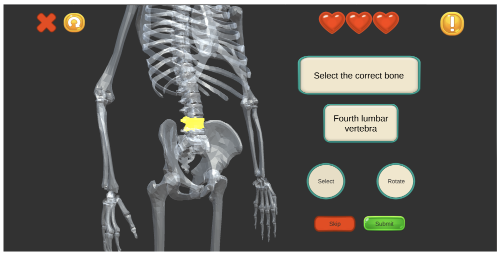


Figure 4.4: BioLab Quiz - Select correct bone

Augmented Reality I also developed an augmented reality aspect of BioLab; this was only available using the mobile app, as I utilised the phone's camera to place the skeleton into the room. User interactions and accessibility were very important during this implementation, allowing the user to rotate the skeleton without moving around. This offered a 360-degree view of the human skeleton inside your room.

The implementation process with the AR was very difficult. The technical work that went into it required a lot of C++ coding and raycasting technology. It was developed using Unity AR. The most challenging aspect of this function was detecting where the user was trying to place the skeleton, it would have to detect the floor and then raycast the touch input into where the user was tapping. This would place the skeleton in this position. However, it was very temperamental at first and would change the size of the skeleton to various sizes. Once I mapped the object file's origin onto the skeleton's feet, it became more easily manipulated.

The next challenge was detecting the difference between the user touching the skeleton to find the name of the bone and swiping to rotate the skeleton. To simplify this, I added a button on the side to enable the user to switch between the 'Rotate' and 'Select' modes. This allowed the user to switch between the two modes as they wanted. Allowing for a fully interactive and personal learning experience using BioLab.

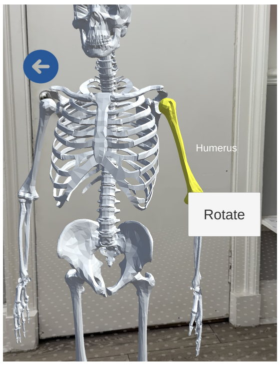


Figure 4.5: AR BioLab Mobile

Multiplayer I also completed a multiplayer game where users could match up against each other and compete in a race to answer as many questions as possible in a timed environment.

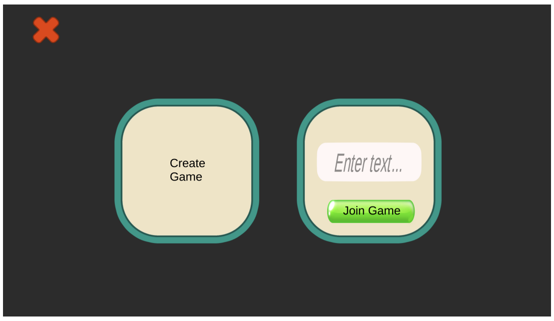


Figure 4.6: Multiplayer feature on BioLab

This feature led to many difficulties in creating, most notably enabling real-time communication between the two players, which was the most difficult. I used Pusher [37] to create this communication between the two players using a generated random game code. When the first player presses create a game, the code generates a game code, which is sent to the database using PHP. At the same time, a Pusher channel is created and subscribes to the channel using the same game code. It then waits for the second player to join. When the second player inputs the game code, an event in Pusher notifies the original player that they can start the game, enabling real-time communication syncing. As this is taking place a PHP script lets the Heroku database know that the game has been joined by both players and it can generate the game questions which will be sent to both of the players at the same time.

Using Pusher allowed for real-time communication between both players. Each player could see their score and their opponent's score in real-time. This feature played into education pathology around peer learning and collaboration. [31]

Database Another feature of BioLab that is not visible on the front end but is the most vital feature of the project is its database. The database stores the user's information, progression statistics, questions asked to users, teachers' information, and multiplayer games. Without this database, BioLab would not be able to function.

There were many difficulties creating the database at the start. I used Heroku to store the online information, and a MySQL database was attached. This enabled a web platform for users to use the game and allowed for the simultaneous storage of needed database information. Integrating the C++ scripts from the Unity game to call PHP scripts from the Heroku web platform proved troublesome. The scripts would have to be called at the correct instance, error checking would have to take place, and authentication in the C++ script would have to ensure they are accessing the correct PHP script and not requesting unnecessary information. These asynchronous calls to the database allowed large amounts of data to be sent to the user and back to the database, where it could be stored later.

This database was essential in developing BioLab, enabling personalised learning experiences for users. It allows users to access their progress and continue where they left off.

4.2 Participants Data Gathered

Name	First Test Score		Second Test Score	
	BioLab	**Traditional**	**BioLab**	**Traditional**
Mean	6.45	4.25	7.33	7.00
Standard Deviation	1.37	2.05	1.15	1.84
Cohen's D	1.25		0.22	
T-statistic	3.06		0.51	
P-value	0.00643		0.61390	
Statistical Measure	**% Improvement**			
	Traditional		**BioLab**	
Mean % Improvement	6.58		127.22	
Standard Deviation % Improvement	15.03		133.31	
Cohen's D	1.24			
T-statistic	3.11			
P-value	0.00957			

Table 4.1: Statistical Analysis Summary

4.2.1 Cohen's d

Cohen's d measures the effect size (the significance of the difference) between the two groups using standard deviations. In this case, how significant the score differences and improvement rates of the learning tools are. It does this by measuring the mean and the difference in standard deviations away it is. The effect size tells us how strong and significant the difference is between the two groups that are being compared. A higher Cohen's d means a more statistically significant difference there is.

Cohen's d starts at 0 (this is no effect size, meaning that there is no difference between the two groups) A Cohen's d value of 0.2 is a small difference between the two groups, 0.5 is a medium difference, and 0.8 up is a significant difference between the two groups. The higher the cohen's d the more significant the difference in their mean.

Looking at the data from the table, we can see that Cohen's D from the first test is 1.25. This shows that there is a large and strong effect size. Highlighting a significant difference between the group using BioLab and the group using traditional methods. This indicates that the group using BioLab had significantly higher scores than the group using traditional methods.

Moving on to the second test, we see that Cohen's D score goes down to 0.22. This shows us that there is no significant difference in the participants' scores. This indicates that the order in which the user takes 'BioLab' compared with traditional methods did not matter.

Moving onto the percentage of improvement of the individual participant. This again showed a substantial Cohen's D value, meaning that the improvement of participants using BioLab was much higher than that of the participants who used Traditional methods. Participants who used traditional methods first experienced a significantly higher improvement level when switching to BioLab, compared with participants who used BioLab first and then switched to traditional methods. This supports the hypothesis that BioLab is effective in helping improve users' understanding of anatomy. It could also indicate that the experience with traditional methods first enhances the impact of subsequent BioLab use.

In conclusion, we can see that users first gain an advantage from using BioLab in test scores. However, when the groups switch over to the opposite set of materials, we can see that this advantage is diminished, which leads us to assume that the order of using BioLab does not affect the user's ability to score highly in the second test as long as they are exposed to BioLab at some point. This is evident from the percentage improvement levels as the improvements from using traditional first and then moving onto BioLab are statistically significant. This finding supports our initial hypothesis that gamification can significantly improve students' knowledge retention.

4.2.2 T-test's

The t-test for Test 1 produced a t-statistic of 3.06 with a p-value of 0.00643. This p-value is well below the conventional threshold of 0.05, indicating that the difference is statistically significant. This means that we have evidence to suggest that the scores from BioLab in Test One are significantly higher than those from the traditional methods. This confirms that Cohen's D that we also got from the same group, further indicating the efficacy of BioLab.

For Test 2, the t-test did not show statistical significance, as the P-value of 0.61390 is much more than 0.05. This indicates that any differences in the mean scores between the two groups are not statistically significant. This leads us to a new hypothesis: once users are exposed to both learning materials, they will have a better and more well-rounded knowledge of anatomy, as long as they have exposure to using BioLab.

The percentage increase from Test One to Test Two shows that the T-test is 3.11 with a P-value of 0.00957. This shows a large significance in the percentage increase users get from using BioLab compared to Traditional methods, indicating that users using BioLab improved at a much more significant rate.

In summary, the effect size and statistical tests suggest that in Test 1, BioLab significantly impacted participants' scores significantly more than Traditional methods. In contrast, Test 2 did not show a significant difference between the methods, which might imply that the impact of the learning methods was more similar in the second test once the participants had used both sets of learning materials.

4.3 Analysis

4.3.1 Test One

BioLab has a median score of 7.0, this is a high median score by users which means that participants on averaged scored highly. The mean was close at 6.45, this means that there where a few outliars as seen from the graph which scored slightly lower than expected. This is seen from the narrow interquatile range of 0.5, this means that the scores were very consistent among participants. There are three outliers, two of which scored five and three, respectively, and one scored above average at eight.

Traditional methods has a median of 5.5, this score still means that most of the participants were able to score more than 50% on the test. As many participants had no background in anatomy, they picked up all of this knowledge from the traditional teaching materials. However the Interquatile range seen from using the traditional materials was very large at 4.0, this indicates that there was significant variability in the scores among the participants. However, there were no outliers, meaning no one could score above 6.

4.3.2 Test Two

BioLab has a median score of 8.0, which is 100% in the test. This means most users who used BioLab second scored full marks in the test. The interquartile range is, this time, 1.0; this is still relatively narrow and shows us that most participants in this group could score within this range. The scores were very consistent between all of the participants. There was a single outlier which scored five in the second test, however this participant

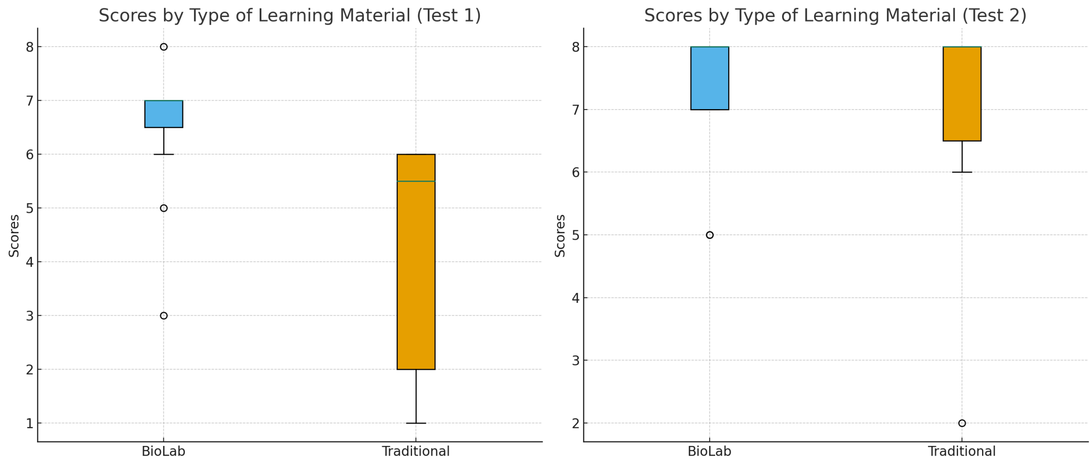


Figure 4.7: Boxplots representing the scores achieved in tests one and two depending on the participants' learning materials.

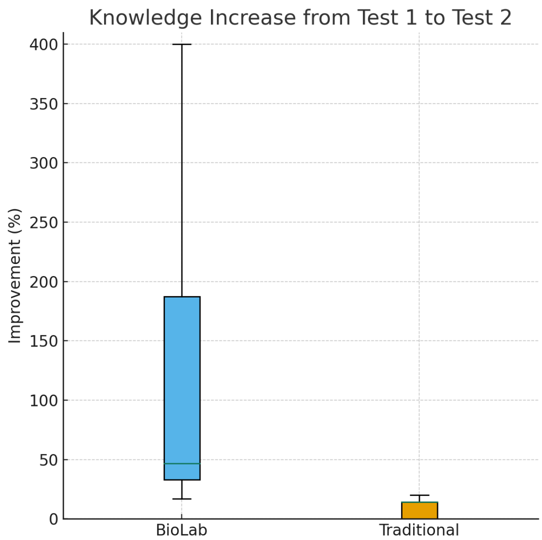


Figure 4.8: Boxplots representing the knowledge increase from tests one to two, split up by the learning materials used in the second test.

scored two in the first test using traditional teaching methods, which highlights that this participant still improved their score by 150% since using BioLab.

Traditional methods has a median score of 8.0 as well, this again shows that a majority of participants where able to score a full marks. The interquartile range, however, is slightly more extensive at 1.5, showing somewhat more significant variability in the scores compared with BioLab. There was again a single outlier, which scored two. This was less than their score in the first test, this seems like an anomaly in the test taking as opposed to the standards of the tests.

4.3.3 Percentage Increase

I also wanted to compare and visualise the spread of the improvement in scores between the different learning materials, so I created a box plot with whiskers. This will show the percentage increase in score that each participant got after taking either one of the two learning materials.

The participants who learnt using BioLab before taking the last test showed a very large rate of improvement. The median improvement was 46.665%. This is a very large increase in scores from taking the first test. In relation to the quiz, which is scored out of eight, this means that each participant would get around four extra questions right since using BioLab. The interquartile range for this is very high, which tells us the spread of the participant's average improvement was very high. The interquartile range was 154.50%. This is because many participants got more than 200% improvement in their score, with one participant getting a 400% increase in their score after using BioLab. The lowest improvement for any of the participants using BioLab second was 17%. This participant scored very highly on the first test and only improved their score by one.

Looking at the traditional method improvement, this had a median of 14.290%. This is a relatively low improvement from test one to test two. The interquartile range seen was also 14.290%. This shows that the spread was not very high, indicating that most participants could not improve their scores as much as the users using BioLab. This could indicate either the materials were less accurate than BioLab or that the participants could learn more efficiently using BioLab. This is evident from the most significant improvement from a participant using traditional materials, which was 20%, compared to the lowest from using BioLab at 17%. This highlights a stark contrast in improvement between the use of the two materials.

4.3.4 Conclusion

In conclusion, these graphs highlight that the difference between the two learning materials enables the participants to gain more knowledge using BioLab. However, again, we can see that it doesn't matter in the order of usage by the participant, as in both cases (if they use BioLab first or second), they can score very similar scores.

4.4 Confounding Variables

This section will explain the implications of the different confounding variables explored in this study. I will compare different attributes that could change the tests' results. I will compare the two sets of materials, how the confounding variables behaved with them, and their implications.

4.4.1 University Students vs Teachers

In this part of the study, I looked at university students and teachers and how they interacted differently with the learning materials they were given. I also looked at the

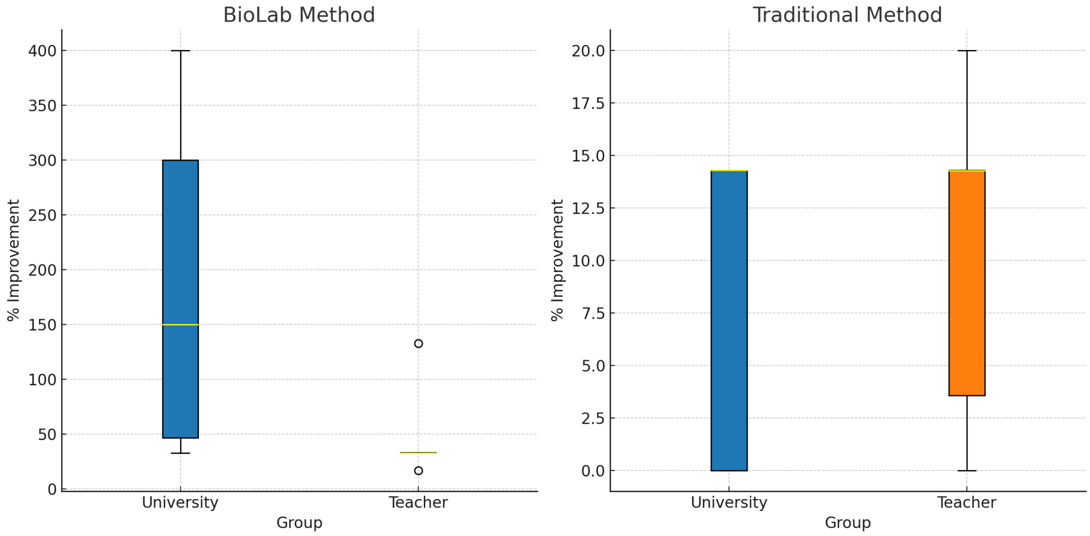


Figure 4.9: Boxplot showing the improvements made comparing university students and teachers with BioLab and Traditional methods

improvement rates of their scores and compared this to whether they were students or teachers and what learning material they were using. For the participants who are currently studying at university and using BioLab, a median percentage improvement of 150% was shown. This was a very high score, signifying that university students gained a lot from using BioLab. This alludes to one of my research questions regarding the efficacy of digital learning for users who are well-versed in digital learning. As well as this, there was an IQR of 253.5%. This is a very substantial spread. This suggests high variability, which implies that the BioLab method has been very effective for some students, but its effectiveness may vary from student to student.

Some reasons for this high spread could be variations in learning styles, with some students preferring a more interactive and digital environment. More hands-on, self-guided learning will take place in BioLab. It is also important to consider a novelty effect, where introducing a new learning method boosts performance for some time as users are excited to try something new.

The teacher group has a very different pattern from the university students. The median is 33% and has a very small IQR of 0.33. This much smaller median improvement and a very small IQR mean that the BioLab method provides teachers with a consistent but smaller boost. This slight boost of around 33% contrasts with the 150% observed by the university students.

The very tight spread could be because additional novel teaching methods do not mean significant performance gains due to time or cognitive load constraints. It could also be that many teachers have become accustomed to teaching a subject rather than having to become a student in this example. They are taking on this new role, which could lead to a steep learning curve that they will have to overcome.

For the university students group, the traditional materials showed a median improvement of 14.29%. This is a much more modest improvement compared to what was seen during the university students using BioLab. The IQR was also 14.29%. The traditional learning materials, being much more textbook-based and less interactive than BioLab, give students a more modest improvement on their original test scores. The IQR highlights that this experience is relatively consistent throughout. This shows how traditional teaching methods are somewhat effective for students to acquire knowledge on the subject. However, the inability to cater to diverse learning styles, such as in BioLab, where students can learn in their style, makes this method less effective.

For the teacher group, we also observe the same median improvement as the university group, at 14.29%, but with an IQR of 10.72. This means that while the median is the same for both groups, the spread among teachers is less, signifying that teachers are able to use these materials more effectively on average than university students.

This could be due to the teaching experiences teachers have grown up using and teaching. They are familiar with these textbook-style materials and can learn more effectively from them. At the same time, university students who went to school and university during the pandemic, when they had to start learning online using more dynamic and digital materials, have become less familiar with these types of materials.

Overall Implications

The implications of these findings highlight how effective BioLab is at teaching university students compared to traditional methods. We can observe a considerable difference in the mean of the two sets of materials for university students. We can see the efficacy of BioLab from this and infer that university students can benefit from more interactive learning methods. Teachers might require a combination of learning strategies to push for higher improvement.

The high spread in the university group's results means that university students are not a 'One size fits all' group. They will require more personalised materials to achieve higher scores. Being able to cater to self-learning was done using BioLab. However, there are also opportunities for adapting more learning styles to bring the spread down.

Conclusion

This study shows how important it is to cater the learning methods to the student. The success of BioLab with university students suggests that digital learning can improve educational outcomes for students. However, it is also important to note that the effectiveness has a high spread across this could, this calls for even more personalised teaching methods to cater towards more students enabling for a more inclusive learning experience for anyone that wants it.

Conversely, the modest and consistent gains observed among teachers using both BioLab and Traditional methods suggest that professional learners may benefit from a more balanced mix of digital tools and established materials.

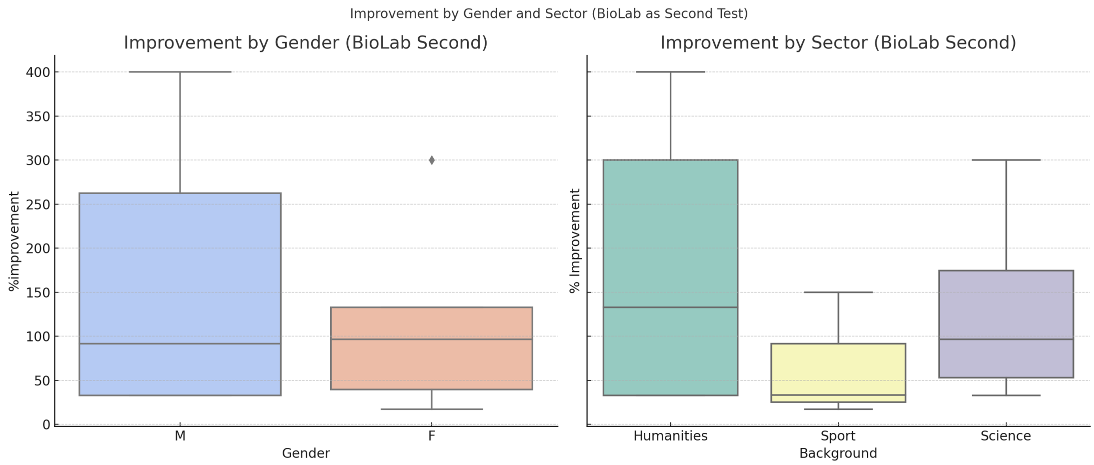


Figure 4.10: Boxplot showing the improvements made comparing university students and teachers with BioLab and Traditional methods

4.4.2 Gender and Sector

4.4.2.1 Gender-Based Performance Improvement with BioLab

In the introduction chapter, I discussed gender roles in technology and gamification. It was previously highlighted that gender inclusivity is neglected in educational methods. Gender inclusivity is one of the most critical pillars in teaching a topic effectively. This was important during the development of BioLab, as previous papers had seen disparities in students learning outcomes based on gender, particularly a paper by Jared R and Peter J [16] where they found different gamification attributes caused different outcomes based on the sex of the participants.

One specific gamification tool they looked at was feedback on questions. This was an essential attribute of a gamification experience to determine if the questions were correct or incorrect. They found that feedback was more beneficial for female participants compared to males. An essential attribute that gamification can offer is a well-rounded and multifaceted approach to tackling many different tools to cater to a wide range of student's needs and learning styles. Gamification aims to bridge the gaps in gender-specific education.

The findings of this study on gender-based performance improvement with BioLab shed some light on this discussion. The study observed significant median improvements for female and male participants using BioLab. This reiterates the adaptability of gamified learning environments.

Splitting up the percentage improvements from female and male participants who used BioLab led to the following results. For female participants using BioLab, they had a median percentage improvement of 96.5%. This shows that female participants were able to benefit massively from the gamified learning environment. The IQR was 93.25%. This somewhat medium IQR means that there isn't much spread in the change of improvement rates and provides a consistent experience among the majority of female participants compared with the male participants.

The study's findings indicate that the interactive and self-guided learning materials in BioLab aligned well with the participants' learning styles. The consistent results also suggest that the gamified elements effectively engaged the female learners, and this translated into tangible results, as seen in their improvement from test one to test two.

The male participants' findings gave us a very similar median improvement of 91.665%; this shows that most participants using BioLab could improve at the same rate regardless of gender. However, male participants had a much wider IQR at 229.4175%. This significant IQR indicates that while male participants generally benefitted from BioLab, the spread of the individual participants was much more critical. The efficacy of BioLab depended greatly on the individual male participant.

The large spread in outcomes for the male participants could be due to various factors. For example, it could be due to the diverse range of engagement levels seen with the gamified content, which would be due to individual differences in learning techniques. Also, this previous gaming experience could factor in how comfortable the participant was with gamification and the interactivity of the game design.

Age

The study also looked at the variance of age. Age played a significant factor in the efficacy of the gamified learning experience. As the participants' age increased, there were more pronounced gender differences than in younger students. Younger students displayed no gender variance in efficacy, while older participants saw this variance increase slightly; however, this had a minimal impact on the overall differences in the study.

The study showed the benefits of BioLab's inclusive and personalised gamified learning. Offering a more personalised and inclusive design led to higher improvement rates in both Female and Male participants. Inclusivity plays a significant role in the effectiveness of learning tools and aids in educational pathologies.

4.4.2.2 Sector-Based Performance Improvement with BioLab

When looking at BioLab's efficacy, it becomes important to examine users' different backgrounds and how they react differently to the gamified tools. In creating a personalised environment, it's important to cater to different learning styles from various sectors. The different sectors exhibit unique learning strategies with the gamified tools. It is also important to note the different prior exposure to digital tools that the other disciplines may have used.

Humanities Sector Performance

Students from the Humanities sector showed a median improvement of 133%. This shows a significant improvement rate for participants from tests one to two. However, there was a large IQR of 267%. The broad IQR indicates a high variability, which shows that some students in this sector achieved remarkable improvement while others experienced much less improvement. The variability of the results could be inducive to the subjective nature of humanities subjects.

Science Sector Performance

The science sector displayed a median improvement of 96.5%. While this is a lower median improvement from the humanities participants, this still reflects a very high improvement rate of nearly doubling the participant's score. The IQR was 121.5%, this is much lower than the humanities subjects, this shows that there was less of a spread and more consistent results from the science sector.

One important note that comes from this sector would be students studying medicine. Many of the students were in their fourth year of study of medicine, where they had to learn the anatomical components of the human skeleton. Their improvement scores would have been affected by this, as they would have been able to score highly on both tests.

However, the balance between the efficacy and the spread of the results in the science sector indicated that the gamification learning styles pair very well with the disciplines of correctness and the unambiguous nature of these subjects.

4.4.2.3 Sport Sector Performance

The sports sector had the lowest median improvement rate, at 33.33%, while still improving from the first test, an interesting difference from the other two sectors. The interquatile range, however, was much lower, at 66.5%. This is the most consistent of all the sectors in the study. This suggests that while the gamification tools were less successful, they provided a much more uniform experience across all individuals in the sector.

While this sector had much more consistent results, the overall improvement percentage was not close to improvements made by other sectors. This could be down to the more intrinsic nature of sports, which focuses less on the repetitive nature of this sort of study.

Conclusion

The reasons behind these differing responses likely stem from each discipline's intrinsic engagement with gamified learning, BioLab's ability to emulate the learnt characteristics of each sector's learning methods, and the overall engagement each individual found with BioLab.

One way to improve the overall improvement for each of the sectors could be numerous, it would need further study to understand each sectors most suitable learning methods fully. For example, if a gamified learning tool were to add more of a Narrative Pedagogy, this could help boost users who are used to humanity sector learning.

4.4.3 Conclusion

The study has underscored the differential impact of BioLab across genders and academic sectors. The platform has been able to adapt to different learning methods across all of the participants' backgrounds. However, there is room for improvement in different sectors and age groups, such as improving the success rate of older learners who are less well-versed in unambiguous learning styles. This would help maximise the educational efficacy of BioLab.

Future research that can be conducted on the confounding variables should investigate the specific elements of gamification that conform with different sectors' current learning styles. Exploring how they could be adapted into gamification to help enhance the potential for user learning outcomes. This could help combat sectors with low median and high spread of improvement rates. In addition, including a longitudinal design would help measure the long-term retention rates of the participants. Being able to measure the long term retention rates would better measure the efficacy of any learning tool, however, in this study this was not possible due the time constraints.

4.5 K-means clustering

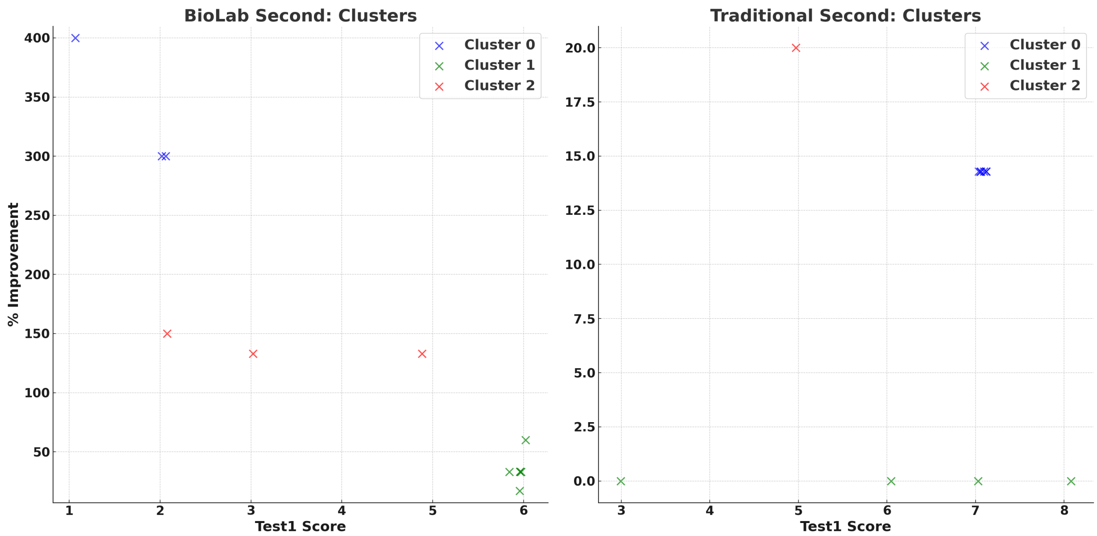


Figure 4.11: K-Means clusters, showing participants improvement from initial scores. Split up based on the learning materials used before the second test.

In this study, I also conducted K-means clustering on the data. I split the data into two groups: the group that took BioLab for the second test and the group that took Traditional for the second test. Doing it this way meant that each participant would act as their control when evaluating their improvement levels. For each group, the individuals were clustered according to their Test One score and the percentage of improvement they got from taking their second test.

In the BioLab group diagram, we observe three different clusters. The distribution of points shows us the different levels of improvement and how they relate to their original scores. Looking at the clusters it reveals that a lower initial score correlates with higher improvement level, this is seen from all of the clusters visible. As seen in the diagram the improvement levels range from 0% to 400% improvement, these participants initially used traditional methods to learn the anatomy and then moved onto BioLab, gaining these improvements. One notable cluster seen in this graph is cluster 0, which initially had very low scores from traditional teaching materials. Still, it had a much higher improvement rate after using BioLab. This cluster represents a group in which innovative learning materials such as gamification have allowed them to gain a much higher improvement rate.

In the traditional group, there are also three clusters; however, they have distinctly different distributions. However, the spread of the %improvement is less varied, indicating that using traditional learning methods gives the user a more consistent learning improvement. The improvement however is much less impressive, as these participants originally learnt using BioLab, this implies that the added learning using Traditional materials meant that the participant gained much less than the participants that used BioLab after using Traditional materials.

Comparing the results and the clusters from the two diagrams, we can see the effectiveness of the learning materials in allowing for improvement in scores. We can infer that BioLab materials are more effective for specific individuals. However, the traditional group shows much less variability in improvement, which could benefit users seeking more consistent improvement rates.

These results offer a new perspective on learning tools. Students who have not previously responded to traditional learning materials may better understand topics using gamification. Personalised learning materials for these students could benefit them greatly, allowing them to excel using the material that better suits their personalised learning style.

4.6 Participant Feedback

4.6.1 Quantitive feedback

Feedback Question	Avg Score
Satisfaction with the game	4.67
Ease of navigation and play	4.56
Understanding anatomical concepts	4.67
Aiding memory retention of anatomy information	4.67
Effectiveness in demonstrating the practical application of anatomy	4.56
Engagement	4.56
Willingness to play again for learning anatomy	4.89
User interface design rating	4.67
Appropriateness of difficulty level	4.44
Game's speed and performance rating	4.78
motivation to learn about human anatomy from traditional	3.00
Increase in motivation to learn about human anatomy	4.56
Preference for BioLab over Traditional (1 - Traditional 5 - BioLab)	4.86

Table 4.2: Average Scores from participants' feedback out of 5

The feedback from the participants reveals that BioLab has been well received by the 23 participants in the study. The very high average scores out of five reveal the usefulness of the game and the ease of navigation and use of BioLab. Particularly, the average score was 4.89 when participants were asked if they would use BioLab again if they needed to learn about anatomy. This shows the motivational pull of the game design, which enables a more effective understanding and willingness to learn, which translates into better results. [23] By motivating the users and driving them to return, it subsequently improves the long-term retention of knowledge and increases the motivation and engagement of students. [18, 13, 44]

The feedback indicates that BioLab has been able to merge the instructive expertise of traditional learning materials with the engagement factor intrinsic to gamification. This means that the users were not only exposed to all of the necessary anatomy knowledge but also enjoyed the learning experience. This combination has proved a very effective tool in increasing the rates of improvement from the study participants. The participants' recount shows they have also enjoyed the experience.

The data indicates a positive correlation between BioLab's interactive features and advanced memory retention. The high score of 4.56/5.00 for easy navigation and play shows that the game is intuitive to play and use. This reduces the cognitive load of having to work out how to use the game and instead allows users to focus on being able to focus on the materials used to learn.

The inclusive ability to cater to different learning styles offered by BioLab's features allows the user to feel comfortable and go at their own pace, increasing engagement and motivation.

The slightly lower but still high score of 4.44 for the appropriateness of difficulty level expresses that there could be some improvement to the levels in BioLab. Some participants found the game hard when they first started using it, this could be improved with future development of BioLab. They added more levels and easier sections for the users to learn.

In conclusion, this study has shown that the motivation, engagement, and improvement obtained from using a gamified learning tool specifically for BioLab can provide tangible benefits for the user. Some users are able to see higher improvement scores than others, but allowing the user to learn using the self-guided and engaging nature of games highlights a new potential for digital tools in education pathology.

4.6.2 Qualitative Feedback

In this section, I also received qualitative feedback from some of the participants about features they would like to see or improve in the game. This meant that I would be able to split up the results into similar categories that were received and discuss the categories seen.

4.6.2.1 Improve the clarity of correct and incorrect answers

This section was the most requested feature from the feedback; the participants used the quizzes in the game and noticed that the feedback gained was not as straightforward as it could have been. 66.7% of participants who took this survey said the same thing.

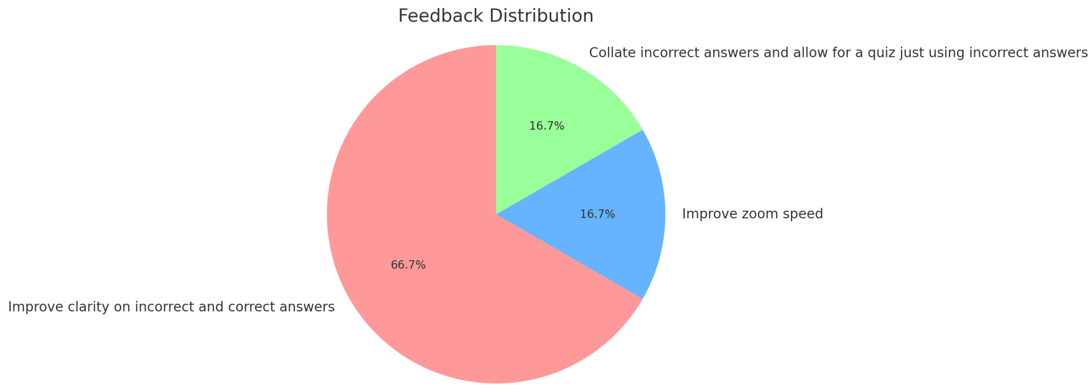


Figure 4.12: Qualitative feedback from participants split up with the main point they would like to see added to the game

This is a large chunk of the participants. The ambiguity of questions' feedback should be avoided to ensure that the users are aware of the correct and incorrect answers. This ambiguity could lead to less memory retention as users could still be unsure of the correct answers to certain questions. A clearer system for displaying accurate and incorrect answers could increase the learning environment's efficacy.

4.6.2.2 Improve zoom speed

Nearly a sixth of participants felt the zoom speed function could be improved. The technical functions of the application are essential in keeping up the engagement. If something doesn't work the way people would expect it to, this causes the cognitive load to deviate away from learning and instead focus on different parts of the application's technical functionality. Improving this attribute in BioLab would ensure fewer distractions caused by my technical functions and allow the user to focus entirely on the learning materials.

4.6.2.3 Collating incorrect answers into one quiz

Again, nearly a sixth of participants wanted a function allowing them to practice their incorrect answers. Implementing this feature into BioLab would reinforce knowledge acquisition from the user and cater to individual learning curves. This level of personalisation in the application would allow for a new level of self-guided learning [47, 49, 39], which would help the users to be able to work on their own mistakes. This helps enable the education pathology of mastery of the topics, and it also prioritises the customised learning tools that gamification tries to emulate, tackling the 'One-size-fits-all' approach of traditional learning tools.

4.6.2.4 Conclusion

All of this qualitative feedback points to the users being engaged with the learning materials presented to them. For example, the desire for clear feedback and improved technical performance indicates that the users are involved in the learning process.

They want to get more out of the application, and their engagement and motivation are increased by using BioLab.

By implementing these suggestions, BioLab would be able to improve and offer a more personalised design. This improved design would be able to increase the improvement rate of users due to the increased time spent using the learning materials.

However, as this feedback is self-reported, it's hard to tell if the provided feedback is how the user truly felt. Without technology such as tracking eye movements, for example, the study cannot say for sure the true motivation and engagement that BioLab enabled. Though this data is subjective, it comes immediately after the participants have used the application, providing valuable insight into the user's sentiment and experience. Future studies could incorporate different features, such as eye movement tracking, to measure the engagement levels of the participant. This would offer insight into the participant's account of self-reported feedback.

4.7 Subsequent Improvements to BioLab

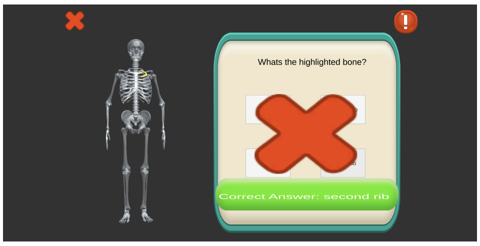


Figure 4.13: Quiz one indicating that the user was incorrect and giving the correct answer.

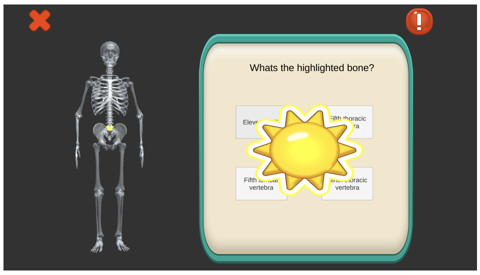


Figure 4.14: Quiz one indicating that the user was correct.

In response to the study participants' feedback, I built new features into BioLab. As the most requested feature was clearer feedback from correct and incorrect answers, I added this feature to the quizzes to let users know whether their answer was correct, wrong, or close. A new response was added for text input, allowing the game to tell if the user's input was misspelt. This was done using Levenshtein Distance from the correct answer. This would enable the user to try again, allowing for typographical leniency and prompting the user to try again. This reinforces the objective of memorising anatomical bone names through a forgiving and iterative environment.

Adding this feature has enabled a much clearer feedback mechanism for the user to know the outcome of their answer. It also reduces the cognitive load of working out if they got the correct answer, enabling the user to focus fully on the learning materials. In addition, the zoom feature was improved to allow users to zoom into their choice of bones more easily. This again allows for the cognitive load to be focused on learning materials.

In conclusion, these additions aim to improve the user experience with BioLab. By enabling users to focus more on the learning materials, it enables more engagement and motivation, which links back to increased user achievement.[31]

Chapter 5

Conclusions

5.1 Research Questions

RQ1: How does using a gamified learning app impact people's short-term retention of anatomical knowledge compared with traditional learning methods?

The study in chapter four of this paper was able to answer this question in relation to BioLab. The participants in the study using gamified learning tools achieved better improvement rates overall than participants using traditional learning materials. This was done by analysing the participant's improvement rate from test one to test two and comparing this with the order of learning materials used. After using one of the learning tools, their first test acted as a control to gauge their current knowledge of the subject. The study then compared this score with the score they received after using the opposite learning materials. This showed us that using a gamified learning app improved people's short-term retention of anatomical knowledge compared with traditional methods. The study analysed them using K means clustering, boxplot with whiskers, Cohens'd, and t-tests from these results. This improvement rate was found to be significant enough to reject the null hypothesis that traditional learning materials give a better improvement rate.

RQ1.1: To what extent does a gamified learning tool increase knowledge retention over traditional learning materials?

This question was somewhat answered using the research from the study in chapter four. We could see that short-term retention from a gamified learning tool was improved by, on average, 127.22% compared with only 6.58% improvement using traditional learning materials. However, we could not determine whether long-term retention was improved through gamification.

RQ1.2: Is there a difference in knowledge retention using gamification, which can be measured based on gender?

The study participants showed little to no difference in the gamification knowledge retention measured by gender. This was demonstrated at an early age. However, as the participants' age grew, there was a slight difference in knowledge retention based on their gender. However, the sample size was relatively small for female participants over 40, so this could be skewed.

RQ2: How does student engagement with anatomical content differ between gamified learning environments and traditional paper study?

From the feedback obtained from the participants in the study, it was observed that engagement levels compared with traditional learning materials massively improved. The participants overwhelmingly preferred using BioLab over traditional learning materials. This self-reported motivation and engagement meant that the participants in the study found much higher levels of engagement with the anatomical content.

RQ3:Does the integration of gamification into anatomy education improve student motivation?

Using the feedback from the study and the results, we find that the motivation to use a gamified learning tool improved the participant's motivation to learn. The participant's motivation improvement meant they ultimately scored higher after using the gamified learning tool. The self-reported higher motivation score of 4.56 on average meant that the motivation was higher than the motivation to learn using traditional learning methods, which was 3.0 on average.

5.2 Limitations and future work

The research had several limitations due to the time frame of creating the application and testing on the participants. The study was only able to look at the short-term retention of knowledge instead of measuring retention of knowledge over a long time and multiple uses. The study was only able to provide answers to immediate learning outcomes. In addition, the small sample size of 23 meant that the results from the study might not have much external validity. The study also relied on self-reported measures for engagement and motivation which means that there was subjectivity in the results of the feeback. Due to the time needed to participate in the study, which was 1 hour, a lot of the participants had to use the application on a video conferencing platform while I watched them use the application and answer any questions that they had on it. This meant that they could not test the augmented reality feature of BioLab as this was only available on the mobile version of the app.

The results and feedback from this study indicated what future work could take place. For example, exploring the long-term retention of knowledge gained through gamification. Doing this would be able to show the change in motivation and engagement over time. In addition, having more participants would benefit the study as one group could be the control and use traditional learning materials. The other group could just use the gamified learning tool. This would build an overall picture of the improvement rate over time using gamification instead of traditional learning tools. The groups would have to be assessed beforehand to ensure that the participants had similar knowledge in the area to start with. In addition, including a more diverse number of participants across different educational backgrounds and sectors would allow for a better and a deeper understanding of the efficacy of gamification. Future work could also be done using biometric data to assess the participants' true measures of engagement and motivation. This would remove any subjectivity from the self-reported feedback.

Another avenue that could be considered is artificial intelligence and its integration into the game. Not only could AI be used to find new educational pathways for each individual student to learn that best suits them. However, it could also be used in the multiplayer section; playing against the computer would allow for individual play while still accomplishing the educational pathology present through multiplayer activities.

In conclusion while this research has helped shed light on the benefits of gamification in education, particularly in anatomy learning. It has also become clear that future work can fix shortcomings in this study. The original work sets a foundation for future work to be able to work from and help to develop new digital learning tools and materials to help motivate and engage learners around the world.

5.3 Conclusion

This dissertation explains how gamification could increase the learning effectiveness of anatomy using gamification. I created the BioLab app and web application to teach learners about anatomy. It is aimed towards medicine students who wish to learn more about anatomy as well as learners from all backgrounds who also wish to learn more. I also conducted a study in which I trialled the use of BioLab and compared it to traditional teaching materials. I then compared the scores users got from each learning material and the increase in the scores they got from using the other set of learning materials. The study showed that gamification was able to help engage and motivate students more than traditional learning materials could.

I also showed how gamification helped engage and motivate learners from different backgrounds. I also demonstrated how gender did not play a role in the efficacy of the application. This highlights how educators worldwide should use gamification to teach learners. Participants in the study who struggled with traditional teaching materials benefited from using a gamified learning environment. Gamification can be used to teach a bigger range of students' needs, abilities, and learning approaches by offering a more personalised learning experience to the users. This work highlights the need for more research into the potential uses of gamification to teach students in the digital age.

Bibliography

[1] José L. Arco-Tirado, Francisco D. Fernández-Martín, and Miriam Hervás-Torres. Evidence-based peer-tutoring program to improve students' performance at the university. 45(11):2190–2202.

[2] Merilyn Armstrong and David Boud. Assessing participation in discussion: An exploration of the issues. 8(1):33–44.

[3] Army and Navy Academy. How gen z is changing higher education, 2024. Accessed January 6, 2024.

[4] Gabriel Barata, Sandra Gama, Joaquim Jorge, and Daniel Goncalves. Engaging engineering students with gamification. In *2013 5th International Conference on Games and Virtual Worlds for Serious Applications (VS-GAMES)*, pages 1–8. IEEE.

[5] Ernst Bekkering and Ted Ward. Class participation and student performance: A tale of two courses. 18(6):86–98.

[6] BL00DDR1V3. Human skeleton.obj, 2022. Accessed: 2023-10-05.

[7] D.M. Bressler and A.M. Bodzin. A mixed methods assessment of students' flow experiences during a mobile augmented reality science game. 29(6):505–517.

[8] Encyclopædia Britannica. human skeletal system. *Encyclopædia Britannica*, 2024.

[9] Encyclopædia Britannica. human vertebral column. *Encyclopædia Britannica*, 2024.

[10] Encyclopædia Britannica. human vertebral column. *Encyclopædia Britannica*, 2024.

[11] Giovanni Bronzetti, Graziella Sicoli, and Dominga A. Ippolito. The digital divide in the world of education at the time of COVID-19:. In Patricia Ordóñez De Pablos, Mohammad Nabil Almunawar, Kwok Tai Chui, and Maniam Kaliannan, editors, *Advances in Educational Technologies and Instructional Design*, pages 77–92. IGI Global.

[12] Ryan Brydges, Heather Carnahan, Oleg Safir, and Adam Dubrowski. How effective is self-guided learning of clinical technical skills? it's all about process. *Medical education*, 43(6):507–515, 2009.

[13] Stefan Büchele. Evaluating the link between attendance and performance in higher education: the role of classroom engagement dimensions. 46(1):132–150.

[14] Jared R. Chapman and Peter J. Rich. Does educational gamification improve students' motivation? if so, which game elements work best? 93(7):315–322.

[15] David Codish and Gilad Ravid. Gender moderation in gamification: Does one size fit all?

[16] Mouna Denden, Ahmed Tlili, Fathi Essalmi, Mohamed Jemni, Nian-Shing Chen, and Daniel Burgos. Effects of gender and personality differences on students' perception of game design elements in educational gamification. 154:102674.

[17] Reeta Dhote. The evolution of classroom technology: From chalkboards to digital learning, 2023. Accessed: 2023-03-22.

[18] Naomi N. Duke. Adolescent adversity, school attendance and academic achievement: School connection and the potential for mitigating risk. 90(8):618–629.

[19] Educational Technology Program, Universitas Sebelas Maret, Indonesia, tp11010.yanuaridwi@gmail.com, Yanuari Dwi Puspitarini, Muhammad Hanif, and Educational Technology Program, Universitas Sebelas Maret, Indonesia, muhammadhanif@student.uns.ac.id. Using learning media to increase learning motivation in elementary school. 4(2):53–60.

[20] Cagdas Erbas and Veysel Demirer. The effects of augmented reality on students' academic achievement and motivation in a biology course. 35(3):450–458.

[21] Yassanne GarrawayLashley. Integrating computer technology in the teaching of biology. 3(2).

[22] Karl M. Kapp. *The gamification of learning and instruction: game-based methods and strategies for training and education*. Pfeiffer.

[23] Amna Khan, Farzana Hayat Ahmad, and Muhammad Muddassir Malik. Use of digital game based learning and gamification in secondary school science: The effect on student engagement, learning and gender difference. 22(6):2767–2804.

[24] Georgios Lampropoulos, Euclid Keramopoulos, Konstantinos Diamantaras, and Georgios Evangelidis. Augmented reality and gamification in education: A systematic literature review of research, applications, and empirical studies. *Applied Sciences*, 12(13):6809, 2022.

[25] Digital Learning. Impact of technology in educating gen z: A global perspective, 2024. Accessed January 6, 2024.

[26] Zayar Linn, Yasura Tashiro, Kunimasa Morio, and Hiroki Hori. Peer evaluations of group work in different years of medical school and academic achievement: how are they related? 22(1):102.

[27] Julio Litwin Lima, Carla D. Castanho, Luiz Belmonte, Veronica Slobodian, Guilherme Gomes Carvalho, Gabriela Hirata, and Julia Klaczko. Skull hunt: An educational game for teaching biology. In Xiaowen Fang, editor, *HCI in Games*,

volume 14047, pages 43–60. Springer Nature Switzerland. Series Title: Lecture Notes in Computer Science.

[28] Randi Q Mao, Lucy Lan, Jeffrey Kay, Ryan Lohre, Olufemi R Ayeni, Danny P Goel, and Darren de Sa. Immersive virtual reality for surgical training: A systematic review. *Journal of Surgical Research*, 268:40–58, Dec 2021.

[29] Sharan B Merriam et al. Andragogy and self-directed learning: Pillars of adult learning theory. *New directions for adult and continuing education*, 2001(89):3, 2001.

[30] Ahmed Hosny Saleh Metwally, Maiga Chang, Yining Wang, and Ahmed Mohamed Fahmy Yousef. Does gamifying homework influence performance and perceived gameful experience? 13(9):4829.

[31] Sharples Mike. *Practical Pedagogy: 40 New Ways to Teach and Learn*. Routledge, 1 edition.

[32] Tracey Muir. Google, mathletics and khan academy: students' self-initiated use of online mathematical resources. *Mathematics Education Research Journal*, 26:833–852, 2014.

[33] Pilar Munday. THE CASE FOR USING DUOLINGO AS PART OF THE LANGUAGE CLASSROOM EXPERIENCE. 19(1).

[34] Ernesto Panadero and Maryam Alqassab. An empirical review of anonymity effects in peer assessment, peer feedback, peer review, peer evaluation and peer grading. 44(8):1253–1278.

[35] Pearson. What do generation z and millennials expect from technology in education?, 2018. Accessed January 6, 2024.

[36] Mary Peat and Anne Fernandez. The role of information technology in biology education: an australian perspective. 34(2):69–73.

[37] Pusher Ltd. Pusher real-time communications platform. `https://pusher.com`, 2023. Accessed: 2024-01-05.

[38] Muhammad Puspitarini, Yanuari Dwi Hanif. Using learning media to increase learning motivation in elementary school. 4(2):53–60.

[39] Intan Sinta Dewi Rahayu and Pupung Purnawarman. The use of quizizz in improving students' grammar understanding through self-assessment. In *Proceedings of the Eleventh Conference on Applied Linguistics (CONAPLIN 2018)*. Atlantis Press.

[40] L. Gayle Rayburn and J. Michael Rayburn. Impact of course length and homework assignments on student performance. 74(6):325–331.

[41] Carolyn S. Ryan and Nancy S. Hemmes. EFFECTS OF THE CONTINGENCY FOR HOMEWORK SUBMISSION ON HOMEWORK SUBMISSION AND QUIZ PERFORMANCE IN a COLLEGE COURSE. 38(1):79–88.

[42] Kew-Cheol Shim, Jong-Seok Park, Hyun-Sup Kim, Jae-Hyun Kim, Young-Chul Park, and Hai-Il Ryu. Application of virtual reality technology in biology education. 37(2):71–74.

[43] Yoonhee Shin, Dongsik Kim, and Donggil Song. Types and timing of scaffolding to promote meaningful peer interaction and increase learning performance in computer-supported collaborative learning environments. 58(3):640–661.

[44] Subotica Tech –College of Applied Sciences, Robert Pinter, Sanja Maravić Čisar, Zoltán Balogh, Helena Manojlović, and Constantine the Philosopher Universityin Nitra. Enhancing higher education student class attendance through gamification. 17(2):13–33.

[45] İlkay Aşkin Tekkol and Melek Demirel. An investigation of self-directed learning skills of undergraduate students. *Frontiers in psychology*, 9:410879, 2018.

[46] Times Higher Education. How digital platforms are changing learning, 2021. Accessed: 2023-03-22.

[47] Despina Vasileiadou and Konstantinos Karadimitriou. Examining the impact of self-assessment with the use of rubrics on primary school students' performance. 2:100031.

[48] Aman Yadav, Chris Mayfield, Sukanya Kannan Moudgalya, Clif Kussmaul, and Helen H. Hu. Collaborative learning, self-efficacy, and student performance in CS1 POGIL. In *Proceedings of the 52nd ACM Technical Symposium on Computer Science Education*, pages 775–781. ACM.

[49] Zi Yan, Xiang Wang, David Boud, and Hongling Lao. The effect of self-assessment on academic performance and the role of explicitness: a meta-analysis. 48(1):1–15.

Appendix A

First appendix

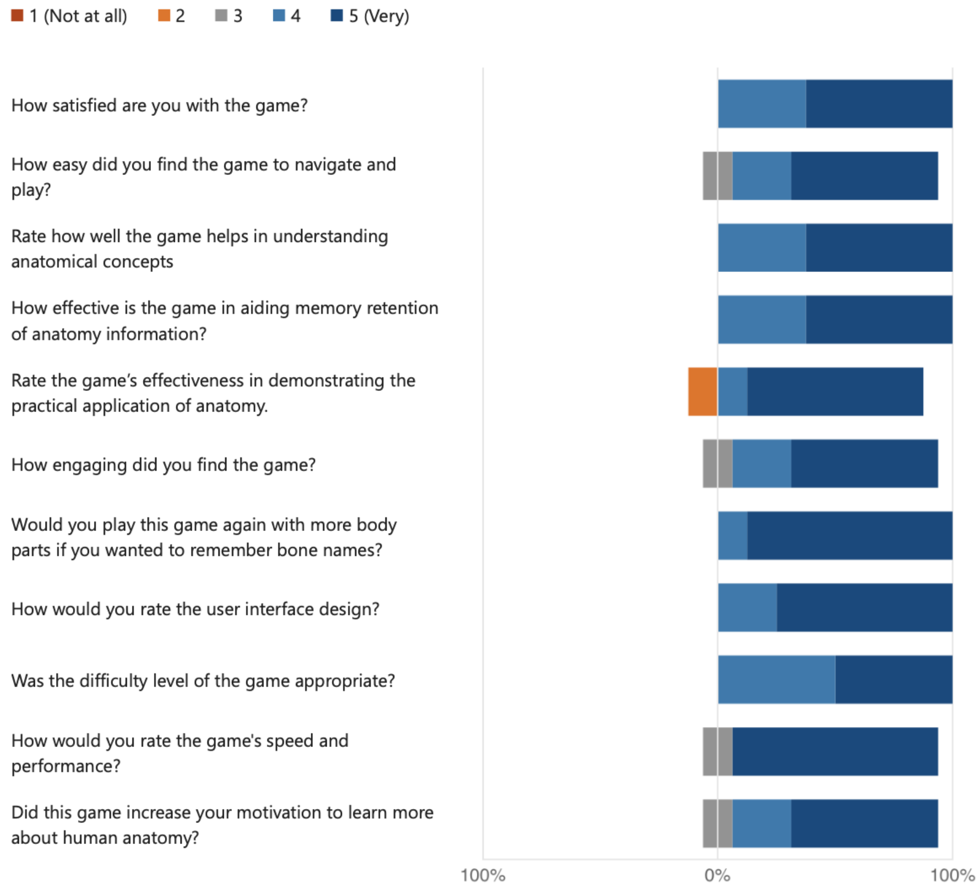


Figure A.1: Results from feedback

A.1 First section

Any appendices, including any required ethics information, should be included after the references.

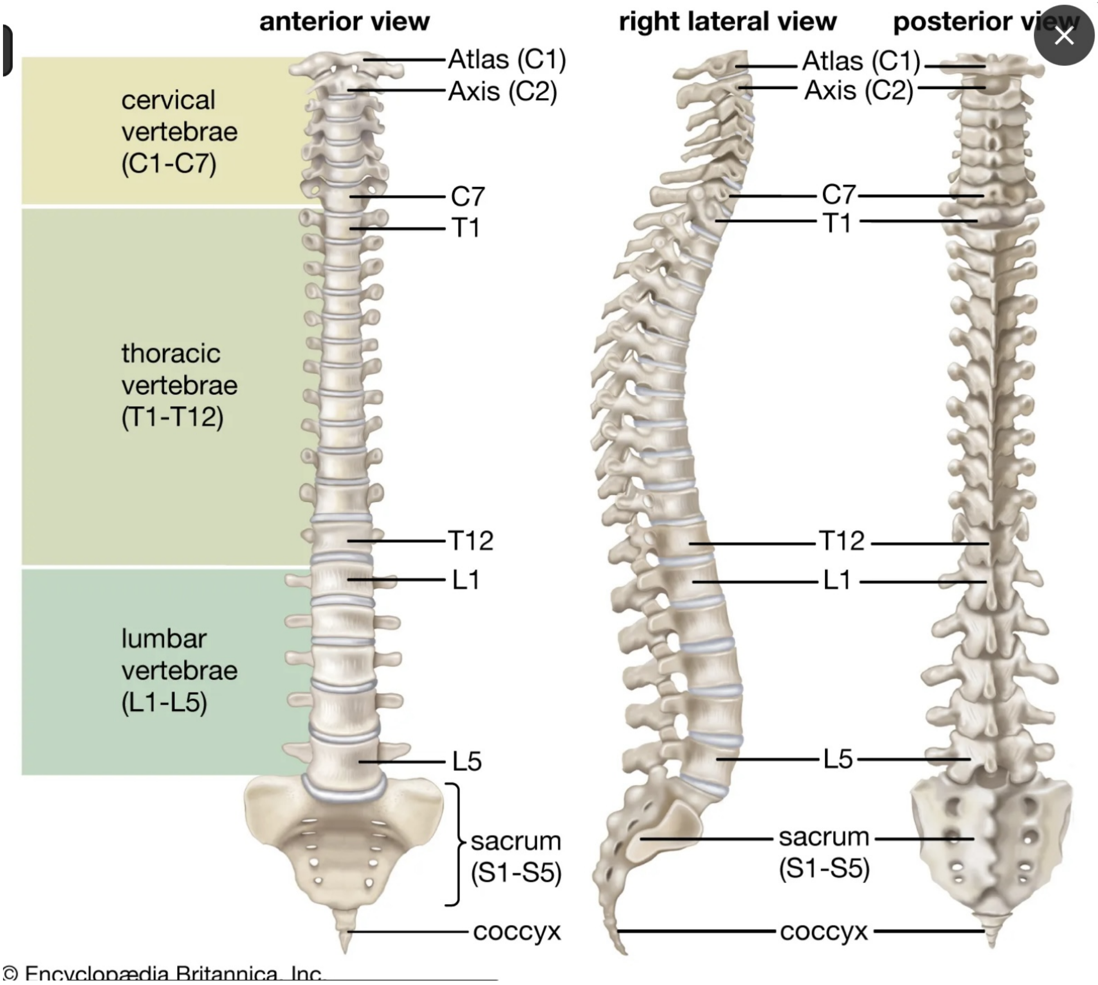


Figure A.2: Annotated skeleton spine taken from Encyclopedia Britannica

Markers do not have to consider appendices. Make sure that your contributions are made clear in the main body of the dissertation (within the page limit).

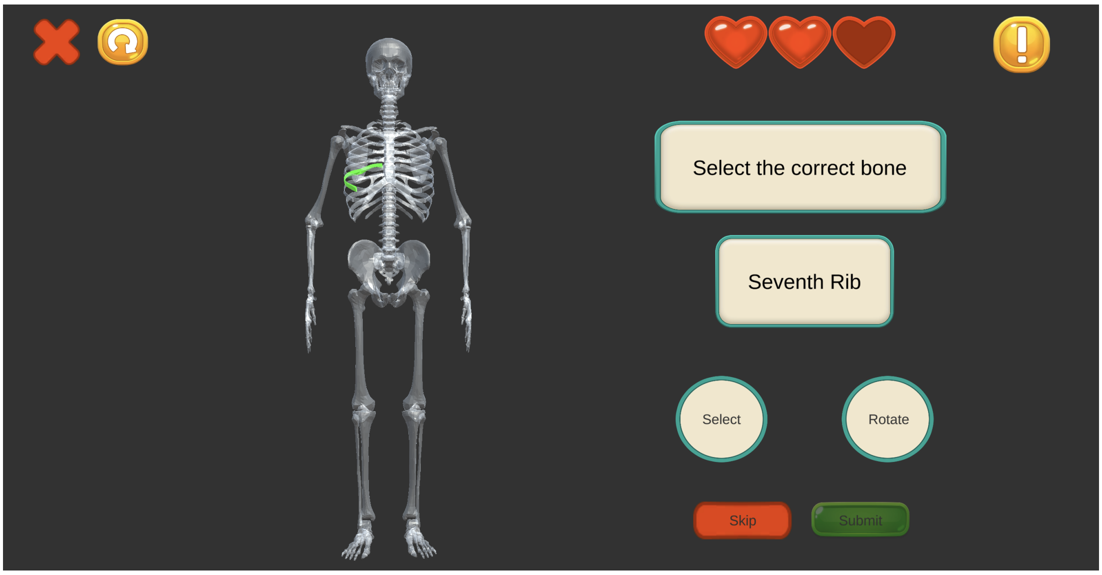


Figure A.3: Quiz two indicating the correct bone when the user skips.

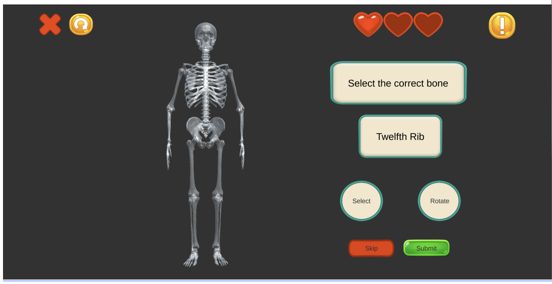


Figure A.4: BioLab Select correct bone

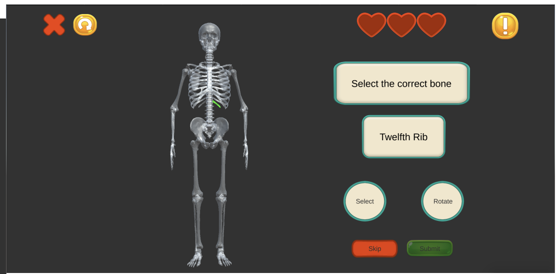


Figure A.5: Green bone to show the user where it is

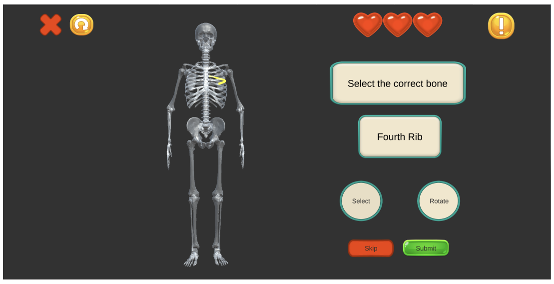


Figure A.6: Bone selected by user

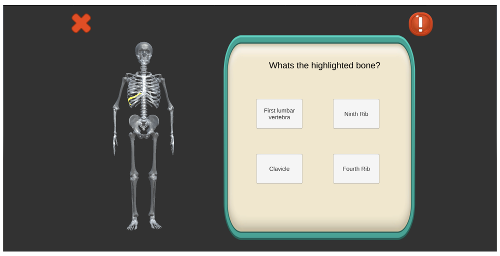


Figure A.7: Multiple choice bone selection

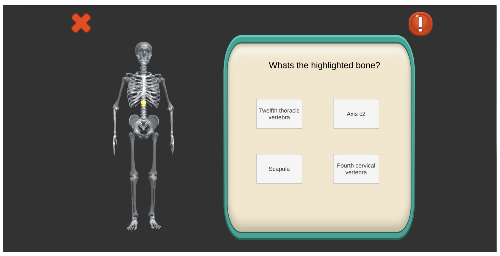


Figure A.8: Multiple choice bone selection

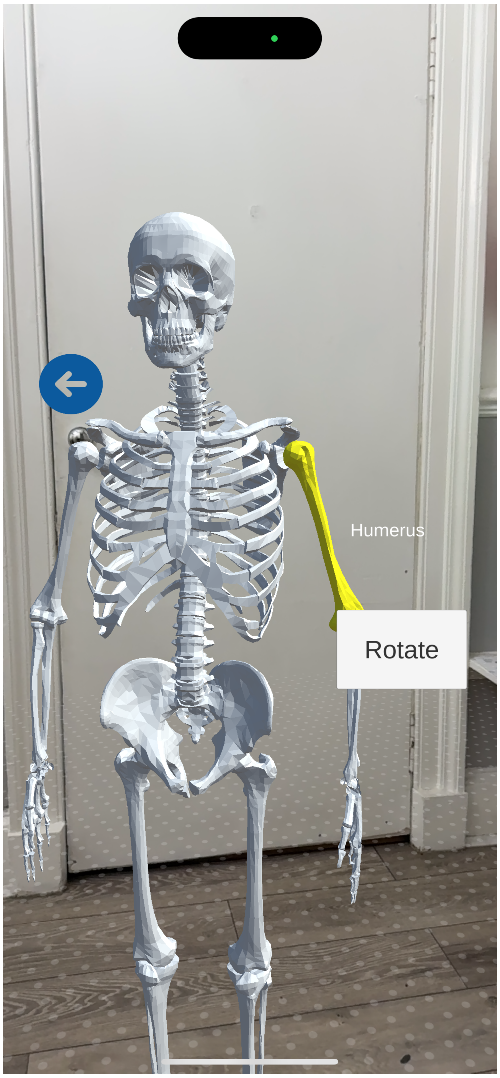


Figure A.9: Full size AR

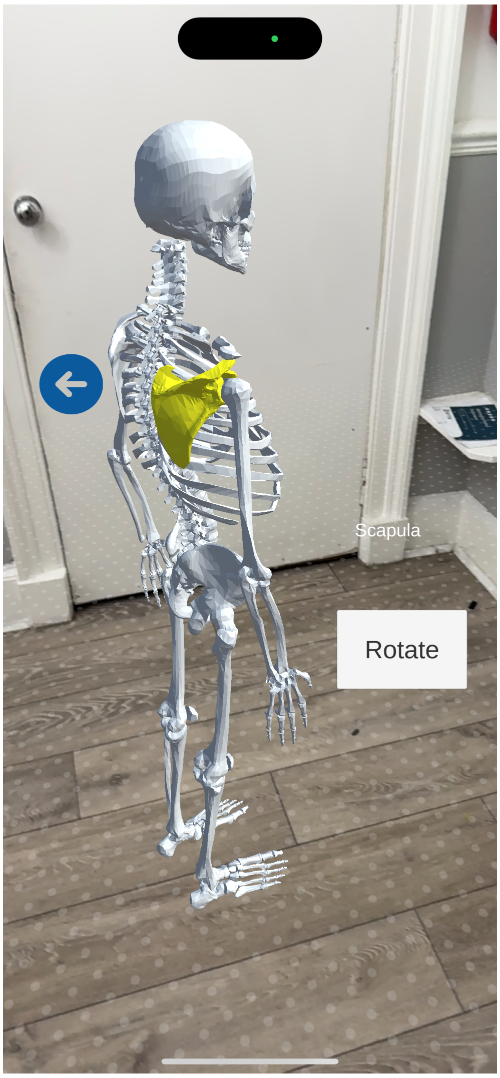


Figure A.10: Full size AR

Appendix B

Participants' information sheet

Participant Information Sheet

Project title:	Using gamification to teach students anatomy
Principal investigator:	James Garforth
Researcher collecting data:	Eric Baldwin
Funder (if applicable):	

This study was certified according to the Informatics Research Ethics Process, reference number 219854. Please take time to read the following information carefully. You should keep this page for your records.

Who are the researchers?

The researchers are myself (Eric Baldwin), I'm a student at the university of Edinburgh studying artificial intelligence and computer science, and my supervisor James Garforth.

What is the purpose of the study?

This study is used to help with designing and evaluating a gamified learning platform which will be used to help the understanding of human anatomy. This project will focus on students who are studying medical and science's fields, considering their education and experiences. It will also have an easier level for students that are not experts in the field and want to learn more about human anatomy.

This study will determine whether gamified learning, using quiz mini games, effectively helps with understanding the topic. We also want to explore how this gamified environment could help with motivation and engagement.

I want to identify what could be improved to help with engagement and learning. I also want to evaluate the usability of AR technology in education, and the potential constraints of integrating such a learning tool.

THE UNIVERSITY of EDINBURGH informatics

Why have I been asked to take part?

You have been asked to take part because you either are an expert in this field, learning this field, take an interest in human anatomy for any other reason, or you are a teacher or teach students in any other field of study.

Do I have to take part?

No – participation in this study is entirely up to you. You can withdraw from the study at any time, up until 1st March 2024 without giving a reason. After this point, personal data will be deleted and anonymised data will be combined such that it is impossible to remove individual information from the analysis. Your rights will not be affected. If you wish to withdraw, contact the PI. We will keep copies of your original consent, and of your withdrawal request.

What will happen if I decide to take part?

You will be invited to experience the gamified learning environment through the stages of development. This will be three times; you will then answer a short questionnaire about the experience that you had with the game. These questions will range from how much the game promoted motivation, to different aspects of the game such as AR, and the quizzes themselves. This should not take more than an hour altogether for each session. No audio or video will be recorded while the participant takes part in the survey.

These can happen in person with Eric Baldwin, or the game can be sent to the participant to be done at their own pace.

Are there any risks associated with taking part?

There are no significant risks associated with participation.

Are there any benefits associated with taking part?

There's an indirect benefit by taking part by using a new type of learning experience first hand.

What will happen to the results of this study?

The results of this study may be summarised in published articles, reports and presentations. Quotes or key findings will be anonymized: We will remove any information that could, in our assessment, allow anyone to identify you. With your consent, information can also be used for future research. Your data may be archived for a maximum of 4 Years. All potentially identifiable data will be deleted within this timeframe if it has not already been deleted as part of anonymization.

Data protection and confidentiality.

Your data will be processed in accordance with Data Protection Law. All information collected about you will be kept strictly confidential. Your data will be referred to by a unique participant number rather than by name. Your data will only be viewed by the researcher/research team Eric Baldwin.

All electronic data will be stored on a password-protected encrypted computer, on the School of Informatics' secure file servers, or on the University's secure encrypted cloud storage services (DataShare, ownCloud, or Sharepoint) and all paper records will be stored in a locked filing cabinet in the PI's office. Your consent information will be kept separately from your responses in order to minimise risk.

What are my data protection rights?

The University of Edinburgh is a Data Controller for the information you provide. You have the right to access information held about you. Your right of access can be exercised in accordance Data Protection Law. You also have other rights including rights of correction, erasure and objection. For more details, including the right to lodge a complaint with the Information Commissioner's Office, please visit www.ico.org.uk. Questions, comments and requests about your personal data can also be sent to the University Data Protection Officer at dpo@ed.ac.uk.

Who can I contact?

If you have any further questions about the study, please contact the lead researcher, Eric Baldwin – S2025952@ed.ac.uk

If you wish to make a complaint about the study, please contact inf-ethics@inf.ed.ac.uk. When you contact us, please provide the study title and detail the nature of your complaint.

THE UNIVERSITY of EDINBURGH
informatics

Updated information.

If the research project changes in any way, an updated Participant Information Sheet will be made available on http://web.inf.ed.ac.uk/infweb/research/study-updates.

Alternative formats.

To request this document in an alternative format, such as large print or on coloured paper, please contact: Eric Baldwin – s2025952@ed.ac.uk

General information.

For general information about how we use your data, go to: edin.ac/privacy-research

Appendix C

Participants' consent form

Participant number:______________________

Participant Consent Form

Project title:	Using gamification to teach students anatomy
Principal investigator (PI):	James Garforth
Researcher:	Eric Baldwin
PI contact details:	jgarfort@ed.ac.uk

By participating in the study you agree that

- I have read and understood the Participant Information Sheet for the above study, that I have had the opportunity to ask questions, and that any questions I had were answered to my satisfaction.
- My participation is voluntary, and that I can withdraw at any time without giving a reason. Withdrawing will not affect any of my rights.
- I consent to my anonymised data being used in academic publications and presentations.
- I understand that my anonymised data will be stored for the duration outlined in the Participant Information Sheet.

Please tick yes or no for each of these statements.

1. I allow my data to be used in future ethically approved research.

Yes	No

2. I agree to take part in this study.

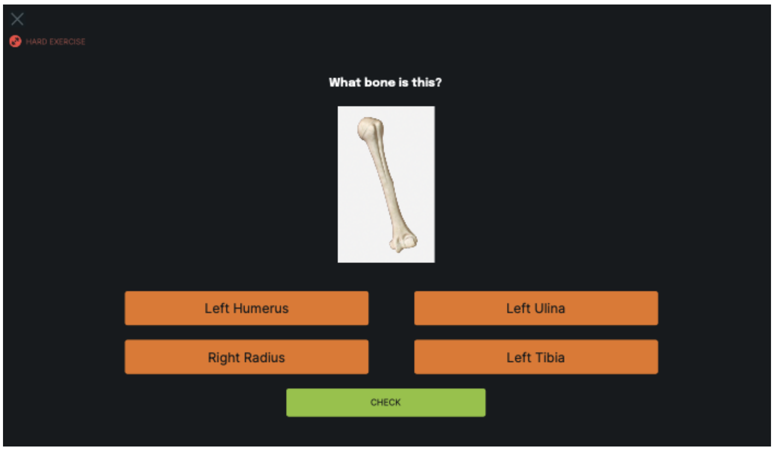
Yes	No

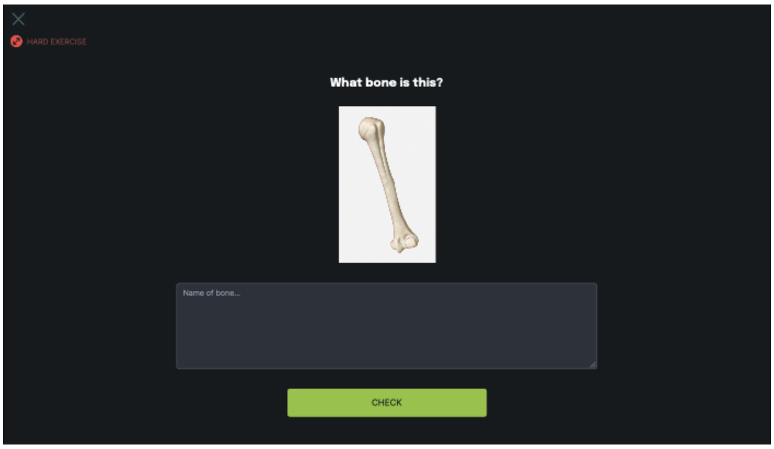
Name of person giving consent	Date dd/mm/yy	Signature
Name of person taking consent	Date dd/mm/yy	Signature

THE UNIVERSITY of EDINBURGH informatics

Appendix D

Original Design Mockup




HARD EXERCISE
What bone is this?
Name of bone...
CHECK

Appendix E

Test Questions

Figure E.1: Example taken from Test 1

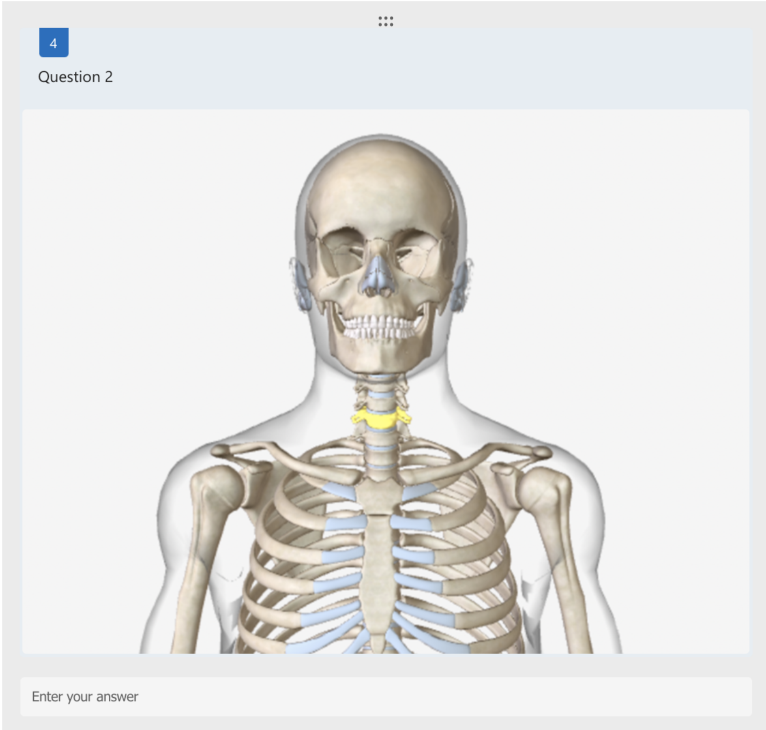


Figure E.2: Example taken from Test 1

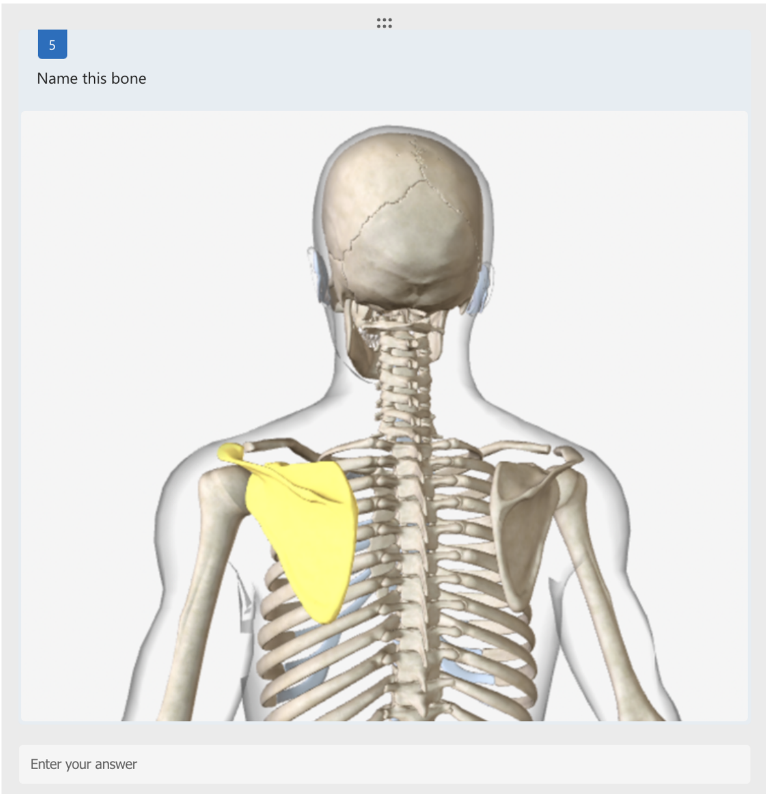


Figure E.3: Example taken from Test 1

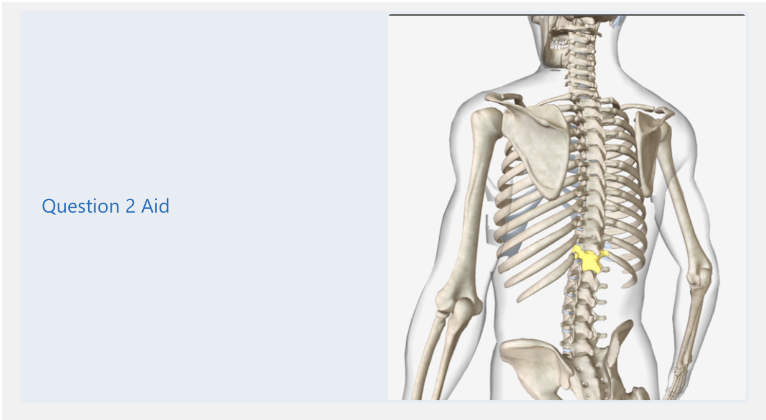


Figure E.4: Example taken from Test 2

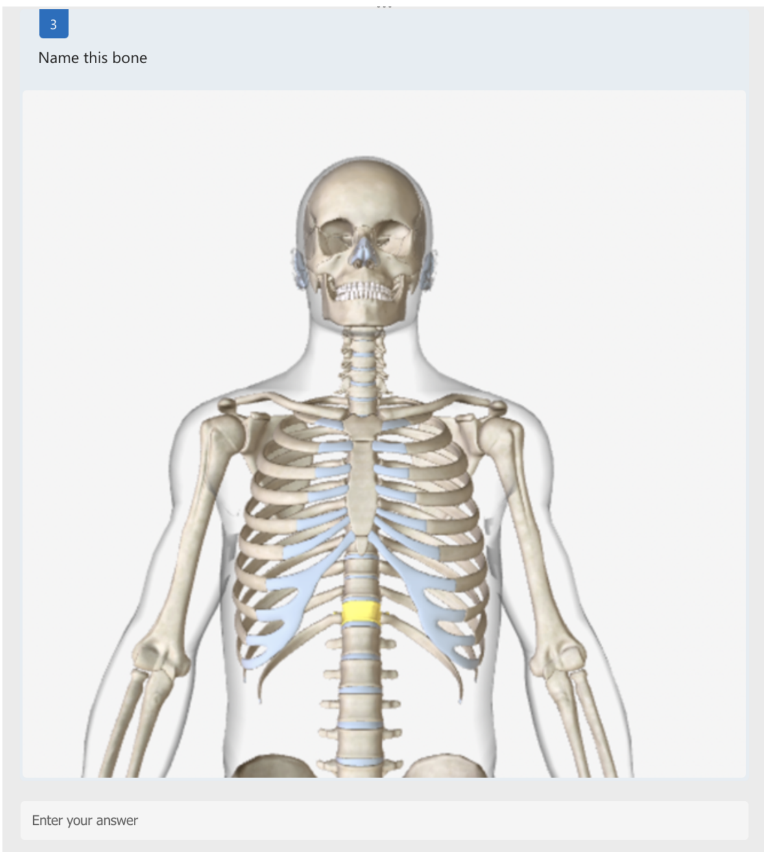


Figure E.5: Example taken from Test 2

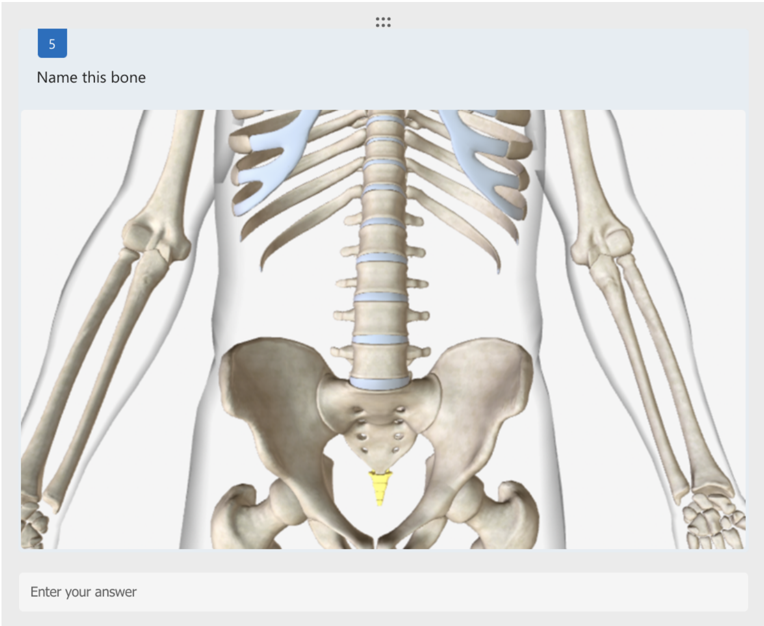


Figure E.6: Example taken from Test 2