Biological Realistic Educational Technology (BRET)

A thesis presented in partial fulfilment of the requirements for the degree of Masters in Human Interface Technology

New Zealand, 2014

Written by

NATALIE EUSTACE

Supervised by

CHRISTOPH BARTNECK
The University of Canterbury
HIT LAB NZ

Contents

Abstract

The aim of this project was to develop and evaluate an interactive Augmented Reality interface for teaching children aged 8 to 15 about biological systems present in the human body. The interface was designed as one component of a "human body scanner" exhibit, which is to be featured at the ScienceAlive! Science Centre. In the exhibit, the interface allows visualization and interaction with the body systems while being moved across a human male mannequin named BRET.

Prior research has shown that Augmented Reality, Visualization applications, and games are viable methods to teach biology to university aged users, and Augmented Reality and interactive systems have been used with children and learning biology as well.

BRET went through three iteration phases, in the first phase, prototypes were evaluated by ScienceAlive! and designs and interactions were implemented, while the use of Augmented Reality through a transparent display was rejected. Iteration two included integration of the non-transparent touch display screen and observational evaluation of six children from 9 to 15 years old. This evaluation resulted in design and interaction changes. Iteration three was the last iteration where final interface and interaction modifications were made and research was conducted with 48 children from the ages 8 to 15. This was to determine whether learning, fun, and retention rates were higher for children who interacted with BRET versus those who watched video clips, or read text. Each child used one learning method to learn the three different body systems: skeletal, circulatory, and digestion. The results of the final evaluation showed that overall there was no significant difference in the children's rating of fun or the amount of information they retained between the different learning methods. There was a positive significant difference between some of the expected fun scores and the actual fun scores. It was also found that learning with text was higher than the interactive condition but there was no differences between learning with video and interaction, or with text and video.

1 Introduction

ScienceAlive!, a Christchurch not-for-profit charitable trust Science Centre, are currently re-designing exhibits and re-building after the damage sustained in the 2011 Christchurch Earthquake. As part of the whole exhibit there will be a health section, and in this section ScienceAlive! were looking to include an interactive teaching tool that children could physically interact with, which would help them learn and to complement the other health activities. ScienceAlive! completed research into the different solutions, and decided upon a mannequin and touch screen combination where the screen would move across the body and act as a "body scanner". This was inspired by Computed Tomography (CT) and Magnetic Resonance Imaging (MRI) technology. The aim of this thesis was to elaborate and refine this idea, and design and develop the interface and underlying software. The hardware aspect impacted the design of the software, therefore the hardware was designed and developed along side the software. The software needed to allow users to view different body systems, involve a good user experience for the younger audience, and allow the users to process and retain the information presented as the exhibit is primarily a learning tool.

1.1 Requirements Analysis

The main goal of this research was to develop and build a system that allowed children and young adults to interact and learn about some of the body systems. The secondary goal was to find an interesting and fun way to teach this younger audience about the body. Currently there is no system that specifically targets this age group for learning body systems in a museum environment. There were a few different technological ideas that could be explored for the software design, including Augmented Reality (AR), games, and website based designs where applications were used.

To create a solution, there were several problems which needed to be addressed. The main problem was:

Interactively teaching young users in a fun way about the body systems in a museum environment using a body scanning idea.

This was the main problem that ScienceAlive! and this project were focused on. The goal of ScienceAlive! is to promote the value of science and technology through interactive experiences. This meant the solution to this problem needed to be interactive and work in such a way that the users enjoyed their learning experience. The solution needed to be easy for the younger audience to use, and take into account the different factors that were involved with this target audience. For example: height, attention span, technological knowledge, and how many users would be around the display at one time.

These are the main requirements obtained through discussion with ScienceAlive!:

The solution needed to be robust and maintainable.

Due to the project solution being part of a museum exhibit, it needed to be able to handle large volumes of use. It also needed to be easily maintained, switching power on and off with very little to no maintenance required. The technology and solution needed to be reliable.

It needed to be able to cater to groups as well as individuals.

People who visit museums do not always move around the exhibits by themselves. The solution needed to be viewable by multiple people at once. This way if people stayed longer, it would not stop the flow of people through the exhibit, as multiple people could experience it and then move on.

It needed to build on or teach the body systems that children in the age range of 8 to 15 years are taught in school.

Curriculum varies from school to school, therefore the project would be basic enough to teach beginners while including information for older users, high school onwards. The end goal was to have an exhibit that complements knowledge that the user already had, using it as scaffolds to increase their knowledge on the body, or to help create an interest and desire to learn more in users who had very limited knowledge about the body.

2 Background

Due to the wide range of topics in this project, research areas have been split into sections. This allowed a review of different aspects that were investigated for possible implementation solutions including: Augmented reality and learning; Learning body systems with current technology; Learning with traditional methods, interaction, animation and multimedia; Children, museums and interactive displays; Intuitive interaction design for children.

2.1 Augmented Reality and Learning

AR is a technology that has become a popular learning tool and allows for increased interaction with the world around us. AR is where virtual objects are superimposed over the real world (Azuma, 1997), either using a tangible interface like a tablet or spatially through methods such as projection. With AR we can now use a new method of learning that allows interactivity as well as more immersion than learning from books. AR captures reality and depicts it with augmented information such as shown in Figure 1, where a mobile phone is used to overlay real-estate information over a building.

Figure 1: Using Augmented Reality with mobile phone to view building information - retrieved from Google Images

AR and learning has a reasonable amount of research behind it. AR is advantageous as it enhances perception and interaction through conveying more information that would otherwise not be viewable (Azuma, 1997). Chien et al. (2010) found that AR technology helps learning with visual support and enhancing spatial memory. Spatial memory is the ability that humans have to generate, retain, retrieve, and transform our visual environment (Korakakis et al., 2009), and it is a necessary feature that we use with learning. An example of this includes being able to navigate through and remember different sections of the skeleton. There is disagreement on the improved spatial memory in 3D, as some studies have shown no difference when research was repeated (Cockburn, 2004), whereas later studies suggest positive results (Chien et al., 2010).

Research has found both positive and negative learning effects associated with AR. Positive effects can include increased content learning such as spatial structure, better long-term memory retention, increased student motivation, and improved collaboration (Radu, 2012). Due to this, the use of AR technology in this research would be appropriate, as collaboration can increase the number of users interacting with the exhibit at one time and increased retention suggests that more of the information learned would be retained long term. Learning and retention are both important factors in this project and could have practical use outside of museum exhibits, while collaboration could also be beneficial.

Negative AR effects included attention tunnelling, usability difficulties, and learner differences (Radu, 2012). Attention tunnelling would not be a problem in this research as focus on one aspect could increase learning in that area, although usability difficulties and learner differences could all impact the children's learning and retention levels.

Studies show that most AR systems reduce cognitive load and encourage exploration while still being able to highlight important features (Radu, 2012). Exploration can be a successful way to capture a younger audience's attention. Exploration, games, and activities that use fun and interaction to teach were incorporated into this project to try to increase fun, learning, and retention. The AR concept can be used to entice children to look into an area more deeply (Billinghurst and Duenser, 2012). When interacting at a deeper level, the user's attention will be held, helping the learning process. All these factors suggested that developing BRET with AR technology could be beneficial to the overall goals of this research, while keeping in mind that conventional learning methods work well, therefore complementing traditional learning with AR instead of replacing them (Asai et al., 2005). Use of markers and AR technology was one possible solution for displaying the body system models, another was through displaying information over the mannequin on a transparent screen.

When designing AR systems there are normally two approaches, optical versus video blending (Azuma, 1997). Optical uses a see-through device, where the virtual image is displayed in the foreground, and video is where the real and virtual images are blended together then displayed to the user (Azuma, 1997). Each type encounters problems, for example, in video blending a camera is required, which can introduce image distortion (Azuma, 1997). Just as there are multiple image approaches, there are also different forms of devices that can be used, for example in AR you can use hand held devices such as mobile phones and iPads, head mounted displays (HMD), or projection mapping. Projection mapping is where the virtual objects are projected on to the real environment. Due to the possible dimensions and limited spce between the mannequin and display screen, there would not have been enough room to implement projection mapping.

The form of display has positives and negatives depending on the main use and environment. These factors mean that when designing a system, you need to take into account the tools, what the system's main purpose will be, who the users are, and how long it will be used, among many other factors (Figueroa et al., 2005). For instance, hand held devices have been found to be more suitable than the HMD for a presentation tool of augmented instructions and long term use(Asai et al., 2005). Positives for HMDs can include the level of immersion, but this can also cause simulation sickness. HMDs do restrict use to one person, which would not be suitable for a museum environment, therefore this form of AR was not included in the design.

2.2 Learning Body Systems with Current Technology

An influx of technology available online helps teach biology through text, images and animation, to online tools such as websites and applications. One of the most difficult tasks in learning anatomy is comprehending the visual spatial relations of anatomical structures (Temkin et al., 2002). This is where AR and other 3D modelling applications can help make spatial learning easier. Juan et al. (2008) developed an AR system using a HMD and a tangible interface where the user opens zips to view organs. This system is for one user at a time and only shows parts of the interior of the human body, whereas this research was developing a solution for multiple users to see the whole of the body. Results from Juan et al. (2008)'s study suggested that the children liked this form of interactive learning, which had positive implications for using an AR interactive system with children. Tangible AR has been found to be particularly suited to museums (Sinclair and Martinez, 2004). In this research, the interface was designed to be tangible to allow interaction with the mannequin.

Chien et al. (2010) have also created an interactive AR system for medical students to learn about the skull in detail. Despite a similar subject area of studying human anatomy, the audience in this research is younger and will be focused on basic body system knowledge. Unfortunately, this paper was lacking results or a discussion, therefore it was difficult to pinpoint areas of interest and commonality.

AR has been useful to doctors as a visualising and training tool, for instance allowing the doctors to see virtual instructions and objects (i.e. organs) that identify medical steps without having to look away from the patient (Azuma, 1997). This technology has made a significant contribution to medical education and training (Chien et al., 2010).

There are also online tools which are run on desktops or ipads. Chien et al. (2010) describe “The Visible Body”, which is a 3D visualisation tool used by healthcare professionals, patients, and students, as well as the “Bio Digital Human” which is a virtual 3D body tool for education, personal or business use. These tools provide very similar functionality to a similar audience group as in this research.

Web-based Three Dimensional Virtual Body Structures (W3D-VBS) is another online medical tool, where users don’t need medical background or knowledge (Temkin et al., 2002). This aligns well with this research, as they both aim to provide a learning tool that did not require the users to have prior knowledge, or allowed users to have only basic knowledge from which to build on.

Figure 2: Anatomage table used for medical learning

Figure 2 shows the Anatomage Table developed by Anatomage USA for medical purposes. It includes two large touch screens that display detailed 3D images that can be interacted with, to the point where dissections can even be practised (Anatomage USA, 2005). It was developed for medical students and professionals, therefore was far too complex for what this project covered. Similarities between the Anatomage and BRET included that they both allowed viewing from large groups, are interactive, and are used for learning purposes.

The different applications and AR tools above gave ideas into the development of the software for BRET. Although each had positives and negatives that needed to be taken into account in the design of the solution to reduce the negative factors.

2.3 Learning With Traditional Methods, Interaction, Animation and Multimedia

Learning is a complex topic that involves many different variables. There are two different ways that learning can occur, actively or passively. Passive learning is when the learner receives new information but does not engage, whereas active learning is when the learner actively engages in the material. Learning has been found to have better success when the student actively learns, although an overload of information can be bad (Meo et al., 2013). People learn at different rates, therefore leaving the control of the learning to the students allows for them to process at their individual pace (Korakakis et al., 2009). Due to the range of ages in this project, this needed to be taken into account. The content and interaction also needed to be interesting enough to actively engage the user while allowing them to move through at their own pace. Age differences in children also show a wider range of variability in task performance when compared with adults (Korakakis et al., 2009). This needed to be factored in when evaluation took place.

When learning new things, schema are used. Schema are how information is organised and related to things that are already known. This determines how knowledge is structured, and how old knowledge is used to interpret and process new knowledge (Park and Michael, 1991). Knowledge is best integrated when unfamiliar concepts can be related to familiar concepts (Park and Michael, 1991). This means that knowledge that can not be integrated is harder to understand, process, and recall as current schema can not be used.

This explains why users find interfaces easier to process and understand when they are associated with visual and procedural metaphors (Park and Michael, 1991). The design becomes more intuitive because if the users can relate it to schema, they can already interpret how the system should act and react to their interactions.

Learning now uses a variety of different methods, from text to images to animations. There are positive and negative factors to be taken into account for every learning method. For instance, when retaining information over a longer period of time, text has been shown to be better than audio (Najjar, 1998).

Computers have forever changed the way information is conveyed and retained. We have computers from primary school through to university, as well as in most careers, and it has been found that having computer technology in classrooms can enrich teaching and learning (Billinghurst and Duenser, 2012). Studies also show that interactive and computer based learning can reduce learning time while still being as effective as traditional methods (Dewhurst and Williams, 1998). From this it is no surprise that as technology advances, so does learning methods, from computer based learning, to multimedia, animation and interactive tools. This has helped to increase interest of students while making the material more appealing (Korakakis et al., 2009).

With the influx of animation in learning material, studies have found that there are both positives and negatives to their use. Animations can either facilitate or hinder learning depending on the person's spatial abilities (Münzer et al., 2009). They can provide external support for internal visual-spatial processing (Münzer et al., 2009). This reduces the cognitive load of the learner, therefore making the learning process easier. Although animations can also hinder learning because they require visual short term attention and retention (Münzer et al., 2009). If the person does not have very adept short term memory or attention spans, they can miss important details, therefore hindering their learning due to not gaining a full understanding of the topic being presented.

Location is also very important aspect of learning. The location should determine to some degree how the information is portrayed, which factors into the usability of solutions as well. Both internal and external constraints need to be satisfied in order for a solution to work well and fit with the location (Cuendet et al., 2013). In a primary and secondary school study, five principles for designing a learning environment, such as a classroom, where found. These included integration, awareness, empowerment, flexibility and minimalism (Cuendet et al., 2013). As museums are also learning environments, these are principles that were very beneficial. Integration is where the solution is not stand alone, but fits into the topics that are being covered in class, or for a museum, it is one part of an exhibit that has multiple facets that all connect to convey information about the general topic. Awareness is how aware the teacher is of the student (Cuendet et al., 2013), therefore was not important for this project as children can move around exhibits by themselves or with others. Cuendet et al. (2013)'s empowerment is in relation to the teacher keeping a central role in the learning. This research took empowerment more literally, where the users were empowered to explore, learn and grow their knowledge themselves. Flexibility again is more for a classroom environment in terms of time factors and unexpected events (Cuendet et al., 2013). Time was not big factor as this project's solution was not complex, and allowed multiple users to view the display while one user interacted with it. Designing for a museum environment also meant that time was not able to be factored in as different users spend different periods of time at each section in the exhibits.

2.4 Children, Museums and Interactive Displays

The learning environment in museums differs to those in classrooms. People move from section to section inside exhibits, learning at their own pace and in their own way. For children in particular, holding attention is difficult. This is where the type of exhibit can be very helpful. Research showed that hands on experience is crucial for children to learn, and that visitors spend more time with interactive exhibits, which are associated with better learning and recall of information (Van Schijndel et al., 2010). Interactive exhibits can be a great way of grabbing the attention of children, helping them to engage more actively which also increases the chance of learning. Children learn differently to adults, they use games, observations and senses as they gain knowledge (Unal, 2012). Integration of activities and games in the exhibits can help the learning process in younger audiences. This was taken into account by including games and interactive animations to help the users learning process and increase enjoyment. Any educational activities that include sensory explanation, role playing, storytelling, creative drama and other methods work well for children 6 to 11 years old, ensuring that they learn effectively (Maccario, 2012).

Learning in informal spaces, such as museums, is fluid, sporadic, social and participant driven (Yoon et al., 2012). There are many different factors that result in learning to different degrees. One of these is collaboration. Collaboration is a prominent feature in building knowledge and understanding, allowing children to reflect more deeply on what they know (Yoon et al., 2012). Children can share ideas, working together for solutions in different ways then they may have done if by themselves. Although collaboration does not just have to involve other people of similar ages. Children can also collaborate with their parents. Children that engage in exhibits with parents tend to explore longer, broader and deeper than those that explore by themselves or in peers (Van Schijndel et al., 2010).

AR and digital devices have also been found to improve engagement, enhancing museum experience while increasing conceptual knowledge (Yoon et al., 2012).

2.5 Intuitive Interactive Design For Children

Learning for children is different to how adults learn, for example a person's brain is not fully developed until around early adulthood (Korakakis et al., 2009). This is definitely an important factor that needed to be considered, as most products similar to this project were designed for an adult audience, therefore they could only be used for guidance to a certain degree. With such a young age group, this project needed to focus on simplicity and basics while being intuitive and as natural as possible.

There are three stages of learning, two in particular that this research focused on. Ages 7 to 11 is called the concrete operational stage, where the children are old enough to use relatively sophisticated software, but young enough to appreciate a playful approach (Bruckman and Bandlow (2007);Markopoulos et al. (2008)). Ages 11 and up is called the formal operational stage, where designers can assume the child's thinking is similar to that of adults, but asking them to interact with technology designed for younger children can be seen as an affront or as boring (Bruckman and Bandlow (2007);Markopoulos et al. (2008)). As this research included 8 to 15 year olds, these were the two stages being designed for. Factoring in limitations or differences in the way interactions needed to be approached was very important. It was also important for the solution to be simple enough for a younger audience to enjoy, while still creating challenge and information that a more mature audience would enjoy.

Children have different design specifications that need to be taken into account. For instance, reading should involve larger fonts, such as 14pt over 12pt, and simpler words (Bruckman and Bandlow, 2007). The font in particular was a factor this project took into account. As children will not necessarily be right in front of the display, it needed to be easily discernible to those reading from a small distance away. Children's background knowledge also needed to be taken into account. Metaphors need to either be familiar to children or made clear and consistent (Bruckman and Bandlow, 2007). For instance, a child might not be aware that a picture of a folder means that if you go into the folder you will find documents. Dexterity was not taken into account as much, as a touch display device was used. If a mouse was to be used, the fact that children can use point and click easier than drag and drop and that some children have problems using a mouse would have needed to be factored in (Bruckman and Bandlow, 2007). Interaction design studies have found that giving children multiple input devices increases satisfaction, productivity and collaboration (Bruckman and Bandlow, 2007). In terms of interaction, the software component of this research used "injections" and the idea of feeding the mannequin.

3 Design

Designs were created and implemented in each iteration of this research, and the overall process and rough schedule are described below.

3.1 Design Process

The design process used in this research was adapted from Bruckman and Bandlow (2007)'s educational design process. The modified steps included:

- Requirements analysis → where ScienceAlive! requirements were gathered
- Tool and Learning Selection → where the media, technology, and pedagogy were evaluated and decided on
- Iteration 1 → including prototyping, design, implementation and business evaluation
- Iteration 2 → design change implementation and observational testing with children

- Iteration 3 → final design implementation and final evaluation with children on learning methods

This particular design process was chosen over basic waterfall, extreme programming, or other agile methodologies because it kept the fact that the project is based as a learning tool. The environment, users, and purpose are all based around learning, therefore it was believed that this should hold focus or at least be an important influence in the methodology.

This methodology is one of learner-centred design, and in Bruckman and Bandlow (2007)'s study they took this approach and adapted it to suit their needs. This methodology was taken and adapted to suit the needs of the this research.

The needs analysis, in this case requirements, were discussed and agreed upon with ScienceAlive! (see Section 4.1.2 and Section 4.1.4). This was because they were the main stakeholders, and the main design was to tie in with their health exhibit. Further requirements centered around the users would be taken into account after observational testing occurred.

The tools and learning types are covered in Section 4. These decisions were based off of research into tool capabilities, the learning curves, what was required for the project to succeed, and the types of learning required.

Prototyping in this case revolved around initial paper prototypes and design concepts, working out all of the data that had to be displayed at what point, as well as how the system would be assessed (which is covered in Section 7.4). Prototyping and design changes occurred in each iteration.

The first iteration's business evaluation and second iteration's observational evaluation cover Bruckman and Bandlow (2007)'s usability evaluation. The third iteration's final evaluation covered the summative evaluation that Bruckman and Bandlow (2007) said helps document the effectiveness of the design, where most researchers use quantitative and qualitative methods.

3.2 Schedule

This section outlines a rough schedule that was followed in order to complete the research within the time allowed. The schedule began on May 27th 2013 and continued until 16th February 2014. There were a total of 34 workable weeks.

Month	Task	Duration
May - Jun	Lit Review, children's body knowledge, evaluation methods	2 Weeks
Jun	Requirements, design process, design ideas	1 Week
	Meeting ScienceAlive!, research software tools	2 Weeks
Jul	Prototype interface and interaction	1 Week
Jul	Evaluation design, ethics submission	3 Weeks
Aug - Sept	Design and build interface and interaction	8 Weeks
Oct	Iterative Business evaluation, refinement and fixing bugs	4 Weeks
Nov	Setup and complete evaluations	4 Weeks
Dec	Data analysis	2 Weeks
Jan - Feb	Thesis write up and submission	6 Weeks

4 Tool and Learning Selection

In this section the beginning tools and learning methods were selected and evaluated, giving a starting point in development.

4.1 Prototype Tool Components

There were two components to the prototype that were delivered at the end of this research. There was a software component and a hardware component. This research was mainly focused on the former, but also involved helping ScienceAlive! create a hardware solution to work with the software.

4.1.1 Software Design

There were several options for environments to develop the software in. These included Flash, C++ using the OpenFrameworks library, and the Unity game engine. The main tool for the software component needed to fit certain requirements. Below is a table showing a comparison of the different tools, the requirements, and how the tools measured up to the criteria required.

Requirements	Flash	C++ Openframeworks	Unity
Basic GUI* features such as buttons	✓	✓	✓
Allows advanced aesthetic design	✓	×	✓
3D modeling	×**	✓	✓
Learning curve	High	Med	Med
Ease of use in general	Med	Med	Med
Ease of use for research requirements	Hard	Hard	Med

* GUI - Graphical User Interface
** flash was not designed for use with 3D modelling, instead based off of 2D modelling, although has added features now that include 3D usage.

After completing this evaluation from testing and anecdotal evidence, Unity was the tool that was finally selected to use in the main building of the software.

Unity is a game engine that has the ability to make a GUI, work with 3D models and animation, and be exported as a standalone application on different platforms. There was also more support in the form of comprehensive documentation and tutorials, an active community, and people available that had some experience with this product.

For this project, visual content in the form of 3D models and animations was required. The table below shows the comparison of various tools designed for content creation.

Requirements	Flash	Blender	3D Max	Maya
3D Modeling	×	✓	✓	✓
Particles	×	✓	✓	✓
Support	Medium	Medium	Medium	Medium
Animation	✓	✓	✓	✓
Learning Curve	High	High	High	Medium

Animation and 3D modelling was completed in Maya then imported into Unity. Maya was chosen because it fulfilled all of the basic technical requirements needed for 3D modelling, and there was information available about particle flow using this software that was required for blood flow simulation. There was also a comprehensive text book on the program.

4.1.2 Software Requirements

The software requirements of this project included:

- Interaction for each system
- Learning/teaching tool
- Information on body systems
- Use of images and animation
- Displaying images in relation to where the display is over the mannequin

Due to the research time frame, the body systems included in the application were:

- Skeletal System
- Digestive System
- Circulatory System

These three systems were selected from systems taught in schools (mentioned in Section 4.2). The skeletal, digestive, and circulatory systems were viewed as some of the main body systems and would tie in well with ideas that ScienceAlive! had for other parts of the health exhibit.

Interactive elements were added into the circulatory and digestive systems. In the circulatory system, the user could inject a radioactive dye into the mannequin through selecting a button, and then the touch device would display the path the liquid travelled around the body from that point. In the digestion system, the user could feed the mannequin, and if in the digestive system, then the user could watch the food work its way through the body.

4.1.3 Hardware Design

The hardware component was researched and developed with ScienceAlive!, the idea included a pod-like chamber that a mannequin would lie inside. The goal was to make the mannequin realistic including a heart beat, which is shown in the software. To ensure users could not manually manipulate the mannequin, a clear plastic cover was to be added. A touch device was attached to the side of the pod allowing it to move along, giving the scanning functionality. For this research, the final prototype included the basic mannequin with the sliding interface. Additional aesthetic design was to be conducted at a later date.

The goal was to allow the hardware to interact with the software component, while adding a fun Sci Fi experience through the aesthetic appeal of the device. Figures 3 to 4 are some of the inspirations for where the design of BRET came from.

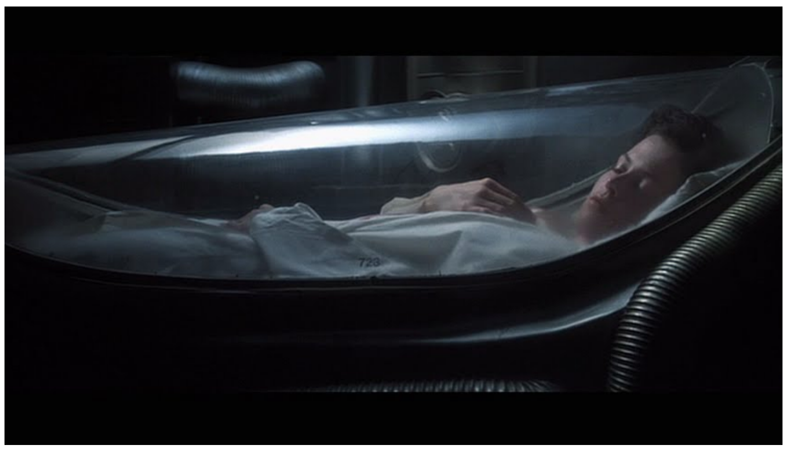

Figure 3: Hypersleep pod from Aliens movie - retrieved from Google Images

Figure 4: Hypersleep pod from Promethius movie - retrieved from Google Images

The prototype was to look similar to Figure 5, being relatively basic but also giving the impression of a medical table.

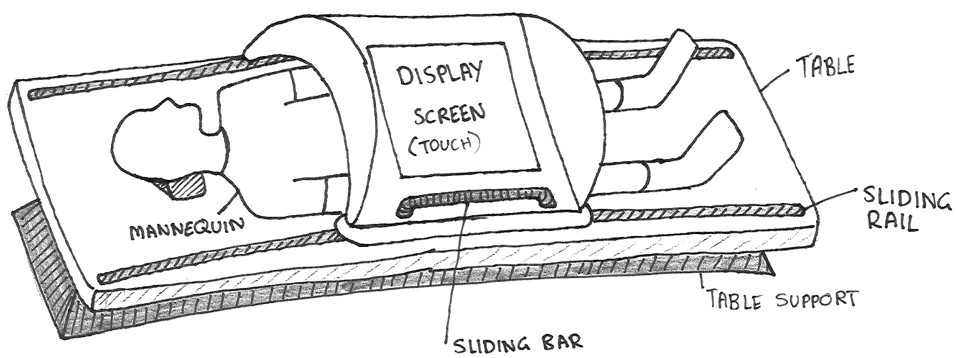

Figure 5: Sketch of prototype design

4.1.4 Hardware Requirements

The hardware requirements of this project included:

- Mannequin
- Slide-able touch display panel
- Absolute rotary encoder for display position
- Information poster

4.2 Learning Styles

Due to the fact that the solution would be a learning tool, pedagogy was an important aspect that needed to be explored and be factored in to the design. The pedagogy of anatomy varies between countries, schools and ages. In New Zealand there is a program called the Life Education Trust bus that runs programs for primary and intermediate schools. It teaches 5 to 12 year olds in the basics of muscles, bones, heart, brain, respiratory, circulatory, digestive and nervous systems as well as other health related topics (Life Education Trust, 2010).

The New Zealand curriculum standards have one section called Health and Physical Education. One strand under this section is called Personal Health and Physical Development where students develop knowledge, understanding skills and attitudes that they need in order to maintain and enhance their personal well-being and physical development (The New Zealand Government, 2007). It is under this section that the body systems fall, although again the pedagogy varies, and is more directed by the children than the teachers, as explained by an expert in the area of teaching, Heather Bartlet.

Body systems are generally an area for older students and can be quite complex and tedious to learn, due to human biology being a very complex subject, as shown by Figure 6[1].

[1](Anatomage USA, 2005)

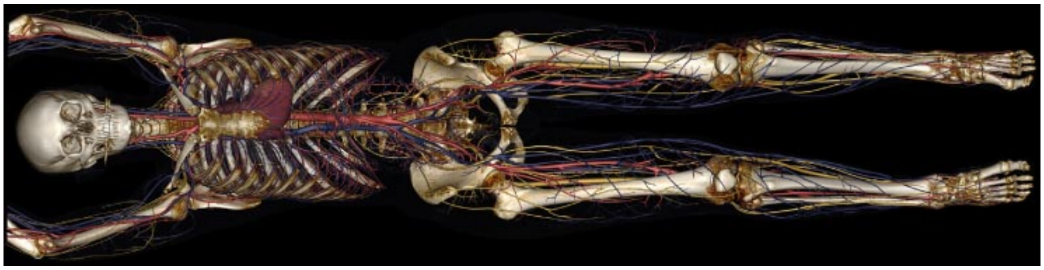

Figure 6: Image of the skeleton with the circulatory system

In order to line up with the information taught in school and in this age range, the teaching method and information taught through BRET needed to be relatively simple, easy to understand, and cover some of the body systems mentioned above.

The pedagogy approach followed in this research was discovery learning and game-based learning. This was due to the users learning body system information through interaction with BRET. Game-based learning was achieved with a bone game, while discovery learning was achieved through selecting the different system parts and reading the information that was displayed.

5 Iteration One

5.1 Initial Design

After investigation into the different software and hardware solutions, several design decisions were made. AR was one of the possible development solutions. One reason for investigating AR was due to the fact that AR applications have shown positive results in classroom learning, but it is thought that they are best used to complement traditional learning (Billinghurst and Duenser, 2012). Looking at the component requirements and the physical space behind the screen of the hardware, HMD and projection AR were not used for BRET.

The second idea was to use a transparent screen for AR. It allowed users to view the body of the mannequin with the additional information transposed over it. It allowed for a more realistic feel than having a normal monitor with a model image of the body. Due to the way the prototype was built, the original transparent LCD touch screen did not work. It required too much lighting to get a clear, visible image. Also, depending on the viewing angle, the discrepancies between the mannequin body and the body system model were too apparent. In the end an LCD touch screen of 22" dimensions and 1680x1050 resolution that was not transparent, was used.

Due to the restrictions mentioned earlier, it was decided that instead of using cameras and AR technology, the images would be within the system and rely on the coordinates of the touch display in relation to the mannequin, simulating a 'Wizard of Oz' situation.

Although AR technology was not used, the technology is similar to AR, and thus there was a possibility that the benefits and limitations of AR in learning would still apply, therefore they were still taken into account. This meant that spatial memory may or may not be enhanced when users interact with BRET. Users can still learn even if spatial memory is not enhanced, but not necessarily to the same degree. If spatial memory was enhanced, BRET could be a valuable learning aid that could be used outside of a museum. If not, the users could still learn the body systems to a degree while enjoying an interactive system, reaching the project's main goal.

Taking AR negative effects into account, the system needed to be developed so that it was easy to use, had a small learning curve, and was an effective teaching method to users of all learning types. Attention tunnelling might not necessarily be a factor in this project as the user would be interacting with the system as a whole.

5.2 ScienceAlive! Evaluation

The first form of evaluations were iterative and completed with help from ScienceAlive!. Business evaluations began with interface design decisions. ScienceAlive! were given a few pages with initial interface design paper prototypes after the requirements discussion. They then selected Figure 7.

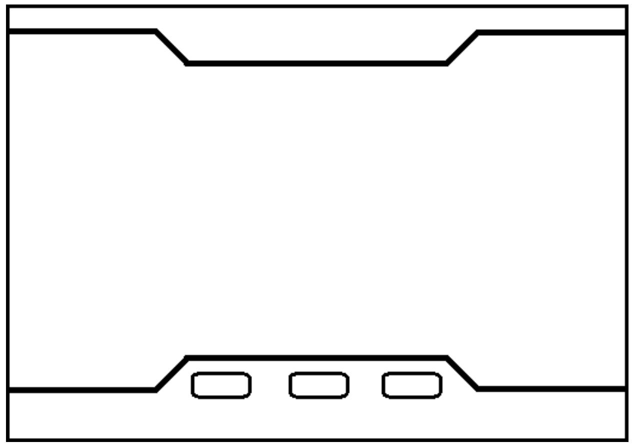

Figure 7: Image of basic paper UI design chosen

Some of the reasoning behind this selection included the fact that the business wanted the interface to be as simple and minimalist as possible so as not to detract from the information being presented. From a design point of view, the interface was to have a Sci Fi feel while still being medical. In this way it would increase interest in the display while suggesting the technology was advanced and "from the future". The hexagon pattern and the blue based colour scheme for lighting and background are all associated with technology and Sci Fi in interface design, especially for games. The interface design iterated with feedback from ScienceAlive! and other designers, moving through many different prototypes. Figure 8, Figure 9 and Figure 10 show some of the interim designs.

Figure 8: Image of first UI design

Figure 9: Image of second UI design

Figure 10: Image of later UI design

6 Iteration Two

6.1 Observational Evaluation

The observational evaluation method would be close observation of some children interacting one on one with the prototype. This ensured that if changes were required after observations were gathered from the interactions, then they would be relatively inexpensive in terms of time, resources and money as it would not require all of BRET to be completely rebuilt. Markopoulos et al. (2008) suggests that early qualitative input by a few children can have a bigger impact on the quality of the product, therefore input before the final stages could be beneficial for BRET.

With the evaluation involving children, the methods needed to be tailored so that they returned useful results. Investigation into methods of evaluation with this age range was completed, as well as finding out the general knowledge base of the body that this age range might have.

During this process there was one administrator watching the child to make sure that they did not act in a way that would cause harm to themselves, or the prototype, while an observer was recording observations, and the child's parent was sitting at the back of the room watching their child.

The Canterbury University Ethics committee required the child's parent or caregiver to be in the room at all times during the experiment. Due to this they were placed at the back of the room where the child could not take visual or verbal cues from them while interacting with the system or answering the questions.

6.1.1 Method

Participants
The participants were 6 children within the ages of 9 to 15 recruited from children of ScienceAlive! employees. There were 3 females and 3 males. The child's parent or guardian also needed to be available to sit in while they participated.

Procedure
How to use BRET was explained to each participant, and they were told that they had a maximum time of 10 minutes to interact with BRET. Afterwards, they took part in a semi-structured interview where they were asked what they liked and what they did not like or thought could be improved. They were also asked about any other factors the observer had noted during the testing.

6.1.2 Results

The results of the observational tests were quite instructive. In total 6 children were tested, from the ages 9, 10, 11, 12, 12 and 15. Overall findings were that the sliding movements of the display was fine, the touch wasn't as intuitive as was expected, most tried to select more than once when it did not appear to work, some children took time to read the information displayed while others did not, the feed and dying interactions were not obvious and most did not watch until the end, the children switched between the body systems a lot and they all explored the different body systems.

When asked what the children liked about the prototype, the main responses were:

- The scanning idea

- The movement of the display over the body, and its correlation to the image displayed
- The touch interface
- The definitions that appear when parts are selected
- The different things that you can do
- The colours
- The game
- The skeleton
- That they learned a lot about those three body systems

When asked if there were any improvements they would like to make the prototype better, the responses were:

- The visibility of fractures in the game
- Pressing did not always work
- More instructions in the game
- More features such as the brain
- Lettering size too small
- The prototype was not tall enough

6.2 Design and Interface Changes

The feedback received meant that changes could be made to improve the interaction and interface of the prototype. The changes that were implemented in response to the testing included:

- The ulna fracture in level three of the game was rotated to be more visible
- All fractures in the game were lit with a soft yellow light to distinguish them from the surrounding bones

- More informative instructions were added to the game
- Instructions on selecting parts to get more information were added to the fracture, skeleton and digestion screens
- Instructions explaining what happens were added to the dying and feeding screens
- Lettering was increased from 16pt to 20pt
- The prototype height was increased by around 30cm

The touch problems were not able to be resolved as this was an issue with the hardware, and the children had to become familiar with how to select things, such as putting more pressure when tapping the display. More features were not included as the three body systems were just the skeleton, circulatory, and digestive system, but this could be addressed in future improvements after this research is completed.

7 Iteration Three

Iteration three covers the final hardware and software specifications, what tools were used in which sections of the project and the functionality and features of the finished result. It also covers the final evaluation, and the results and discussion that resulted from that research.

7.1 Final Design

The final result that was at museum standard is shown in Figure 11 and Figure 12. ScienceAlive! decided that due to money factors, the pod covering would not be included. There were happy with the prototype design and did not think that anything additional needed to be added.

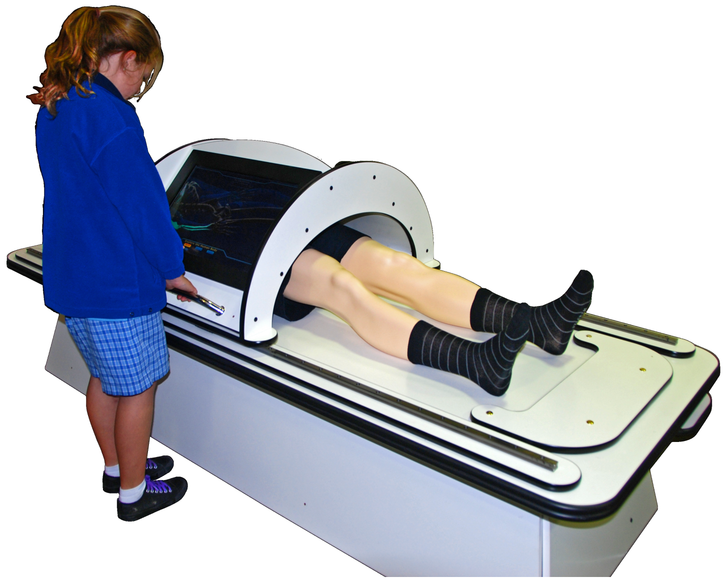

Figure 11: Photo of final Prototype being used

Figure 12: Photo of final Prototype being used

Figure 13 shows the final UI design at the end of the project. This screen is the first visible screen when the project executable is run.

Figure 13: Final UI design

The final design functionality and features for the different body systems included that in each body system, parts of the models were selectable. Once selected, an information box would appear with the part's name and information about it, as shown in Figures 14 , 15 and 16. The part selected is also highlighted a green colour. When the screen moves or another selection is made, the information box disappears or is replaced and the colour returns to normal.

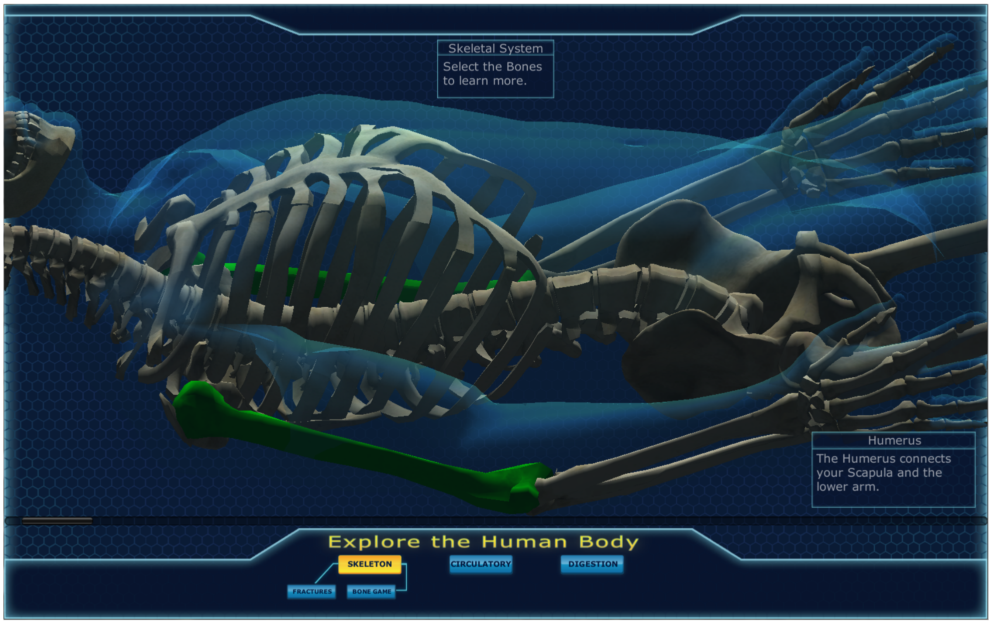

Figure 14: UI showing Skeleton with Humerus selected

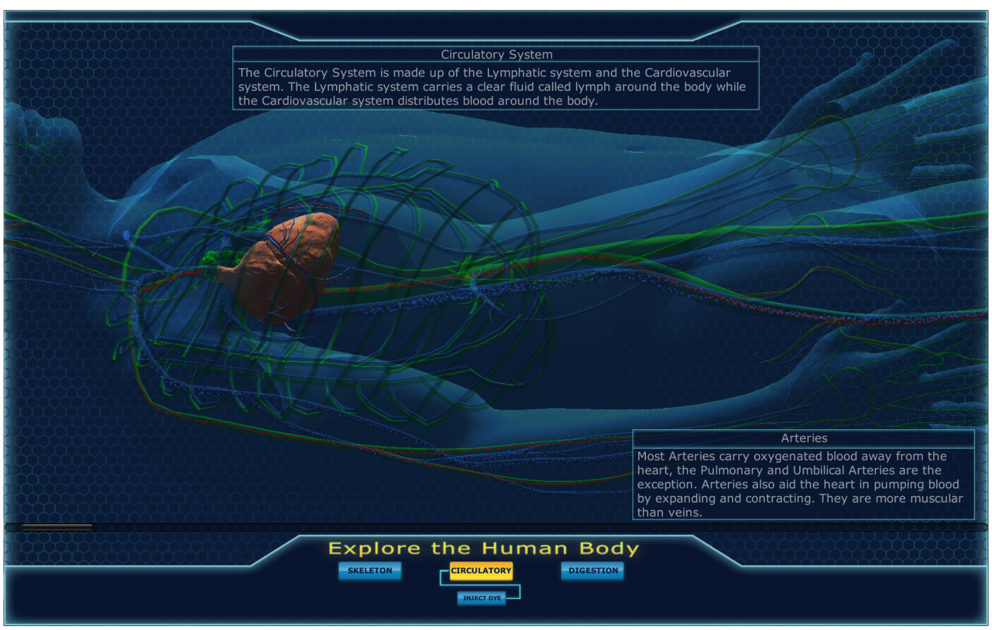

Figure 15: UI showing Circulatory System with Arteries selected

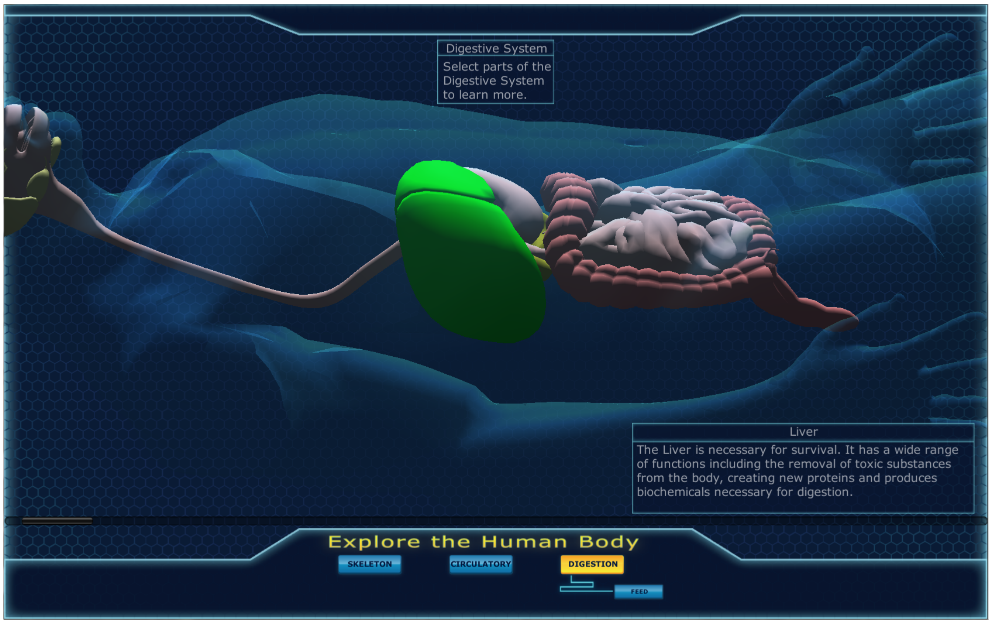

Figure 16: UI showing Digestive System with Liver selected

Skeletal System - Fractures

This button shows four common fractures that children acquire. These include greenstick, torus or buckle, bowing or plastic deformity, and complete displaced fractures. The model shows each of the fractures surrounded by pulsing blue light to draw attention to them. Once selected they display an information box with a title, a definition on that type of fracture, and a closer image of the fracture. This is shown in Figure 17. Research was conducted into common fractures found in children, and due to limiting the complexity of the models, these four fracture types were selected. They were created on bones that are also common for each of those types of fractures.

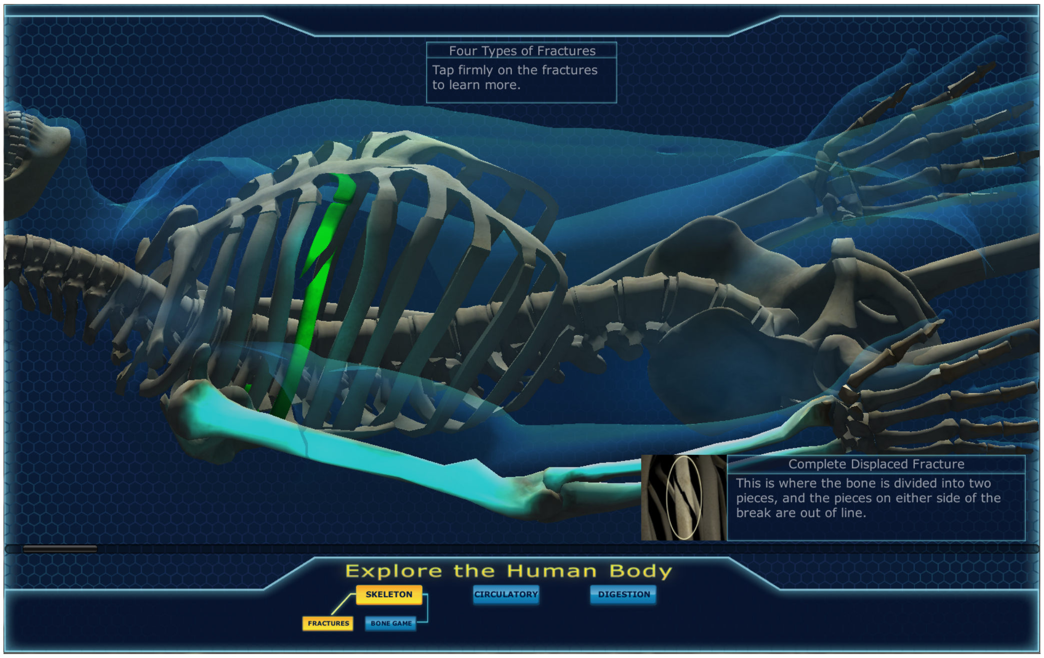

Figure 17: UI showing Fractures with Complete Displaced Fracture selected

Skeletal System - Bone Game

This has three different levels. In each level four bones have fractures that are selectable. The fracture to find is randomly chosen and an information box appears telling the user to find this particular fracture. Each fracture is softly highlighted to make them more visible, as shown in Figure 18.

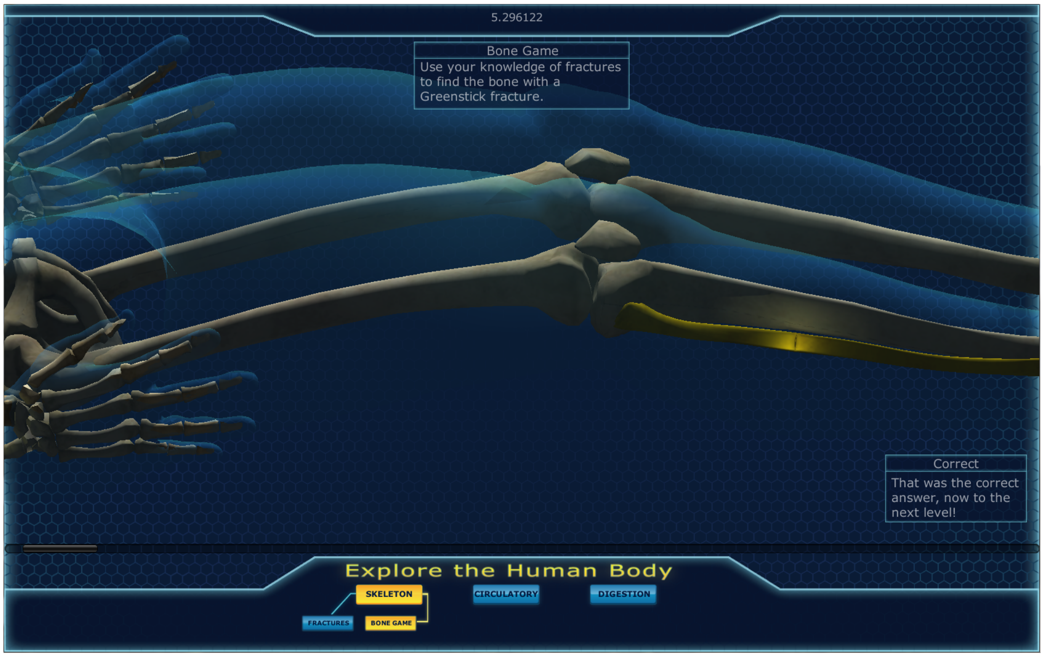

Figure 18: UI showing Bone Game with Greenstick Fracture selected

This was done in response to the observational testing which showed children had difficulty finding the fractures with no lighting. If the incorrect fracture is selected, a message box appears giving the user a clue on what to look for. There is also a timer that times how long the user takes to move through each level. If the correct fracture is selected then a message box informs the user that that is the correct answer and moves them to the next level. At the last level, when the user selects correctly, a message box is displayed saying congratulations and it gives them the time it took for them to finish all of the levels. The interface then moves back to the main skeleton view.

Circulatory System - Inject Dye
This changes the body model to have a white tinge to give the impression that system is in a different view to be able to pick up the dye, as shown in Figure 19.

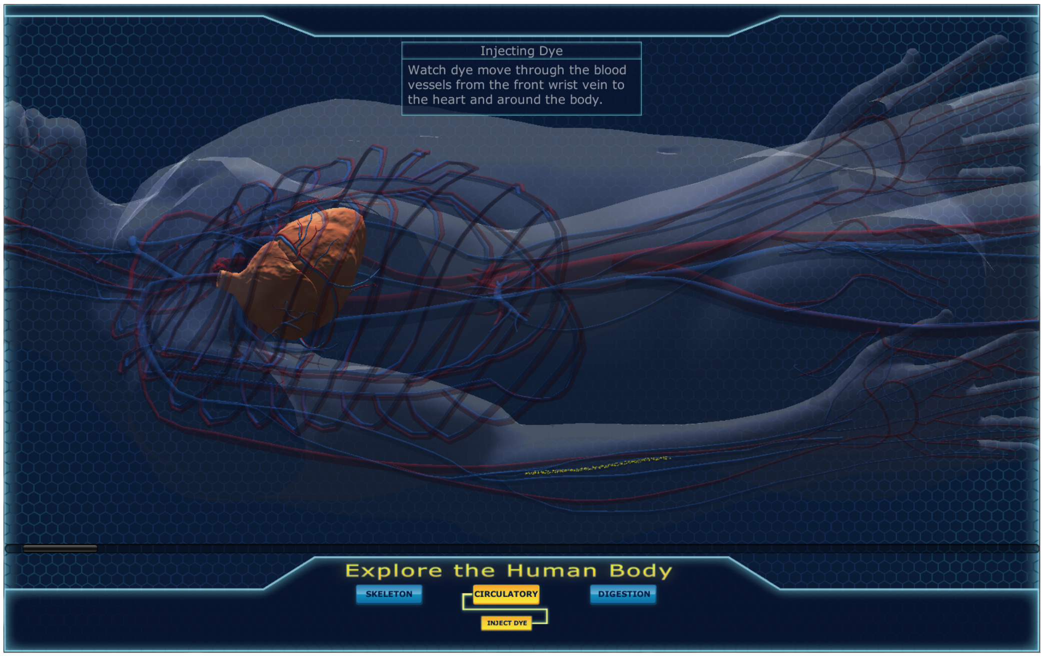

Figure 19: UI showing the start of Dye moving up front arm

The dye, in this case yellow particles, travel up on of the veins in the front arm, all the way to the heart. Once it reaches the heart, the dye is then circulated around the arteries and then back via the veins. When the dye reaches the heart in the end the body model returns to the normal colour and the blood flow is visible again, shown in Figure 20 .

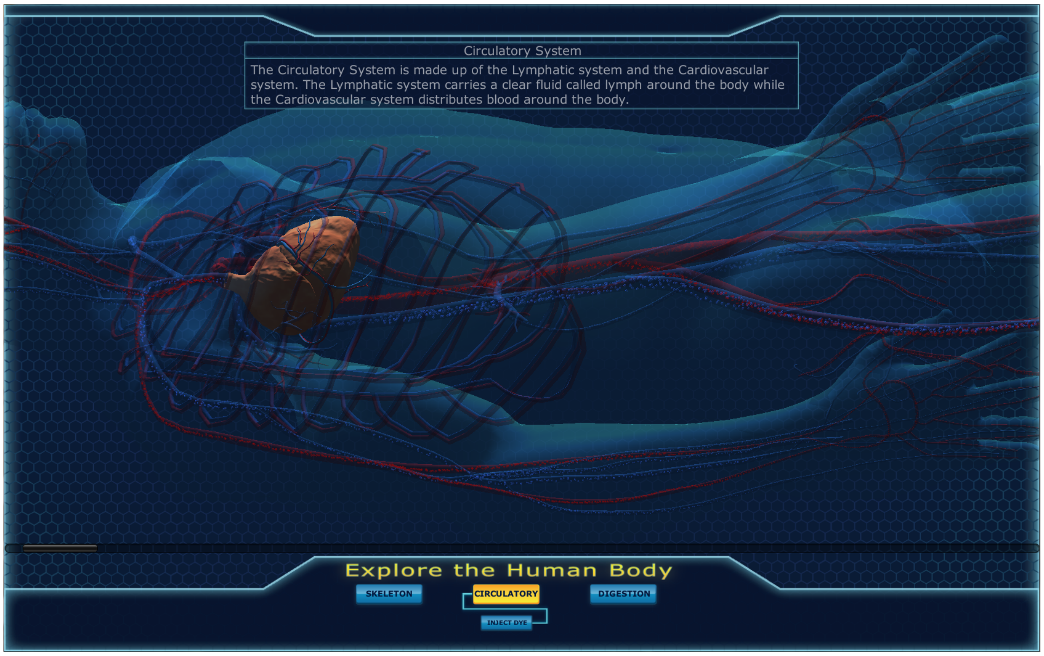

Figure 20: UI showing the Circulatory System with blood flow

The dye was sped up to travel more quickly through the body, decreasing the overall time. This was due to observational testing results that showed the children did not spend very long watching the dye before they got bored and moved on. In both the Circulatory and inject dye modes the heart is animated to pump at around 80 beats per minute, which is within the 60 to 100 range for a normal adult resting heart rate.

Digestive System - Feed
This starts off an animation where particles flow down the esophogus, through to the stomach, then the small intestines, and finally to the large intestines. When the animation is occurring, these parts of the digestive system become transparent so that the yellow particles can be seen, as seen in Figure 21 . After the animation is finished, the model becomes non-transparent again. The particles were designed as yellow to be more visible in their journey through the body.

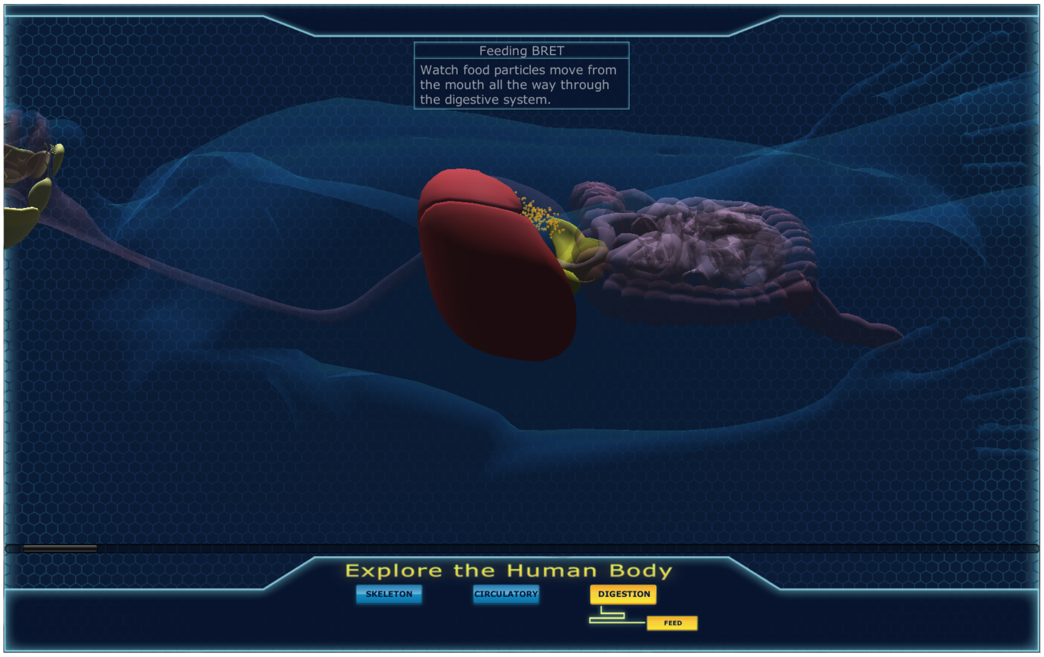

Figure 21: UI showing the Food Particles now in the Stomach

7.2 Final Software

Unity was used which works with scripting in Javascript and C♯. Figures 22 and 23 demonstrate the software development component of the program and the GUI development component respectively. The GUI allows you to place cameras, objects and lighting, while controlling other aspects through scripting.

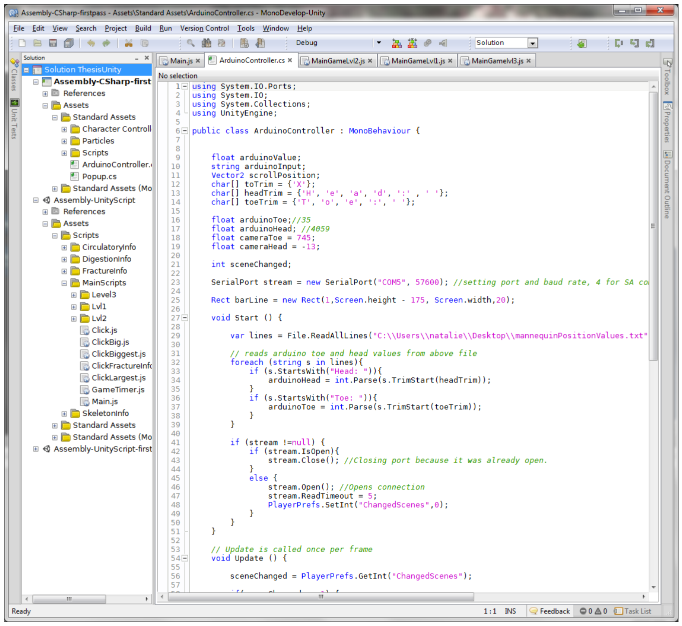

Figure 22: Screenshot of C♯ code for getting the rotary encoder values into Unity

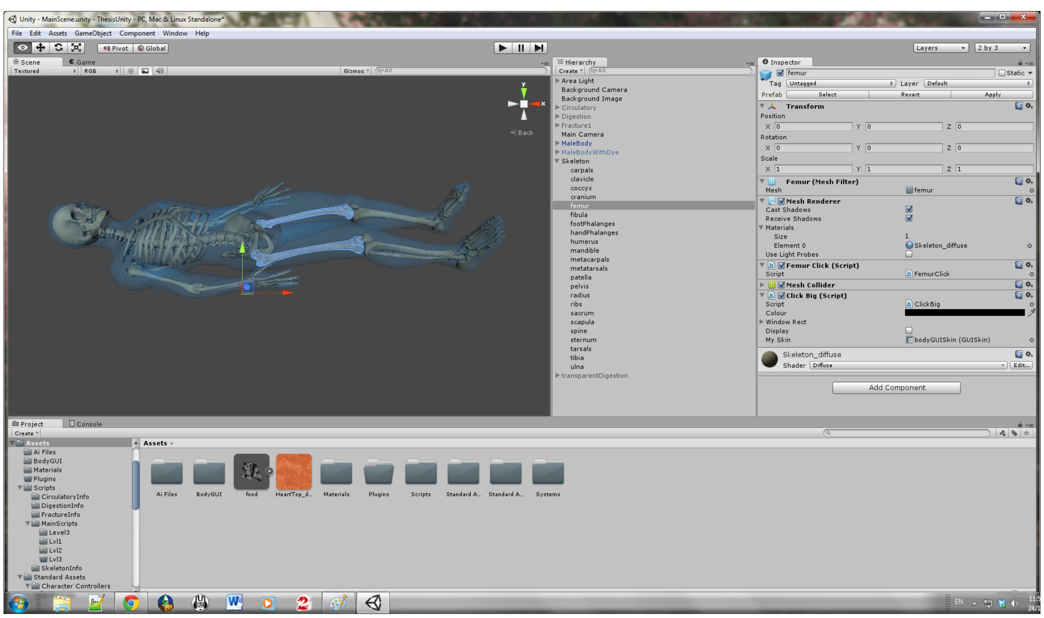

Figure 23: Screenshot of project in Unity

Maya was used to animate the heart, which was then imported back into Unity. It was also used for the skeletal, circulatory, and digestive models. The models were purchased from an online model marketplace TurboSquid[2] and then imported into Maya so that the arms could be moved from a horizontal away from body position to vertically against the sides of the body. Fractures for both the fracture view and for the bone game were also created in Maya. Unity was used to add particles for realistic blood flow, to show the dye movement, as well as food through the digestive system. Maya was to be used for this, but support for importing particle systems from Maya to Unity was not available in the Unity software.

7.3 Final Hardware

The computer specifications needed to be able to have sufficient processing power to work with the display and give good quality output. The main specifications were:

- Intel Core i5-4670 3.4 GHz
- 2 x 4GB RAM Corsair Vengeance
- Samsung 840 Series MZ-7TD250 Solid State Drive 250 GB
- Corsair CX series modulare CX500M, 500W ATX PSU, Active PFC, power supply unit
- GeForce GTX760 Video Card, 2048MB
- Asus Z87M-Plus Motherboard

The computer settings in the BIOS were changed so that if power was lost, when power returned, the computer would return to the last state that it was in. This allowed BRET to be turned off via mains power and to turn on automatically when power was switched back on. On loading the operating system, login was removed so that it would go automatically into the desktop screen. Then all the museum supervisor would need to do is double select the unity BRET executable to run the program. All popups and notifications

[2] http://www.turbosquid.com/

were disabled, and without access to a keyboard, the only way to exit out of the application was to turn off the power. This meant that users would not be able to stop the program and get into any of the main computer systems.

7.4 Final User Evaluation

The second evaluation had three different types of learning for each of the three body systems. So each participant would learn information on the skeletal, digestive and circulatory systems via either text, a short video clip, or through interacting with BRET. Each participant would have a different method but the body systems were counterbalanced to avoid any bias in the results due to order of the body systems. The participants were then sent the last questionnaire to complete through email, one week later. This provided an idea of the possible retention of learned information, as well as how much they learned from one use.

Some aspects that were investigated were the issues of trust and a good testing environment for the participants; how questions were posed and presented to the participants; and how the answers were measured and recorded.

7.4.1 Hypothesis

This study was intended to investigate the relationships between three different methods of learning commonly found in museums (text, video clips, and an interactive display - BRET) and learning, learning retention, and fun. It was also intended to find observational information and data about BRET that could be used to improve the system. The main hypotheses tested were:

- H1: Children who learned with BRET would score higher on the learning retention test than children who learned through text or video clips.
- H2: Children learning with BRET would score higher on the learning test than children who learned through text or video clips.
- H3: Children learning with BRET would find it more fun than those who learned through text of video clips.

7.4.2 Method

Participants
Primary, Intermediate and High schools in Christchurch were approached to

recruit students for participation over the Christmas break. Social media, Canterbury clubs, and the HIT Lab NZ were also used to advertise the need for participants when there was not enough response from the school children. No compensation was advertised to reduce the bribe effect on the children. In the end, 48 participants based in Christchurch and Palmerston North, 24 males and 24 females took part in the study. Their age ranges were from 8 to 15 years old. Participants in Palmerston North were randomly assigned either video or text by gender so that there would be the same number of females and males for each condition. The same number of males and females were also selected for the interactive condition. All participants had the order of the body systems counterbalanced.

Design of the Pre and Post Test for Learning

To determine learning, paper-based pre and post-tests were made from information found within the text, video clips and BRET. The text and video clips were made, and only showed information that was also found when interacting with BRET.

The paper learning tests were the same for pre-test and post-test with a few minor tense changes, for example "how much fun do you think this activity will be to do?" and "how much fun was this activity to do?". The questions included an opinion question on how much knowledge the child thought they had for each system, and a combination of multiple choice and recall questions. They also included a Smileyometer (Figure 24) and "Again Again" table for each system (Read and MacFarlane, 2006), to determine the level of fun experienced and whether the child would like to repeat the activity.

The retention questionnaire asked the child's knowledge level and had the same knowledge questions as the pre and post-test, but did not include the fun Smileyometer or "Again Again" table.

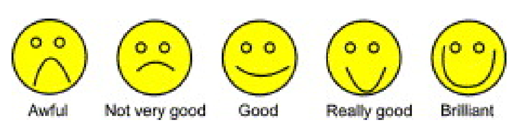

Figure 24: Smileyometer used to record children's fun

As the questions were based for a large age range (8 to 15) they had to be designed so that they were mid-range difficulty (Taylor and Druin, 2002), although reading and understanding was more difficult for the younger children due to the nature of the information. It was difficult to adapt the biology information so that it fit well with all ages tested as some words were quite high level, for example deoxygenated blood. The questionnaire was designed to avoid negative questions, used questions with 3 to 4 options, as well as being as simple, short, and clear as possible. These are some factors that have been suggested as part of completing questionnaires with children (Read and MacFarlane (2006);Bell (2013)).

Design
The design was mixed within-between, where the within factors included three body systems: skeleton, circulatory, and digestive. The between factors included three learning material conditions: text, video, and interactive (BRET).

Materials
A video recorder was used during the interactive condition with BRET, and a log file was also saved. A laptop was used to show the video clips and paper with skeleton information was used for the text condition. All conditions had pre and post-tests. Participants were required to have access to email and a computer one week later.

Procedure
Children in Palmerston North were randomly assigned to one of two conditions (watching video clips, or reading text). This testing either occured at the house of the researcher or the house of the participant. Due to the location of BRET, only Christchurch participants were able to take part in the BRET condition, as they needed to come to the ScienceAlive! building.

Timeslots of 30 minutes were organised where both the parent and child were available, as this has been suggested as a reasonable length for activities with children and fit well with the current activity (Bruckman and Bandlow, 2007). On arrival both adult and child were given information sheets to read over and consent to. The experimenter then gave a short explanation of the condition that the participant was completing. All children learned about the skeletal, circulatory, and digestive systems, but the order was varied using counterbalancing to reduce the effects of the repeated measures design.

After the beginning explanation, the children were given a before knowledge questionnaire to find out how much they already knew about the body systems, and to gauge their fun expectations. The experimenter emphasized to the children that they were helping test the learning material, that they themselves were not being tested (Bruckman and Bandlow, 2007). This was to help decrease anxiety over the questionnaire. Then in the first body system the child either watched, read, or interacted with BRET. When the maximum time of 10 minutes was over, or the child indicated they were finished, they were given the knowledge questionnaire for that body system. This was repeated for all three body systems. The children and parents were then thanked for their participation, and reminded that they would receive the last knowledge questionnaire in an email, a week later.

7.5 Analysis

Learning
The pre-test and post-test resulted in a percentage score for each system: skeletal, circulatory, and digestion. The learning effect was calculated based on the difference between the post and pre test scores. The same was completed with the retention scores, so the retention effect was calculated based on the difference between the retention and pre test scores.

Observed Fun and Usability - BRET
Observational information was taken from the video recordings of the participants where smiles, explanations, or signs of boredom were counted. Interaction data was also recorded to gather information about the system, how much the children explored, and how long they spent actually reading the information displayed.

Reported Fun
The Smileyometers were coded in an ordinal way 1 - 5, where 5 represented brilliant, and 1 was awful. The Again Again tables were also coded in an ordinal way 0 - 2, where 0 indicated they did not want to do the activity again and 2 indicated that they did.

7.6 Results

Observational Data
From observing video footage of the 16 participants who took part in the BRET condition, it was found that 9 participants smiled at least once while

interacting with BRET. Two participants fidgeted a majority of the time, while 5 showed obvious signs of boredom such as yawns or looking away from the display for periods of time, or actually saying that they were bored. Three participants showed signs of confusion through body language or frowns, and 2 showed open exclamation, one being "yes" at finishing the game, and the other saying how cool BRET was. Lastly, all participants showed concentration, in terms of facial expressions and body language, a majority of the time while interacting with BRET.

BRET Interactivity Log Data
The interactivity logs showed that overall the participants selected 60.1 % of the skeletal system, 82 % of the circulatory system 70.6 % of the digestive system and 62.5 % of the fractures. Out of the participants who started the bone game, 79 % finished it.

The average amount of time spent watching the dye was 1 minute 29 seconds, while the whole animation took 6 minutes and 20 seconds. The average time watching the food particles going through the digestive system was 1 minute 7 seconds, while the whole animation took 2 minutes and 24 seconds.

The value for the rotary encoder that is tied to the position of the display over the mannequin starts at 35 at the head of the mannequin, and can be moved to 4059 which is where the toes of the mannequin are. Overall, the slider was moved an average of 2603.4, which is 65 % of the total movement available. The average minimum value was around BRET's ribs at 892.8, while the average maximum value was 3496.2, which corresponds to just above BRET's feet.

Statistical Analysis of Learning Methods
The retention results for the different learning methods are shown in Table 1. An ANOVA test revealed that there was no significant difference between retention and media for the different body systems $F(2,48) = 1.281$, $p = 0.288$.

Table 1. Mean scores for Retention

Media	Skeleton	Circulatory	Digestive
Interactive	0.21	0.23	0.19
Video	0.22	0.26	0.28
Text	0.42	0.13	0.35

The learning effect results are shown in Table 2. An ANOVA test revealed that there was a significant difference between the amount learnt for the different body systems $F(2{,}48) = 3.606$, $p = 0.035$. A post hoc Bonferroni test showed the significant difference was between interactive and text learning ($p = 0.030$), suggesting participants learnt more through the text learning method than the interactive learning method.

Table 2. Mean scores for the pre-test, post-test, and learning effects

Media	System	Pre-test	Post-test	Learning effect
Interactive	Skeleton	0.14	0.35	0.20
	Circulatory	0.13	0.37	0.24
	Digestion	0.26	0.43	0.17
Video	Skeleton	0.21	0.54	0.34
	Circulatory	0.19	0.46	0.27
	Digestion	0.18	0.45	0.27
Text	Skeleton	0.17	0.71	0.54
	Circulatory	0.34	0.54	0.21
	Digestion	0.24	0.64	0.39

Table 3. shows the mean scores for fun before and after the activity was completed, as well as Wilcoxon test significance values. This test shows expected fun ratings for the interaction condition and the digestive system were less than the actual fun ratings (Z = -2.179, $p = 0.029$). This also occurred for the video condition and the skeletal (Z = -2.828, $p = 0.005$) and circulatory systems (Z = -2.449, $p = 0.014$). Lastly the fun score for text was higher than the expected fun in both the skeletal (Z = -2.000, $p = 0.046$) and digestive systems (Z = -2.530, $p = 0.011$). There was no significant difference between the fun scores for the different body systems and learning methods $F(4{,}48) = 2.437$, $p = 0.053$. A post hoc Bonferroni test also confirmed that there was no significant difference.

Table 3. Mean scores for fun responses with Smileyometer and results of Wilcoxon test.

Media	System	Before Use	After Use	p-value
Interactive	Skeleton	3.25	3.50	0.285
	Circulatory	3.12	3.69	0.058
	Digestion	3.25	3.81	0.029
Video	Skeleton	3.13	3.38	0.005
	Circulatory	2.94	3.19	0.014
	Digestion	3.00	3.56	0.157
Text	Skeleton	3.13	3.63	0.046
	Circulatory	3.00	3.38	0.102
	Digestion	3.19	3.44	0.011

The Figures 25, 26, 27, 28, 29, 30, 31, 32, and 33 are regression graphs showing system learning vs age for the different learning methods. This was also completed with retention and fun but the results also were not significant. These graphs were included to show general weak trends to give an idea of the impact of age. Overall in learning vs age, there is a weak positive correlation between the learning score and the participant's age.

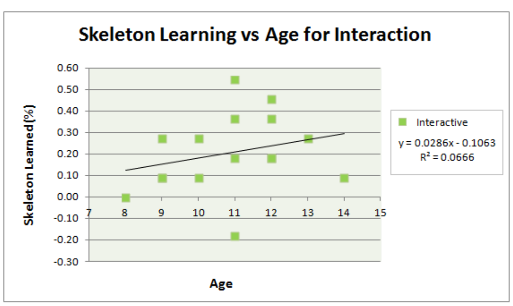

Figure 25: Skeleton Learning vs Age for Interaction

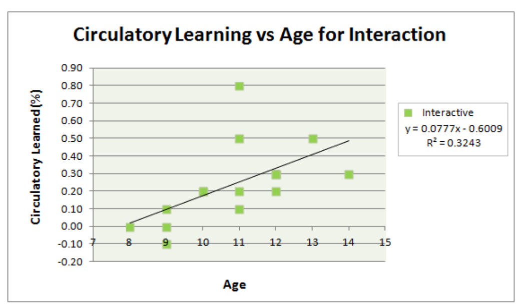

Figure 26: Circulatory Learning vs Age for Interaction

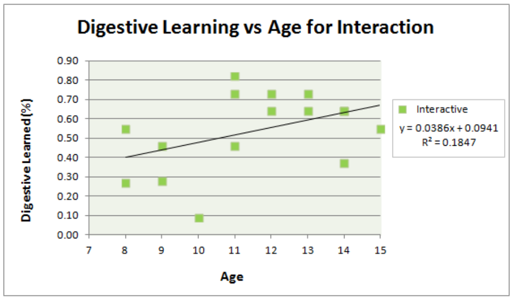

Figure 27: Digestive Learning vs Age for Interaction

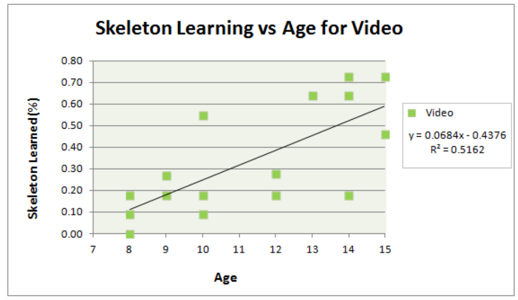

Figure 28: Skeleton Learning vs Age for Video

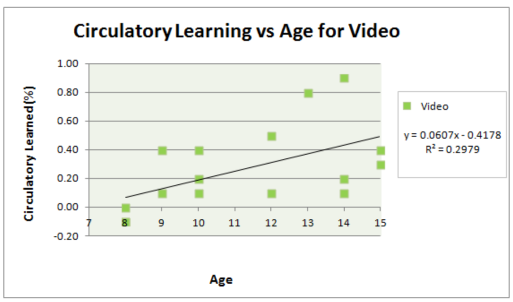

Figure 29: Circulatory Learning vs Age for Video

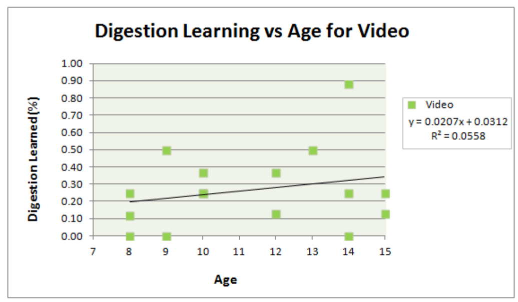

Figure 30: Digestive Learning vs Age for Video

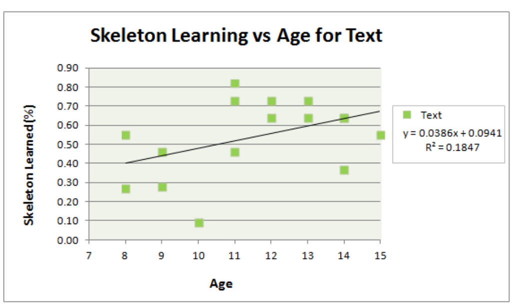

Figure 31: Skeleton Learning vs Age for Text

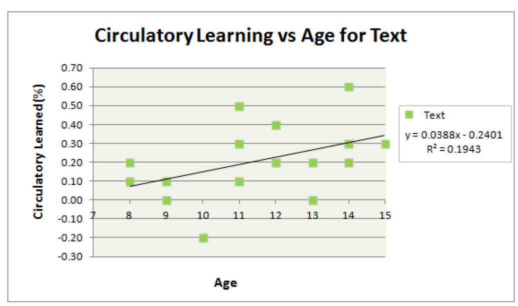

Figure 32: Circulatory Learning vs Age for Text

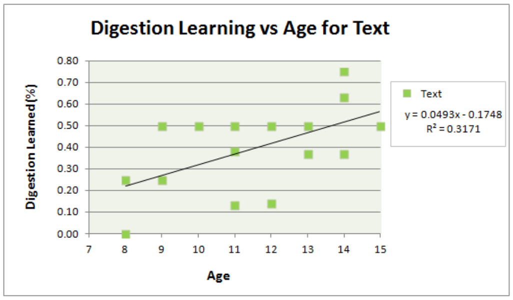

Figure 33: Digestive Learning vs Age for Text

7.7 Discussion

7.7.1 Observational Data

The observational data from interaction with BRET indicates that overall the interactions were positive ones. Notes from observations made during testing can be crucial, as smiling or sighing are often used to judge the quality of interaction (Diah et al., 2010). Only one child openly stated that they were bored, while others remarked after recordings were finished and final testing had concluded, that they thought it was “cool”. The concentration could be from one of two things, either they were actively engaged in the activities, or the information may have been at a high level, requiring more effort to understand. Due to the overall mean post test scores, it is suggested that it may have been both. This conclusion is drawn from the fact that even though the mean scores were low, fluctuating from 17 % to 54 % learned, they were not at the very bottom of the scale. This suggests that although the information may have been difficult to understand, they still managed to process some of it.

Another factor to mention is the fact that some of the participants, during the course of the experiments, asked whether they could scan themselves (or family members). This was taken as a sign of success as it meant that the children truly thought BRET scanned bodies in real time. They thought it was real, which shows that the combination of interface design, movement of the image with sliding the display over the mannequin, and the angle and display of the system model, made the concept believable.

7.7.2 BRET Interactivity Log Data

With around at least three fifths of each system being explored, this leads to a few possible explanations. These could include the ability to actually select the parts, the number of parts selectable, how the children actually viewed the task, and the engagement in the activities.

The gallbladder and pancreas in the digestive system were almost never selected as they were not very visible, hidden behind the stomach and liver. Due to this the stomach was moved left, and the liver right so that you actually see the other two organs and can select them. The anus was also not selected as often as there was less surface area to select.

The lower selection in the skeletal system was probably due to the number of bones actually selectable. There were much more than the circulatory system for instance, which only had 3. This suggests that the children selected a few but then moved on to the other areas, such as the game or fractures. Some of the children approached the activity as a more serious learning task, while others just played around or selected out of interest. This would also have impacted how thoroughly they explored each system.

In terms of engagement in the activities, purely looking at the observational data, it appears that the participants may have had a combination of active and passive engagement, as they showed concentration combined with a high level of exploration. It appears to be a combination of both types of engagement, as the only tasks that force active engagement was the bone game, where they had to process and apply what they had learnt about the fractures. The influence of the knowledge survey after each activity could also have increased their engagement levels. Some of the children may have actively engaged in learning the information displayed while trying to understand it, rather than just reading the information. The combination of both type of engagement is shown through the scores, as if the scores were higher it would imply more active learning (Meo et al., 2013).

Due to 79 % of people finishing the bone game suggests that the lighting and other improvements that were made after observational testing were successful. This implies that although it may still provide a challenge, a majority of the users will be able to successfully play the bone game. One factor to mention was that some participants started playing, then went back to the fractures to read up more, while others did not actually play properly. The latter tapped all of the bones that they could see until they found the bone with the correct fracture. Generally it was the younger participants who were observed doing this, or the ones who did not read the instructions at the top of the game.

The amount of time watching both the dye particles and food particles suggested that they both were too slow, taking too long to complete. The children lost interest in the case of the dye, or could not actually tell what was happening in the case of the food. As a result of this, the dye particles were sped up to only take 1 minute 16 seconds, and the food particle speed was increased to only take a time of 1 minute 9 seconds. The food particles were also changed to a yellow colour and the surrounding digestive sections were made more transparent so that you could actually see the particles all the way from start to finish.

Finally, the amount of movement that the children covered shows that their main areas of focus were from the shoulders to the bottom of the legs. This is understandable as most of the content is visible in this area. The only aspects missed are the skull and the feet, but these are only important in the skeleton view. Therefore the children explored visually most of the models.

7.7.3 Statistical Analysis of Learning Methods

The Questionnaire

After completing the research it was obvious that the content and questions were difficult for the younger participants. Younger participants need clearer parameters to respond to (Diah et al., 2010), and after collating and marking the questionnaires it was obvious that some of the questions were difficult to understand. The way they understood and responded to the questions was different to what would have been expected from an adult. Although children have been known to take things literally, and their responses cannot always be predicted (Read and MacFarlane, 2006). A case of this occurring in the questionnaire was the responses to the question "what is the last part of the digestive system?" a lot of the children responded with anus. This was not what the question was supposed to mean, it was actually asking for the last organ that officially has a function in digestion. The correct answer is the colon/large intestines. If the children read/listened/selected colon/large intestine then they would have found the exact wording "it is the last part of the digestive system". Instead it was found that participants responded with what they already believed was correct logically. In this way the study found that asking good questions was difficult, and that for some children, understanding and interpreting the question can be difficult (Read and MacFarlane, 2006).

One interesting point noted was that most of the children left questions blank when they did not know the answer, even the multiple choice questions. Adults generally try to get the most marks by guessing the answer when they don't know what it is, even with some recall questions. The fact that most of the children did not do this suggests that it is a developmental skill gained later on in high school. This makes sense as answering skills of children are based off developmental effects such as language ability, reading age, motor skills, as well as temperamental effects such as confidence, self-belief and the desire to please (Read, 2007). In this case, the children may not have developed the skill or might have preferred to not put an answer at all then put an incorrect response. As sometimes children can be nervous of getting answers incorrect when they assume the adult knows the correct answer (Bell, 2013).

As stated in the experimental procedure section, the experimenter tried to stress the value of the children's input, as it is extremely important for them to feel at ease (Diah et al., 2010). Nerves can effect responses for fun levels, as well as their ability to answer the questionnaire. It was noticed that quite a few participants appeared nervous during the beginning questionnaire task.

This study did try to follow suggestions in order to decrease methodological challenges, although the methodological challenges associated with conducting research with adolescents is not a main area addressed in textbooks (Christian et al., 2010).

It is not a belief that the questionnaire was too easy, as the ceiling effect did not occur. If it had, it would mean that the majority of participants would have had high scores for both pre and post tests, with little variance as they were already at/near the highest point of learning (Judson, 2011).

It also does not appear that the prior knowledge effect occurred either. Prior knowledge effect is where students with higher pre-test scores make greater gain due to a stronger schema of understanding to use as scaffolding for the new learning (Judson, 2011). Although there is a weak positive correlation between participant's before score and how much they learned. It is possible that with more participants, the effect may have occurred.

Participant Retention

The tests showed that there were no significant differences between the amount of information retained and the three learning methods. Judging from the mean scores shown in Table 1, text showed the highest scores overall, with the exception being the circulatory system. It was interesting that the interactive retention rate did not differ, as AR has been shown to result in better long term retention (Radu, 2012). This suggests that although BRET succeeded in giving a realistic experience to the children, it may not have mimicked real AR interaction. Retention scores indicate that when the participants were reading the text, they were actively engaged, allowing the information to pass from short term to long term memory and creating connections with their schema to allow access to the information.

Participant Learning

The tests indicted a significant difference in the amount learned from text versus the interaction condition. There was no significant difference between text and video or interaction and video. This means that the null hypothesis for H2 was rejected, recognising there was a difference between learning methods. Looking at the mean learning differences in Table 2, it suggests that the children learned the least well with the interactive condition and generally better with text.

It is unlikely that the methodology caused the difference in scores as any factors from the questionnaire would impact all conditions.

Environment may have been a factor in terms of concentration level and how nervous the participants were. For instance it has been found that children learn more on a field trip when the learning environment is familiar than if it was novel (Flexer and Borun, 1984). In this research, the video and text conditions were conducted in a home environment in Palmerston North, so either at the researcher's home or at the participant's house. In contrast to this the interactive condition was conducted at the ScienceAlive! building in Christchurch where their displays are stored. The novelty and busy nature of the latter environment may have effected how comfortable and nervous they were while taking the tests and participating in the activities.

It could also be a factor that those reading were more actively engaged as it was their only task to focus on, therefore increasing the amount learned (Meo et al., 2013). As participants with the interactive condition may have been more interested with playing around with BRET rather than learning. One last point on the difference of the interactivity condition is that what the participant explored and selected would determine what information they read. In comparison, both text and video participants were exposed to all of the information.

The learning results could also have been influenced by the fact that children learn at different rates, have different learning styles that work best for them, and that performance in children covers a wide range (Korakakis et al., 2009). This means that some of the children who got the text learning condition may have had learning styles suited to this, whereas some of the children who had the interaction condition may have learned better with audio and visual information that comes with videos.

The lower scores overall could have been due to the limited time spent on each activity, the fact that recall in children is lower than with adults (Bell, 2013), that the questionnaire and biological content was quite difficult for the age group of 8 to 15 to understand, or that during activities users experienced points of engagement, a period of sustained engagement, disengagement, and re-engagement (O'Brien and Toms, 2012). Depending on when each of these phases occurred could result in them missing or not fully remembering answers to the questions they later had to answer.

Participant Fun

As tests showed that there was no significant difference in the level of fun found between the different learning methods this means that the results are inconsistent with the hypothesis that interaction would score higher than the other learning methods. The Smileyometer has also been used to measure expectation, and whether children were let down or surprised (Sim et al. (2006);Read (2007)). Although there was no significance between fun and the different forms of media, there were some significant differences between expected fun and the actual fun scores for participants. Overall it appears that interactivity with the digestive system, video with the skeletal and circulatory systems, and text with the skeleton and digestion showed significant differences. This means that they found these activities more fun then they thought that they would.

The means shown in Table 3 for the actual fun experienced fluctuate between 3.19 to 3.81, which on the Smileyometer signifies between Good and Really Good. The actual fun results of participating in the activities are higher than the expected, with the interactive condition showing higher (but not significant) means than video or text for all body systems apart from the skeleton.

One thing found when children rate the fun level for technology, is that they either want to please the adults by giving high ratings (Read et al.

(2001);Bell (2013)), they are likely to always find the novel technology to be a good experience, and children find almost all things fun (Read, 2007).

This research did not appear to have a problem with this as the fun scores were not around the really good (4) to brilliant (5) marks. So the complication around measuring fun did not seem to occur for the interaction condition. Although the honesty in the reading text fun scores was questioned, and speculation was that these scores were inflated due to the children wanting to appease the researcher. Alsumait (2008) also found that when children had trouble completing an activity, the Smileyometer rating was low. This could indicate that the harder an activity is, the more truthful the children are in their ratings, which seemed to be the case for this research.

Learning versus Age
Although there was an interest in comparing the learning versus Age, lack of participant numbers meant there was not enough information to investigate significance. The regression graphs displayed show an indication that there might be a positive correlation between learning and age, which you would expect, but due to the weakness it is not a good indicator currently. It would have been more interesting to compare the fun scores for the different learning methods over the ages, to see if different age groups enjoyed certain learning methods over others.

8 Conclusion and Future Work

Overall the main goal of this project, and BRET was successful. An interactive display was created that ScienceAlive! can use as part of a health exhibit in their upcoming museum. BRET contains biological information on three body systems which children have been shown to learn from. They may not learn more than if the children were watching video clips, or as much as the more traditional text, but learning did still occur. A majority of the children also found interacting with BRET fun, which was the other main goal. So BRET is a fun interactive biological learning tool that the business are very happy with as well.

ScienceAlive! are happy with the finished result, and there has been discussion into making BRET available for different countries, therefore including different languages versions. Improvements could also include more body systems, more interactive features, and better visual features. It is believed that the current number of body systems is good, as it keeps the content small and manageable, and that adding too many would start to clutter the interface and detract from the experience as a whole. Improvements, such as a dye pack connected to the mannequin's arm that works in simulation with the dye button, as well as adding other interactions could potentially increase the enjoyment level of the children. Visual features, such as increasing aesthetic appeal of the mannequin could also occur.

The age factor could be expanded upon with research, so finding out whether there is actually significant differences between age groups, but this would require a lot more participants. It would also be interesting to conduct research on the impact on learning and retention when BRET is used in conjunction with teaching in a class, as opposed to just class learning. Lastly research could be conducted on the learning impact when BRET is set up in the health exhibit, as there is not a lot of current research on health learning in museums with interactive displays.

9 Acknowledgments

I would like to acknowledge the support and help that I have received from the HIT Lab NZ, my friends, but most importantly my family. The willingness to listen and lend opinions, to help with technological solutions, and support me through the Masters has meant the difference.

I would also like to specially thank Robert Silberbauer for his help with conducting the research, and with helping out with understanding the masters process.

References

Alsumait, A. (2008). Use of Survey Techniques as Usability Evaluation for Child e-Learning Programs. *Proc. Conference ICL2008*, 1(3):1–3.

Anatomage USA (2005). fullbody.

Asai, K., Kobayashi, H., and Kondo, T. (2005). Augmented instructions - a fusion of augmented reality and printed learning materials.

Azuma, R. T. (1997). A Survey of Augmented Reality. *Media*, 6(4):355–385.

Bell, A. (2013). Designing and testing questionnaires for children. *Journal of Reseach in Nursing*, 12que(5):461–469.

Billinghurst, M. and Duenser, A. (2012). Augmented Reality in the Classroom. *Computer*, 45(7):56–63.

Bruckman, A. and Bandlow, A. (2007). HCI for Kids. In Sears, A. and Jacko, J., editors, *The Human-Computer Interaction Handbook*, chapter 40, page 1384. CRC Press, USA, 2 edition.

Chien, C.-h., Chen, C.-h., and Jeng, T.-s. (2010). An Interactive Augmented Reality System for Learning Anatomy Structure. *Computer*, I.

Christian, B. J., Pearce, P. F., Roberson, A. J., and Rothwell, E. (2010). It's a small, small world: data collection strategies for research with children and adolescents. *Journal of pediatric nursing*, 25(3):202–14.

Cockburn, A. (2004). Revisiting 2D vs 3D Implications on Spatial Memory. *Reproduction*, 28:25–31.

Cuendet, S., Bonnard, Q., Do-Lenh, S., and Dillenbourg, P. (2013). Designing augmented reality for the classroom. *Computers & Education*, pages 1–13.

Dewhurst, D. and Williams, A. (1998). An investigation of the potential for a computer-based tutorial program covering the cardiovascular system to replace traditional lectures. *Computers & Education*, 31(3):301–317.

Diah, N. M., Ismail, M., Ahmad, S., Dahari, M. K., Teknologi, U., Shah, M., Selangor, A., and Nielsen, J. (2010). Usability Testing for Educational Computer Game Using Observation Method Faculty of Computer and Mathematical Sciences. *IEEE*, pages 157–161.

Figueroa, P., Bischof, W. F., Boulanger, P., and James Hoover, H. (2005). Efficient comparison of platform alternatives in interactive virtual reality applications. *International Journal of Human-Computer Studies*, 62(1):73–103.

Flexer, B. K. and Borun, M. (1984). The impact of a class visit to a participatory science museum exhibit and a classroom science lesson. *Journal of Research in Science Teaching*, 21(9):863–873.

Juan, C., Beatrice, F., and Cano, J. (2008). An Augmented Reality System for Learning the Interior of the Human Body.

Judson, E. (2011). Learning about bones at a science museum: examining the alternate hypotheses of ceiling effect and prior knowledge. *Instructional Science*, 40(6):957–973.

Korakakis, G., Pavlatou, E., Palyvos, J., and Spyrellis, N. (2009). 3D visualization types in multimedia applications for science learning: A case study for 8th grade students in Greece. *Computers & Education*, 52(2):390–401.

Life Education Trust (2010). Life Education Trust – Registered Charity.

Maccario, N. K. (2012). Stimulation of Multiple Intelligence by Museum Education at Teachers' Training. *Procedia - Social and Behavioral Sciences*, 51:807–811.

Markopoulos, P., Read, J. C., and MacFarlane, S. (2008). *Evaluating Children's Interactive Products: Principles and Practices for Interaction Designers*. Morgan Kaufmann, 1st edition.

Meo, S. A., Shahabuddin, S., Al Masri, A. a., Ahmed, S. M., Aqil, M., Anwer, M. A., and Al-Drees, A. M. (2013). Comparison of the impact of powerpoint and chalkboard in undergraduate medical teaching: an evidence based study. *Journal of the College of Physicians and Surgeons–Pakistan : JCPSP*, 23(1):47–50.

Münzer, S., Seufert, T., and Brünken, R. (2009). Learning from multimedia presentations: Facilitation function of animations and spatial abilities. *Learning and Individual Differences*, 19(4):481–485.

Najjar, L. J. (1998). Principles of Educational Multimedia User Interface Design. *Human Factors: The Journal of the Human Factors and Ergonomics Society*, 40(2):311–323.

O'Brien, H. and Toms, E. (2012). Examining the generalizability of the User Engagement Scale (UES) in exploratory search. *Information Processing & Management*, 49(5):1092–1107.

Park, I. and Michael, J. (1991). Empirically-Based Guidelines for the Design of Interactive Multimedia. *ETR and D*, 41(3):63–85.

Radu, I. (2012). Why should my students use AR? A comparative review of the educational impacts of augmented-reality. In *2012 IEEE International Symposium on Mixed and Augmented Reality (ISMAR)*, pages 313–314. IEEE.

Read, J. and MacFarlane, S. (2006). Using the fun toolkit and other survey methods to gather opinions in child computer interaction. *... conference on Interaction design and children*, page 81.

Read, J., MacFarlane, S., and Casey, C. (2001). Expectations and Endurability-Measuring Fun. *Computers and fun*, (2000):2000–2001.

Read, J. C. (2007). Validating the Fun Toolkit: an instrument for measuring children's opinions of technology. *Cognition, Technology & Work*, 10(2):119–128.

Sim, G., MacFarlane, S., and Read, J. (2006). All work and no play: Measuring fun, usability, and learning in software for children. *Computers & Education*, 46(3):235–248.

Sinclair, P. and Martinez, K. (2004). Adapting Information Through Tangible Augmented Reality Interfaces. *New Review*.

Taylor, P. and Druin, A. (2002). The role of children in the design of new technology. *Behaviour & Information Technology*, 21(1):1–25.

Temkin, B., Acosta, E., Hatfield, P., Onal, E., and Tong, A. (2002). Web-based Three-dimensional Virtual Body Structures: W3D-VBS. *Journal of the American Medical Informatics Association*, 9(5):425–436.

The New Zealand Government (2007). The New Zealand Curriculum.

Unal, F. (2012). Observation of Object Preferences of Interest by Children Aged Between 4 and 8 in Museums: Antalya Museum Examples. *Procedia - Social and Behavioral Sciences*, 51:362–367.

Van Schijndel, T. J. P., Franse, R. K., and Raijmakers, M. E. J. (2010). The Exploratory Behavior Scale: Assessing young visitors' hands-on behavior in science museums. *Science Education*, 94(5):794–809.

Yoon, S. a., Elinich, K., Wang, J., Steinmeier, C., and Tucker, S. (2012). Using augmented reality and knowledge-building scaffolds to improve learning in a science museum. *International Journal of Computer-Supported Collaborative Learning*, 7(4):519–541.