Alexandre Heitz

# Improving Clinical Education Through the Use of Virtual Patient-based Computer Simulations

# Improving Clinical Education Through the Use of Virtual Patient-based Computer Simulations

By

Alexandre Heitz

A Thesis Submitted to the University of Canterbury's Postgraduate office

in Partial Fulfillment of the

Doctor of Philosophy in Human Interface Technology

Approved by the Supervision Committee:

________________________________

Dr. Andreas Dünser, Thesis Advisor

________________________________

Dr. Christoph Bartneck, Thesis Co-advisor

University of Canterbury

Private Bag 4800

Christchurch 8140 (New Zealand)

19 April 2013

**Abstract**


Alexandre Heitz. Improving Clinical Education Through the Use of Virtual Patient-based Computer Simulations.

The term Virtual Patient (VP) refers to the use of virtual characters which embody patients in a virtual environment. They are implemented in computer simulations to create realistic clinical encounters. VPs have been used successfully in health education to promote and foster clinical communication skills. Additionally, computer simulations offer the advantage of being standardized, safe, repeatable, and do not require as much resources as role-play simulations which rely on actors.

This thesis addresses the design and evaluation of a VP-based system aimed for clinical trainees, and uses the field of audiology as a case study. The system is designed to simulate real client encounters and allows students to practice using a standard set of procedures that they have to master in their profession. A wide range of VPs have been implemented for this purpose. The system was evaluated with audiology students, reinforcing the ecological validity of the research. The design of the system was guided by an iterative process of implementation, usability testing, and experiments focusing on students' learning outcomes.

The Clinical Audiology Simulator (CAS) was evaluated during five experiments, assessing students learning gains following exposure to the CAS. Learning gains have been assessed through the use of role-play simulations and paper assessments. The procedures evaluated are clinical history taking, pure tone audiometry, and speech audiometry. A further experiment assessed the impact of additional formative feedback on students learning gains, using the pure tone audiometry procedure as an example. The results of these experiments suggest that the system has a great potential to foster students learning, with measurable gains in some of these procedures. They also indicate that feedback and its delivery take an important role in this process.

This thesis elaborates how VP-based simulations can reinforce young clinicians' ability to learn procedural skills. I highlight some of the challenges a researcher faces in designing and evaluating such systems, focusing on the implementation of interaction scripts for the VPs, the assessment of learning gains and transfer of skills, and the evaluation of computer simulations as part of a curriculum. VPs have the potential to promote clinical trainees' learning of skills, and to provide students with more opportunities for safe practice in a field where beginning trainees often have few opportunities for actual hands on experience.

## Acknowledgements

First, I would like to extend my deepest thanks to my senior advisor, Dr. Andreas Dünser, who provided me with guidance and always supported me over the past few years. The same regards go toward my co-advisor, Dr. Christoph Bartneck, for his expertise and direction in preparing this thesis.

Further, I would like to thank the HIT Lab NZ's Director, Professor Mark Billinghurst, and the HIT Lab NZ's manager, Ken Beckman, who both trusted in me and provided me with the opportunity and funding to conduct this research. I would also like to thank Dr. Thomas Furness, HIT Lab founder and a a great inspiration for us all, and those from the HIT Lab NZ I have had the pleasure to work alongside with over the course of my research

I sincerely thank my colleagues as well, Jonathan Grady, Dr. Catherine Moran, Dr. Philippa Seaton, and Elizabeth Hanley who helped me to conduct my research. Furthermore, I would like to use this opportunity to thank all the students from the University of Canterbury's Communication Disorders department without whom this research could not have been possible. Among them, I had the pleasure to work directly with Lynda Guard, Gerard William, Sarah Howland, and Libby Sanderson.

I would like to extend a special thank you to my fellow PhD candidate, and dear friend, Hazel Bradshaw. Having friends like this who embarked on a similar journey and help each other out makes it a lot easier to reach the destination safely.

My sincere gratefulness and thanks also go towards my friends, Tyler J. Brown, Will Pearce Loe, Dave White, Kieran M. Tibble, Miles Burford, and James Wigzell to only name a few.

Finally and most importantly, I owe the greatest gratitude to my sister, Benedicte Heitz, and my parents, Marie-Cecile and Michel Heitz who over the course of my life always encouraged me to take on challenges and carry on the tasks I set myself to.

# Contents

# List of Figures

# List of Tables

# 1 Introduction

## 1.1 Introduction

In traditional clinical education, training practice can be a significant burden on resources both for educators and their establishments. These practices include, but are not limited to, clinical placements where an experienced clinician mentors a pupil in actual settings, role-play simulations with either an educator, a peer, or a trained actor taking on the role of a patient, or the use of mannequins.

This thesis is written for clinicians and clinical educators who wish to implement computer simulations in their programs to support trainees' education and clinical experience. The primary goal of this thesis is to answer whether computer simulation using virtual characters can adequately support clinical trainees' skills development. It also explores the effect such systems can have on students from different level of skills, as well as feedback.

Funded by a grant from the New Zealand Tertiary Education Commission (TEC) [1], parts of this this thesis were realised in a collaboration project be-

[1] TEC website http://www.tec.govt.nz/

tween the HIT Lab NZ [2], the University of Canterbury [3], the Christchurch Polytechnic Institute of Technology (CPIT) [4], and the University of Florida[5].

This thesis focuses on my work with the simulation platforms created to supplement the clinical education of audiology trainees from the University of Canterbury's Communication Disorders department, and the students of CPIT's school of nursing. It is a multidisciplinary endeavour as clinicians and course coordinators are necessary to integrate the adequate knowledge components and logics into the systems. Associate Professor Catherine Moran, and Clinical Audiology course coordinator Jonathan Grady act on behalf of the University of Canterbury's Communication Disorders department and contribute to the audiology field knowledge. Principal nursing lecturer Dr. Philippa Seaton and Elizabeth Hanley from CPIT offer their expertise in regards to the nursing field knowledge. I offer my perspective by designing, implementing, and assessing the educational impact and benefits of the computer simulations produced.

In order to understand the rationale underpinning this thesis, the literature review has been spread over the following sections:

1. Clinical education
2. Simulations
3. Virtual characters
4. Second life for clinical situations

[2]HIT Lab NZ website http://www.hitlabnz.org/

[3]University of Canterbury website http://www.canterbury.ac.nz/

[4]CPIT website http://www.cpit.ac.nz/

[5]University of Florida website http://www.ufl.edu/

5. Virtual Patient based systems for clinical situations

The review starts with identifying a need for computer artefacts to support clinical education. I then introduce simulations, presenting their advantages as well as their primary fields of use. The following section introduces the use of virtual characters. I then narrow our focus to the technologies that have been the most widely used over the past decade and use virtual characters, Second Life [6], and Virtual Patient (VP) systems. This section discusses their intended use as well as the challenges educators, students, and programmers are facing using those technologies.

Finally, I present the approach I used in the implementation and assessment of the computer simulations, the main technical and educational aims, before finishing with a summary of the main research hypotheses.

## 1.2 Literature Review

### Clinical education training

Healthcare education is a field which requires students to undertake a considerable amount of practical training. An audiologist or speech therapist for instance is expected to reach at least 300 hours of supervised clinical experience in order to be considered for a certification of competence by the American Speech and Hearing Association (Shaw, 2013), and 200 hours for the certificate of Clinical Practice awarded by the Audiological Society of Australia (Wayne et al., 2010).

[6] Second Life website http://secondlife.com/

In healthcare education, Simulated Patients (SPs), who are actors trained to interact with students in role-play settings are often used as a way to provide more training opportunities. However, access to SPs for healthcare education is limited due to their cost and availability. SP training is expensive. Additionally, SPs work for a low incentive, which can reach $10USD per hour. Considering this, SPs use in clinical scenarios to produce valid interactions for clinician and student training can be limited (Rizzo et al., 2010). In turn students take on themselves to assume those roles and often practice on each other, while not being trained for it.

Access to classroom is another limitation, as after hours access for students who wish to practice is difficult. Students would benefit from having an environment to support learning that would be available at their convenience (Minocha and Reeves, 2010; Kyle Johnsen and Andrew Raij, 2005a). Consequently there is a need in healthcare education for computer artefacts that support learning and provide students with more opportunities for hands on practice.

### Simulations

The use of simulations in education can be attributed to the work of an international group of enthusiasts who have been involved in this area since the mid-1960s. Simulations gained popularity not so much as a result of research reports, than through the experience of participants, who responded positively to the approach. Computer supported simulations started to appear in classrooms and increased from the 1990s.

A simulation is a working model of reality, which is distinguished from game or gaming simulation where the notion of competition and/or cooperation is included. Educational simulations are usually simplified representations of actual settings. They: "allow students to explore situations which would be too dangerous, expensive, time-consuming, or overwhelming to deal with 'for real', like a nuclear reactor, a breeding experiment, or a general election" Husén and Postlethwaite (1994, p.4577).

In addition, simulations often use role play in which participants may assume a different identity during their experience. Through the use of role play, simulations allow participants to experience someone else's situation, to imagine themselves into new experiences and anticipate their response to them, and to make decisions according to these events (Van Ments, 1999). Simulations are becoming more common as a teaching tool, and are being used throughout academia, the military, health care and emergency systems. For instance, virtual battlefield simulations and flight simulators can be used for a safe practice and development of tactical skills. Some research suggests that simulation-based training could be used to train interaction skills by assessing for post traumatic stress syndrome in military professionals before their return from the battlefield (Hubal and Frank, 2007).

Large scale simulations using virtual environments and virtual characters are also used in other areas. Wilkerson et al. (2008) demonstrated through the recreation of the aftermath of a terrorist explosion at a sportive event that a fully immersive environment could be used to train first responders for mass casualty incidents. This simulation provided a realistic experience without the expense of real time and live victims. Moreover, the virtual exercise allowed for repetition and the opportunity to record the users' actions for immediate feedback and decision analyses. At the same time, the experience had a high level of realism, and multiple users reported being stressed and feeling an excitation similar to what they would feel in a real life situation. The study also showed that while subjects had been able to identify appropriate actions in a classroom or another relatively stress-less environment, many more errors were committed when immersed in this high fidelity scenario. The participants appreciated the opportunity to experience this kind of chaos, make mistakes, have them identified and go through the scenarios again. The study concluded that simulations using immersive environments could be a promising way to train for high acuity, low frequency events and allow for the practice of necessary skills to avoid errors due to stress and distractions.

Simulations have been used within the health care sector in a variety of ways. Applications can range from teaching doctors or working with disabled people in order to increase their confidence, to practice for people diagnosed with autism in order to get them accustomed to specific social situations simulated through virtual agents, and learn by making mistakes but once again without real-world consequences (Fabri et al., 2007; Shea et al., 2003; Lee and Lee, 2004; Thwaites et al., 2008). Hildner (2007) stated in an interview focusing on the use of a patient based simulation in nursing education: "What this does is allow us to bring students into a lab setting as realistic as we can get", "You get to see more acute (conditions)", and "You usually don't give the truly sick patients to students, but you can do that here."

The first reported advantage of simulations in Education concerns motivation. Teachers declared that their students found the experience motivating, enjoyable and compelling. More surprising, it still seemed to be the case when simulations were used repeatedly, as each experience is at least slightly different. Participants' involvement appeared to relate to the duration of the simulations, their degree of realism, and the initial experience and attitudes of the participants. Moreover, the fact that participants had to draw knowledge from a variety of disciplines during simulations to resolve problems makes this learning approach especially valuable for interdisciplinary studies in secondary and higher education. Simulations have also become widely established as a major effective and motivating technique in various subjects (Husén and Postlethwaite, 1994). Repeatability is another major argument put forward by more recent research as simulation using computer assisted technology can also easily be altered to quickly provide a new experience (Heiphetz and Liberman, 2008; Wilkerson et al., 2008)

On the other hand, a simulation's main constraint is its heavy demands in resources, both on the students and on the educators. While they have benefits in education, simulations can still present a real challenge as they require a large investment of time to be implemented, and conducted. Some materials can be expensive and, more importantly, it puts a heavy burden on the teachers. Teachers are the ones who have to initially spend time to become familiar with simulations' concepts to be able to use them efficiently in their classrooms. Some teachers, however, appreciate the role of facilitator that those simulations allow them to take (Husén and Postlethwaite, 1994). Nowadays, with computer enabled simulations, the time necessary to become able to use a simulation system has been particularly reduced; generally requiring only a short amount of time for the participants to get accustomed to the interface. Costs depend on the technologies used and increase with the level of realism intended.

Another risk of simulations is that students can emerge with knowledge or procedure misconceptions, unless they are followed up by a discussion or debriefing to discuss the ways in which the exercise was and was not realistic. It is a vital part of the exercise which allows the participants to: 1)talk about what happened, 2)relate the simulation to reality, and 3)step back and discuss how the simulation could be modified or improved (Husén and Postlethwaite, 1994). A pilot study from Thwaites et al. (2008) inspects the use of two real life simulation-based scenarios and their impact on doctors in their first year of training. The study suggests that scenario training was acceptable and highly valued by the participants, but also that the debriefing could be a key factor in learning. But, more than just a debriefing, feedback in itself is a highly valued component in simulations and education in general (Havnes et al., 2012; Ong, 2007), as long as it is delivered adequately (Kneebone and Nestel, 2005; Shute, 2008). Astwood et al. (2008) state that feedback is the component that differentiates training from practice. It can take different forms but research shows that formative feedback, which occurs during a task, has the highest potential for new users. In a similar fashion, detailed feedback, providing corrections and leads on the next steps students could undertake seems also more beneficial for novice users compared to a more basic correct/incorrect type of response (Billings, 2012; Schartel, 2012; Kulhavy and Stock, 1989; Bangert-Drowns et al., 1991; Hsieh and O'Neil Jr., 2002). Feedback during a task as also the advantage of allowing students to deal and learn one unit at a time (Billings, 2012). Furthermore, it appears that positive feedback has a higher impact than negative feedback (Saraf et al., 2012).

Finally, the other issue concerns the problematic nature of evaluating simulations. It is particularly difficult to assess the effects and value of simulations, as each session is unique and depends on its actors (e.g. teacher who facilitates it and/or participants taking part). It was also suggested that, alternatively to performance and outcomes, importance could be accorded to the simulation process itself, through observations methods (Husén and Postlethwaite, 1994).

### Virtual characters

Virtual characters have been introduced in virtual environments in many forms. They can be used as avatars. An avatar refers to a 3D humanoid character in a virtual environment, which can have varying degrees of animation and behavioural capacities (Fabri et al., 2007). They are generally used as a representation of the users in the Virtual World, allowing them to explore it, socialize with each other, and participate in all kind of activities (Ye et al., 2007). In addition, virtual characters can also be used as computer interfaces, representing the computer in a human-like way (e.g. using body language or typical facial expressions to complete speech) to interact with the users. In this case they can be called Embodied Agents, Embodied Conversational Agents or even Pedagogical Agents (Foster, 2007; Mahmood and Ferneley, 2006).

Avatars and embodied agents were first introduced in simulations in the area of Edutainment which combines entertainment and education. They are used to achieve specific learning outcomes through the support of an entertaining interactive multimedia program (Foroughi, 2006). Garzotto and

Forfori (2006) illustrate this application in their project focused on hyperstories and social interaction for children. The research was based on the development of the Fate2 platform which was designed for kids to learn collaboration using their avatars to interact with each other in a virtual world, and to learn grammar and how to construct sentences by creating a common story. In addition to this main purpose, Fate2 also used avatars for educative games in the Virtual World, including ability (e.g. moving an avatar through a labyrinth), problem solving (e.g. treasure hunt in 3D space) or content oriented games (e.g. quizzes).

The use of avatars and embodied agents in an e-learning context had an increasing interest for educators because of the positive effects it could have on human cognition and motivation, perception of presence, learning, interactions or even confidence (Baylor and Kim, 2009; Callaghan et al., 2008; Fabri et al., 2007; Foster, 2007; Franceschi et al., 2008; Mahmood and Ferneley, 2006; Mikropoulos and Strouboulis, 2004). Concerning embodied agents, Foroughi (2006, p.2) writes that "one can say that the triangle consisting of the 'learner', the 'instructor' and the 'educational content', present in classroom is reconstructed" as they give a human face to the instructors or advisers whose roles they assume. They can also provide a more responsive feedback than a static or asynchronous environment, thus increasing the quality of learning and performance.

Studies have shown that the appearance of avatars can play a large role in the relationship users have with a virtual environment. For pedagogical agents, personalisation of the virtual characters by manipulating variables like the virtual character's ethnicity, gender or even attractiveness could enhance the learner's confidence and beliefs (Baylor and Kim, 2009). Hubal and Frank (2007) agree and suggest that the users should have the possibility to alter additional parameters such as the role (e.g. level of support of an agent during an activity) or the personality (e.g. humor, politeness) of the conversational agents used. The same phenomenon applies to virtual characters used to represent users in virtual environments where collaborative activities are commonplace (Pan et al., 2004).

### Second life

Second Life (SL)[7] is the most widely used three dimensional multi-user virtual environment (3D MUVE) (Salmon, 2009). Those environments are characterized by allowing the synchronous presence of multiple users embodied by their avatars. This allows them to interact and collaborate with each other in real time (Gorini et al., 2008). SL's supports a varied range of multimedia and interaction. Users can chat which in turn allows them to access their history, use voice, embed pictures, sounds and even videos (Burgess and Caverly, 2009; Boulos et al., 2007). Using chat has the advantage of allowing clinical students to review and assess their own interactions after assessing an SP's avatar for instance. They can download their own history to engage in self-criticism. This medium also provides educators with a way to review their students' work (Sweigart and Carlton, 2009).

Burgess and Caverly (2009) introduce SL by stating that it has a great potential for higher educational activities due to its immersion and collaboration possibilities, as well as its creation and scripting tools. Supplementary

[7] http://secondlife.com/

tutorial sessions for maths and biology courses are cited as examples, where tutors would be able to start by giving a problem example, discussing the possible solutions with their students, and making them engage cognitively in the problem solving. The advantage being that collaboration is created and facilitated in a risk free environment, meanwhile allowing distant actors to take part in the endeavour.

In an article, Bruck (2008) qualifies SL as an amazing virtual world for healthcare professionals. He introduces this medium under the light of its varied range of uses. Ongoing work on replicating real life stress management classes into the virtual world is mentioned. He also reports the existence of major institutes' virtual space. In this fashion, the US Centre for Disease Control (CDC) has its own island, their location in SL, which allows private interviews with some of their personnel. But SL is also employed as a way of giving virtual tours of healthcare establishments, conduct health surveys, provide medical students training, or even holding virtual charity events.

Salmon (2009) states that his research suggests that SL has high potential for learning, all the while remaining low cost. The avatars' interaction within the virtual environment contribute to the feeling of presence and immersion, while allowing discussion and encounters. Another advantage of the use of SL mentioned in this article, further than creating virtual classroom, is the ability of the environment to involve its actors to partake in its creation. Moreover, SL is also adaptable to almost any discipline and allows interaction and manipulation of virtual artefacts in ways that would not be possible in the real world. They also present examples of some SL projects. In this light, SL supports countless educational medical projects, such as a nutritional education game held on the virtual grounds of the University of Ohio, where users can learn the health consequences of improper nutrition and fast-foods. An instructional designer at San Jose state University created a virtual world for cardiac auscultation training allowing medical students to tour a virtual clinic and try to identify different types of heart murmurs based on sounds. Gene Pool is another educational endeavour where visitors can learn extensively about DNA and human chromosomes.

As part of a collaboration between London's universities, the second life virtual environment has been used as a way to train paramedics and allow them to explore more open ended questions in their decision making process. The researchers concluded that a virtual environment can offer realism impossible to reach in a classroom environment, allowing users to experience the consequences of the different choice they make (Conradi et al., 2009). The University of California also created a location in SL to educate people about schizophrenic hallucinations (Boulos et al., 2007). In addition, SL has also been used to familiarize students with transmissive disease, train first emergency responders, or run round tables on bioterrorism awareness (Skiba, 2007). In another article, Gorini et al. (2008) explore the possible applications of 3D environments and more specifically SL to support clinical psychology via therapeutic interventions. The aims highlighted in their article is to reach remote patients for cyber-consultation, but also to create communities of patients and allow them to meet each other, socialize, and engage in group activities. They also point out that clinicians and researchers have to create protected environments which have to meet the needs of the patients.

According to Wang and Braman (2009), SL can contribute to learning, assuming it is used adequately. Their research suggests that to support learning in SL, specific attention must be given to the approach used: 1) Students should first be educated on how to navigate in SL before starting learning activities, 2) A wide range of activities should be integrated in combination in order to improve students engagement. 3) The opportunity should be given to students to form their own groups. 4) The groups of students collaborating should remain small to ensure that both students and their tutor will not be overwhelmed by the learning activities.

Second life is ideal for distance education, as it allows people to meet with their peers, but also to learn virtually through the presence of a tutor in the environment as well. In a study accross european health-centers, Melús-Palazón et al. (2012) reported that 74% of the participants assessed estimated that SL eliminated the need to travel. SL also enables to access elderly people, or people with physical disabilities (Boulos et al., 2007).

However, SL is still multi-user centred. It allows users to reproduce a typical role-play situation in the virtual environment via their avatars (Walker and Rockinson-Szapkiw, 2009), but it still is implemented for multiple actors to engage together. Moreover, SL displays a range of limitations. Sanchez (2009) puts forward the steep learning curve involved with using SL as the main barrier to its educational use. Overall, they point out that students report four main issues: the complexity of the interface, the technical difficulties involved with a resource intense application, both for the computers' processors and with the internet bandwidth, a possible let down in terms of user expectations with students reporting that the environment lacked an actual purpose, and the fact that the experience is particularly time consuming. In a qualitative study from Melús-Palazón et al. (2012), 91% of the sample, 76 healthcare personals who attended clinical sessions in SL, reported technical problems as being the main weakness of the system.

Moreover, while one of the initial benefits of SL for online collaboration is an emotionally risk free environment, this is not really the case across this world on a daily basis. SL, like other collaborative online games gave birth to a societal structure of its own where long term users have advantages over novices. In addition, the issue of griefing also exists. This terms refers to any behaviour in a virtual world that aims to disrupt the experience of others. Most of the content in SL is publicly accessible to all users (Bell, 2009). This could translate into anything from verbal harassment of students, graffiti and tags of virtual structures used in learning activities, or to other behaviour having the potential to disrupt an online tutoring session.

Cost is also an ongoing concern, while the CDC initially invested $73USD for their online presence in SL, the creation of their own island was $2000USD. On top of these there are maintenance, content creation, and scripting costs as well. In February 2008 the cost of an Island in SL was $1675USD, with a monthly maintenance cost of $295USD (Gorini et al., 2008).

Consequently, while SL offers medical practitioners more opportunities for practice, other systems might provide more advantages, and fit the needs of clinical students better by not relying on expensive hardware, an internet access, and additional people to participate in the experience.

### Virtual Patients-based simulations

Another training method in healthcare education is Virtual Patients (VPs). VPs have the potential to improve clinical competence and decision making of healthcare trainees (Round et al., 2009). They are computer generated patients that allow realistic training in controlled settings. Either virtual characters can represent the patients, or other media can, but they are always based on realistic clinical cases. One reality of healthcare practical education is that students can spend days doing clinical work when they only see repetitive cases, when patients come in with the same pathologies. VP simulations can be used to provide a varied and standardized training among students, assuring that every one of them also has the chance to train for low frequency events, not just the cases that are the most commonly seen in clinics (Collins and Harden, 1998; Kyle Johnsen and Andrew Raij, 2005a; Triola et al., 2006). However, at the same time, students can choose to engage with the same VP as many times as they choose; for instance if debriefing made them realize they missed critical pieces of information for instance. VPs also enable students to practice within safe boundaries, preparing them through repetition and development of their clinical reasoning to interact with real patients where mistakes would be costly (Round et al., 2009). This technology can also enable students to practice within the comfort of their own homes (Stanton, 2008).

Previous research shows that VPs can effectively be implemented as an alternative to SPs in various types of scenarios to represent a wide range of illnesses. Students can use them to train for standard medical exams, or practice eye or breast exams (Deladisma et al., 2009; Kotranza et al., 2009; Rizzo et al., 2011). Cases aimed at paramedic students have also been implemented (Conradi et al., 2009), so have cases for pharmacy students to train continuity of care (Fuhrman Jr et al., 2001). Tan et al. (2010) describe in an article a case aimed for students in geriatrics as well. Surgery cases can also employ VPs (Vash et al., 2007). At the University of Munich in Germany, a VP-based system allowed students to train gynaecological sonography. The conclusions showed that training with virtual patients seems comparable to live practice with the advantage of a standardized consistent output, which is not the case with a real patient (Heer et al., 2004). Finally, VPs are also used to replicate psychiatric disorders such as disorderly conduct, post traumatic stress syndrome (PTSD) and phobias (Gorrindo and Groves, 2009; Kenny et al., 2008; Rizzo et al., 2010; Triola et al., 2006; Guise et al., 2012a,b). VPs have also been successfully used to explain medical documents to patients with inadequate health literacy (Bickmore et al., 2009, 2010), or to promote healthy behaviours such as adherence to medication and exercise (Bickmore and Picard, 2005; Hayes-roth and Saker, 2003; Ruttkay and Welbergen, 2008).

While VPs are often valuable, this is not the case for every clinical speciality. Edelbring et al. (2011)'s article assessed how rheumatology students perceived encounters with VPs. While VPs replicate meaningful patient-encounters, students still reported that the VP lacked emotions. They, however, perceived this experience as a valuable preparatory work before meeting real patients. Forsberg et al. (2011) aimed to evaluate the value of VPs for assessing clinical reasoning in nursing education. Students involved in the study had a high level of acceptance of this system as an assessment method. These results are, however, not supported by the findings of Courteille et al. (2008) in a similar study, who identified weaknesses in this technology when used as an assessment method. Indeed, they point out that, if the system is not able to be autonomous and a human has to be involved to pilot the VP, then standardization cannot be reached. In a comparative study between students practicing on a VP system and students following traditional methods, Botezatu et al. (2010) found that students who trained with VPs had better long term retention. They also point out the necessity of a robust software and case design method.

The research from the following groups is considered the most relevant to this thesis, as the technologies they use are similar. The University of Florida's Virtual Experiences Research Group (VERG) [8], which provided our research team with the core of our simulation platform, is one of the leaders in VP research. The early studies of this group focused on the possibilities of creating a system with patient-doctor interactions, focused on assessing whether current technology enabled simulation of such experience with sufficient immersion and fidelity. They focused on implementing medical history interviews for different disciplines as well as investigating different ways to improve user-VP exchanges, thus contributing to the field. Studies exploring different ways to affect interactions have also been conducted, using projectors to display life size VP for instance, investigating natural interactions with VPs such as hand gestures (both real hand with infra red cameras, or virtual hands), tablet PC and audio (Ferdig et al., 2007; Kotranza et al., 2009; Kyle Johnsen and Andrew Raij, 2005a). Real size displays, using projectors, were found to allow users to display more empathy than head mounted displays (Johnsen and Lok, 2008). The use of synthesized speech in comparison to natural recorded speech for

[8] VERG website http://verg.cise.ufl.edu/

the VPs' voices has also been investigated. It was concluded that if the intent was to teach what questions to ask, then both methods are equally effective. However, if the purpose was to teach how to ask questions, a high level of realism and expressiveness from the VP would be need, and pre-recorded speech would be ideal (Dickerson et al., 2006).

Rather than assessing its continuous long term use by medical trainees, VERG's evaluations often focused on comparing the quality of the VP to an SP using the Maastrich assessment of Simulated Patients questionnaire, student posture, tone of voice and speech content for different exams. The learning outcomes measured were limited to the ability of participants to retrieve the full range of information a VP could dispense in a one time experiment (Deladisma et al., 2009; Kotranza et al., 2009; Rizzo et al., 2011). The research group also investigated how racial and social disparities can affect participants' relationships with a VP. In the different conditions explored, research showed the significant result that participants were able to interact with a VP equally well as they would with an SP (Ferdig et al., 2007; Kyle Johnsen and Andrew Raij, 2005a).

The Institute for Creative Technologies of the University of South California is another group that has heavily invested in VP research. It primarily investigates the use of VPs with military technicians and clinicians to treat soldiers, which can be dangerous both mentally and physically. VPs are considered for PTSD patients, patients with suicidal tendencies, or even patients with traumatic brain injuries. Participants were able to detect trauma, and its duration, or to present its origin. However, symptoms requiring the students to have a deep level of exchange with the VPs in order to identify them were almost not detected. This can be explained by the quality of the interaction scripts implemented, and current performances of speech recognition (Kenny and Parsons, 2010; Kenny et al., 2008). VPs were also used as an online guide to promote access to psychological healthcare information and assist as well as encourage military personel and their families to seek care if necessary.

When this research is evaluated, the main disadvantage of the VPs-user interactions appears to be that their interactions are scripted; thus, a necessary time consuming step is to implement a script with sets of clues or triggers for each of the VP's answers. This is the main limitation from which VPs currently suffer from and that can prevent them from getting closer to SPs in terms of realism and interactions. In addition, an underdeveloped script can be frustrating for participants (Kyle Johnsen and Andrew Raij, 2005a). Development of an interaction script for one VP is a time consuming process and often needs to be modified following pilot studies to make it usable. Work has been done on implementing ways to create more robust scripts efficiently for VP interactions but it is still an area that needs to be improved (Halan et al., 2010; Rizzo et al., 2010; Rossen et al., 2010). As a consequence, VP simulations used in health education commonly focus on a limited number of VPs, and often employ a technician to ensure the answers are relevant to the questions asked by the users (Bickmore and Picard, 2005; Deladisma et al., 2009; Fuhrman Jr et al., 2001; Gorrindo and Groves, 2009; Hayes-roth and Saker, 2003; Heer et al., 2004; Kenny et al., 2008; Kotranza et al., 2009; Rizzo et al., 2010, 2011; Ruttkay and Welbergen, 2008; Stanton, 2008; Tan et al., 2010; Triola et al., 2006; Vash et al., 2007).

Our research focuses on the fields of Audiology and Nursing education.

In Audiology, simulations have been used with success to teach clinical skills. Simulators allowing practice of procedural skills, such as pure tone audiometry are available on the market (e.g. Otis Audiology simulator [9], Parrot Software's Audiology Clinic [10]) for Universities to use to supplement traditional course work. There is, however, a lack of simulators incorporating such procedural skills with clinical history taking training possibilities and immersing students in realistic experiences from meeting a patient up to and including pathology assessments. In nursing education, simulators either used as standalone or with mannequins exist (e.g. SimMan[11], VitalSim[12]). However, the standalone simulators often have only a limited level of immersion and do not allow for interviewing, while the ones used alongside a mannequin still need experienced personnel to operate them.

### Simulations in Audiology

Studies aiming to implement and assess the impact of new training methods in Audiology have previously been conducted, however to our knowledge no study has directly compared the impact of computer simulation on students learning of the clinical procedures, including history taking. Lieberth and Martin (2005) present a comparative study on audiology students training pure tone audiometry through an audiometer, and a virtual audiometer. Their findings show similar grades for students in both groups at the end of the training period, suggesting that the virtual audiometer implemented is a

[9] INNOFORCE creative solutions website http://www.innoforce.com/

[10] Parrot Software website http://www.parrotsoftware.com/

[11] SimMan website http://www.laerdal.com/simman/

[12] VitalSim website http://www.laerdal.com/nz/doc/247/VitalSim

valid tool to learn pure tone audiometry. They however note that students who practiced on the virtual audiometer had overall better technique while lacking in comparison to the control group in interactions and feedback with the patient. The control group scored better in those two areas, while scoring less on technique. Another study from Wayne et al. (2010) investigates the addition of computer simulations and SPs in audiology students' curriculum. The study covered a wide range of procedures and the results suggest that students perceive those two addition as a strong positive impact on their learning. Unfortunately, the study focuses on subjective measures, as well as using both form of training on the whole sample of students.

## 1.3 Approach

### Design-based research

Husén and Postlethwaite (1994) suggest that the design of simulations should consist of a large amount of trial and error. Researchers should study pilot versions and listen to participants' reactions, then polish and improve the simulations as a whole. This description relates to design based research. Research conducted for this thesis will follow this approach, which proved during the past decade its potential for both research and design of technology-enhanced environments.

Design-Based Research (DBR) is a methodology that blends both empirical educational research and the theory-driven design of learning environments. It goes beyond designing and testing interventions as they embody theories about teaching and learning. They reflect a relationship between the artefacts produced, theory, and practice. Design-based research aims to improve educational practice and to contribute to the theories it relies on, by testing educational interventions in different areas of instruction and under multiple settings. The interventions are linked with the theory and designed to test it. In addition, it also aims at producing contextually sensitive design principles Reeves et al. (2005, p.109-110). This type of research requires to:

1. "Explore significant educational problems, rather than conduct research for its own sake.

2. Define a pedagogical outcome and create learning environments that address it.

3. Emphasize content and pedagogy rather than technology.

4. Give special attention to supporting human interactions and nurturing learning communities.

5. Modify the learning environments until the pedagogical outcome is reached.

6. Reflect on the process to reveal design principles that can inform other instructors, other researchers, and future development projects."

DBR aims to investigate what makes an intervention successful in a particular context, which links back to ecological validity and can therefore limit generalisation of the findings (Sandoval and Bell, 2004). What would typically be called generalisation in other research approach then occurs through design recommendations. In a case study implementing a web-based multistoryline system for learning purposes, Zeng and Blasi (2010) assessed learning gains for the students implicated in the process, and provided design recommendation as an outcome. Those recommendations were that such a system should provide authentic cases, arouse curiosity and interest, provide explanatory feedback, and adapt to different learners.

People conducting design based research have the functions of both designers and researchers, drawing procedures from the two fields. Compared to other social sciences approaches the researchers are not concerned about contaminating the research, instead they have a direct involvement as they will manage research processes in collaboration with their participants in real world settings by designing, and by refining intervention systematically (Wang and Hannafin, 2005). Hoadley (2004), and Joseph (2004) agree that it is not only about the research but also on how the interventions played out in actual practice. How are they used? How do they succeed? How do they fail? This design process is described as cyclic following iterations made of analysis, design, and evaluation stages (den Akker et al., 2000).

DBR tends to use mixed methods taking from both quantitative and qualitative research (PeerGroup, 2006). Mixed methods are used to maximise the credibility of ongoing research and vary according to the different phases of the research as new needs and issues emerge and the research focus evolves. However, Wang and Hannafin (2005) alerts researchers to the large amounts of data produced and the time and resources needed to analyse that data routinely. They conclude that data will regularly have to be disregarded and encourage design-based researchers to share those with their peers to prevent them from being wasted.

Another potential issue in design-based research is the role division between development and research. den Akker et al. (2000) suggest that in the earlier stages of the project the researcher is more influenced by his designer's perspective while a shift towards a more critical researcher's perspective will take place at the later stages. Moreover, Tabak (2004) warns researchers implementing DBR against focusing only on the design, to the extent of missing other factors that could impact on learning.

### Design based research application

DBR will be used with the goal of designing the main interventions, both for audiology and nurses trainees. This includes both simulation systems, and their method of assessments. As Joseph (2004) states, investigations are centred around the evolution and implementation of the artefact produced. This will result in building a first prototype, then testing and assessing some of its functionalities in real settings, the classroom. If learning gains are not observed during this evaluation stage, we will investigate how students used it. The prototypes will then be refined in order to improve functionalities.

Usability questionnaires will be used to assess each stage of the implementation as well (Appendix A), prompting students to assess the technical qualities of the systems as well as offer advices for further implementation or comments about frustrating aspects. As Wang and Hannafin (2005) mentioned, this generates a huge amount of data, it will consequently not be entirely reported in this thesis, however, it contributes in drawing conclusion when learning outcomes are evaluated.

On the other hand, if learning gains are observed, then support for further clinical procedures will be implemented, and once again tested. This process will be repeated until all necessary functionalities and simulated procedures are integrated into the system and students can use it adequately in their curriculum.

Brown (1992) also raises the issue of maintenance. A new learning procedure or artefact needs to be used by educators after a study. For this purpose, the audiology and nursing simulation systems will be built alongside educators, involved in the entire process. Moreover, tools adapted to the educators will be implemented to allow them to change existing content, as well as creating new content, allowing them to easily implement new clinical cases.

## 1.4 Aims

This section first presents the technical aims for implementing the simulation systems in order to benefit of all their real life practical advantages while training. I will then introduce the research aims on which this thesis focuses. This research focuses on the field of audiology and nursing education as case studies and aims to generalize the findings to general clinical education in the conclusions section.

The main goal of this research project was to evaluate the simulation systems designed using Virtual Patients (VPs), and aimed at clinical practitioners to supplement and improve their clinical skill development. The target audience for this research are the field of audiology, and nursing education.

The technical aims of these two simulation systems aimed were:

1. Integrating clinical audiology and nursing expertise within immersive

simulation systems.

2. *For audiology trainees, creating various set of clinical cases for trainees to practice in real time with VPs, practicing the range of procedures respective to their field.*

3. *For nursing trainees, creating a single non linear scenario replicating mannequin training.*

4. Implementing a safe environment to supplement traditional training, and foster trainees' knowledge and clinical reasoning skills before facing real patients.

5. Allowing clinical trainees to train at their own pace, engage in self-assessment, and repeat particular situations if necessary.

6. Increasing control and standardization of practice among trainees.

Fulfilling these goals is necessary to provide us with a platform to conduct educational research using the simulation system.

Due to lack of resources and logistical issues, the nursing simulator was put to a halt. Moreover, following the February 2011 earthquake that took place in Christchurch, access to students was impossible for an extended period of time, and once re-opened recruiting students was difficult due to the changes in teaching curricula.

Therefore, the nursing simulator was not evaluated as it was initially planned. I will present its implementation in the following chapter, however the research findings presented in this thesis will only reflect the results of the audiology simulator's experiments.

### Evaluation

The research aim was to evaluate whether using the VP-based simulation system would impact positively on learning outcomes of young audiology students, when used in addition to traditional trainings.

This objective was broken down to the following objectives:

1. Improve audiology trainees' medical history taking skills through exposure to the VP-based simulation system.

2. Improve audiology trainees' procedural knowledge through exposure to the VP-based simulation system.

3. Accurately assess audiology trainees' transfer of skills in role-play situations from the exposure to the VP-based simulation system.

## Hypotheses

During the implementation of the simulation system, a series of studies were conducted with audiology and speech language therapy students. The first section is on medical history taking, we then move to other clinical procedures, and finish with the simulation system used as a primary teaching tool. This section will summarize the overarching focus of these studies as well as their respective hypotheses.

### Virtual patient based simulation to improve medical history taking & pure tone audiometry

The main hypothesis is that when tested for transfer of skills in role play situations, trainees exposed to the VP-base simulation in addition to traditional training will significantly improve their clinical history taking skills compared to trainees only following traditional training.

This hypothesis was divided down into the following sub-hypotheses:

1. The trainees confidence in conducting a clinical history will improve as a result of using the system in addition to traditional methods.
2. The trainees ability to retrieve information in clinical history taking will improve as a result of using the system in addition to traditional methods.
3. The trainees will improve their efficiency, requiring less interaction with patients to retrieve relevant information, as a result of using the system in addition to traditional methods.
4. Students ability to conduct a Pure tone Audiometry (PTA) procedure accurately will increase as a result of using the system in addition to traditional methods.

### Virtual patient based simulation to improve speech audiometry

When tested for transfer of skills in role play situations, trainees exposed to the VP-base simulation in addition to traditional training will significantly improve their performance conducting a speech audiometry exam, and their knowledge of the procedure. In addition, students exposed to the system will have better retention of the procedure when assessed after not accessing the system for a period of four weeks. Finally, we intend to explore whether, if, independent of training format, novice students should benefit more from the system, when compared to intermediate students. To our knowledge, no previous research explored this aspect in similar settings.

These hypotheses were divided into the following sub-hypotheses:

1. Novice students will benefit more from being exposed to speech audiometry than intermediate users.
2. Novice students will retain knowledge better than intermediate students when assessed for retention four weeks after the post-test.
3. Students training with the Clinical Audiology Simulator (CAS) will perform better than students who were only exposed to traditional methods.
4. Students training with the CAS will retain knowledge better than students who were only exposed to traditional methods following four weeks of not practicing using the CAS.

### Virtual patient based simulation and novice tailored feedback to learn procedural skills

The main hypothesis is that when using the system as a standalone (without traditional course work or placements) to teach novice students procedural skills, additional cues and feedback will significantly increase learning.

This hypothesis was divided into the following sub-hypotheses:

1. Novice trainees exposed to the system with additional feedback will score significantly higher than novice trainees exposed without this additional feedback when tested for retention following training.

2. Novice trainees exposed to the system with additional feedback will require less mental effort than novice trainees exposed without this additional feedback when assessed following exposure.

3. Novice trainees exposed to the system with additional feedback will require more mental effort than novice trainees exposed without this additional feedback while training.

## 1.5 Ethical approval

The University of Canterbury's Human and Ethics committee as well as Christchurch Polytechnic Institute's ethical committee granted their approval to conduct the research outlined in this thesis and use students as participants.

Participants were approached first in the classroom and then via email, following approval of their respective course coordinators. They were provided with an ethical agreement as well as a summary of the studies. The participants were not penalized if they decided to not undertake part in this research or wished to stop the study early.

Anonymity was maintained using alias. The data retrieved was stored in locked desk, for the hard copies, while electronic data was on password protected computers.

# 2 Design

## 2.1 Introduction

The initial platform provided by the University of Florida's Virtual Experiences Research Group [1] was implemented by this group for research study purposes. Each component used in VERG's platform was originally a separate application, primarily to allow researchers to pilot the VPs responses from another computer, but also for flexibility purposes. At this stage the system was only allowing to support the interview of virtual characters in a 3D environment. Moreover, no scripts existed that were adequate for the two fields we aimed to explore.

This chapter presents the design of the two simulators. First, I will give an overview of the system. I will then go over the design step by step of the Clinical Audiology Simulator (CAS). Each iterative design is accompanied by the results of a usability assessment. The final section presents the first prototype of the nursing simulator, its functionalities, as well as some preliminary feedback from medical experts.

[1] http://verg.cise.ufl.edu/

### System architecture

To conduct this research, the first step we undertook was to adapt the platform provided by VERG so that its components could run as a standalone application. Both the CAS and nursing simulator have then been implemented in Visual studio 2010 using C#/C++, with the first standalone prototype as a basis to build upon. The applications make use of the open source 3D library Ogre [2] for graphics.

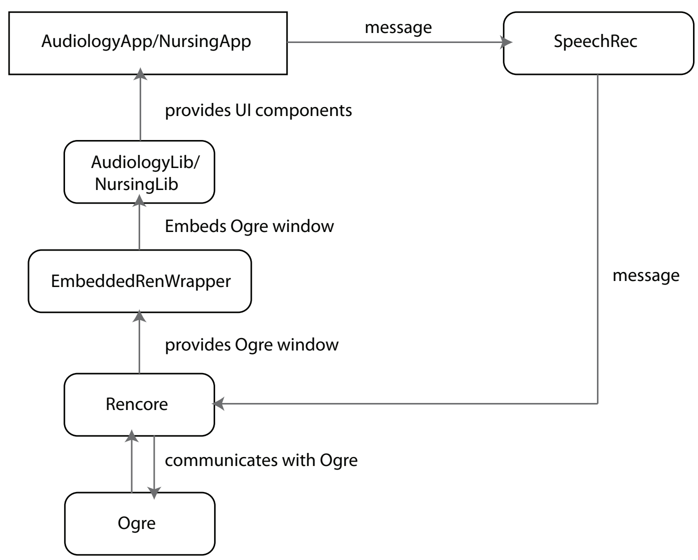



Figure 2.1: Main system architecture

As seen in Figure 2.1, *Ogre* provides the necessary 3D libraries, and using *Rencore* and the *EmbeddedRenWrapper* the VPs as well as the room where the consultations take place are rendered, within an embedded window in

[2]http://www.ogre3d.org/

the application. Animations are managed through the same components. *SpeechRec*'s primary function is to match users' input to potential VP answers. The questions users ask the VPs are matched to a list of triggers contained within a script. If the match is positive, the associated answer will be sent, if it is not, one of the generic answers where the VP ask the users to reformulate a question is returned. AudiologyLib and NursingLib are the main user interface components that are used in their respective application.

The following sections presents our work on building the CAS and nursing simulators.

## 2.2 Clinical Audiology Simulator

The Clinical Audiology Simulator (CAS) is used to practice procedural skills. This takes the form of the standard range of tests used in an audiological exam. It includes history taking, otoscopy, puretone audiometry, speech audiometry, and pathology assessment.

The CAS was implemented following a series of steps. I will first present the initial prototype we designed before moving to its refinement, and the subsequent additional features implemented.

### Prototype 1

#### Design

This section presents the first prototype of the CAS. It focuses on allowing students to practice clinical history taking, as well as pure tone audiometry.

In addition, eardrum pictures were presented to provide more information to the case, this falls under the otoscopy procedure.

When launching the CAS, a student starts with selecting one of the VPs among the different cases offered. The application then starts and students have access to the different features of the software, labelled *Interview*, *Otoscopy*, *Tone test*, and *Submit Results*. The students can choose the order of the procedures but would typically start by interviewing the patient.

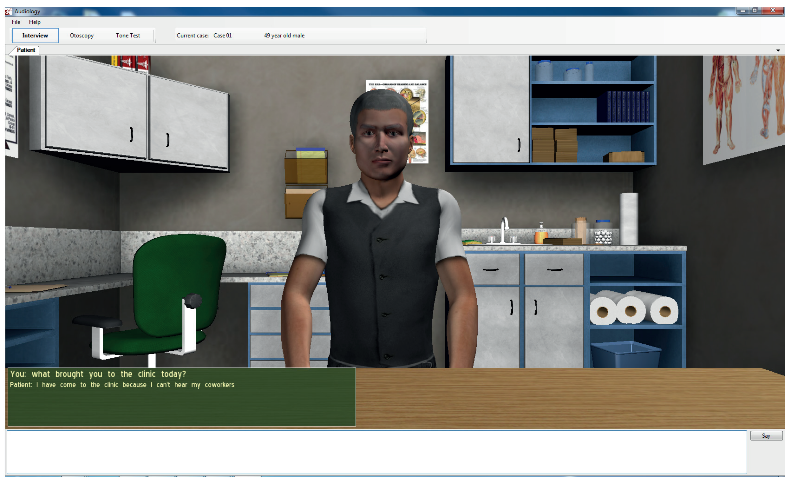



Figure 2.2: Prototype 1: clinical history taking interface

The history taking component of the consultation takes the form of a loop where the students ask questions to the VP which will be answered according to the VP's script. The students interact with the VP, who is sitting in a room, by using the keyboard to type the questions they wish to ask. (see Figure 2.2 History taking interface of the CAS). While the system does support speech recognition, it is not as efficient in a lab with multiple computers setup and more than one student practicing at the same time. The VP will either answer the students' question, if it was understood through the script, or hint to the student to reformulate and/or ask another question (e.g. "Could you please rephrase that?"). The VP will speak the answer out-loud and it will also be presented in the bottom left of the screen. The students aim to collect answers from the VPs. They then need to sort information significant to assess their diagnostic and report them on a Diagnostic Adult History Form (DAHF). The DAHF is commonly used in clinic when interviewing patients (Appendix B). The students can stop this process at any time if they consider that all the relevant information got retrieved from the VP.

Students are then able to check the VP's ears if they consider it necessary. This will display two eardrum pictures from the collection of pictures retrieved from real patients, these can present additional elements to help identify the correct diagnosis.

The students will then follow with the pure tone audiometry procedure, which takes the form of another loop. This procedure is about determining the VP's hearing thresholds. Hearing thresholds refer to the softest sounds a patient's ears can detect. Hearing thresholds are typically displayed on audiograms, which are graphs of the hearing levels in decibels for a set of frequencies ranging from 250 Hz to 8 kHz, range which covers speech understanding. To determine hearing thresholds in the CAS, a student will have to first select a transducer between supra-aural, insert, and bone conduction before deciding which ear to test. This represents the type of headphone used, each having its specificities.

Typically, a clinician looks for when the patient stops responding at the softest levels. The audiologist, following a threshold seeking method, will then mark the softest response at which the patient detects in 2 out of 3 presentations. During practice with the CAS, when a tone is submitted, the VP can either answer or not. Once a VP answers to a particular tone the student should mark the response level on the virtual audiograms. After obtaining these basic thresholds and looking at the results, a student has to decide if masking is required. Masking aims to seek a patient's true thresholds while eliminating possible conduction from one ear to another. This is necessary for patients with significantly bad hearing. The process is similar to the original threshold seeking task just completed, with additional white noise (the masking) presented into one ear as an extra variable. This procedure involves more calculations and reasoning from the students. This process is however only repeated for the range of frequencies the students decides to test, where one might suspect conduction. Figure 2.3 represents the pure tone audiometry interface of the CAS, with the virtual audiograms on the left side of the VP. The bottom right section embeds the audiometer interface, used to present tones, and the bottom left is a reminder of the symbols used to report the different thresholds on the audiograms.

Following those assessments, students have to submit their results. They have to determine the pathology(ies) associated with their patient. In addition, students can choose to add a comment to explain their diagnostic decision. This information can be recorded and used for assessment. Finally, once the diagnosis is submitted the students will be given feedback in the form of the VP's actual audiograms, and the correct diagnosis.

This first prototype of the CAS was deployed on a total of seven computers to allow participants to practice during their free time. During their

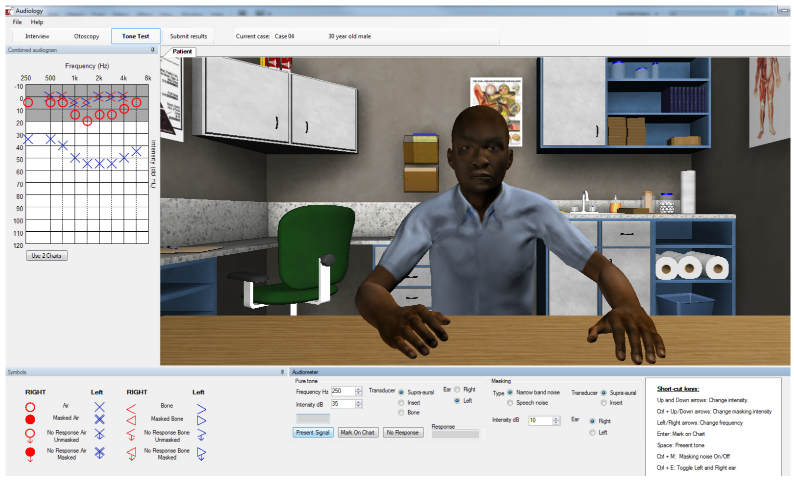



Figure 2.3: Prototype 1: pure tone audiometry interface

designated training period, participants were able to access the computers and practice on the CAS at any time. As part of this pilot studies detailed in chapter 3, students had to complete a usability test to assess the quality of this first prototype. The following section presents these results.

### Usability

Table 2.1: Descriptive statistics for the USE questionnaire*(n=12)*

| Questionnaire category | Mean | Std. dev. |
|---|---|---|
| Usefulness | 3.88 | 1.08 |
| Ease of Use | 4.57 | 1.01 |
| Ease of Learning | 5.77 | 1.21 |
| Satisfaction | 3.72 | 1.12 |

Twelve students were provided with a questionnaire to assess the system's usability during the study presented in chapter 3. A copy of the USE questionnaire is attached in Appendix A. This questionnaire uses 7 points Likert-scales to assess the students' perception of *Usefulness*, *Ease of Use*, *Ease of Learning*, and *Satisfaction*, in reference to the CAS prototype. Table 2.1 presents average results for each category, with *1* being *Strongly Disagree*, and *7* being *Strongly agree*. We note that the scores for the *Usefulness* and *Satisfaction* categories are only slightly above the Likert-scales' midpoint of 3.5, while higher scores were attributed to *Ease of use* and *Ease of learning*.

Table 2.2: Negative and positive aspects of the CAS, by number of time if was reported by students, organised in categories*(n=12)*

| | | Times reported |
|---|---|---|
| Negative aspects | Scripting | 19 |
| | Insufficient feedback | 5 |
| | Stability | 4 |
| | Realism | 2 |
| Positive aspects | Self-practice | 4 |
| | Ease of use | 3 |
| | Help to learn procedures | 10 |
| | Practical(fast,safe) | 8 |

Table 2.2 presents coded qualitative data from open ended questions listed at the end of the questionnaire. Students were asked to report three negative aspects, and three positive aspects of the CAS. The answers from the students have been categorized and suggest that the main negative aspect of the CAS is

related to scripting and does not allow to conduct a clinical history procedure adequately. The second downside reported is the lack of feedback. Stability is another issue, as students reported the simulator to crash repeatedly. Finally, realism was also reported as a negative aspect, with VPs not appearing as realistic as some students wished it to.

However, students also reported that the CAS helped them to learn procedures, pointed out its ease of use, the way it allows one to engage in self-practice, and other practical advantages such as going through cases fast, or practicing when they wanted to. These information helped us to prioritize the next design decisions.

### Prototype 2

#### Design

Following the first set of pilot studies, we refined the prototype according to the study limitations and the students' feedback. The main changes were as follows:

1. The CAS has being re-implemented using Windows Presentation Fundation (WPF). This allowed us to stop embedding an Ogre window into the application and running it in parallel. This process greatly improved the stability of the CAS, as students had reported crashes during the pilot studies. It also produced a more appealing interface.

2. Interaction scripts have been refined, extending the understanding of the VPs by adding more triggers for each of their answers. Participants from the pilot studies were asked to repeatedly interview the VPs. For

every instance where the VPs did not answer, new triggers were incorporated into the interaction scripts. This process was repeated three times, using two students for each intervention.

3. A new feedback component has also been implemented, which presents a completed DAHF once students submit their results. The purpose of this was to reinforce students' ability to adequately record information from a clinical history taking.

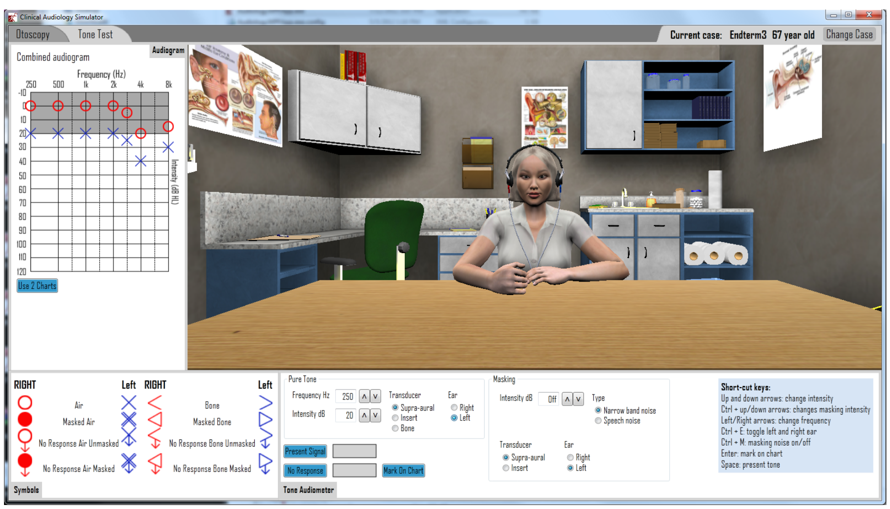
4. The VPs' models have being upgraded with a wider ranger of virtual characters covering both genders, different ethnicities, and a wider age range.

5. 3D models representing the different type of headphones used have been added to the system. Calculations for puretone audiometry are dependent of the type of headphones used, and having a visual reminder could reinforce students into applying the right formulas.

6. Idle animations, and facial expressions were also implemented, to add to the VPs behaviour which was judged as robotic by some of the students in the pilot studies.

Figure 2.4 presents the new interface. This prototype was assessed for both its clinical history and its pure tone audiometry component in a combined study presented in chapter 4. Usability was again assessed in this study and the results are presented in the next section.


(a) Prototype 2: clinical history taking interface




(b) Prototype 2: pure tone audiometry interface

Figure 2.4: Prototype 2: interface

### Usability

Six students who used the CAS were asked to fill out the USE questionnaire. Table 2.3 presents their results. Compared to the results on the first prototype, these results are better for each category of the questionnaire.

Table 2.3: Descriptive statistics for the usability questionnaire *(n=6)*

| Questionnaire category | Mean | Std. dev. |
|---|---|---|
| Usefulness | 5.35 | 0.50 |
| Ease of Use | 5.57 | 0.73 |
| Ease of Learning | 6.54 | 0.43 |
| Satisfaction | 5.36 | 1.21 |

In the qualitative items of the questionnaire (Table 2.4), the main downside identified was still the scripting of the interaction between students and the VPs during the clinical history taking procedures. However, while students pointed out that the scripting did not allow them to perform the procedure adequately in the pilot studies, this time around students were able to accomplish this task, but still asked for more improvements as some items appeared more difficult to retrieve from the VPs. As presented in chapter 4, a study using this prototype suggests that students significantly increased their ability to perform clinical history taking when assessed before and after exposure to the CAS. Some students still estimated that there is a lack of feedback, however they point out that they would like to see more options in the way feedback is presented. One student mentioned insufficient realism as a downside. Finally, no student reported stability as a problem anymore, which suggest that one of the main issues of the previous prototype has been resolved. The main advantages of the CAS listed by the students are still its practicality and how it supports them in learning procedures. Students still reported self-practice as an advantage as well, and one student reported ease of use as an advantage.

Table 2.4: Negative and positive aspects of the CAS, by number of time if was reported by students, organised in categories*(n=6)*

| | | Times reported |
|---|---|---|
| Negative aspects | Scripting | 6 |
| | Insufficient feedback | 3 |
| | Stability | 0 |
| | Realism | 1 |
| Positive aspects | Self-practice | 3 |
| | Ease of use | 1 |
| | Help to learn procedures | 5 |
| | Practical(fast,safe) | 8 |

Based on these results, it was decided to start the implementation of the next component which would allow users to practice speech audiometry. This is presented in the *Prototype 3* section. In addition, a separate prototype was being implemented at the same time, with the intent to test an additional formative feedback mechanism that could provide guidance to students while they are assessing a patient, and potentially make their learning of pure tone audiometry easier. This will complete the existing summative feedback component which presents correct results once students submit their results.

This implementation is presented in the *Prototype 4* section.

### Prototype 3

#### Design

The third prototype focused on implementing a new speech audiometry procedure. This procedure aims to assess a patient's ability to detect and understand speech stimuli in quiet and noisy environments. It provides information on how a patient performs at quiet and normal levels of conversation, and also cues whether a patient might need some kind or amplification or hearing aid. Figure 2.5 shows the speech audiometry interface. The virtual graphs and answer sections are located on the right of the VP. The bottom left section again presents the symbols used for this exam on the graph. The remaining of the bottom section allows the user to play lists of words, at different intensity.


Figure 2.5: Prototype 3: speech audiometry interface

In practice, a clinician presents to a patient lists of 10 words following the pattern consonant-vowel-consonant (CVC), one at a time, and asks the patients to repeat them. The patients are scored according to the number of correct phonemes repeated in each word. This process is repeated at different intensities in order to evaluate both the PI-Max, and the Half-peak level, of each ear. The PI-Max is the minimum level of intensity where the patient obtains maximum scores and is measured as a percentage. The Half-peak level is the intensity for which a patient would score 50% of the PI-Max. Students are asked to provide the estimated Half-peak level for each of the VP's ears, and whether these results are consistent with the audiograms determined during the pure tone audiometry procedure. In a similar fashion to pure-tone audiometry, once students submit their results they are presented with the correct graph representing the VP's speech results.

#### Usability

As part of the study focusing on speech audiometry (Chapter 5), the 18 students exposed to the CAS were asked to complete the USE questionnaire. Their results are presented in Table 2.5. Overall, students scored each category of the questionnaire higher than 5, which suggests that they were satisfied with the prototype.

Table 2.6 shows the qualitative responses of the students coded by category. The *Scripting* category has been removed as clinical history taking was not included. A *Usability* category has, however, been added. It represents students who did not think access to the simulator was sufficient in the classroom where it was setup, or layout issues for the speech audiometry feature.

Table 2.5: Descriptive statistics for the usability questionnaire*(n=18)*

| Questionnaire category | Mean | Std. dev. |
|---|---|---|
| Usefulness | 5.15 | 0.48 |
| Ease of Use | 5.49 | 0.19 |
| Ease of Learning | 5.96 | 0.29 |
| Satisfaction | 5.28 | 0.61 |

The main downside reported was insufficient feedback for the speech audiometry feature, the second was a lack or access or problems with the layout of the simulator. It is important to point out that half the students assessed were not familiar with the procedure at all, so problems with the layout or operation of the CAS was anticipated. Finally, three students reported minor stability issues.

However, overall, students suggested the the CAS's ease of use was its main advantage. Its practical advantages such as practicing within a safe environment, or having a wide range of cases available, and its ability to support the learning of the procedure, were also reported as being an advantage. Finally, only one student mentioned self-practice as being an advantage of using the CAS.

### Prototype 4

#### Design

During testing of the second prototype of the CAS presented in Chapter 4, no clear statistical evidence was found to suggest that students who were exposed

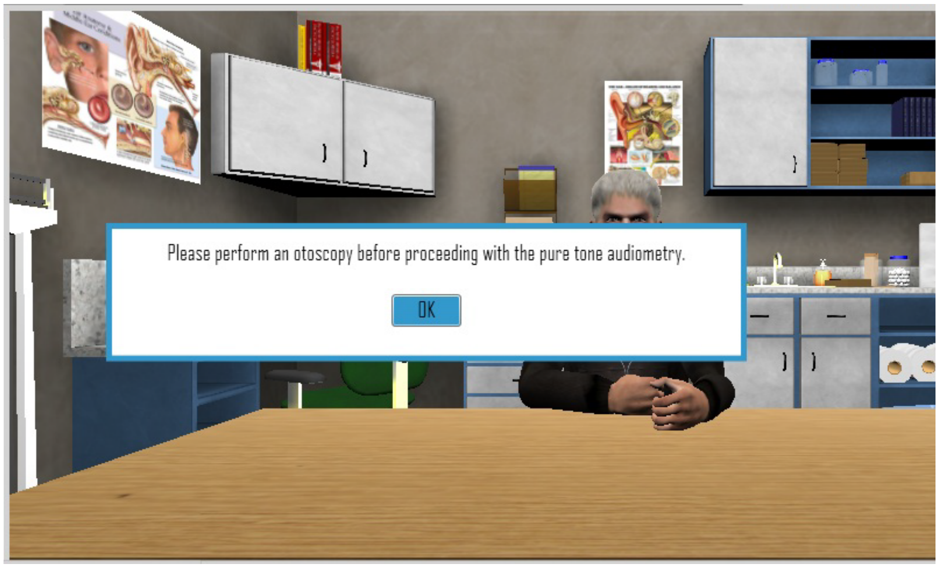
Table 2.6: Negative and positive aspects of the CAS, by number of time if was reported by students, organised in categories*(n=18)*.

| | | Times reported |
|---|---|---|
| Negative aspects | Insufficient feedback | 9 |
| | Stability | 3 |
| | Realism | 5 |
| | Usability | 10 |
| Positive aspects | Self-practice | 1 |
| | Ease of use | 17 |
| | Help to learn procedure | 11 |
| | Practical(fast,safe) | 9 |

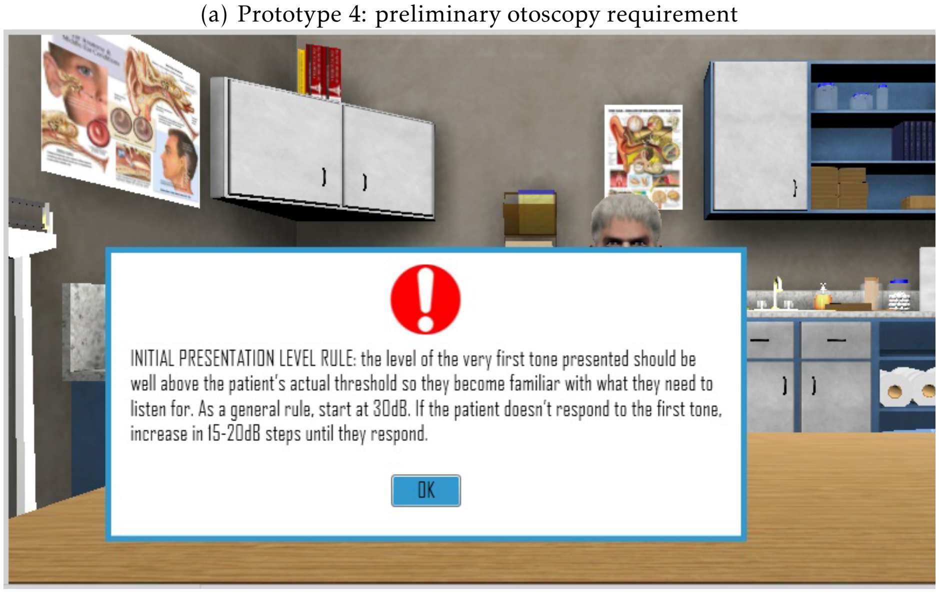
to the system increased their ability to perform the pure tone audiometry procedure when compared to students in a control group. Two explanations of this were that either the method of assessment was not adequate, or that the students who practiced pure tone audiometry on the CAS might have lacked guidance and feedback on their performance.

For this purpose, we decided to implement an additional feedback module which would present users with messages while they are conducting a pure tone audiometry assessment with a VP. This feature was meant to be optional and integrated in the whole system at a later stage as a tutorial mode. This feature's evaluation is presented in Chapter 6.

Four cases were implemented, with both directing feedback giving cues to students towards the next steps of the assessment, and validating feedback at each milestone. This feedback was provided through message boxes. As




(a) Prototype 4: preliminary otoscopy requirement




(b) Prototype 4: adequate intensity/frequency required

Figure 2.6: Prototype 4: directing feedback

seeing in Figure 2.6, students can for instance be directed into, (a) performing otoscopy before assessing the VP for pure tone audiometry, or (b) using the adequate intensity levels during the procedure.

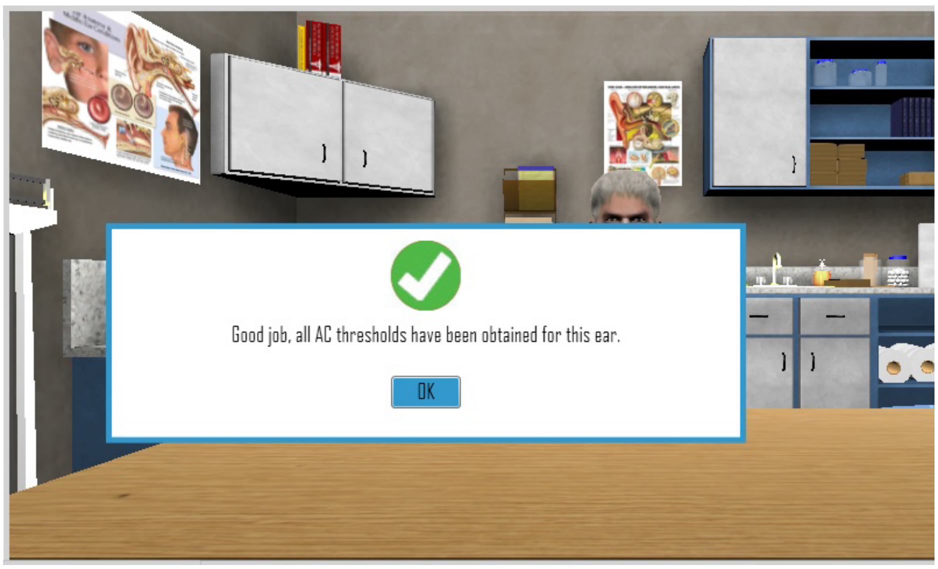



Figure 2.7: Prototype 4: validation feedback

Figure 2.7 presents one of the validation message, which is displayed to students upon acquiring all unmasked air threshold for the right ear. Validation messages appear once a complete threshold has been acquired, or upon completing more difficult tasks such as obtaining a masked threshold for a particular frequency.

**Usability**

In the study focusing on feedback presented in chapter 6, while all students practiced using the CAS, only 25 students received directing and validating

Table 2.7: Descriptive statistics for the usability questionnaire*(n=25)*.

| Questionnaire category | Mean | Std. dev. |
|---|---|---|
| Usefulness | 5.63 | 0.80 |
| Ease of Use | 4.84 | 1.02 |
| Ease of Learning | 5.33 | 1.08 |
| Satisfaction | 5.39 | 1.01 |

feedback. Their scores on the USE questionnaire are presented in Table 2.7. Overall, the scores for each category are well above the Likert-scales midpoint of 3.5. We notice however that the Ease of Use category is lower than the others. This could be explained by the complexity of the formative feedback messages, or by usability issues.

Table 2.8: Negative and positive aspects of the CAS, by number of time if was reported by students, organised in categories*(n=25)*

| | | Times reported |
|---|---|---|
| Negative aspects | Insufficient feedback | 10 |
| | Stability | 1 |
| | Realism | 0 |
| | Usability | 17 |
| Positive aspects | Self-practice | 4 |
| | Ease of use | 14 |
| | Help to learn procedure | 14 |
| | Practical(fast,safe) | 13 |

The coded qualitative answers from the students presented in Table 2.8

provide more insights on this issue. According to the data, students reported usability as been the main downside of the CAS. The students reported one recurring problem, where after making specific mistakes in attempting to retrieve the VPs' hearing thresholds, it was particularly difficult to carry on performing the procedure due to the way the feedback messages were implemented. One solution to fix this issue would be to implement an option to allow students to activate/deactivate these feedback messages at will. The second downside reported was insufficient feedback. Indeed, some students reported they would have wished to see more definition of the terms involved, and also more specific feedback at times. While I agree with their first point, I do not believe that the feedback should be more specific as to not provide students the answers to easily. Students should reflect through the feedback information before applying its content.

The ease of use, the ability to learn pure tone audiometry, and the practicality of the CAS were still reported as the main advantages of the system. Meanwhile, self-practice was only reported four times as been an advantage of using the CAS with formative feedback messages to learn pure tone audiometry.

### Case editor

Another goal was to provide audiology educators with an authoring tool that allows them to manipulate cases and implement new ones. This section describes the use of the Audiology case editor. The different features have been integrated to meet the needs of the Clinical Observation and Practice course coordinator, by following his recommendations.

This editor allows for the creation and modification of cases which are stored as *.xml* files. The user can choose to start with a blank case, or load an existing one which will serve as a basis.

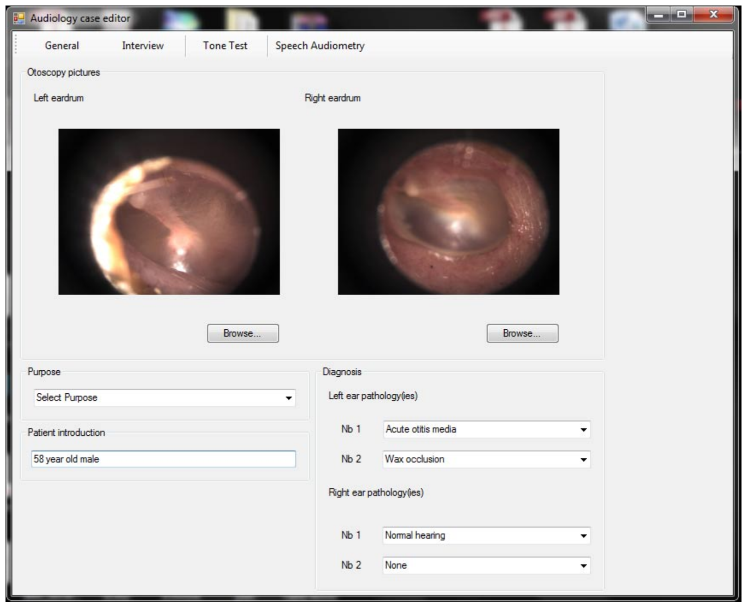



Figure 2.8: Case editor: general panel

The case editor is split in 4 sections, each located on a different panel of the application. The *General* panel presented in Figure 2.8 allows to set the otoscopy pictures by selecting pictures of clients' eardrum from our sample which covers the range of different pathologies. The diagnosis section also allows to associate up to two pathologies to each ear. An educator can also choose to add a brief patient introduction which will be presented to the user at the start of the case. The pictures are not linked to the pathologies associated, consequently creating the cases do require some expertise in audiology. Finally, the purpose dropdown button allows to select the mode of use of the CAS associated with this case. Selecting ToneTest would for instance only associate this procedure to this virtual case.

Further work will include another option to select interview scripts from an existing set to associate with this case for the clinical history taking component of the exam.

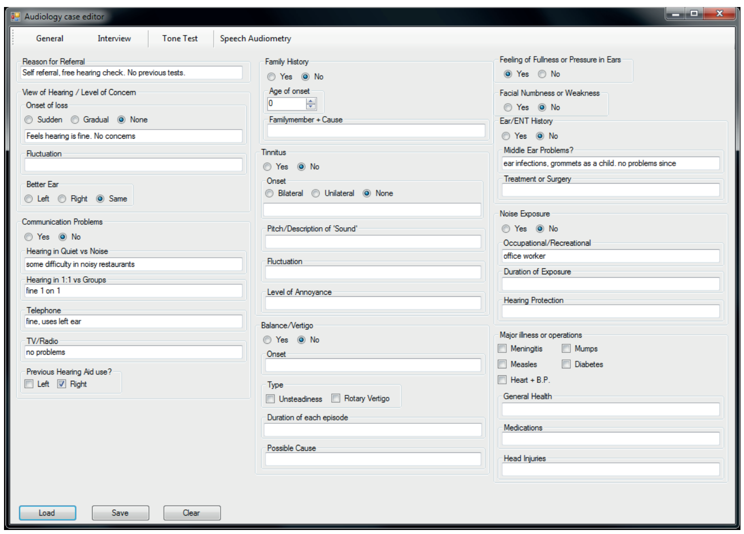



Figure 2.9: Case editor: interview panel

The *Interview panel* (Figure 2.9) allows to set the answer that would have to be reported onto a DAHF once the clinical history part of the exam is completed. Moreover, it enables the system to present realistic answers in the form of a DAHF once users submit their results.

Figure 2.10 named *Tone Test* panel is the interface that allows the user to set the hearing thresholds associated with the case. While typing the values for both ear's Air and Bone thresholds, the associated audiograms are

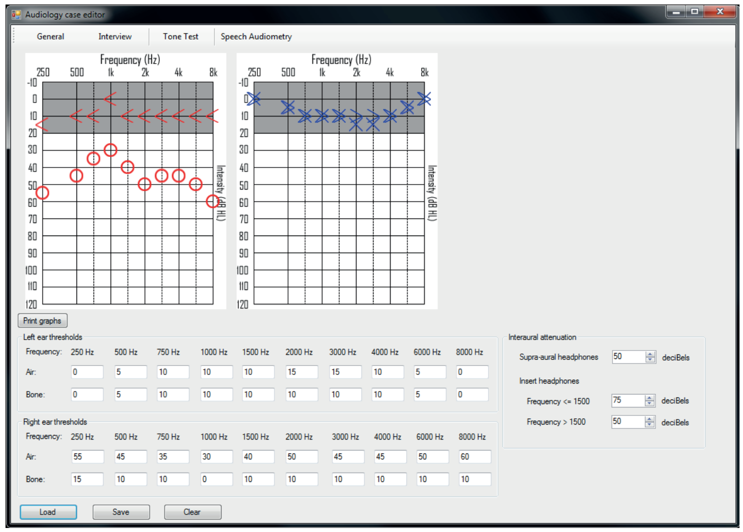



Figure 2.10: Case editor: tone test panel

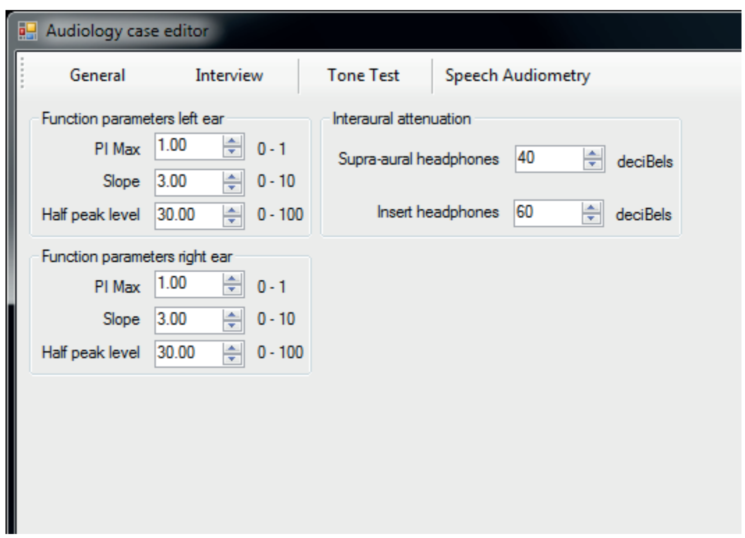



Figure 2.11: Case editor: speech audiometry panel

automatically completed. An option allows the user to print those audiograms for archiving purposes or future reference.

The interaural attenuation for both types of headphones, Supra-aural, and Insert, can also be altered from their standard values. This can provide students with more atypical cases.

The fourth panel, *Speech audiometry*, allows the user to set the the parameters of the function that calculates the speech audiometry graphs. For each ear, the user has to provide the PI Max, the slope of the curve, as well as the Half-Peak level. The Interaural attenuation section allows again to modify the values of the interaural attenuation for each type of headphone used, but only for the speech audiometry procedure.

## 2.3 Nursing simulator

### Design

Other than the CAS, the nursing simulator is focusing on the treatment of a single patient. The goal is to have nurses start caring for a patient who has been admitted in rapid respiratory distress. The procedure involves assessing the patient over a period of time and making the right decisions depending on how his condition evolves.

A nurse can ask the VP questions, and use various tools to assess his condition before providing him with different treatments and medication/drugs. The nurse can then take the decision to go and care for other patients before returning at a later stage to see how the situation evolved.

To keep this scenario simple, the VP has been implemented so that his vitals will keep on deteriorating as time passes. A mysql database has been set in place where for each *time* a nurse assesses the VP, a different set of vital information is presented, and the VP gives different answers. The nurses' objective is to apply the appropriate treatments until she concludes that nothing can be done in the current ward, and that the patient has to be transferred to the Intensive Care Unit (ICU).

This section will present how this scenario is used, based on the first prototype of the nursing simulator.

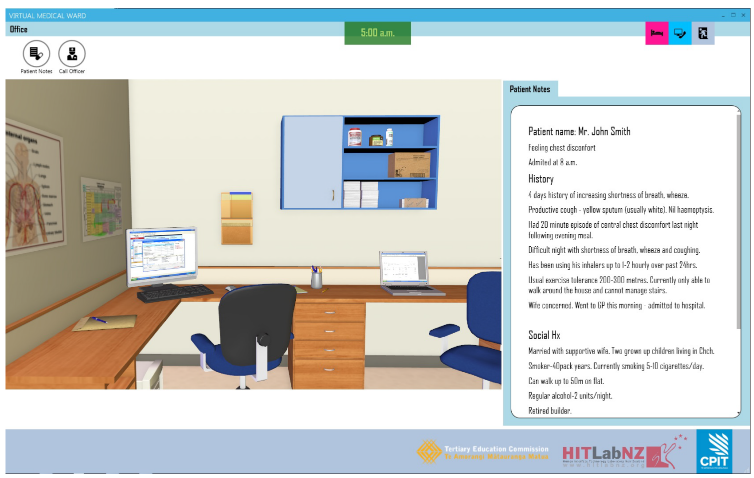



Figure 2.12: Start of the scenario

The nurse starts after the VP has been admitted. Consequently, she would typically start the care of this new patient by consulting the notes another clinician has left for her.These notes are presented in Figure 2.12 in the office.

Following the assessment of the clinician notes, the nurse is able to ask the VP questions to assess his level of pain for instance, or if he is disoriented. Basic observational tools are available and can be used to retrieve more information on the patient's vital functions. These features are enabled when seeing the patient (Figure 2.13).

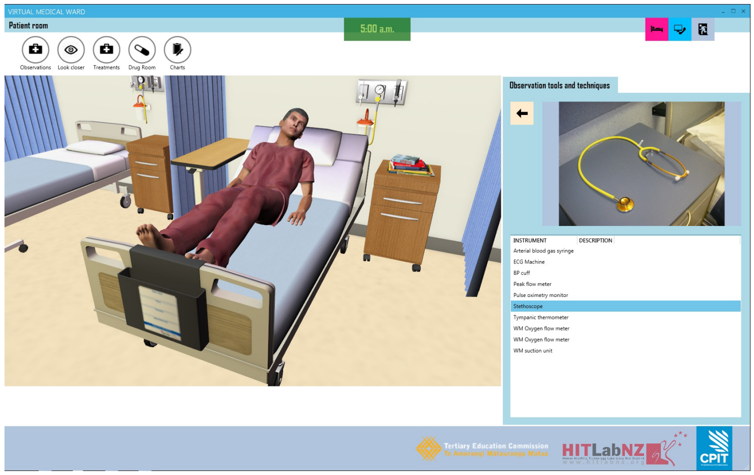
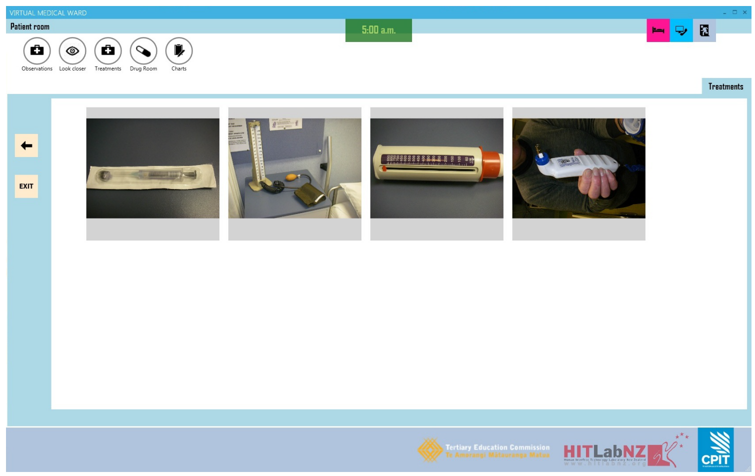


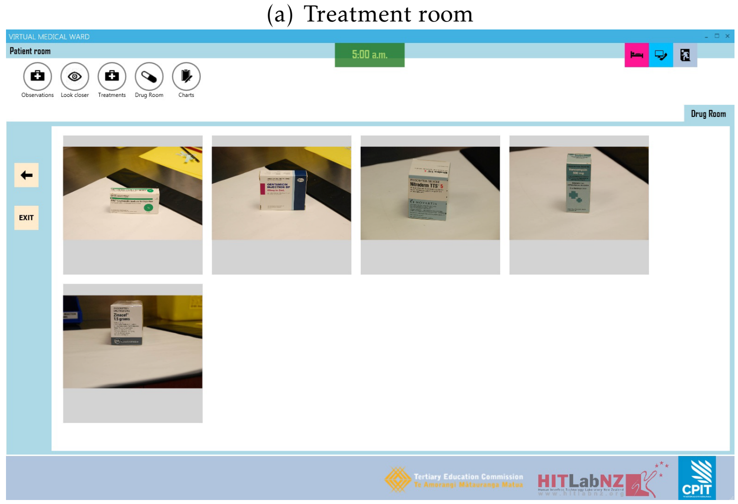

Figure 2.13: Patient with obervation tools

Further treatments can also be applied to the patient by visiting a treatment room, where tools are organised in shelves, each including a set of treatment options (see Figure 2.14 (a)). This room contains the equipment that would not be immediately available in a patient wards, or that a nurse would not have on her person. The information gathered through observation, treatments, and dialogue with the VPs are then reported on an Observation Chart (Appendix C). The user has then to identify *Neuro*, *Early Warning Score (EWS)*, and *Pain score*, on the bottom of the chart. These scores are stored each time the patient is being assessed and directly correlate with the need to transfer the patient to the ICU.

Another task a nurse has to take care of is providing the right medication recommended by a doctor, the drug treatment sheet (Appendix D) contains




(a) Treatment room




(b) Drug room

Figure 2.14: Treatment and drug rooms

the name of the drugs as well as the quantities and the number of times they should be given to the VP. The nurse can then use the drug room, to choose the right drug from organised shelves and deliver the correct amount to the VP. The interface associated with this task is presented in Figure 2.14 (b).

These steps are the ones typically followed when assessing a patient who is treated for rapid respiratory distress. At this stage a nurse would leave the patient rest, and return at a later stage to re-assess the situation. In this simulator, the process described above can be repeated up to five times, with the VP's vitals, ability to communicate, and skin colour deteriorating each time. The ideal course of action is that by the third visit, a nurse would come to the conclusion that nothing can be done to improve the VP's condition in the current ward. She should phone the ICU, requesting the VP to be transferred using a phone located in the office.

All information, such as questions asked, treatments or drugs provided at each step can be stored by the nursing simulator for a later debriefing on the nurse performance.

### Evaluation

A preliminary evaluation of the system took place with three educators from CPIT's school of nursing as participants. These participants used the prototype of the nursing simulator and were asked to complete a questionnaire following this experience.

An instrument from the National League of Nurses (NLN)[3] was used to assess this simulator. The NLN instrument focuses on three distinct aspects;

[3] http://www.nln.org/index.cfm

Students' satisfaction and self-confidence in learning, simulation design scale, and educational practices.

The questions were answered on five point Likert-scales, where participants were asked to answer statements between *1-Strongly disagree*, and *5-Strongly agree*. In addition, for each answers given to the simulation design scale, and educational practices questionnaires, participants had to rate how important they perceived each item was, between *1-Not important*, and *5-Essential*.

The results of the satisfaction and self-confidence in learning questionnaires are presented in Table 2.9. The scores for both categories are above the midpoint of 2.5, but it also shows that further refinement is necessary to obtain the users' satisfaction. This refinement would primarily focus on providing more guidance. The system should present hints to the users on the next steps of the scenarios, as participants assessed required some direction during the experiment.

Table 2.9: Descriptive statistics for the educators satisfaction and self-confidence in learning NLN questionnaire*(n=3)*

| Questionnaire category | Mean | Std. dev. |
|---|---|---|
| Satisfaction with learning using the VP | 3.92 | 0.52 |
| Self-confidence in Learning | 3.67 | 0.08 |

As seen in Table 2.10, the three categories of the simulation design scales are rated with a similar level of importance of 3.75-3.80. Again, the scores of the participants are still above the 2.5 midpoint, with a mean of 3.00 for the *Obective and Information* category, 3.56 for the *Support* category, and 3.63 for the *Feedback/Guided Reflection* category. Similar results showed from analysing the educational practices questionnaire (Table 2.11). Participants rated the importance of each category between 3.83 and 4.00, and while all means were above the 2.5 midpoint, only the results for the *Diverse Ways of learning* can be considered high at 4.17. Overall, participants rated the *Importance* of those two questionnaires' categories similarly, with the lowest being *Support* (3.75), and the highest being *Diverse Ways of learning* (4.00).

Table 2.10: Descriptive statistics for the simulation design scale NLN questionnaire*(n=3)*

| | Score | | Importance | |
|---|---|---|---|---|
| Questionnaire category | Mean | Std. dev. | Mean | Std. dev. |
| Objectives and Information | 3.00 | 0.20 | 3.80 | 0.87 |
| Support | 3.56 | 0.09 | 3.75 | 0.00 |
| Feedback/Guided Reflection | 3.63 | 0.53 | 3.75 | 1.06 |

Table 2.11: Descriptive statistics for the educational practices NLN questionnaires*(n=3)*

| | Score | | Importance | |
|---|---|---|---|---|
| Questionnaire category | Mean | Std. dev. | Mean | Std. dev. |
| Active Learning | 3.48 | 0.67 | 3.95 | 1.03 |
| Diverse Ways of learning | 4.17 | 0.76 | 4.00 | 0.00 |
| Expectations | 3.17 | 0.29 | 3.83 | 1.04 |

The analysis of the nursing simulator prototype using the NLN instrument shows that the system could have a great potential, with positive scores for a first prototype. These results are encouraging and suggest that further refinement could really meet the users' need in clinical nursing. While most measures are only slightly above the medians, users seem to already respond positively to the diverse ways of learning that the simulator provide, and they are also satisfied learning using the VPs.

Unfortunately, I did not have the opportunity to design follow up versions of this system due to the reasons mentioned in this thesis' introduction. I hope to resume work on the nursing simulator upon the completion of my PhD, and to have the opportunity to conduct experiments using this simulator as part of the nursing curriculum of CPIT, to try and evaluated its impact on students' learning gains.

# 3 Clinical history taking and pure tone audiometry pilot studies

The content of this chapter has been accepted for presentation at the 18th Annual CyberPsychology and CyberTherapy Conference.

## 3.1 Introduction

Virtual Patients (VPs) have been used in a wide range of fields to train clinical reasoning and patients' assessment in different settings. However, there is a gap in the body of research when it comes to VP systems for clinical education. Rather than assessing the continuous use of virtual patient based systems by medical trainees and their educational impact, evaluations often focuses on comparing the quality of the VP to a Simulated Patient (SP) in terms of user's engagement, and realism (Kyle Johnsen and Andrew Raij, 2005b; Vash et al., 2007; Tan et al., 2010). Moreover, most systems research offered a one time experience to students focusing on a single case (Bickmore and

Picard, 2005; Deladisma et al., 2009; Fuhrman Jr et al., 2001; Gorrindo and Groves, 2009; Hayes-roth and Saker, 2003; Heer et al., 2004; Kenny et al., 2007; Kotranza et al., 2009; Rizzo et al., 2010, 2011; Ruttkay and Welbergen, 2008; Stanton, 2008). This finds an explanation in the fact that developing VP interaction scripts working as standalone, without the assistance of a facilitator puppeteering in the shadows, is a resource intensive process. Such a script usually starts small with a basic set of questions and answers, before being strengthened through pilot studies and refinement. Work has been done on implementing ways to create more robust scripts efficiently for VP interactions, but it is still an area that needs to be improved (Halan et al., 2010; Rizzo et al., 2010; Rossen et al., 2010).

Our research uses the field of Audiology as a case study, but findings could extend to other clinical fields. In audiology, simulations have been used with success to teach clinical skills. Simulators allowing practice of procedural skills, such as pure tone audiometry are available on the market (e.g. Otis Audiology simulator [1], Parrot Software's Audiology Clinic [2]) for Universities to use in supplementing traditional course work. However, at the time of our research we are not aware of simulators supporting clinical history taking for trainee audiologists as well as other procedures.

This research explores the use of a computer simulation based on a varied range of VPs, with audiology trainees as our primary target audience. The first prototype of the Clinical Audiology Simulator (CAS) we implemented

[1] INNOFORCE creative solutions website http://www.innoforce.com/

[2] Parrot Software website http://www.parrotsoftware.com/parrotstore/Products/120-audiology-clinic-full-version.aspx

provides students with the opportunity to practice history taking and pure tone audiometry skills at their own pace on 25 virtual cases, each representing a different patient.

This chapter presents the findings of two pilot experiments, using first year audiology students as our sample. The first experiment focuses on pure tone audiometry, the second on clinical history taking. These pilot experiments also aim to assess the research setup and the evaluation instruments.

## 3.2 Pure tone audiometry pilot study

This section presents the early research on evaluating the impact of using the CAS as part of the master of audiology course to support first year student's training of the pure tone audiometry procedure. This research was conducted using our first working prototype of the CAS. The experiment aimed to answer the following hypotheses while also testing the functionalities of our system presented in chapter 2, and the design setup:

1. Students' ability to conduct a pure tone audiometry procedure will increase as a result of using the CAS, compared to students who were only given traditional instruction.

2. Students' perceived level of learning will increase as a result of using the CAS to practice pure tone audiometry, compared to students who were only given traditional instruction.

### Evaluation

This experiment assessed students on the pure tone audiometry exam. This exam is centered around gathering patients' hearing thresholds, based on their response from varied sound stimuli. Students have to learn the procedure as well as the reasoning behind this process. These results have then to be reported appropriately on a clinical form to be added to the patient's file and contribute to the overall diagnosis.

#### Participants

The University of Canterbury's Master of Audiology Degree (MAud) is spread over two years, with the bulk of theoretical teachings in the first year. Entry into the MAud is very restricted and competitive and only 10 to 12 students per year are accepted into the program. The primary reason for this is that audiology training, just like any healthcare profession, is resource intensive, particularly in term of practical training. However, within the audiology field in a small country such as New Zealand, training opportunities are even more limited. According to audiology teachers, the first year students are the ones benefiting the most from extra training opportunities as they are relatively new to the field, starting with virtually no previous domain knowledge. First year students undertake as part of their course one weekly day in clinics, to observe and sometimes practice. In addition their academic year ends with a three months full time placement in a clinic. From the educators perspective, having a supplementary training tool could contribute to students having the set of required skills to make the best of their summer placement.

A group of 12 students was recruited, 10 females and two males, aged between 22 and 55 years old. All participants held at least a Bachelor degree. Each student was enrolled in the Clinical Observation and Practice course which includes one day a week of observation and practice at various clinics in the Christchurch area. Each group was made of 5 female participants and 1 male participant, while the mean gpa (grade point average) of the two groups was counterbalanced (Group A mean gpa=7.48, sd=0.97, and Group B mean gpa=7.46, sd=0.87).

**Design**

The Pure tone audiometry pilot study took place in the first semester of the academic year. Participants were split into two groups, Group A and Group B. Following a presentation of the CAS as part of one of their course's tutorial session, participants in Group A got access to the simulator in addition to their in-class teaching for a period of two weeks. During that time, Group B was only training using traditional means of learning. After this first training phase, both groups had access to the simulator for a week in order to prepare for Test 1, their midterm assessment and extra role-play assessment. In addition, participants answered a questionnaire on their experience with the CAS.

After these assessments, the conditions were reversed and Group B had access to the CAS for a period of two weeks in addition to their courses while Group A was only using traditional learning methods. After these two weeks, both groups had again access to the simulator in order to prepare for their end of term assessment (Test 2). This assessment also got followed by a second role play assessment, and another CAS questionnaire (Figure 3.1).

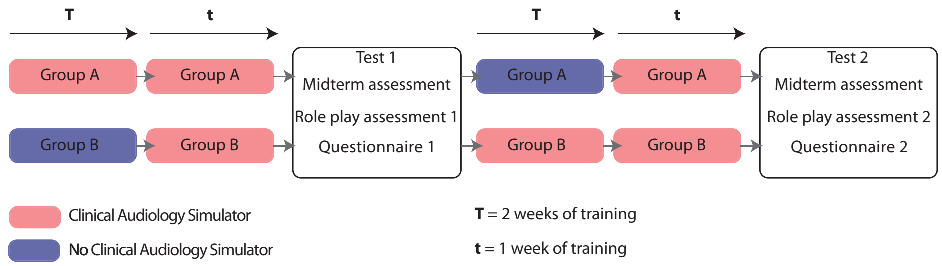



Figure 3.1: Design of the pure tone audiometry pilot

### Measures

Midterm and end of term assessment grades were used as a measure for this experiment. For both these assessments, participants had to determine and report on paper the hearing thresholds of three different patients using the CAS. Marks were allocated depending on how close the hearing thresholds reported were to the actual threshold of the patients implemented, out of 100.

Role-play assessments took place in the audiology clinic located on campus where a research assistant played the role of Simulated Patient (SP) in realistic clinical settings. Participants were asked to perform a pure tone procedure on the SP and report their results. Marks were allocated to participants following typical clinical work marking criteria where in addition to actual threshold validity, points were also granted for method, pace, confidence, and explaining results (Appendix E). A questionnaire was given to the participants following the role play situations. The main measure of interest in this questionnaire was a subjective evaluation of students' perceived level of learning (Appendix F), out of five points.

### Results

For the perceived level of learning scores, we used a nonparamatric independent sample *u*-test. This analysis showed significant differences for perceived level of learning between the two groups at Time 1 ($U$=1.50, p=.007), but not at Time 2 ($U$=15.50, p=.684).

For the learning outcomes, we used t-tests to analyse the differences in role-play assessments and simulator assessments between the two groups at Time 1, at Time 2, and on the difference off scores between Time 2 and Time 1. No statistical differences were found between the two groups for these measures.

The means and standard deviations for these variables are presented in Table 3.1.

### Discussion

Preliminary data was analyzed to explore if the additional exposure to our VP based simulator could increase audiology students' perceived level of learning and performance. Students were assessed using both the CAS and with SP based role-play scenario. First, as expected, the students did improve their score independently of their learning group on perceived level of learning, Simulator assessment, and Role-play assessment over the course of the experiment.

At the start of the experiment we hypothesized that students' perceived level of learning would increase following extended exposure to the CAS. The results support our hypothesis as Group A showed at Time 1 significantly higher perceived level of learning, following their dedicated period

Table 3.1: Descriptive statistics for perceived level of learning, role play assessments, and simulator assessments at Time 1, Time 2, and the difference between Time 2 and Time 1*(n=12)*

| | | Time 1 | | Time 2 | | Time 2 - Time 1 | |
|---|---|---|---|---|---|---|---|
| | Group of the student | Mean | Std. dev. | Mean | Std. dev. | Mean | Std. dev. |
| Perceived level of learning | Group A | 3.57 | 0.50 | 3.70 | 0.55 | 0.13 | 0.33 |
| | Group B | 1.80 | 0.90 | 3.60 | 0.49 | 1.80 | 1.26 |
| | Total | 2.68 | 1.16 | 3.65 | 0.50 | 0.97 | 1.24 |
| Role play grade | Group A | 74.31 | 6.32 | 78.94 | 17.33 | 4.63 | 15.46 |
| | Group B | 68.29 | 14.55 | 80.56 | 9.42 | 12.27 | 13.87 |
| | Total | 71.30 | 11.15 | 79.75 | 13.33 | 8.45 | 14.56 |
| Simulator assessments | Group A | 82.22 | 11.26 | 88.62 | 7.31 | 6.40 | 6.66 |
| | Group B | 84.92 | 7.08 | 90.71 | 6.02 | 5.79 | 7.68 |
| | Total | 83.57 | 9.08 | 89.66 | 6.47 | 6.09 | 6.86 |

of exposure to the CAS. Following group B's exposure, this difference is not significant anymore which could indicate that they caught up with their peers (Figure 3.2). Statistical equivalence has however not been calculated.

The second hypothesis was that students would get better grades when assessed following their exposure to the CAS. The results, however, show that while there is an increase in students grades when assessed over time, for both types of assessments, both groups score increase similarly.

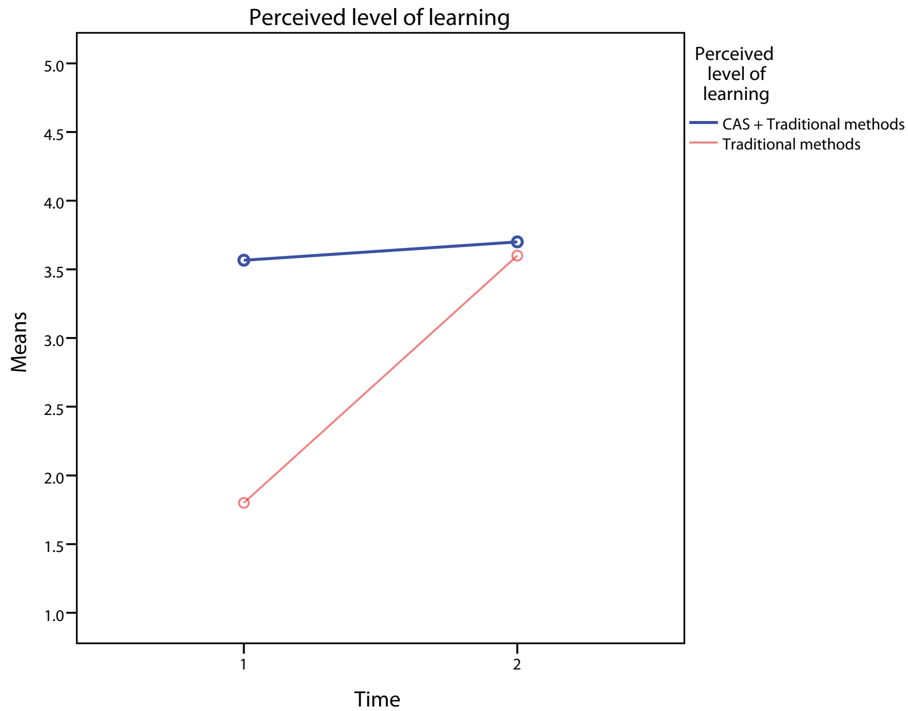



Figure 3.2: Perceived level of learning scores in function of time

### Limitations

The main limitation of this study is that both groups received one week of training with the CAS right before each assessments. This was one of the restriction of testing students as part of the Clinical Audiology course, and one of the course requirements; every students had to have access to the CAS before those course assessments. However, this allowed us to use both students' midterm and end of term assessments as measures.

This week of common training, however, seemed to have leveled potential differences between the two groups. A few students reported during informal discussions that they got the most of their training done during the week prior to each assessment, whether in group A or group B. Additionally, it is possible that students generally spent more time training on the CAS to prepare for the end of term assessments than for midterm's. We concluded that there are two ways to remedy this problem.

First is to remove the common week of training both groups had. Second, is to designate specific training times for students in both groups, and have them book specific time slots to ensure that the student group exposed to the CAS receives a sufficient amount of additional training and control CAS exposure times between the two groups. This would also mean to abandon mid-term assessments and end of term assessment as measures. Indeed as mentioned earlier, to guarantee fairness for the students, they should all have access to the CAS before any of their course assessments. This means that the study needs to run before their midterm, while leaving enough time for the control group to practice as well following the study, in preparation of their exam.

In addition, it seems that an increase in students' perceived level of learning does not correlate with an increase in results. This measure will be removed as well from the following studies which will focus on the transfer of skill assessed with role-play situations.

## 3.3 Clinical history taking pilot study

This experiment was conducted using the same prototype of the CAS. While testing the functionality of our system which are presented in chapter 2, and the design setup, this experiment also aimed to answer the following hypotheses:

1. Students' confidence when conducting a clinical history interview will improve as a result of using the CAS, compared to students who were only given traditional instruction.
2. Students' ability to retrieve information in clinical history situation will improve as a result of using the CAS, compared to students who were only given traditional instruction.
3. Students' ability to accurately report information in clinical history taking will improve following exposure to the CAS, compared to students who were only given traditional instruction.
4. As a result of using the CAS in addition to traditional instruction, students will take clinical history more efficiently and will require less interaction with patients to retrieve relevant information, when compared to students who were only given traditional instruction.

### Evaluation

This study focused on assessing students on one of the main clinical exams our system supports, clinical history taking. Clinical history taking is centred around learning how to interview patients to retrieve key information that help determining an accurate diagnosis and follow up procedures. Assessments are conducted using role-play with an SP in order to test for transfer of skills from use of the CAS. While SPs are not real patients they allow for a simulated experience which is the closest to real practice. In addition, SPs enable to standardize the assessment, presenting a similar patient to each participant.

#### Participants

The same participants as in the pure tone audiometry pilot study were recruited for this experiment. Nine students reported English as being their native speaking language, two reported German, and one reported Chinese. All students are fluent in English. The 12 students declared having adequate vision and hearing to use the simulator, as well as conduct and record a medical history from a patient, as well as having adequate hearing. The twelve students answered they considered having the necessary level of skills required to operate the CAS.

#### Design

Participants for this study were split into two equal groups according to the results of a pre test role play situation. Following discussions with the course coordinator and other audiology expert, Students' grade point average had been determined to not be a valid representation of clinical interview taking abilities, thus could not be considered as a means to group students for this study. Indeed, clinical history taking skills somehow differ from other procedures as students need to ask questions accordingly, all the while interpreting each of the patients' answers to contribute towards a diagnosis.

Following the pre-test (Test 1), participants were split into two groups, CAS and no CAS balanced based on the accuracy, confidence, and efficiency results of this test (see the following section for more details on these measures). Students in the CAS group then practised for a period of two weeks in addition to traditional teaching while students in the no CAS group only followed the traditional instruction. Students practising on the CAS agreed to train at least two hours. At the completion of this training phase both groups undertook a second role play assessment (Test 2), with a similar assessment method to Test 1 (see Figure 3.3).

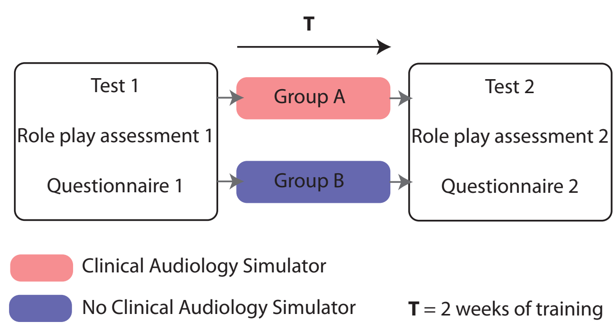



Figure 3.3: Design of the clinical history taking pilot

### Measures

Interview taking skills are not typically graded as part of the first year audiology students' curriculum. Our main focus was to measure students' abilities during role play situations. While quantifying the improvement of the students is outside the scope of this thesis, we assess whether the CAS had a positive impact on students interview taking skills. The role play simulations took place in a realistic setting, with two audiology experts acting as SPs; one for assessment 1, the other for assessment 2. During those experiences students were assessed on accuracy, where students had to retrieve information from the VPs and translate them onto standard history sheets, confidence and efficiency. Assessments were undertaken by two audiology experts to ensure reliability.

**Verbal Accuracy**: The verbal accuracy score is based on the number of answers participants were able to retrieve from the SP while talking during the interview. This measure was assessed using a transcript of the interaction between participants and SPs. Transcripts were recorded during student-SP interactions by one of the audiology experts. It is presented as a percentage.

**Written Accuracy**: The written accuracy score points to the number of adequately reported pieces of information on a history diagnosis form used in clinics to report information critical to diagnose the patients. Students were asked to fill in the diagnosis form while interaction with the SP, as they do during clinical exams. Clinical forms were then checked by both audiology experts and marked. This measure is presented as a percentage.

**Confidence**: for each role play situation, participants rated pre and post simulation confidence in their performance using three seven point Likert scale items.

**Number of questions**: for each role play situation, this measure is the number of questions a student asked the SP before considering the clinical history complete.

**Efficiency**: the efficiency score is calculated by using the verbal accuracy raw scores divided by the number of questions asked to an SP.

### Results

We used a multivariate analysis of variance (MANOVA) to analyze the results of the clinical history taking study. Group was used as a between factor and the Time 2 confidence, verbal accuracy, written accuracy, number of questions asked (to SP), and efficiency measures were used as dependent variables. Time 1 confidence, verbal accuracy, written accuracy, number of questions asked (to SP), and efficiency measures were used as covariates. The means and standard deviation for these variables are presented in Table 3.2.

The MANOVA shows that there were no significant differences between students in the CAS and no CAS group at Time 2 for any of the dependent variables tested.

Students who received additional training with the CAS had a somewhat higher increase in verbal accuracy between Time 1 ($m$=42.64, $sd$=8.91) and Time 2 ($m$=72.48, $sd$=8.75), compared to students following only traditional instruction, Time 1 ($m$=42.33, $sd$=8.92); Time 2 ($m$=60.93, $sd$=7.05). However this difference is not significant.

### Discussion

This pilot study tested the first prototype of the CAS and explored how the additional use of the CAS to traditional methods of teaching impacts on students' confidence, accuracy and efficiency when practising clinical history taking.

Table 3.2: Descriptive statistics for confidence, verbal & written accuracy, number of questions, ant efficiency at Time 1, Time 2, and the difference between Time 2 and Time 1*(n=12)*

| | | Time 1 | | Time 2 | | Time 2 - Time 1 | |
|---|---|---|---|---|---|---|---|
| | Group of the student | Mean | Std. dev. | Mean | Std. dev. | Mean | Std. dev. |
| Confidence | CAS | 3.78 | 0.73 | 4.86 | 0.77 | 1.08 | 0.59 |
| | No CAS | 3.63 | 0.99 | 4.33 | 1.25 | 0.70 | 1.01 |
| | Total | 3.71 | 0.81 | 4.62 | 1.00 | 0.91 | 0.79 |
| Verbal accuracy | CAS | 42.64 | 8.91 | 72.48 | 8.75 | 29.85 | 8.63 |
| | No CAS | 42.33 | 8.92 | 60.93 | 7.05 | 18.60 | 14.43 |
| | Total | 42.49 | 8.46 | 67.23 | 9.73 | 24.74 | 12.45 |
| Written accuracy | CAS | 65.65 | 16.85 | 81.09 | 13.17 | 15.44 | 21.15 |
| | no CAS | 59.03 | 21.42 | 76.60 | 14.86 | 17.57 | 24.55 |
| | Total | 62.64 | 18.37 | 79.05 | 13.44 | 16.41 | 21.59 |
| Number of questions | CAS | 24.67 | 8.04 | 44.50 | 6.09 | 19.83 | 5.56 |
| | No CAS | 22.40 | 6.35 | 38.00 | 6.52 | 15.60 | 6.27 |
| | Total | 23.64 | 7.06 | 41.55 | 6.86 | 17.91 | 6.01 |
| Efficiency | CAS | 1.33 | 0.21 | 1.43 | 0.14 | 0.10 | 0.14 |
| | No CAS | 1.24 | 0.30 | 1.45 | 0.13 | 0.20 | 0.24 |
| | Total | 1.29 | 0.25 | 1.44 | 0.13 | 0.15 | 0.19 |

We hypothesized that students' confidence level in their performance would increase following a dedicated training period with the CAS as an additional practice tool. The results show that both groups had a similar increase in confidence. We could find no evidence that practicing with the CAS increases confidence levels.

For accuracy measures, we hypothesized that both verbal and written accuracy would increase for students in the CAS group. Results show that verbal accuracy, which represents the number of answers students were able to retrieve from the SP, increased for all students and no significant differences between groups were found. Also for written accuracy, which is the percentage of adequately reported information from the verbal accuracy scores, we could not find significant differences between groups. Students seem to perform better than in the initial assessment whether or not they had access to the simulator.

Our final hypothesis was that efficiency would improve for students who used the CAS in addition to traditional methods. We characterize efficiency with two measures: the number of questions students ask the SP in role play, and the relation between the number of questions asked and answers retrieved from the VP. The results did not show a significant effect of CAS training on the number of questions asked to SPs in role play. It seems that over time, students generally started by asking only a small number of questions to the SP during the assessment, then experimented with asking the SP more questions during the second assessment. Our main efficiency measure, the number of questions asked divided by verbal accuracy, however, decreased between the first and the second assessment. This can be attributed to students experimenting with the number of questions they could ask SPs, as mentioned above. No significant differences between groups could be found to support our hypothesis.

### Limitations

The main limitation concerning the clinical history study is the VPs' scripting itself, which was still early stage. While students did have positive results using the VPs, some reported having difficulties in retrieving specific information from them. This was due to the way they asked questions to the VP, that was not adequately recognized by the system. We believe that by improving patients' interaction-scripts, we could obtain more positive results in further studies. It is also important to mention that scripts' implementation and robustness are a crucial element in VP based systems used in those studies (Halan et al., 2010; Rizzo et al., 2010; Rossen et al., 2010). Despite not having captured data on students' frustration, it is clear that a student is unlikely to spend a large amount of time practising if he finds the task daunting due to low responsiveness from the VPs.

Another method that could be followed in further studies could be the approach designed by Rossen et al.(2010), which presents a system aiming to improve virtual agents scripts implementation. It starts by recording interactions between multiple humans taking turn interacting with a VP, before formulating an aggregate encompassing all the previous interactions within a single script. This process should be iterative, and undertaken alongside field experts (Rizzo et al., 2011). It is also a process that takes considerable time commitment in order to capture the wide range of conversational stimuli available (Halan et al., 2010).

We suggest that interaction scripts should start being tested as early on as possible when implementing a VP simulation system, to allow enough time for sufficient refinement.

A further limitation of this study was the lack of control over students' commitment to training using the CAS, and the size of the sample. The Master of Audiology is already an intensive degree and asking students to train in their free time, while dealing with an already full schedule and assessments due for the different courses they were enrolled in proved difficult. Students admitted after the study ended, during an informal meeting, that they did not feel as though they had enough practice time on the CAS. This could explain the lack of significant differences between groups for this study. Another concern was the limited number of participants available due to the small number of students enrolled in this degree.

## 3.4 Conclusion

These pilot studies aimed at testing our first prototype, the research setup and our measurements' instruments. Our current studies have found no clear statistical evidence on the effectiveness of using a VP simulator as a training tool for audiology trainees' history taking or pure tone audiometry skill acquisition. However, the data suggests that practicing on the CAS improved perceived level of learning. In addition, we identified limitations which have been taken into account for the next studies.

As discussed in chapter 2, students identified the lack of response from the VPs, which was due to the state of the interaction scripts, the lack of feedback, and the lack of stability of our system as the main downsides of this prototype.

The step undertaken to address those limitation are presented in the introduction of chapter 4.

While we need more in depth studies to evaluate the effectiveness of the CAS as part of the Clinical Audiology course, at this stage we conclude that additional training with the CAS had some impact with additional opportunities to practice. Taking this into account, using the CAS as a supplementary tool for audiology trainees has a number of practical advantages. Students were able to practice at their own pace during their free time, as the computer room where the CAS was installed could be accessed 24/7.

Then, within the few hours of training to which students' were exposed, they had access to 25 cases covering most of the pathologies they encounter during their professional career. This would not have happened during any of their clinical placements as patients are limited and pathologies can be redundant with students conducting similar diagnoses throughout the day.

Finally, the number of patients and opportunities for actual hands-on experience can vary greatly from one clinic to another. This results in students having very different learning experiences. The CAS, on the other hand, standardizes the training, offering the same opportunities to each student.

# 4 Clinical history taking and pure tone audiometry study

The content of this chapter has been submitted to the International Journal of Audiology.

## 4.1 Introduction

Learning from the pilot experiments' limitations, both the design of the system and set-up of the experiments have been adjusted. The main changes are as follow.

As presented in chapter 2, first, the design and code of the CAS have been refined. This allows in turn for a more stable and realistic system. Moreover, new more realistic virtual characters models have been integrated in order to embody a wider range of possible patients.

Special attention as been given to improving VPs' interaction scripts. The typical method is to ask participants to interact with a patient, record those interactions, and then add the ones that did not result in a VP response, in the form of a new trigger. A trigger can be either an entire sentence or a set of keywords (Halan et al., 2010; Deladisma et al., 2008). As Rossen et al. (2010) described, creating a free form interaction script for a virtual patient can take up to hundreds of hours spread over months. However in audiology, the clinical history taking procedure only revolves around 30 questions, and students are instructed to ask them one at a time, and to keep it as concise as possible. With that in mind, we believe we made sufficient progress over three iterations, each using three participants to add to the lexicon of the VPs. Interaction scripts were improved using this method, with new sets of key-words and key-phrases to make the VPs more responsive, to facilitate students' interaction with them, and to limit frustration.

In addition, a feedback component has been implemented. It displays the critical information for a VP's diagnosis on a history form upon the completion of an interview with a VP. We implemented this feature to help students with their reporting skills.

Concerning the design of the experiment, the clinical history taking and pure tone audiometry training have been merged. This improves the design of the pure tone audiometry pilot by removing the week of common training students had before each assessment, at the cost of giving up on students midterm and end of term grades as a measure. The training period was also extended by one week, to allow students more time to fit in their training. In addition, they were asked and agreed to practice for a minimum of two hours using the CAS.

Moreover, the number of VPs assessed and time spent using the system were recorded for each participants this time. This allowed to check if they complied with the requirements of training at least 2 hours, and to learn if they are eager to practice more than the requirements.

Finally, we also recorded students training opportunities during their clinical placements to investigate just how many occasions for actual hands on experience they had as part of their regular training.

## 4.2 Evaluation

This experiment focused on assessing students on the pure tone audiometry and clinical history taking procedures.

This experiment aimed to answer the following hypotheses:

1. Students' ability to conduct a pure tone audiometry procedure will increase as a result of using the CAS, compared to students who were only given traditional instruction.

2. Students' ability to retrieve information in clinical history taking will improve following exposure to the CAS, compared to students who were only given traditional instruction.

3. Students' ability to accurately report information in clinical history taking will improve following exposure to the CAS, compared to students who were only given traditional instruction.

4. Following exposure to the CAS, students efficiency will improve in clinical history taking, compared to students who were only given traditional instruction.

### Participants

Participants were recruited among the University of Canterbury's first year Master of Audiology students at the start of the academic year, same as in the pilot experiments from Chapter 3. Second year audiology students have not been recruited as they took part in the pilot experiments. In addition, first year students represent the population that has the most to gain from the CAS.

This sample consisted of nine participants, two males and seven females, constituting the entire class of first year audiology master students. They were aged between 20 and 47, with all listing at least a bachelor degree as their most recent academic achievement. One held a Master of Arts, and one a PhD in a non related field. All participants considered having a sufficient level of computer skills to use the CAS.

For the purpose of this study the group was initially split based on their grade point average and gender before the pre-test. Following this pre-test, the groups were adjusted to also be balanced based on the results of this test. There were five participants in the CAS group and four participants in the control group.

### Design

The design of this controlled experiment as seen in Figure 4.1 replicates the clinical history pilot's design at the exception of the training period which has been increased to three weeks. Access to the CAS was available at any time during the three weeks period and monitored to record students' practising habits. Participants in the CAS group were asked to practice at least two hours during these three weeks.

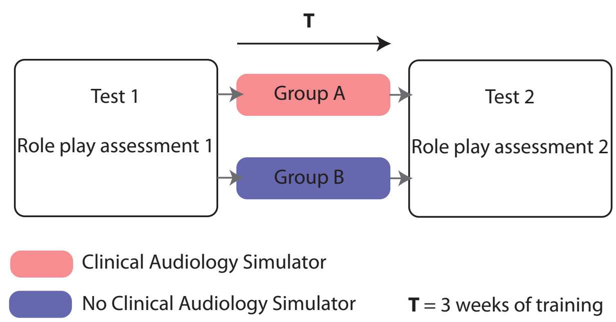



Figure 4.1: Design of the pure tone audiometry and clinical history taking experiment

### Measures

Pure tone audiometry was assessed solely during role play situation. This measure both assess students procedural skills when conducting the pure tone audiometry test and their ability to report the information obtained adequately on clinical forms.

Clinical history taking was assessed with the following measures, which are similar to the pilot experiment apart from the confidence score which has been removed.

***Verbal accuracy***: The verbal accuracy score is based on the number of answers participants were able to retrieve from the SP during the clinical history procedure, as a percentage.

***Written accuracy***: The written accuracy score refers to the number of adequately reported pieces of information on a standard history diagnosis form. This measure is marked on a maximum of 36 points.

***Number of questions***: for each role play situation, this measure is the number of questions a student asked to the SP before considering the clinical history complete.

***Efficiency***: the efficiency score is calculated by using the verbal accuracy raw scores divided by the number of questions asked to an SP.

The following measures have been added to this experiment to investigate the amount of training both using the CAS (one group), and in clinical placement (whole sample).

Students' practice with the CAS was automatically monitored by the system, recording practice time and number of VP assessed.

Students' practice in clinical placement was monitored through their clinical records book. The book contains information about clients seen, procedures performed, and direct contact & observation hours. For the purpose of this study, direct contact hours and number of clients seen were our focus.

### Results

We used a multivariate analysis of variance (MANOVA) to analyze the results of this experiment. The method of teaching was used as a between factor and the Time 2 pure tone, verbal accuracy, written accuracy, number of questions asked, and efficiency were used as dependent variables. Time 1 pure tone, verbal accuracy, written accuracy, number of questions asked, and efficiency measures were used as covariates. The means and standard deviation for these variables are presented in Table 1.

The MANOVA shows that there were significant differences at Time 2 between the two groups for the verbal accuracy [F(1,6)=9.483, p=.022, $\eta^2$=.612], and written accuracy [F(1,6)=26.163, p=.002, $\eta^2$ =.813] measures. These results indicate strong effects for the use of the CAS on students' ability to gather information from an SP in a role play situation, as well as their ability to report this information adequately (Cohen, 1988; Urdan, 2005; Watson, 2001).

The pure tone, efficiency, and number of questions measures, however, were not significantly different at Time 2 between the two groups (see Table 4.1).

The analysis of the CAS's records revealed that students in the CAS group spent on average 175 minutes practising with the simulator over the course of the study, and assessed on average 7.08 virtual clients (see Table 4.2). The results of a one tailed t-test showed that the average time spent on the CAS by students is not significantly higher than the training time prescribed for the experiment; t(5)=1.309, p=.247.

Finally we investigated the participants clinical records book to analyse clinical contact time with real clients through their weekly placement. It is interesting to note that students from both groups seem to have had similar direct contact hours with patients during the course of this study (approximately one hour per week). While this is not a variable that could have been controlled, this information reinforces that the results obtained are solely due to training with the CAS.

Then, over the course of their first semester, the participants spent a

Table 4.1: Descriptive statistics for the pure tone, verbal accuracy, written accuracy, number of questions asked to SP, and efficiency measures at time 1, time 2, and the difference between time 2 and time 1*(n=9)*

| | | Time 1 | | Time 2 | | Time 2 - Time 1 | |
|---|---|---|---|---|---|---|---|
| | Group of the student | Mean | Std. dev. | Mean | Std. dev. | Mean | Std. dev. |
| Pure tone | CAS | 75.00 | 14.29 | 87.67 | 7.23 | 12.67 | 18.13 |
| | No CAS | 90.00 | 13.07 | 88.52 | 7.84 | 6.85 | 15.76 |
| Verbal accuracy | CAS | 42.78 | 14.38 | 66.11 | 4.97 | 23.33 | 15.04 |
| | No CAS | 45.83 | 6.99 | 52.78 | 8.18 | 6.95 | 1.60 |
| Written accuracy | CAS | 13.40 | 3.97 | 24.60 | 3.21 | 11.20 | 2.95 |
| | No CAS | 13.50 | 2.08 | 15.00 | 2.83 | 1.50 | 3.32 |
| Number of questions | CAS | 20.20 | 5.36 | 32.80 | 5.76 | 12.60 | 3.51 |
| | No CAS | 25.25 | 5.50 | 27.75 | 4.99 | 2.50 | 2.38 |
| Efficiency | CAS | 0.77 | 0.18 | 0.74 | 0.10 | -0.03 | 0.14 |
| | No CAS | 0.67 | 0.09 | 0.69 | 0.10 | 0.03 | 0.07 |

weekly average of 47.53 minutes in direct contact with clients, seeing an average of 1.41 clients (see Table 4.3). We also observed that on average each student had at least one week without any direct contact with clients. Students have recorded spending a whole day observing on some occasions, with one student reaching four weeks over the course of the semester without any direct contact.

Table 4.2: Descriptive statistics for practice time using the CAS and number of virtual clients assessed between time 1 and time 2, and per week of training*(n=5)*

| | Practice time using the CAS between time 1 and time 2 (in minutes) | | | | Number of virtual clients assessed between time 1 and time 2 | | | |
|---|---|---|---|---|---|---|---|---|
| Group | Mean | Std. dev. | Min. | Max. | Mean | Std. dev. | Min. | Max. |
| CAS | 175.00 | 102.91 | 30 | 330 | 7.08 | 3.93 | 2 | 14 |
| | Average practice time using the CAS per week (in minutes) | | | | Average number of virtual clients assessed per week | | | |
| CAS | 58.33 | 34.30 | 10 | 110 | 2.37 | 1.31 | 0.67 | 4.67 |

Table 4.3: Descriptive statistics for time of direct contact with patients in clinic and number of patients assessed over the course of the participants first semester in the master of audiology program *(n=9)*

| | Average time of direct contact with clients per week (in minutes) | | | | Average number clients assessed per week | | | |
|---|---|---|---|---|---|---|---|---|
| | Mean | Std. dev. | Min. | Max. | Mean | Std. dev. | Min. | Max. |
| Total | 47.53 | 24.13 | 20 | 110.63 | 1.41 | 0.27 | 1 | 1.88 |

### Discussion

This experiment's main aim was to explore how the use of a virtual patient based simulator as a supplementary tool could impact on first year audiology students' pure tone audiometry and clinical history skills.

This experiment's first hypothesis was that students' ability to conduct a pure tone audiometry procedure will increase as a result of using the CAS, compared to students who were only given traditional instruction. The results do not support this hypothesis.

The results, however, supported two of our hypotheses for clinical history taking as students in the CAS group significantly increased their ability to retrieve information from clients and report it adequately following training with the CAS compared to the control group. We believe that investing more time on script implementation made the difference. As Halan et al. (2010); Deladisma et al. (2008); Rossen et al. (2010), and Rossen et al. (2010) described, it is a time consuming and resources intense process, but it cannot at this stage be avoided. Further research in the field may still lead to more efficient ways to implement VPs' script interactions, and as a result contribute to increase the use of VP systems in clinical curricula.

Looking at the efficiency and number of questions measured aids understanding the results better. Those two measures were not significantly different when comparing both groups following exposure. However we can see that while the efficiency measure remained steady for the whole sample, students in the simulator group had an increase in the number of questions asked from pre to post-test. We conclude from this that students having been exposed to the simulator became more aware of the full range of information needed to perform the clinical history procedure successfully, asked more questions to retrieve this information, and consequently had also a higher rate in written accuracy, reporting clinical information better than the control group. This may have happened as a result of presenting students with a completed diagnosis sheet once they are finished with a VP. While students in the CAS group increased their scores between pre and post test, the difference between both groups in the post test are not significant. The lack of statistical power in this study could possibly explain this particular result as well.

Our results for pure tone audiometry and clinical history taking support the findings from Vash et al. (2007), which suggest that using VPs as an additional tool could increase clinical students' history taking skills, but did not affect other lab procedures significantly.

A possible limitation of the pure tone experiment could also be the mode of assessment. While we believe that using role-play to assess transfer of skill is the best method of assessment for our objectives, this assessment focused only on one case with one SP, due to limited time and availability of qualified personnel to be used as SPs. However, the CAS increases students experience and opportunities to practice through a varied range of cases. This could explain the lack of positive results for this experiment. In future experiments, role-play would ideally be used to assess every participant using a range of SPs presenting different pathologies as well, not using a single case.

The second objective of this experiment was to look at students' opportunities for practice, both with the simulator and in clinical placement. The data suggests that students only had very few opportunities for actual hands on practice with clients during their weekly clinical placements. There was also a large disparity in these opportunities between students, with some students spending full days only observing during their clinical placements.

The simulator on the other hand was accessible at any time and enabled the students to practice as much as they wanted to. Table 4.2 presents the average number of VPs assessed per week practicing on the CAS ($m$=2.36, $sd$=1.31) as higher than the number of actual clients assessed in clinical settings ($m$=1.41, *sd*=0.27). As Tan et al. (2010) point out that VP based system have the potential of addressing several challenges in clinical education, they deliver a safe environment for patient encounters, and standardize the training. VP systems also take advantage of low cost hardware that is easily available in educational computer labs (Kyle Johnsen and Andrew Raij, 2005b). Our virtual clients were always available when students wished to practice, while clinical placements could not offer similar opportunities to every student.

## 4.3 Conclusion

This experiment suggests that our virtual patient based system has the potential to be a valuable tool to supplement audiology trainees' clinical education. It allowed the students involved in this experiment to foster their clinical history taking skills within a safe environment, with a measurable improvement when tested in real settings with SPs. Students were both able to retrieve more of the information critical to a diagnosis, and to record them according to clinical standards.

Furthermore, it exposes students to a wide range of cases they can choose to practice at their own pace and on their own agenda. This gives each student the same opportunity for standardized training. This is not the case with their clinical placements as the descriptive analysis of their clinical record books shows.

# 5 Speech audiometry for intermediate and novice users study

## 5.1 Introduction

Previous studies have shown that our system, the Clinical Audiology Simulator (CAS) can supplement audiology students' training effectively. Students increased their performance in clinical history taking following exposure to the system. Moreover, while our results did not support our hypotheses on pure tone audiometry performance, they suggested that the CAS contributed to students experience by allowing them a wider range of patients and pathology assessment than their clinical placements do.

As mentioned in chapter 3, computer simulations aimed to supplement audiology training are available on the market and used by some universities (e.g. Otis Audiology simulator [1], Parrot Software's Audiology Clinic [2]). However, to our knowledge, there are no computer simulations allowing the training of the speech audiometry procedure available. Moreover, our system

[1] INNOFORCE creative solutions website http://www.innoforce.com/

[2] Parrot Software website http://www.parrotsoftware.com/parrotstore/Products/120-audiology-clinic-full-version.aspx

embeds the practice of speech audiometry within complete simulated cases using VPs and allows students to perform most clinical procedures.

This research explores the use of a new component integrated in the CAS, it aims to practice speech audiometry. This procedure is typically used late in the diagnostic process as it requires preliminary information gathered through clinical history taking and pure tone audiometry. For the purpose of this experiment, however, participants could only interact with the speech audiometry feature of the CAS when assessing VPs. Consequently, the system was set up so that the other components would present the results of their associated procedures (e.g. the pure tone section displayed the completed audiograms for the VP being assessed), without requiring the students to perform them. The original 25 cases used in the previous studies have been completed to integrate the results of this exam in each of them. Six of those cases were used as the basis of our assessments, leaving students with 19 cases to practice on the CAS.

This component of the CAS was tested on two sample of students, Master of Audiology (MAud) whom represent intermediate level users, and Speech and Language Therapy (BSLP) students as novice users. As Cook and Triola (2009) suggest, novice clinicians' exposure to a wide range of clinical cases through computer simulations has the potential to develop data gathering skills and clinical decision making. We included a wider range of participants, to determine whether students with lower skill level could benefit more from the CAS, as MAud students should be already somewhat familiar with the procedure.

We also intended to assess whether students would experience learning gains following exposure to the CAS, and to assess their retention following a period of non exposure. Knowledge retention can increase due to exposure to more realistic simulations (Sitzmann, 2011). Paper assessments were implemented to assess both sample of students' knowledge. In addition, role-play situations were also used with the MAud students to assess transfer of skill, as they have the expertise to operate the necessary instruments involved with the speech audiometry procedure, while BSLP students do not.

We suggested in the previous chapter that assessing students using a single clinical case might not be representative of their learning gains. Consequently, the paper assessments used in this experiment were more elaborate, assessing students first using two different clinical cases, and then enquiring about general knowledge of the speech audiometry procedure through open ended questions.

## 5.2 Evaluation

This study is made of two parts. First, the speech audiometry component of the CAS was assessed with students from the University of Canterbury's Master of audiology program, then it was also assessed using students from the Bachelor of Speech and Language Therapy. The difference between these two samples of students is that MAud students are intermediate users as they have and still are practicing this procedure on real clients through clinical placements; on the other hand, Speech and Language therapy students are novice users, while being familiar with clinical work, they have never practiced this procedure before the study.

This experiment's aim is to answer the following hypotheses, for the two sample of students:

1. Novice students will benefit more from being exposed to speech audiometry than intermediate users.

2. Novice students will retain knowledge better than intermediate students when assessed for retention four weeks after the post-test.

3. CAS students will perform better following exposure to the CAS than students in the control group.

4. CAS students will retain knowledge better than students in the control group following four weeks of not practicing using the CAS.

### Participants

#### MAuD students

The entire classes, first and second years, of MAud students were recruited for this study. These students had some familiarity with the procedure, and skills using the CAS as they were recruited before in studies to assess other components of the system. The sample was composed of 18 students, four male and 14 female. Eight students were in there first year while ten students were in their second year of the master of audiology course. Students were aged from the 20-29 years old range, to 50+ years old.

Students were split into two groups, based on their gender, whether they were undertaking their first or second year of the master degree, and their results on the pre-test. Each group was composed of nine students, two males and seven female; four first year and five second year.

#### BSLP students

BSLP students from every level have been recruited as novice users for this study. In addition, students with an equivalent level have also being recruited. None of these students were familiar with the speech audiometry procedure. The sample was composed of 18 students, one male and 17 female. Five students were in their first professional year, seven were in their second professional year, one was in her third professional year, and five were at an equivalent level having completed similar papers. Students' ages ranged from the <20 years old category, to the 30-39 years old range.

Students were split into two groups, according to their gender, their level, and results on the pre-test. Each group was composed of nine students. The CAS group had one males, and eight females; two first years, four second years, no third years, two from another major, and one prospective MAud. The Control group had nine female members; three first year, three second year, one third year, two prospective MAud, and one prospective BSLP.

### Design

#### MAud students

Every student first undertook, a role-play assessment followed by a paper assessment in order to split the groups adequately. After being split into groups, students in the CAS group were asked to complete at least 6 cases over the following two weeks. Students in the control group had no requirement during those two weeks. Both groups had access to their course material and were undertaking their weekly clinic day during the study. They were then assessed a second time using a role-play situation followed by a paper assessment.

Following this second assessment both groups did not have access to the CAS for a period of four weeks before being assessed again to test for retention. The final assessment was again made of a role-play situation followed by a paper test.

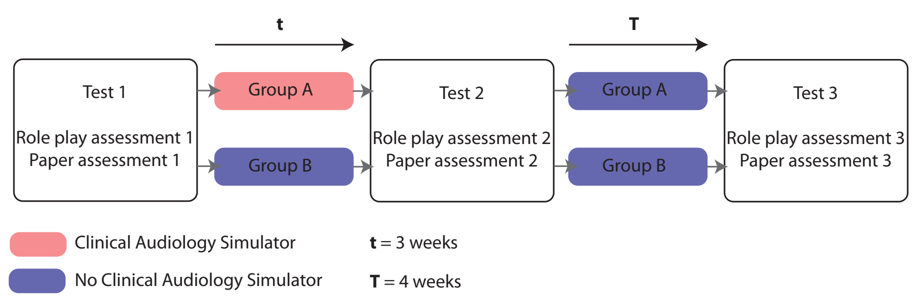



Figure 5.1: Design of the MAud speech experiment

### BSLP students

Language therapy students followed a similar approach to the master students, with two exceptions. First, the students were given a 45 minutes lecture prior to the initial assessment in order to introduce them to speech audiometry. Second, these students were assessed solely using the paper assessment. This sample of students did not have the expertise to use the necessary tools to conduct a speech audiometry exam with SPs, and teaching them how to operate them would have required too much time.

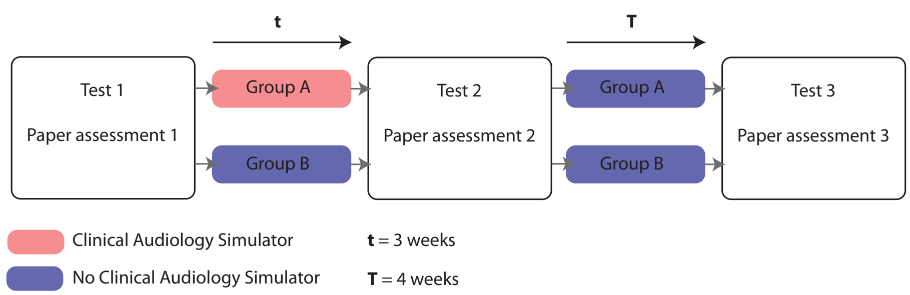



Figure 5.2: Design of the BSLP speech experiment

## Measures

### MAud students

Speech audiometry was assessed using role-play and a paper assessment for each of the three tests. Each test presented a different patient and symptoms.

The role-play assessments used an SP to undertake the role of the patient, while students performed the speech audiometry procedure using the standard clinical tools. Students were marked in accordance with clinical protocol for the procedure and reporting of the information on clinical forms (Appendix G). This measure was recorded as percentages.

The paper assessments were made of three sub-tests (Appendix H). The first presented information on a clinical case and was followed by 16 multi-choice questions, encompassing how to perform the procedure for this particular case, as well as actual results that could be deducted from the information provided.

The second sub-test has an identical structure to the first one, except that the hearing loss presented makes it more difficult for students to perform the procedure adequately and draw conclusions.

The third sub-test took the form of 13 open ended questions, referring to general speech audiometry knowledge.

The paper assessments were structured similarly at each time of the study, while being based on different patients. Each multi-choice subtest was marked on 16 points. The short answer sub-tests were marked on 32 points. Two research assistants were in charge of marking the students to ascertain reliability, both during role play, and for the paper assessments. No disagreement were found between the two markers.

**BSLP students**

BSLP students, as mentioned in the design section, do not have the expertise to perform the speech audiometry procedure in real settings. Consequently, this sample was assessed only using the paper assessments, which were identical to the ones intermediate students received.

Table 5.1: Descriptive statistics for multi-choice 1, multi-choice 2, and short answers at time 1, time 2, and time 3 *(n=36)*

| | | Time 1 | | Time 2 | | Time 3 | |
|---|---|---|---|---|---|---|---|
| | Group of the student | Mean | Std. dev. | Mean | Std. dev. | Mean | Std. dev. |
| Multi-choice 1 | MAud CAS | 13.06 | 1.81 | 13.94 | 0.92 | 13.44 | 1.94 |
| | BSLP CAS | 14.44 | 1.31 | 9.83 | 1.54 | 11.56 | 2.70 |
| | Total CAS | 13.75 | 1.69 | 11.89 | 2.45 | 12.50 | 2.48 |
| | MAud Control | 13.89 | 1.98 | 12.33 | 2.25 | 13.61 | 2.48 |
| | BSLP Control | 13.88 | 1.30 | 10.44 | 2.72 | 11.88 | 2.23 |
| | Total Control | 13.88 | 1.64 | 11.44 | 2.59 | 12.79 | 2.28 |
| Multi-choice 2 | MAud CAS | 13.78 | 1.64 | 12.44 | 1.94 | 14.11 | 1.17 |
| | BSLP CAS | 13.11 | 1.76 | 14.67 | 1.87 | 12.67 | 1.00 |
| | Total CAS | 13.44 | 4.69 | 12.56 | 1.85 | 13.39 | 1.29 |
| | MAud Control | 13.78 | 1.56 | 12.89 | 1.90 | 13.56 | 2.19 |
| | BSLP Control | 12.88 | 1.55 | 11.75 | 1.75 | 12.88 | 1.81 |
| | Total Control | 13.35 | 1.58 | 12.35 | 1.87 | 13.24 | 1.99 |
| Short answers | MAud CAS | 29.11 | 5.40 | 31.78 | 2.63 | 29.89 | 3.06 |
| | BSLP CAS | 25.56 | 3.50 | 28.89 | 3.72 | 27.56 | 4.25 |
| | Total CAS | 27.33 | 4.78 | 30.33 | 3.46 | 28.72 | 3.79 |
| | MAud Control | 25.89 | 5.35 | 30.00 | 1.87 | 30.56 | 4.28 |
| | BSLP Control | 24.00 | 7.29 | 27.50 | 3.74 | 28.25 | 5.75 |
| | Total Control | 25.00 | 6.20 | 28.82 | 3.09 | 29.47 | 5.00 |

### Results

#### Paper assessment

We used a three-way repeated measure ANOVA with training (*CAS*, or *Control*) and the type of student (*MAud*, or *BSLP*) as between factors, and assessment times as within. Time 1 is the pre-test, Time 2 is the post-test three weeks after the pre-test, and Time 3 is the retention test four weeks after the post-test. The paper assessments were divided into three sub-tests, *multi-choice 1*, *multi-choice 2*, and *short answers*. The descriptive statistics are presented in Table 5.1.

There was a statistically significant interaction between time of the paper assessments and the type of students [F(6, 120)=2.69, p=.018, $\eta^2$=.118]. This is a small effect.

The interaction between time and the type of students on the multi-choice 1 assessments is also significant [F(2, 62)=8.00, p=.001, $\eta^2$=.205]. This is a small effect as well. As seen in Figure 5.3, MAud students' scores remain steady between Time 1, Time 2, and Time 3 for the first multi-choice assessment. On the other hand BSLP students' scores decrease between Time 1 and Time 2, before increasing slightly for the retention test at Time 3.

There were no other significant effects of the type of students. Both groups' scores decreased between Time 1 and Time 2 for the second multi-choice assessment, before increasing again at Time 3 for the retention test. MAud scores are overall higher for each assessment, despite not being significant (see Figure 5.4).

For the short-answers assessments, both groups' scores increased between

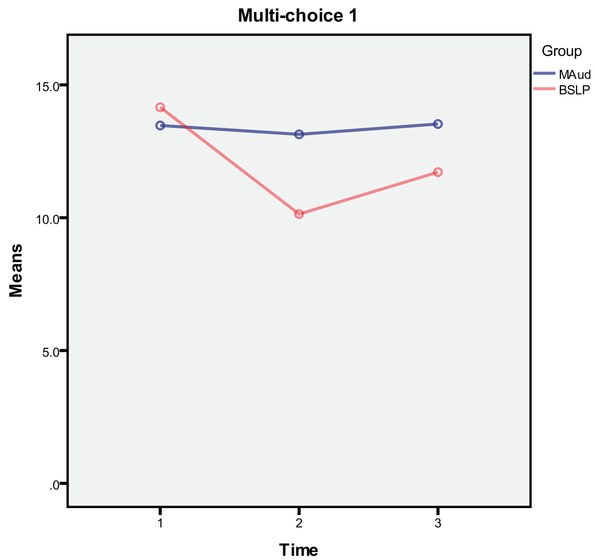



Figure 5.3: Overall results on multi-choice 1 over time for MAud, and BSLP students

Time 1 and Time 2, before remaining stable for the the retention test at Time 3. Again, MAud students scored better overall (see Figure 5.5).

There was however, no statistically significant interaction effect over time for the type of training [F(6, 120)=2.69, p=.811, $\eta^2$=.039]. Overall, both for the multi-choice 1 and multi-choice 2 assessments, CAS and Control students' scores decreased between Time 1 and Time 2. Those scores however increased again for the retention assessment at Time 3, remaining lower than the pre-test' at Time 1 for the multi-choice 1 and reaching their initial levels for the multi-choice 2 assessments (see Figure 5.6 & 5.7).

For the short-answers assessment, both the CAS and Control students'

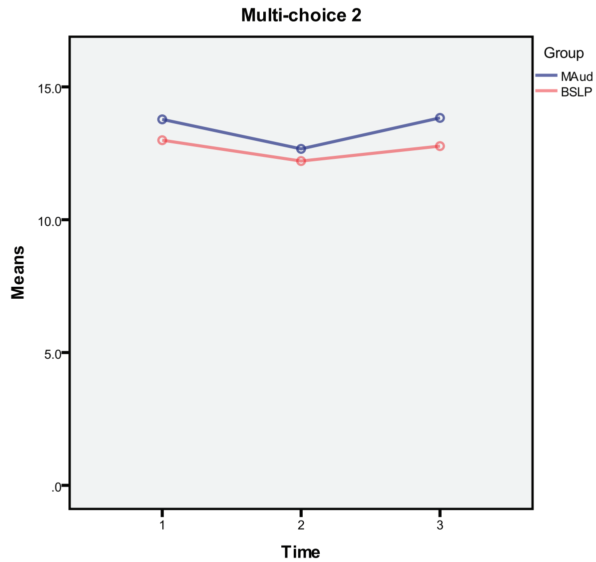



Figure 5.4: Overall results on multi-choice 2 over time for MAud, and BSLP students

increased their scores between Time 1 and Time 2. CAS students' scores then decreased in the retention test at Time 3, while Control students had a very slight increase (see Figure 5.8).

#### Role-play assessment

A three-way repeated measures ANOVA was conducted to compare scores in role-play assessments and paper assessments for MAud students between the MAuD CAS group, and the MAud Control group. No statistical significant results have been found. Descriptive statistics for the role-play measure are presented in Table5.2.

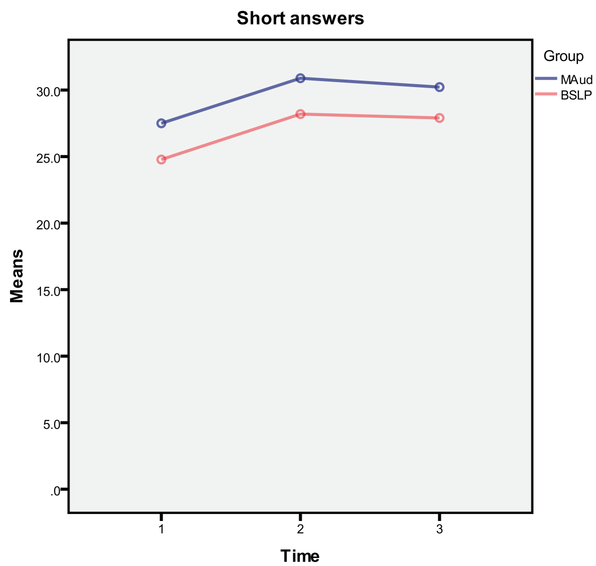



Figure 5.5: Overall results on short answers over time for MAud, and BSLP students

Table 5.2: Descriptive statistics for role-play, at time 1, time 2, and time 3 *(n=18)*

| | | Time 1 | | Time 2 | | Time 3 | |
|---|---|---|---|---|---|---|---|
| | Group of the student | Mean | Std. dev. | Mean | Std. dev. | Mean | Std. dev. |
| Role-play | MAud CAS | 83.22 | 15.37 | 91.00 | 9.15 | 89.78 | 12.32 |
| | MAud Control | 81.67 | 16.64 | 87.00 | 8.15 | 90.11 | 8.66 |
| | Total | 82.44 | 15.56 | 89.00 | 8.66 | 89.94 | 11.02 |

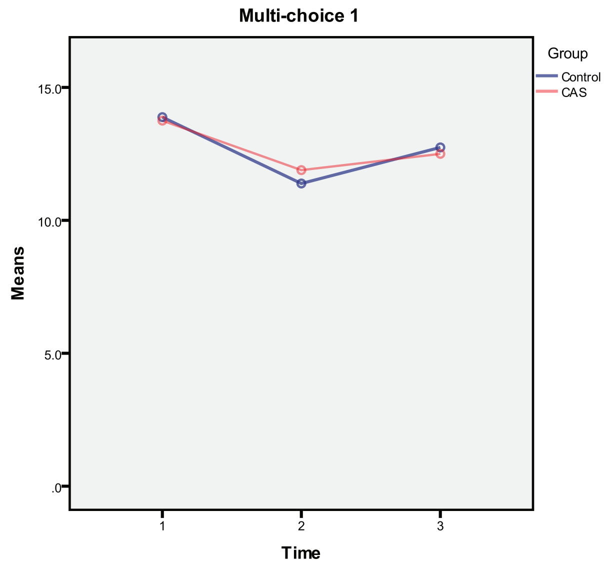



Figure 5.6: MAud and BSLP results on multi-choice 1 over time for CAS, and Control students

A descriptive analysis shows that MAud students overall had an increase in role-play score between Time 1 and Time 2, however students in the CAS group had a higher increase. The control group's scores increased again at Time 3 for retention, reaching CAS students' scores which remained steady since Time 2 (see Figure 5.9).

While not significant, the first multi-choice assessments results show that the MAud students in the CAS group's scores increase between Time 1 and Time 2, while control students had a high decrease in scores over the same period. Control group scores, however, had increased at Time 3 while CAS

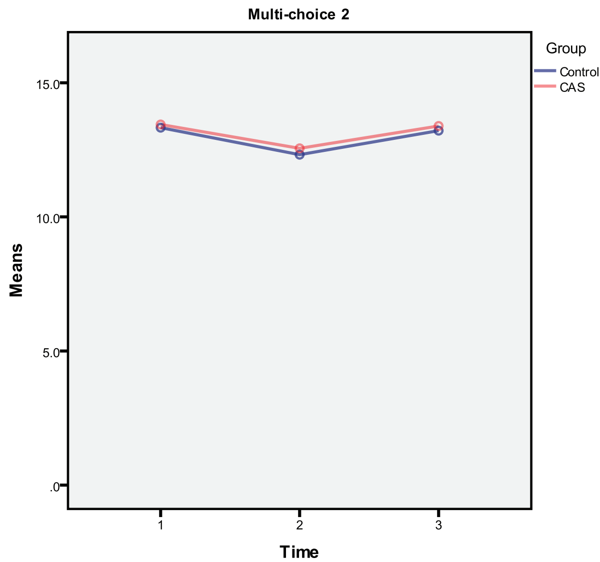



Figure 5.7: MAud and BSLP results on multi-choice 2 over time for CAS, and Control students

group scores decreased slightly (see Figure 5.10).

For the second multi-choice assessment, both CAS and Control students' scores decreased between Time 1 and Time 2, before increasing again when tested for retention at Time 3. CAS students, however, decreased somewhat more at Time 2, and increased more at Time 3 than students in the Control group (see Figure 5.11).

Both CAS and Control students increased their score in the short-answers assessments between Time 1 and Time 2. Control group scores then remain stable for the retention assessment at Time 3, while CAS students' scores decreased to almost their initial level at Time 1 (see Figure 5.12).

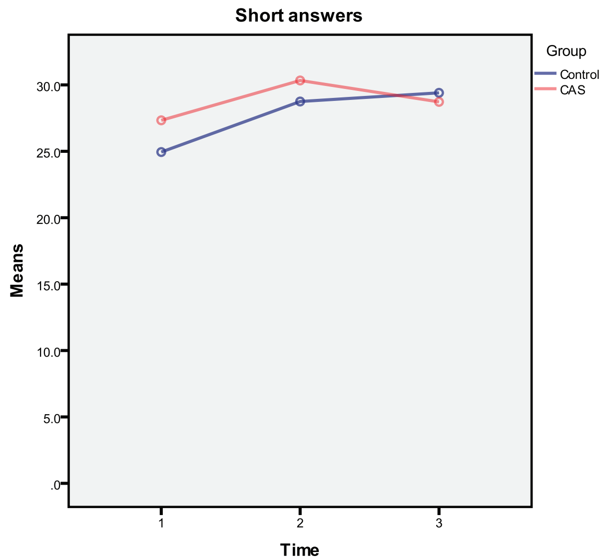



Figure 5.8: MAud and BSLP results on short answers over time for CAS, and Control students

## Discussion

### Overall

The first two hypotheses were that novice students would have a higher increase in knowledge following exposure to speech audiometry, as well as a higher rate of retention than intermediate students. The experiments' results do not support these hypotheses. Novice students scored less than intermediate students on the multi-choice questionnaires. They followed the same increase as intermediate students on the general knowledge assessment both post-test and when assessed for retention, while remaining a few points

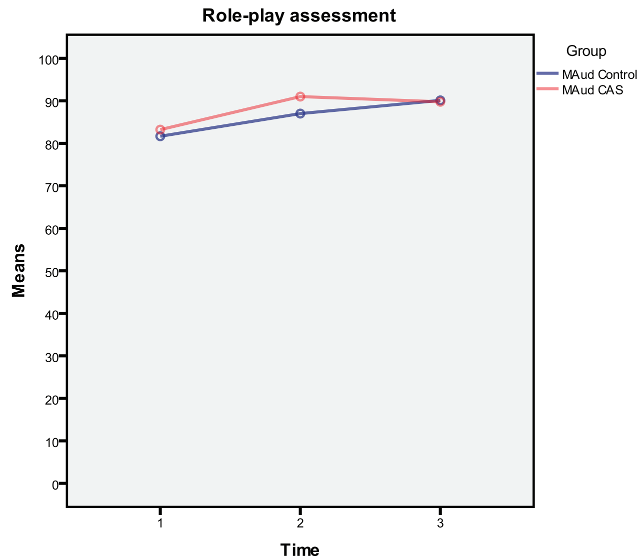



Figure 5.9: MAud results on role-play over time for CAS, and Control students

below them.

The second two hypotheses were that CAS students, independent of which sample they were from, would perform better when assessed following exposure to the CAS, and maintain higher retention rates than students who did not have access to the simulator. Our results do not support these hypotheses. All students, both in the CAS and Control conditions showed similar performance over the course of the experiment. This suggests that the learning effect came from undertaking the paper assessments themselves, or studying the procedure's materials, but not from the CAS.

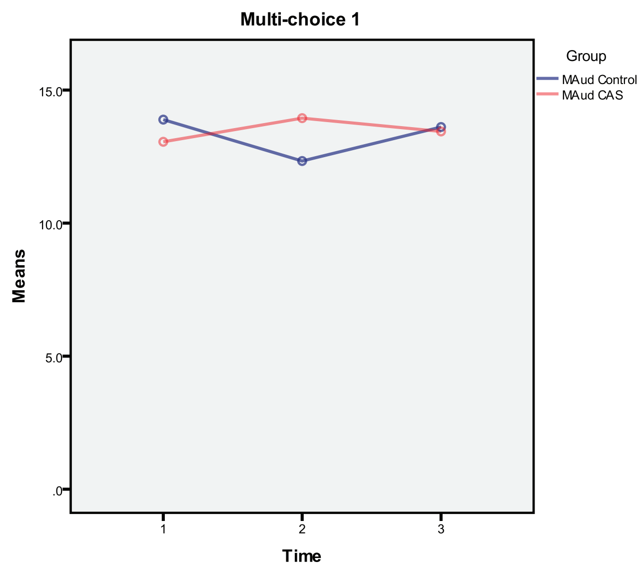



Figure 5.10: MAud results on multi-choice 1 over time for CAS, and Control students

**MAud students**

The first two hypotheses for intermediate students were that the ones practicing with the CAS would perform better than students in the control condition, when assessed through role-play following exposure to the CAS, and when assessed for retention four weeks after the post-test. Our analysis does not support these hypotheses. While results of the statistical analysis are not significant, there is a trend in the data showing that students in the CAS group perform better than students in the control group when assessed through role-play following exposure to the CAS. Their scores then remained steady for the retention test while students in the control condition's scores kept

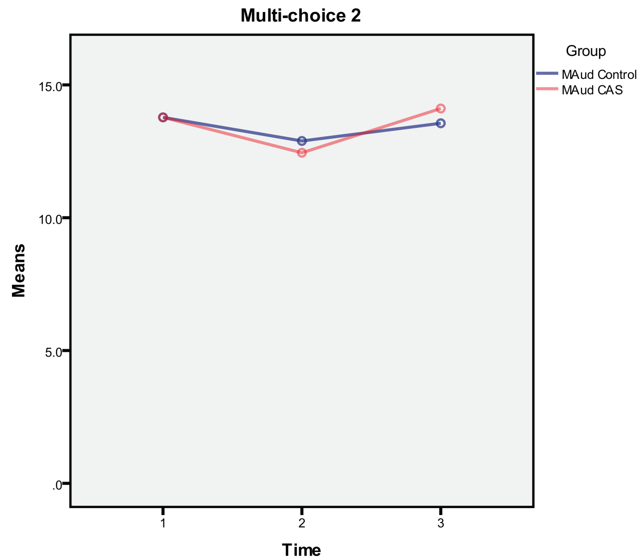



Figure 5.11: MAud results on multi-choice 2 over time for CAS, and Control students

increasing slightly. It could be that students in the control group rehearsed their course materials following the post-test, after noticing the quality of their performance, or lack thereof, at Time 2.

The following two hypotheses predicted that students in the CAS condition would score better than the control group in the paper assessments following expose to the CAS, and when assessed for retention. The results do not support these hypotheses. We note however that both group of students experienced an increase between pre and post-test when assessed on their general knowledge of speech audiometry through a short answers questionnaire, this could be a learning effect of the interventions. Those scores remained stable when

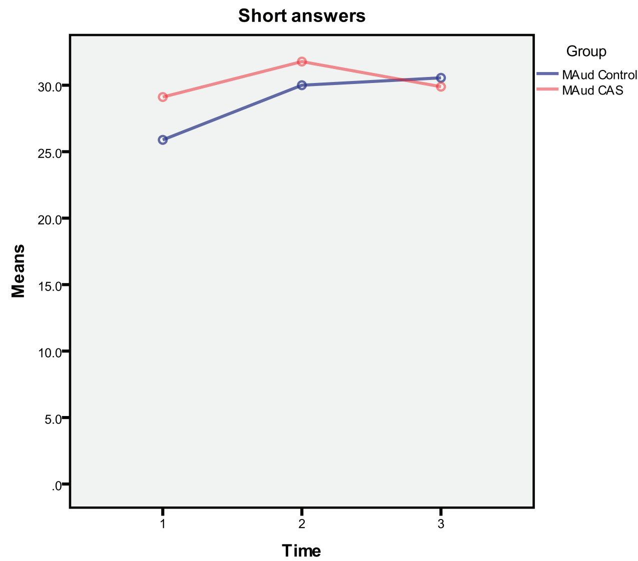



Figure 5.12: MAud results on short answers over time for CAS, and Control students

students were assessed for retention.

For the multi-choice 1 questionnaires, the easiest of the two, students in the CAS condition increased their scores post-test while the control group's scores decreased. However, on the second, more difficult multi-choice 2 questionnaires both groups decreased their scores between pre, and post-test.

This could suggest that the CAS is effective for MAuD students to practice and learn how to cope with more accessible cases. However, it could be that either the CAS does not offer the same advantages on more difficult cases, or that the post-test assessment for multi-choice 2 was overall too difficult for the students to comprehend.

**BSLP students**

For novice students, our two hypotheses were that students in the CAS group would perform better than students in the control group when assessed through paper tests following exposure to the CAS, and when assessed for retention. The results do not support these hypotheses. We note, however, that students in the CAS condition had a smaller decrease in scores than control students between pre, and post-test on the second multi-choice assessment.

It seems that a limitation is, like with the MAuD students, that the multi-choice assessments at Time 2 were too difficult. The CAS does not seem to enable this sample to even assess more simple cases like the ones in the multi-choice 1 assessments. The content of the assessments was left to the field experts. While I do not have the necessary audiology expertise to implement exams, I planned the structure and topics covered.

Again, both group increased their scores on general knowledge of speech audiometry when assessed with the short answer questionnaire post-test, before remaining stable in the retention test. This links back to an intervention learning gain.

## 5.3 Conclusion

This experiment found no clear statistical evidence that the CAS increased students learning of the speech audiometry procedure, with intermediate or novice students. Generally, students learned about speech audiometry over the course of the experiment, whether they practiced on the CAS or not. This could be explained from being exposed to the different assessments themselves, or from rehearsing traditional learning materials, leading to a learning gain for all students.

These results could also find an explanation in the low statistical numbers employed in this experiment. Also, our plan to assess students using two clinical cases and general knowledge questions in the paper assessment did not appear to be successful. Looking at the data, a possible explanation could be that overall, the second paper assessment, post-test, was too difficult in comparison to the pre-test and retention assessment, even though they were validated by field experts. We note, however, that there might be a slight trend in the data on the general knowledge part of the assessment, with CAS students scores increasing more than control students following exposure to the CAS.

More studies will be necessary to fully understand how to assess this students for transfer of skills. In the meantime, students who practiced on the CAS did have access to a wide range of patients. Moreover, they could practice their skills within a safe environment and at their own pace to build experience. As Cook and Triola (2009) mentioned, this exposure to diverse patients and pathologies within a safe environment could be the key in developing students decision making and data gathering skills.

# 6 Feedback in virtual patient based computer simulation study

## 6.1 Introduction

Feedback has been identified as one of the key elements that impacts positively on students' learning both in traditional teachings and in simulation, whether through debriefing or during the learning experience (Havnes et al., 2012; Ong, 2007). In order for feedback to be effective is has to be delivered correctly (Kneebone and Nestel, 2005; Shute, 2008). Feedback has also been mentioned as the element that differentiates training from practice (Astwood et al., 2008). While there is a wide range of types of feedback and different delivery methods, research suggests that formative feedback, and detailed feedback have the greatest potential for beginners in opposition to summative feedback (Billings, 2012; Schartel, 2012; Kulhavy and Stock, 1989; Bangert-Drowns et al., 1991). Formative feedback refers to feedback delivered during a learning experience, being a tutorial session, or a simulation for instance. From this delivery, formative feedback allows students corrections during the task. Formative feedback has also the advantage of allowing the user to deal with one unit of learning at a time (Billings, 2012). Summative feedback on the other hand, refers to feedback delivered at the end of a learning experience, this could take the form of a grade with corrections where a student made mistakes, or be more elaborated like a clinical debriefing.

The CAS has been previously tested as a supplementary training tool for first year audiology students (refer to chapter 3 & 4). The research presented in this chapter intends to evaluate the potential of our system as a teaching method for beginning students to learn procedural concepts. It also aims to indicate whether additional feedback, delivered as in the form of an embedded tutoring session, impacts on students knowledge retention.

For the purpose of this experiment, an additional feature presenting students with formative feedback messages while conducting a pure tone audiometry procedure has been implemented (see Chapter 2: Implementation, section 4. feedback for more information). This additional feedback module adds messages to the traditional system, providing students with validation by referring back to course material & rules. The feedback delivered also offers leads on how to pursue the VPs' assessment.

Following exposure to the system, either with additional formative feedback or without, students were assessed through a paper test. This assessment integrated the core concepts of pure tone audiometry, which is the procedure students' practiced during the training session on the simulator. Students were probed on each item with both multi-choice and open ended questions.

Students' workload was also recorded both following the simulator session and following the paper test. This is to monitor the effect of the additional feedback on students' workload, as feedback has also the potential to overwhelm students during a learning process if it is not regulated adequately

(Schartel, 2012). We then try and evaluate workload during the assessment to determine if the initial investment during training impacts positively on students, potentially facing less workload during this part of the experiment.

This experiment is intended to contribute on improving the potential of the CAS as a teaching tool, by reaching a larger audience, including novice users without course work or access to real clients, as well as validating a tutoring mode, using formative feedback as the students' training wheels.

## 6.2 Evaluation

This study aimed to answer which feedback delivery method is the most appropriate for novice students to learn clinical procedures and field knowledge. The study's hypotheses were:

1. Students in the group with additional formative feedback will score better in the post experiment paper test than students who only received summative feedback.

2. Students in the group with additional formative feedback will be subjected to less workload with the post experiment paper test than students who only received summative feedback.

3. Students in the group with additional formative feedback will be subjected to more workload while training compared to students who only received summative feedback.

### Participants

Our sample consisted of 51 students, 50 females and one male. The youngest was 18 while the oldest was 49, with an average age of 24 years old across the sample. Forty of those students were undertaking their Bachelor of Speech and Language Therapy while 11 either already graduated or were taking another course. Among the Speech and Language therapy students, 13 were first year, 10 were second year, 11 were third year, and six were in their fourth year. Out of our sample only two participants did not consider themselves as IT proficient on a yes or no question, they were still however able to use the CAS after a demonstration; the others rated their skills as proficient interacting with computers between three and 40 hours a week (m=20.30 hours).

This population was selected because, while familiar with clinical work, they have very limited knowledge of clinical audiology concepts and procedures. Second year Speech and Language Therapy students have to undertake one audiology paper, however this paper does not encompass the knowledge components integrated in our system. In addition, this Bachelor degree is the most standard path into a Master of Audiology and students are eager to participate in extracurricular activities that have to do with this field. All participants were novice to the clinical procedure practiced with the CAS, they never experimented it before.

This sample was split into two groups, balancing their background, academic year in their respective degrees, and ages.

### Study design

This controlled experiment investigated the impact of an additional feedback component implemented within our simulation system, the CAS. Before the start of the experiment, students were assigned to one of the two groups; CAS with extra formative feedback, or CAS without. Assignment to groups was made to balance them in by gender, and by year level. The first day of the experiment focused on training using the CAS; each student assessed the same four VPs. There was no time limit, students could practice at their own pace. Following those assessments students were asked to complete a questionnaire were they rated their perception of the workload associated with this task. On the second day, the students were assessed to measure their learning using a paper test. This assessment was followed by the workload questionnaire again, plus an additional questionnaire to collect students' demographics (see Figure 6.1).

### Measures

Students were assessed on the following measures:

1. Learning outcomes: the learning outcomes of the procedure practice in the simulation were assessed using a paper test. This test was implemented by audiology experts and in accordance with the University of Canterbury's clinical audiology coordinator to guarantee its validity. Reliability was ascertained by having two assessors marking each test. The test itself was made of two parts; multi-choice and short answer questions. The test focused on 13 knowledge items, each item was tested

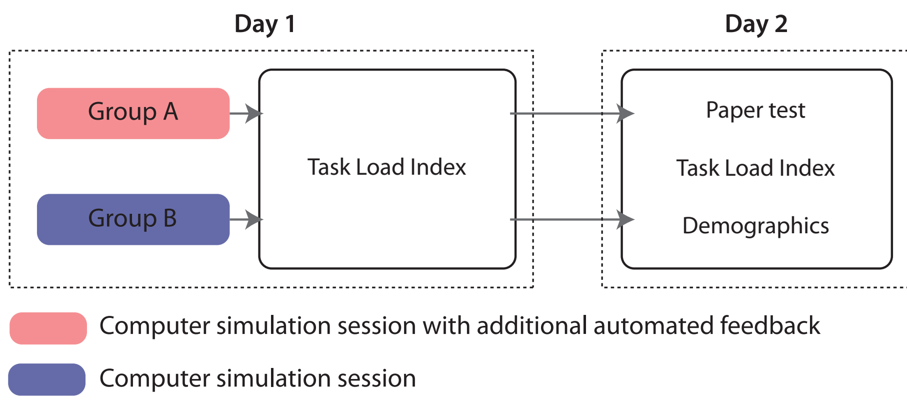



Figure 6.1: Feedback study design

twice, once per type of question (see Appendix I). These measures were reported as percentages.

2. Workload: the workload or effort required by students was assessed both for the training session using the simulation, and for the paper test. This measure was determined using an existing instrument, the Task Load Index (TLX)[1]. This instrument aims to determine workload on specific tasks by asking to rate the following six categories on Likert scales: mental demand, physical demand, temporal demand, performance, effort, and frustration. Each user then weights the different categories.

[1] http://humansystems.arc.nasa.gov/groups/TLX/

### Results

We used a multivariate analysis of variance (MANOVA) to analyze the learning results of this experiment. The feedback method was used as a between factor. The multi-choice, short answers, and total measures were used as dependent variables. The means and standard deviation for these variables are presented in Figure 6.2.

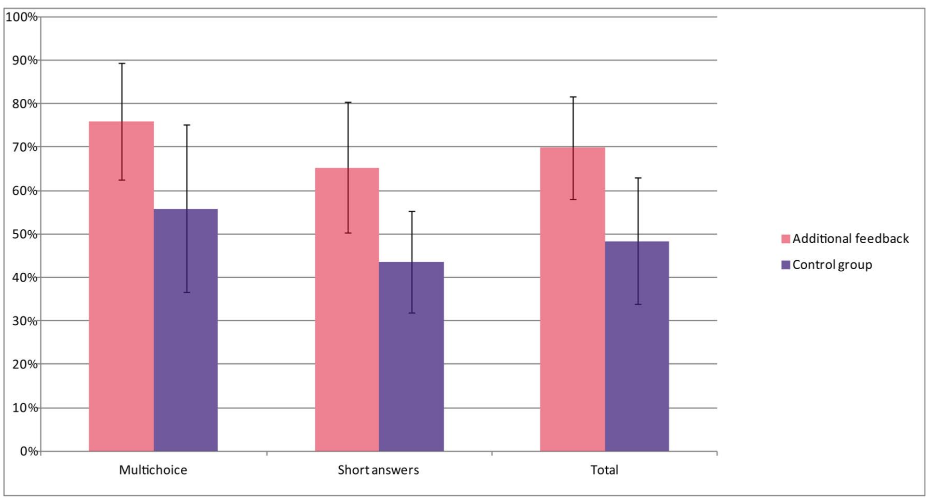



Figure 6.2: Paper test mean scores, by type of questions, and total

In the paper assessment, students with additional feedback scored significantly better than students without for the multi-choice questions [F(1, 49)=18.547, p<.001, $\eta^2$ =.275], the short answers [F(1, 49)=33.140, p<.001, $\eta^2$ =.403], and the total marks [F(1, 49)=33.529, p<.001, $\eta^2$ =.406]. When referring to effect sizes magnitude guidelines (Cohen, 1988; Urdan, 2005; Watson, 2001) these are medium effects.

When looking at students total marks on the paper test and applying the standard grading scale of the University of Canterbury[1] we see that overall students in the additional feedback condition scored better. Table 6.1 presents the results by grade for students in both conditions had they been assessed as part of a course. First, only five students overall would have scored an A grade, and they are all in the additional feedback group. The B grade would have been awarded to 13 students in the additional feedback condition, and to four in the Control group. Six students in the additional feedback group would have been awarded a C grade, 10 from the control group got this grade. Finally, 13 students would have failed overall with marks below 50%. Twelve of those students were in the control group and 1 was in the feedback condition.

Results of the workload measures analysis show a significant difference between the two groups for the simulation training [t(51)=-3.205, p=.002, $\eta^2$ =.173], where the additional feedback group scored higher than the control group. It also shows significant differences for workload during the paper test [t(51)=2.095, p=.041, $\eta^2$ =.082], this time with the control group scoring higher than the additional feedback group. These are small effects (Cohen, 1988; Urdan, 2005; Watson, 2001). Figure 6.3 presents the two TLX scores for both groups, for the training session, and for the assessment.

### Discussion

The experiment explored how adding formative feedback embedded in our simulation system would impact on students' learning and workload.

[1] http://www.canterbury.ac.nz/aqua/grading.shtml

Table 6.1: Grades awarded to students in accordance to the University of Canterbury's grading scale

| Grade | Additional feedback | Control group | Total *(n=50)* |
|---|---|---|---|
| A+ (90-100) | 0 | 0 | 0 |
| A (85-89.99) | 4 | 0 | 4 |
| A- (90-84.99) | 1 | 0 | 1 |
| Total A | 5 | 0 | 5 |
| B+ (75-79.99) | 5 | 0 | 5 |
| B (70-74.99) | 2 | 1 | 3 |
| B- (65-69.99) | 6 | 3 | 9 |
| Total B | 13 | 4 | 17 |
| C+ (60-64.99) | 4 | 5 | 9 |
| C (55-59.99) | 2 | 0 | 2 |
| C- (50-54.99) | 0 | 5 | 5 |
| Total C | 6 | 10 | 16 |
| D (40-49.99) fail | 0 | 5 | 5 |
| E (0-39.99) fail | 1 | 7 | 8 |
| Total fail | 1 | 12 | 13 |

We hypothesized that students who trained in the condition with added formative feedback would perform better when assessed following their training session. The results support this hypothesis with students in the formative feedback condition scoring significantly better when assessed for retention through a paper test. Moreover, as seen in Figure 6.2, students who trained with formative feedback while assessing VPs scored better in each section

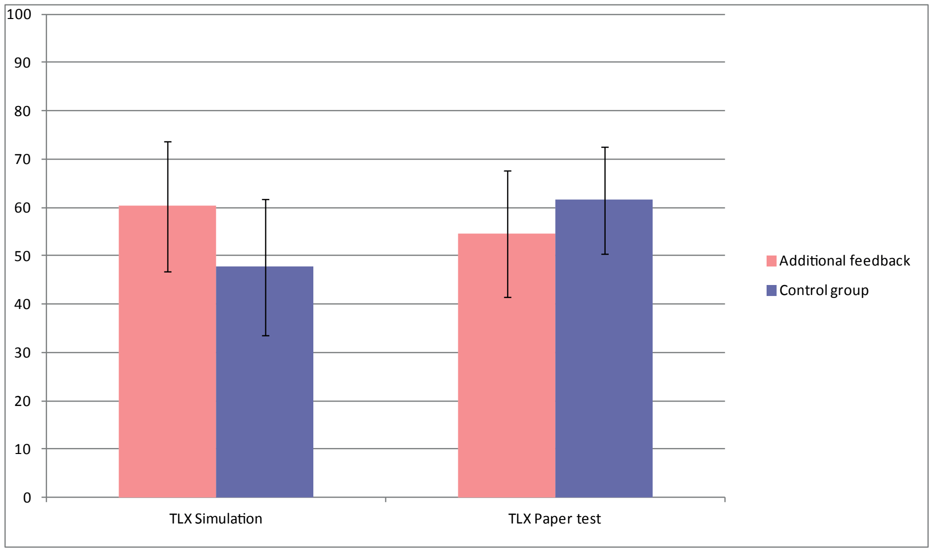



Figure 6.3: Workload measure by group, during the training session, and during the paper test assessment

of the paper test, with multi-choice type of questions, and with questions requiring short answers.

These findings also support the literature that recommends the use of extensive feedback in simulation for novices, and also advocates formative feedback as being more useful for this population, in comparison to summative feedback, or general feedback (Billings, 2012; Schartel, 2012; Bangert-Drowns et al., 1991).

Our second hypothesis concerned workload and estimated that students in the formative feedback condition would find the assessment less demanding than students in the control group. The results of the experiment supported this hypothesis as students who trained with formative feedback rated their workload for the paper assessment significantly less than students in the control group, as the TLX scores suggest. The final hypothesis stated that students subject to the formative feedback condition would, however, associate more workload with the training using our system. The experiment's results supported this hypothesis as well with students in the formative feedback condition rating the workload while training significantly higher than students in the control group.

This joins Schartel (2012)'s findings suggesting that the higher the quantity of feedback, especially formative, the higher the chance to overwhelm students during training, from its constant presence or its richness in terms of information. This could have impacted on students workload during their training session, as our formative feedback delivery was not meant to be regulated, but to always guide students to successfully complete each assessment.

We note, however, a limitation on the TLX results. With this experiment's data, students in the formative feedback group still required more effort as a whole compared to students in the control group, when adding training and assessment. While these results support our hypotheses, we believe it would be interesting to add in a future study another assessment. This test would take place another two weeks following the post training test and assess both students' task load and retention. It would be interesting to see if the TLX results would still stay low for students who were in the formative feedback group, and if these would still correlate with better performance overall.

## 6.3 Conclusion

As Kulhavy and Stock (1989) describe it, good feedback should provide a learner with verification, are they correct or not, as well as Elaboration taking the form of hints and cues leading to the next steps. In addition, feedback aimed to acquire technical skills in clinical education should be offered in real time to allow students the opportunity to correct their behaviour (Schartel, 2012). These are the basis on which the additional feedback integrated in the CAS was implemented.

As a result, this experiment demonstrated that our simulation platform can successfully teach students procedural knowledge when additional formative feedback was introduced. But it also shows how important feedback can be. Students in the formative feedback condition performed significantly better than their counterparts, and showed signs of transfer of the knowledge components embedded in the simulation when assessed following the training session. The results of the students in the control group, however, suggest a lack of understanding of these concepts.

The results of the TLX measures are also encouraging. While students in the formative feedback condition invested more into their training session than the control group when looking at the workload, they reaped the benefits when assessed, by requiring significantly less effort for the paper test than students in the control group.

In conclusion, this experiment showed positive results and validated our formative feedback feature; students assigned to it both were subjected to less workload and demonstrated a better understanding of the knowledge items embedded in the simulation. This component will complete our simulation system as an introduction or tutorial mode and will aim to provide more scaffolding for novice students.

# 7 Conclusions

## 7.1 Achievements

During the course of this research, the following has been achieved:

1. Audiology knowledge, and Nursing field knowledge have successfully been implemented into computer simulations using VPs.

2. The CAS supports practice with a wide range of VPs simulating clients in real settings, within a safe environment and at the students' own pace.

3. Every procedure we initially planned to integrate into the CAS has been implemented.

4. An authoring tool allowing educators to implement new VPs has also been implemented.

5. Different procedures and aspects of the CAS have been evaluated with audiology students, and data suggests positive learning gains with clinical history taking.

6. The nursing simulator prototype, while not having been tested with students from CPIT's nursing degree,showed acceptable usability in a test with nursing educators from CPIT.

This research contributes to the creation of a VP-based simulator for clinical trainees. The simulator gained interest from both the educators and students, who recognized its usefulness as part of a University curriculum. The CAS will now be used alongside traditional teaching methods to support the audiology students' clinical training.

## 7.2 Summary of thesis experiments

### Design & usability

Chapter 2 presents the design and usability testing of the different prototypes. This chapter presents each step of the CAS' design, with the results of the usability testing of each of its components. The design of the first prototype of the nursing simulator is also presented, with the results of preliminary testing with nursing educators. Overall, users were able to report which areas needed to be refined in the simulators, while also presenting their advantages.

This chapter demonstrates how usability testing is an adequate instrument to guide the design and implementation of computer simulations, while ensuring that the expectations of their target audiences are met.

### Pilot studies

Chapter 3 presents the findings of the first pilot studies which aimed to test the setup and collect preliminary data on the impact of the CAS on students' ability to perform clinical history taking and pure tone audiometry. While the findings did not support the hypotheses for learning gains following exposure to the system, we did, however, identify areas for improvement. The pilot studies identified a need to refine the VPs' interaction script and overall to improve the stability of the system.

Students responded positively to what the CAS adds to their methods of learning; its availability and flexibility were appreciated by the first year students who only had a small number of occasions to practice with real clients. Moreover, it allowed them to practice on a wide range of cases, at their convenience.

### Clinical history taking and pure tone audiometry

Using a refined prototype of the CAS, we presented in chapter 4 the findings of a second study focusing on clinical history taking and pure tone audiometry. In addition, we recorded, during the first semester of their course, first year audiology students' opportunities for hands on practice with clients during clinical placements, including time of direct contact and number of clients encountered. Practice time on the CAS and number of VPs assessed have also been recorded.

The study showed significant gains on some measures following exposure to the CAS. Students' ability to retrieve the necessary information from patients and to record it correctly on clinical forms has significantly increased. However, no significant results were found on students' efficiency in conducting clinical history taking or the ability of students to perform pure tone audiometry.

The CAS also contributes to standardizing the students training. The analysis of the students' clinical record books showed that they only had few opportunities to practice with real clients during clinical placements. There was a disparity in opportunities to practice between students, because their placements involved different clinics, and different clinicians supervising them. On the other hand, when comparing these numbers to the students who practiced on the CAS, we noticed that students examined more VP when using this system. Moreover, students could access similar numbers of clients and this provided them with varied forms of encounters at a time convenient for them, which is not the case in clinical placements where the type of patients cannot be controlled.

### Speech audiometry for intermediate and novice users

The speech audiometry study, detailed in chapter 5, expanded the sample of participants from the Master of Audiology students (intermediate students), to include Speech and Language Therapy students who had no previous experience of the procedure (novice students). The results, however, showed no clear statistical evidence that either intermediate or novice users improved their ability to perform speech audiometry following exposure to the CAS. However, while it was not significant, there was a slight trend in the data suggesting that both types of students increased their general knowledge of speech audiometry after practicing with the CAS, when assessed through a short-answer questionnaire.

This study highlighted the difficulty of assessing students for transfer of skills. We concluded that, overall, the post test assessments were too difficult.

Another limitation could have been a potential lack of feedback.

### Feedback in Virtual Patient based simulations

The findings of the study investigating the effects of adding a formative feedback component to the pure tone audiometry module of the CAS are presented in chapter 6. This experiment used novice students as a sample and aimed to determine whether this additional feedback could reinforce students' learning gain following exposure to the CAS. It also aimed to investigate the level of workload to which students were exposed, both while training using the CAS with and without this formative feedback, and while being assessed. Following a training session using four VPs, students were assessed with a paper test.

The results showed that students who had practiced on the CAS with additional formative feedback performed significantly better than students in the control group during the paper assessment; both when tested on multi-choice and short-answer questions. Moreover, students who had practiced with additional feedback, experienced a workload significantly lower than students who were in the control group during the paper assessment. We noted, however, that students in the feedback condition had a higher workload during practice, which can be a side-effect of the additional feedback. We anticipated this last difference and concluded when assessing the results of the post training assessment that the benefits of using additional formative feedback outweighed the training workload.

The results of this thesis' experiments show that using computer simulation to support audiology students training has a direct impact on the acquisition of some of their skills, but not all of them. As the literature suggests (Deladisma et al., 2009; Kotranza et al., 2009; Rizzo et al., 2011), implementing a simulated clinical history taking procedure for the field of audiology is achievable. Moreover, the results presented in chapter 4 present the positive impact on students learning of clinical history taking through practice on the simulator with an increase in their ability to retrieve information and report it adequately.

The results of the experiments did not however find proof that the practice on the computer simulator had a significant impact on improving students acquisition of traditional procedures such as pure tone audiometry, or speech audiometry. This supports the findings from Lieberth and Martin (2005) who demonstrated that students would achieve the same overall grade when training either using a real Audiometer, or a virtual one. However, the benefits of incorporating such computer simulation are still clear from the investigation of students' direct contact hours with patients presented in chapter 4. The computer simulation provided the students with a diverse range of patients and more patients to practice both pure tone audiometry and speech audiometry than their weekly clinical placements.

Finally, the findings presented in chapter 6 also support the use of automated feedback in computer simulation to support learning. The additional formative feedback module resulted in a better understand of the procedure's concepts for the students, when compared to summative feedback, as the literature suggests (Billings, 2012; Bangert-Drowns et al., 1991).

## 7.3 Implications for researchers using Design Based Research

This section elaborates on the implications and limitations of this research. I start with reporting the advantages and disadvantages associated with case study research. Then I will focus on the challenge of implementing functional interaction scripts for VPs. Finally, I will report my perspective on the issues associated with evaluating transfer of skills realistically.

### Case study approach

The research presented in this thesis used the field of audiology as a case study, and was undertaken with clinical educators and students. This type of research offers quantifiable advantages. First, we increased the ecological validity of the research by using students as participants, who are motivated and who are going to use the CAS upon completion of its implementation. The studies were embedded in the curriculum, in conjunction with traditional teachings. In addition, clinical educators are able to provide guidance and insights on how students learn, as well as which methods they are most receptive to. Finally, working alongside students who directly benefit from my research provided a considerable gratification, as well as the motivation to provide students with the tools that best fit their needs.

However, while increasing the external validity of the research, the main downside were the small sample sizes, which made it very difficult to detect small effects. Another issue, is that we had less control over some potential confounding variables and therefore had to sacrifice internal validity.

However, we believe that the benefits outweighed the disadvantages.

Furthermore, another limitation to working in the classroom alongside students, is that the research has to make compromises, primarily to not impact on the students' general learning. An example of this is the study design of the first pilot study presented in chapter 3, which presents the students' results practicing with the CAS to learn pure tone audiometry. The study design had to include an additional week of CAS training for all the students before their midterm and end of term assessments, which were part of the course, in order to be fair on all of them. This was a request from their course coordinator. Another example is that most experiments presented in this thesis had to be conducted either during weekends or in the evenings, as the students' schedule already revolved around their classes, and the studies should not impact on them. In addition, students' have to juggle their traditional examinations and papers to prepare which can, for some of them, result in less focus on the tasks involved in the studies.

Finally, the research conducted in this thesis was part of a bigger project, involving collaborations across disciplines. Consequently, some compromises had to be made in regards of the experiments. This might have affected the results of the first experiments. However, by later conducting more rigorous and controlled studies, clearer results have been found.

The following section presents why the CAS was successful with audiology students and explains the learning gains observed on students' ability to perform the clinical history taking procedure.

### Interaction scripts

As described in the introduction chapter of this thesis, interaction scripts provide VPs with an understanding of the users' vocabulary. Consequently, the more refined an interaction script, the more likely a VP will respond to a user. Clinical history taking is one of the most crucial procedures for a clinician, and arguably the most difficult to implement. As explained in the literature review, creating a working interaction script for a VP is a time consuming process which can take up to hundreds of man hours to refine to a satisfactory level. Moreover, having an underdeveloped interaction script will create frustrated users, and a frustrated user is unlikely to carry on using a system, and even less likely to learn from it.

The case study used in this research had the advantage of incorporating a simple approach to clinical history taking. The amount of information clinicians have to retrieve is limited to approximately 30 types of questions, and those are meant to be asked in a concise manner. This allowed us to create scripts that were usable following refinements of the first CAS prototype, without relying on an operator remotely controlling the VPs' answers. This would have been different in the nursing simulator, which would have incorporated a branching dimension in this procedure as well, with new questions being accessible at a later stage of the interview.

The results collected through this thesis show the potential of having such a functioning system. While still not having 100% of response rate from the VPs upon asking questions, students were able to practice the procedure successfully and learn the range of question they should cover during this exam.

### Evaluating transfer of skills

Another challenge tackled in this research was to assess students for transfer of skills. A *true* assessment of transfer of skills would be to evaluate students in clinic following exposure to the system, with real patients. However, this method is not applicable in practice, as real clients do not allow a standardized assessment and are not controllable.

In order to provide a standardized method of assessing the students while still evaluating transfer of skills, we used role-play situations which provided each participant with the same encounter. While this provides a good assessment method to evaluate transfer of skills to clinical settings, it also has disadvantages.

First, role-play situations are time consuming. They required a significant investment from field experts in order to realistically impersonate the patients, and to provide consistent marking for each of the participants. Another limitation of this research is that it only used one case with those intervention to assess for transfer of skills following exposure to the VPs. This showed effective for clinical history taking, but not for pure tone audiometry and speech audiometry.

We believe that a contributing factor in not having been able to measure significant learning gains on those procedures was due to the limits of our assessments, which were based on a single role-play case. Assessing students on a range of cases simulated in role-play situations could provide a more objective assessment of their skills.

Finally, it is also noteworthy to mention that students were not monitored for potential bad habits they could acquire while training with the simulator in the absence of a clinician's mentoring.

## 7.4 Future Work

The results of this research indicate that it would be valuable to continue. Presentation of the prototypes attracted the interest of educators from overseas and, provided that funding is secured, one aim is to conduct further studies with those Universities. In addition, work on the nursing simulator is planned to be resumed. We still intend to evaluate it in comparison to its mannequin-based replica and assess its validity as an alternative training method.

This research also highlighted the importance to create more efficient ways to implement responsive VPs in computer simulations. It is likely that more findings will be made to support the use of VPs in education while further work on frameworks and approaches to improve scripting is undertaken.

Another goal would be to answer whether it is feasible to use VP-based simulators not only as a supplementary tool but also to replace at least part of the mandatory clinical encounters required in these curricula. By proving the validity of VPs as training tools, clinical organisation could incorporate this form of training in their certifications' requirements. This could contribute to resolve the shortage of qualified personnel clinical professions can face.

# Bibliography


Astwood, R. S., Van Buskirk, W. L., Cornejo, J. M., and Dalton, J. (2008). The Impact of Different Feedback Types on Decision-Making in Simulation Based Training Environments. *Proceedings of the Human Factors and Ergonomics Society Annual Meeting*, 52(26):2062–2066.

Bangert-Drowns, R., Chen-Lin, K., James, K., and Teresa, M. M. (1991). The Instructional Effect of Feedback in Test-Like Events. *Review of Educational Research*, 61(2):213–238.

Baylor, A. L. and Kim, S. (2009). Designing nonverbal communication for pedagogical agents: When less is more. *Computers in Human Behavior*, 25(2):450–457.

Bell, D. (2009). Learning from Second Life. *British Journal of Educational Technology*, 40(3):515–525.

Bickmore, T. W., Pfeifer, L. M., Byron, D., Forsythe, S., Henault, L. E., Jack, B. W., Silliman, R., and Paasche-Orlow, M. K. (2010). Usability of conversational agents by patients with inadequate health literacy: evidence from two clinical trials. *Journal of Health Communication*, 15 Suppl 2:197–210.

Bickmore, T. W., Pfeifer, L. M., and Paasche-Orlow, M. K. (2009). Using computer agents to explain medical documents to patients with low health literacy. *Patient Education and Counseling*, 75(3):315–320.

Bickmore, T. W. and Picard, R. W. (2005). Establishing and maintaining long-term human-computer relationships. *ACM Transactions on Computer-Human Interaction*, 12(2):293–327.

Billings, D. R. (2012). Efficacy of Adaptive Feedback Strategies in Simulation-Based Training. *Military Psychology*, 24(2):114–133.

Botezatu, M., Hult, H. k., Tessma, M. K., and Fors, U. (2010). Virtual patient simulation: knowledge gain or knowledge loss? *Medical teacher*, 32(7):562–568.

Boulos, M. N. K., Hetherington, L., and Wheeler, S. (2007). Second Life: an overview of the potential of 3-D virtual worlds in medical and health education. *Health Information & Libraries Journal*, 24(4):233–245.

Brown, A. L. (1992). Design Experiments: Theoretical and Methodological Challenges in Creating Complex Interventions in Classroom Settings. *The Journal of the Learning Sciences*, 2(2):141–178.

Bruck, L. (2008). Second Life: Test-Driving Real-World Innovations. *Hospitals & Health Networks*, 82(10):50,52,54.

Burgess, M. L. and Caverly, D. C. (2009). Techtalk: Second Life and Developmental Education. *Journal of Developmental Education*, 32(3):42–43.

Callaghan, V., Gardner, M., Horan, B., Scott, J., Shen, L., and Wang, M. (2008). A mixed reality teaching and learning environment. 5169 LNCS:54–65.

Cohen, J. (1988). *Statistical power analysis for the behavioral sciences*. Lawrence Erlbaum Associates.

Collins, R. P. and Harden, R. M. (1998). AMEE Medical Education Guide No. 13: real patients, simulated patients and simulators in clinical examinations. *Medical Teacher*, 20(6):508–521.

Conradi, E., Kavia, S., Burden, D., Rice, A., Woodham, L., Beaumont, C., Savin-Baden, M., and Poulton, T. (2009). Virtual patients in a virtual world: Training paramedic students for practice. *Medical Teacher*, 31(8):713–720.

Cook, D. A. and Triola, M. M. (2009). Virtual patients: a critical literature review and proposed next steps. *Medical Education*, 43(4):303–311.

Courteille, O., Bergin, R., Stockeld, D., Ponzer, S., and Fors, U. (2008). The use of a virtual patient case in an OSCE-based exam - A pilot study. *Medical Teacher*, 30(3):66–76.

Deladisma, A. M., Gupta, M., Kotranza, A., Bittner, J. G., Imam, T., Swinson, D., Gucwa, A., Nesbit, R., Lok, B., Pugh, C., and Lind, D. S. (2009). A pilot study to integrate an immersive virtual patient with a breast complaint and breast examination simulator into a surgery clerkship. *American journal of surgery*, 197(1):102–106.

Deladisma, A. M., Johnsen, K., Raij, A., Rossen, B., Kotranza, A., Kalapurakal, M., Szlam, S., Bittner, J. G., Swinson, D., Lok, B., and Lind, D. S. (2008). Medical student satisfaction using a virtual patient system to learn history-taking communication skills. *Studies In Health Technology And Informatics*, 132:101–105.

den Akker, J., Branch, R. M., Gustafson, K., Nieveen, N., and Plomp, T. (2000). *Principles and methods of development research*, pages 1–14. Kluwer.

Dickerson, R., Johnsen, K., Raij, A., Lok, B., Stevens, A., Bernard, T., and Lind, D. S. (2006). Virtual patients: assessment of synthesized versus recorded speech. *Studies In Health Technology And Informatics*, 119:114–119.

Edelbring, S., Dastmalchi, M., Hult, H. k., Lundberg, I., and Dahlgren, L. (2011). Experiencing virtual patients in clinical learning: a phenomenological study. *Advances in Health Sciences Education*, 16(3):331–345.

Fabri, M., Elzouki, S. Y. A., and Moore, D. (2007). Emotionally expressive avatars for chatting, learning and therapeutic intervention. In *Proceedings of the 12th international conference on Human-computer interaction: intelligent multimodal interaction environments*, HCI'07, pages 275–285, Berlin, Heidelberg. Springer-Verlag.

Ferdig, R. E., Coutts, J., DiPietro, J., Lok, B., and Davis, N. (2007). Innovative technologies for multicultural education needs. *Multicultural Education Technology Journal*, 1(1):47–63.

Foroughi, R. (2006). Work in progress: Giving asynchronous learning platforms a humanoid head for effective communication with the learners. In *Proceedings - Frontiers in Education Conference, FIE*, San Diego, CA.

Forsberg, E., Georg, C., Ziegert, K., and Fors, U. (2011). Virtual patients for assessment of clinical reasoning in nursing - A pilot study. *Nurse Education Today*, 31(8):757–762.

Foster, M. E. (2007). Enhancing human-computer interaction with embodied conversational agents. pages 828–837.

Franceschi, K. G., Lee, R. M., and Hinds, D. (2008). Engaging e-learning in virtual worlds: Supporting group collaboration. *Hawaii International Conference on System Sciences*, 0:7.

Fuhrman Jr, L. C., Buff, W. E., Eaddy, M., and Dollar, M. (2001). Utilization of an integrated interactive virtual patient database in a Web-based environment for teaching continuity of care. *American Journal of Pharmaceutical Education*, 65(3):271.

Garzotto, F. and Forfori, M. (2006). *Hyperstories and social interaction in 2D and 3D edutainment spaces for children*, volume 2006. Association for Computing Machinery, Odense, Denmark.

Gorini, A., Gaggioli, A., Vigna, C., and Riva, G. (2008). A second life for eHealth: prospects for the use of 3-D virtual worlds in clinical psychology. *Journal of medical Internet research*, 10(3):e21.

Gorrindo, T. and Groves, J. E. (2009). Computer Simulation and Virtual Reality in the Diagnosis and Treatment of Psychiatric Disorders. *Academic Psychiatry*.

Guise, V., Chambers, M., and Välimäki, M. (2012a). What can virtual patient simulation offer mental health nursing education? *Journal of Psychiatric and Mental Health Nursing*, 19(5):410–418.

Guise, V. y., Chambers, M., Conradi, E., Kavia, S., and Välimäki, M. (2012b). Development, implementation and initial evaluation of narrative virtual patients for use in vocational mental health nurse training. *Nurse Education Today*, 32(6):683–689.

Halan, S., Rossen, B., Cendan, J., and Lok, B. (2010). High Score ! - Motivation Strategies for User Participation in Virtual Human Development. *Seven*, pages 482–488.

Havnes, A., Smith, K., Dysthe, O., and Ludvigsen, K. (2012). Formative assessment and feedback: Making learning visible. *Studies in Educational Evaluation*, 38(1):21–27.

Hayes-roth, B. and Saker, R. (2003). Next-Generation Automated Health Behavior Coaches. *Symposium A Quarterly Journal In Modern Foreign Literatures*, pages 28–34.

Heer, I. M., Middendorf, K., Müller-Egloff, S., Dugas, M., and Strauss, A. (2004). Ultrasound training: the virtual patient. *Ultrasound in Obstetrics and Gynecology*, 24(4):440–444.

Heiphetz, A. and Liberman, S. (2008). Using Robotic Avatars in Second Life Simulations and Training. 22:19–20.

Hildner, M. (2007). Virtual human boosts nurses' training: Students have to complete most of their clinical training with human patients.

Hoadley, C. M. (2004). Methodological Alignment in Design-Based Research. *Educational Psychologist*, 39(4):203–212.

Hsieh, I.-L. G. and O'Neil Jr., H. F. (2002). Types of feedback in a computer-based collaborative problem-solving group task. *Computers in Human Behavior*, 18(6):699–715.

Hubal, R. and Frank, G. (2007). Embodied tutors for interaction skills simulation training. *RTI International*, 7(1):1–8.

Husén, T. and Postlethwaite, T. (1994). *The International encyclopedia of education*. Number v. 1 in The International Encyclopedia of Education. Pergamon.

Johnsen, K. and Lok, B. (2008). An Evaluation of Immersive Displays for Virtual Human Experiences. In *Virtual Reality Conference, 2008. VR '08. IEEE*, pages 133–136.

Joseph, D. (2004). The practice of design-based research: Uncovering the interplay between design, research, and the real-world context. *Educational Psychologist*, 39(4):235 – 242.

Kenny, P., Parsons, T. D., Gratch, J., Leuski, A., and Rizzo, A. A. (2007). Virtual patients for clinical therapist skills training. In *Proceedings of the 7th* *international conference on Intelligent Virtual Agents*, IVA '07, pages 197–210, Berlin, Heidelberg. Springer-Verlag.

Kenny, P. G. and Parsons, T. D. (2010). Virtual Patients for Virtual Sick Call Medical Training. In *Training, Simulation, and Education Conference*.

Kenny, P. G., Parsons, T. D., Rizzo, A., Reger, G. M., Caroly, P., Pato, M., Sugar, J., and StGeorge, C. (2008). Virtual Patients for Future Leaders.

Kneebone, R. and Nestel, D. (2005). Learning clinical skills âĂŞ the place of simulation and feedback. *The Clinical Teacher*, 2(2):86–90.

Kotranza, A., Johnsen, K., Cendan, J., Miller, B., Lind, D. S., and Lok, B. (2009). Virtual multi-tools for hand and tool-based interaction with life-size virtual human agents. In *2009 IEEE Symposium on 3D User Interfaces*, pages 23–30. IEEE.

Kulhavy, R. and Stock, W. (1989). Feedback in written instruction: The place of response certitude. *Educational Psychology Review*, 1(4):279–308 LA – English.

Kyle Johnsen, R. D. and Andrew Raij, B. L. (2005a). Experiences in using immersive virtual characters to educate medical communication skills. *IEEE Proceedings VR 2005 Virtual Reality 2005*, pages 1–8.

Kyle Johnsen, R. D. and Andrew Raij, B. L. (2005b). Experiences in using immersive virtual characters to educate medical communication skills. *IEEE Proceedings VR 2005 Virtual Reality 2005*, pages 1–8.

Lee, K. and Lee, J. (2004). Benefits of a virtual learning environment in enabling collaborative and constructivist learning in a community of specialist nursing practice. In *Advanced Learning Technologies, 2004. Proceedings. IEEE International Conference on*, pages 672–674.

Lieberth, A. K. and Martin, D. R. (2005). The Instructional Effectiveness of a Web-Based Audiometry Simulator. *Journal of The American Academy of Audiology*, 16:79–84.

Mahmood, A. K. and Ferneley, E. (2006). Embodied Agents in E-Learning Environments: An Exploratory Case Study. *Journal of Interactive Learning Research*, 17(2):143–162.

Melús-Palazón, E., Bartolomé-Moreno, C., Palacín-Arbués, J. C., Lafuente-Lafuente, A., García, I. G., Guillen, S., Esteban, A. B., Clemente, S., Marco, A. M., Gargallo, P. M., López, C., and Magallón-Botaya, R. (2012). Experience with using second life for medical education in a family and community medicine education unit. *BMC Medical Education*, 12(1):30.

Mikropoulos, T. A. and Strouboulis, V. (2004). Factors that influence presence in educational virtual environments. *Cyberpsychology and Behavior*, 7(5):582–591.

Minocha, S. and Reeves, A. J. (2010). Design of learning spaces in 3D virtual worlds: an empirical investigation of Second Life. *Learning, Media and Technology*, 35(2):111–137.

Ong, J. (2007). Automated performance assessment and feedback for free-play simulation-based training. *Performance Improvement*, 46(10):24–31.

Pan, Z. G., Zhu, J. J., Hu, W. H., Lun, H. P., Zhou, X., and Eurographics, A. S. (2004). Interactive learning of CG in networked virtual environments. In *Workshop on Computer Graphics Education (CGE 04)*, pages 273–281, Hangzhou, PEOPLES R CHINA. Pergamon-Elsevier Science Ltd.

PeerGroup (2006). *A PEER Tutorial For Design-based Research*. http://projects.coe.uga.edu/dbr/explain01.htm.

Rizzo, A., Kenny, P., and Parsons, D. T. (2011). Intelligent virtual patients for training clinical skills. *JVRB - Journal of Virtual Reality and Broadcasting*, 8(3).

Rizzo, A., Parsons, T. D., Buckwalter, J. G., and Kenny, P. G. (2010). A new generation of intelligent virtual patients for clinical training. *Proceedings of IEEE Virtual Reality Conference*.

Rossen, B., Cendan, J., and Lok, B. (2010). *Using Virtual Humans to Bootstrap the Creation of Other Virtual Humans*, volume 6356 of *Lecture Notes in Computer Science*. Springer Berlin Heidelberg.

Round, J., Conradi, E., and Poulton, T. (2009). Training staff to create simple interactive virtual patients: the impact on a medical and healthcare institution. *Medical Teacher*, 31(8):764–769.

Ruttkay, Z. and Welbergen, H. V. (2008). Elbows Higher ! Performing , Observing and Correcting Exercises by a. In *Proceeding IVA '08 Proceedings of the 8th international conference on Intelligent Virtual Agents*, pages 409–416.

Salmon, G. (2009). The future for (second) life and learning. *British Journal of Educational Technology*, 40(3):526–538.

Sanchez, J. (2009). Barriers to student learning in Second Life. *Library Technology Reports*, 45:29.

Sandoval, W. A. and Bell, P. (2004). Design-based research methods for studying learning in context: Introduction. *Educational Psychologist*, 39(4):199–201.

Saraf, S., Bayya, J., Minkoff, H., Glick, J., and Fisher, N. (2012). The relationship of type of feedback to learning during simulation: a randomized clinical trial. *American Journal of Obstetrics and Gynecology*, 206(1, Supplement):S307.

Schartel, S. A. (2012). Giving feedback - An integral part of education. *Best Practice & Research Clinical Anaesthesiology*, 26(1):77–87.

Shaw, G. (2013). Cover Story: Simulation Clears Path from Education to Real-World Practice. *The Hearing Journal*, 66(2):12,14,16,17.

Shea, P. J., Pickett, A. M., and Pelz, W. E. (2003). A follow-up investigation of "teaching presence" in the suny learning network. *Journal of Asynchronous Learning Network*, 7(2):61–80.

Shute, V. J. (2008). Focus on Formative Feedback. *Review of Educational Research*, 78(1):153–189.

Sitzmann, T. (2011). A meta-analytic examination of the instructional effectiveness of computer-based simulation games. *Personnel Psychology*, 64(2):489–528.

Skiba, D. J. (2007). Nursing Education 2.0: Second Life. *Nursing Education Perspectives*, 28(3):156–157.

Stanton, T. (2008). Aspiring Medical Assistants Can Now Provide Care for Virtual Patients via Interactive Simulation.

Sweigart, L. and Carlton, K. H. (2009). Volunteer Clients and Second Life Virtual Experiences Change the Landscape of Health Appraisal Acorss the Lifespan. *Clinical Simulation in Nursing*, 5(3, Supplement):e152.

Tabak, I. (2004). Reconstructing Context: Negotiating the Tension Between Exogenous and Endogenous Educational Design. *Educational Psychologist*, 39(4):225–233.

Tan, Z. S., Mulhausen, P. L., Smith, S. R., and Ruiz, J. G. (2010). Virtual Patients in Geriatric Education. *Gerontology & Geriatrics Education*, 31(2):163–173.

Thwaites, J. H., Hilder, P., Sheehan, D., and Smith, S. (2008). Sick patient simulation training for interns - a New Zealand pilot: Does simulation training improve the ability of interns to manage acutely unwell patients? *IJOCS*, 2(1):12–16.

Triola, M., Feldman, H., Kalet, A. L., Zabar, S., Kachur, E. K., Gillespie, C., Anderson, M., Griesser, C., and Lipkin, M. (2006). A randomized trial of teaching clinical skills using virtual and live standardized patients. *Journal of general internal medicine*, 21(5):424–9.

Urdan, T. (2005). Statistics in Plain English, second edition.

Van Ments, M. (1999). *The Effective Use of Role Play: Practical Techniques for Improving Learning*. Kogan Page, Limited.

Vash, J. H., Yunesian, M., Shariati, M., Keshvari, A., and Harirchi, I. (2007). Virtual patients in undergraduate surgery education: a randomized controlled study. *ANZ journal of surgery*, 77(1-2):54–9.

Walker, V. L. and Rockinson-Szapkiw, A. (2009). Educational Opportunities for Clinical Counseling Simulations in Second Life. *Innovate (North Miami Beach, Fla.)*, 5(5).

Wang, F. and Hannafin, M. J. (2005). Design-based research and technology-enhanced learning environments. *Etr&D-Educational Technology Research and Development*, 53(4):5–23.

Wang, Y. and Braman, J. (2009). Extending the Classroom through Second Life. *Journal of Information Systems Education*, 20(2):235–247.

Watson, R. (2001). SPSS Survival Manual by Julie Pallant, Open University Press, Buckingham, 2001, 286 pages, f16.99, ISBN 0 335 20890 8. *Journal of Advanced Nursing*, 36(3):478–478.

Wayne, W., Hill, A., Jane, H., Sher, Auriel, and Arianne, L.-l. (2010). Student Audiologists' Impressions of a Simulation Training Program (IELNZC). *Australian and New Zealand Journal of Audiology, The*, 32(1):19–30.

Wilkerson, W., Avstreih, D., Gruppen, L., Beier, K. P., and Woolliscroft, J. (2008). Using Immersive Simulation for Training First Responders for Mass Casualty Incidents. *Academic Emergency Medicine*, 15(11):1152–1159.

Ye, E., Liu, C., and Polack-Wahl, J. A. (2007). Enhancing software engineering education using teaching aids in 3-D online virtual worlds. In *Proceedings - Frontiers in Education Conference, FIE*, Milwaukee, WI.

Zeng, R. and Blasi, L. (2010). Learning through web-based multistoryline case studies a design-based research. *Quarterly Review of Distance Education*, 11(3):175–182.

# Appendix A: Usability Questionnaire

**Usability questionnaire** **ID**

*Question One - Usefulness*

a) The simulator helps me be more effective

1 - - - 2 - - - 3 - - - 4 - - - 5 - - - 6 - - - 7
strongly disagree strongly agree

b) It helps me be more productive

1 - - - 2 - - - 3 - - - 4 - - - 5 - - - 6 - - - 7
strongly disagree strongly agree

c) It is useful

1 - - - 2 - - - 3 - - - 4 - - - 5 - - - 6 - - - 7
strongly disagree strongly agree

d) It gives me more control over my learning

1 - - - 2 - - - 3 - - - 4 - - - 5 - - - 6 - - - 7
strongly disagree strongly agree

e) It makes the things I want to accomplish easier to get done

1 - - - 2 - - - 3 - - - 4 - - - 5 - - - 6 - - - 7
strongly disagree strongly agree

f) It saves me time when I use it

1 - - - 2 - - - 3 - - - 4 - - - 5 - - - 6 - - - 7
strongly disagree strongly agree

g) It meets my needs

1 - - - 2 - - - 3 - - - 4 - - - 5 - - - 6 - - - 7
strongly disagree strongly agree

h) It does everything I would expect it to do

1 - - - 2 - - - 3 - - - 4 - - - 5 - - - 6 - - - 7
strongly disagree strongly agree

*Question Two - Ease of use*

a) Overall, I am satisfied with the ease of completing the tasks in these scenarios

1 - - - 2 - - - 3 - - - 4 - - - 5 - - - 6 - - - 7
strongly disagree strongly agree

b) Overall, I am satisfied with the amount of time it took to complete the tasks in these scenarios

1 - - - 2 - - - 3 - - - 4 - - - 5 - - - 6 - - - 7
strongly disagree strongly agree

c) I can use it without written instructions
1 - - - 2 - - - 3 - - - 4 - - - 5 - - - 6 - - - 7
strongly disagree strongly agree

d) I don't notice any inconsistencies as I use it
1 - - - 2 - - - 3 - - - 4 - - - 5 - - - 6 - - - 7
strongly disagree strongly agree

e) Both occasional and regular users would like it
1 - - - 2 - - - 3 - - - 4 - - - 5 - - - 6 - - - 7
strongly disagree strongly agree

f) I can recover from mistakes quickly and easily
1 - - - 2 - - - 3 - - - 4 - - - 5 - - - 6 - - - 7
strongly disagree strongly agree

g) I can use it successfully every time
1 - - - 2 - - - 3 - - - 4 - - - 5 - - - 6 - - - 7
strongly disagree strongly agree

*Question Three - Ease of Learning*

a) I learned to use it quickly
1 - - - 2 - - - 3 - - - 4 - - - 5 - - - 6 - - - 7
strongly disagree strongly agree

b) I easily remember how to use it
1 - - - 2 - - - 3 - - - 4 - - - 5 - - - 6 - - - 7
strongly disagree strongly agree

c) It is easy to learn to use it
1 - - - 2 - - - 3 - - - 4 - - - 5 - - - 6 - - - 7
strongly disagree strongly agree

d) I quickly became skillful with it
1 - - - 2 - - - 3 - - - 4 - - - 5 - - - 6 - - - 7
strongly disagree strongly agree

*Question Four – Satisfaction*

a) I am satisfied with it
1 - - - 2 - - - 3 - - - 4 - - - 5 - - - 6 - - - 7
strongly disagree strongly agree

b) I would recommend it to a friend
1 - - - 2 - - - 3 - - - 4 - - - 5 - - - 6 - - - 7
strongly disagree strongly agree

c) It is fun to use

1 - - - 2 - - - 3 - - - 4 - - - 5 - - - 6 - - - 7

strongly disagree strongly agree

d) It works the way I want it to work

1 - - - 2 - - - 3 - - - 4 - - - 5 - - - 6 - - - 7

strongly disagree strongly agree

e) It is wonderful

1 - - - 2 - - - 3 - - - 4 - - - 5 - - - 6 - - - 7

strongly disagree strongly agree

f) I feel I need to have it

1 - - - 2 - - - 3 - - - 4 - - - 5 - - - 6 - - - 7

strongly disagree strongly agree

g) It is pleasant to use

1 - - - 2 - - - 3 - - - 4 - - - 5 - - - 6 - - - 7

strongly disagree strongly agree

*Question Five*

a) Please list five negative aspects of the simulator:

1.______________________________________________________________

2.______________________________________________________________

3.______________________________________________________________

4.______________________________________________________________

5.______________________________________________________________

b) Please list five positive aspects of the simulator:

1.______________________________________________________________

2.______________________________________________________________

3.______________________________________________________________

4.______________________________________________________________

5.______________________________________________________________

# Appendix B: Diagnostic Adult History Form (DAHF)

**University of Canterbury Speech and Hearing Clinic**
**Department of Communication Disorders, Private Bag 4800**
**Phone: (03) 364 2408 Fax: (03) 364 2760**

## Diagnostic Adult History Form

*Client Name:* ______________________________ *DOB/Age:* __________/____*yrs* *NHI#:*__________

*Student:* ______________________ *Audiologist:* ________________ *Date:* __________

**Reason for Referral:**

**View of Hearing / Level of Concern**
Onset of Loss *When:* *Sudden or Gradual*
Fluctuation?
Better Ear? *Right / Left / Same*

**Communication Problems** Yes / No
Hearing in Quiet vs Noise
Hearing in 1:1 vs Groups

Telephone: TV/Radio:
Previous Hearing Aid use? Left / Right

**Family History** Yes / No
Age of onset & cause if known

**Tinnitus** Yes / No
Onset Bilateral / Unilateral
Pitch/Description of 'Sound'
Fluctuation? Level of Annoyance

**Balance/Vertigo** Yes / No
Onset
Type Unsteadiness / Rotary Vertigo
Duration of each episode
Possible Cause

**Feeling of Fullness or Pressure in Ears** Yes / No

**Facial Numbness or Weakness** Yes / No

**Ear/ENT History** Yes / No
Middle Ear Problems?
Treatment or Surgery

**Noise Exposure** Yes / No
Occupational/Recreational
Duration of Exposure

Hearing Protection? Y/N

**Major illness or Operations** *Meningitis / Mumps / Measles / Diabetes / Heart & B.P.*
General Health
Medications

Head Injuries

Version 4c: edited 02/05/08 by J Grady

# Appendix C: Observation chart

Name: Jones Smith

NHI: 1-0937-0343-76

**Adult Observation Chart**

| | | 9:00 p.m. | 1:00 a.m. | 5:00 a.m. | | | | |
|---|---|---|---|---|---|---|---|---|
| Respiratory Rate | 35 | | | | | | | |
| | 30 | | | | | | | |
| | 25 | | | | | | | |
| | 20 | X | X | | | | | |
| | 15 | | | | | | | |
| | 10 | | | | | | | |
| | 5 | | | | | | | |
| | 0 | | | | | | | |
| SaO2 | | 93 | 93 | - | - | - | - | - |
| O2 L/% | | 24% | 24% | - | - | - | - | - |
| O2 method | | Mask | Mask | - | - | - | - | - |
| Heart Rate | 150 | | | | | | | |
| | 140 | | | | | | | |
| | 130 | | | | | | | |
| | 120 | | | | | | | |
| | 110 | | | | | | | |
| | 100 | | X | | | | | |
| | 90 | X | | | | | | |
| | 80 | | | | | | | |
| | 70 | | | | | | | |
| | 60 | | | | | | | |
| | 50 | | | | | | | |
| | 40 | | | | | | | |
| | 30 | | | | | | | |
| | 20 | | | | | | | |
| Rhythm Reg/Irreg | | Irregular | Irregular | - | - | - | - | - |
| Blood Pressure | | 130/65 | 120/63 | - | - | - | - | - |
| Temperature (C) | 39 | | | | | | | |
| | 38 | | | | | | | |
| | 37 | X | X | | | | | |
| | 36 | | | | | | | |
| | 35 | | | | | | | |
| | 34 | | | | | | | |
| Neuro | | Alert | Alert | - | - | - | - | - |
| Urine EWS | | 0 | 0 | - | - | - | - | - |
| EWS | | 0 | 0 | - | - | - | - | - |
| Pain Score 0-5 | | 0 | 0 | - | - | - | - | - |
| Phlebitis Score | | 0 | 0 | - | - | - | - | - |
| Bowel Motion | | 0 | 0 | - | - | - | - | - |
| Weight | | 0 | 0 | - | - | - | - | - |

# Appendix D: Drug Chart

**Drug Treatment Sheet**

## Drug Treatment Sheet

**Name:** John Smith **NHI:** 1-0937-0343-76
**Sex:** male **Age:** 33 **Ward:** Intensive Care
**Consultant/Team:** Dr. Jones

Drug Reaction Sticker here

| | |
|---|---|
| HEIGHT: | 175 cm |
| WEIGHT: | 75 kg |
| LEAN BODY WEIGHT: | 70 kg |
| SURFACE: | 17 m2 |

PATIENT HAS OTHER DRUG TREATMENT SHEETS ☐

CHRONIC RENAL IMPAIRMENT ☐

ABBREVIATIONS PERMITTED ON DRUG TREATMENT SHEETS

| | | | | | |
|---|---|---|---|---|---|
| b.d. | twice daily | q.4.h. | every 4 hours, etc. | i.v. | intravenous |
| t.d.s. | three times a day | a.c. | before food | p.o. | oral |
| q.i.d. | four times a day | p.c. | after food | p.r. | rectal |
| mane | In the morning | p.r.n. | as required | s.c. | subcutaneously |
| nocte | at night | i.m. | intramuscular | | |

| Code | Date | Drug and Dosage | Time to be Given | Route | Prescriber's Signature Surname Print | Time Given | Initials | Second Check |
|---|---|---|---|---|---|---|---|---|
| 0 | 3/4/2013 | Prednisone 40mg daily | 21:00 | 1 | Dr. Williams | 21:05 | JW | none |
| 1 | 3/4/2013 | Aspirin 100mg daily | 21:00 | 1 | Dr. Williams | 21:05 | JW | none |
| 2 | 3/4/2013 | Atrovent mdi. PRN | 21:00 | 1 | Dr. Williams | 21:05 | JW | none |
| 3 | 3/4/2013 | Salbutamol mdi. PRN | 21:00 | 1 | Dr. Williams | 21:05 | JW | none |
| 4 | 3/4/2013 | Beclamethasone mdi. 2puffs BD | 21:00 pm | 1 | Dr. Williams | 21:05 | JW | none |
| 5 | 3/4/2013 | Allopurinol 200mg daily | 21:00 | 1 | Dr. Williams | 21:05 | JW | none |
| 6 | 3/4/2013 | Omeprazole 20mg daily | 21:00 | 1 | Dr. Williams | 21:05 | JW | none |
| 7 | 3/4/2 | Enalapril 5mg | 21:00 | 1 | Dr. Williams | 21:05 | JW | none |

# Appendix E: Pure Tone Audiometry Marking Criteria

***CMDS610 Exam 2009 Student:_________________***

## Pure tone Audiometry

1. *Otoscopy-*
   - Check both ears with otoscope before beginning - (removed) □ (½mark)
   - Use correct technique – brace finger against cheek, etc □ (½mark)

2. **Instructions to client:** should include:
   - will play a series of quiet 'beeps' or tones □ (½mark)
   - client should press the button when they hear the beep □ (½mark)
   - should respond even if they aren't quite sure □ (½mark)
   - it shouldn't matter which ear they hear the tones □ (½mark)

3. **Choosing correct transducers:**
   - Headphones if wax present/Inserts if no wax □ (½mark)
   - Correct headphone placement/Insert depth □ (½mark)
   - Correct left & right □ (½mark)

4. ***Start*** with better hearing ear and use appropriate level □ (1 mark)

5. **AC Frequency order:**
   *1kHz* □ *2kHz* □ *4kHz* □ *8kHz* □ *500Hz* □ *250Hz* □ (Recheck 1k □) _______/3
   - Do intermediate freqs where necessary: (-½ mark if missed) ________

6. Use correct ***Mod. Hughson Westlake*** *method:* (-½ each mistake)_________/4

7. ***Pace:*** out of 3marks: *3* for varied inter-stimulus pace & 1-2s presentations
   *2* for some lack of variation or quick presentations
   *1* for little variation and quick presentations _______/3

8. ***Bone Conduction***: Correct bone conductor placement □ (½mark)
   Instructions to client (for BC and masking) ________/3
   - Recognising need for BC at appropriate freqs ________/2
   - Should do on each AC frequency outside normal range: (500Hz, 1kHz, 2kHz, 4kHz) (-½ for each missed freq) _______/2.
   - Apply correct threshold-seeking protocol (-½ each mistake) _____/2.

9. ***Masking*** Apply correct Plateau method to each frequency that needs it:
   *6 marks total* (4 for BC, 2 for any AC; -1 mark for each mistake) _______/6

10. ***Explaining Results:***
    1 mark each for explaining *Audiogram*; *configuration*; *AC/BC* _______/3

11. ***Recording Results:*** 1 mark for each correctly recorded AC ear & BC ear
    (2 marks for AC; 1 for unmasked BC; 1 for Masked BC) _______/4

***Total:* _____/38**

# Appendix F: Perceived Level of Learning Questions

Name: Student Questionnaire | 1

Consider your interactions with the computer simulation. Please indicate to what extent you agree/disagree with the following statements:

1 – *Strongly Disagree,* 2 – *Disagree*, 3 – *Unsure*, 4 – *Agree,* 5 – *Strongly Agree*

1. My ability to obtain accurate unmasked pure tone thresholds from a client has improved as a result of working with the computer simulation.

   1 2 3 4 5

2. My ability to obtain accurate masked pure tone thresholds from a client has improved as a result of working with the computer simulation.

   1 2 3 4 5

3. My ability to confidently recognise and explain the type and degree of hearing loss from an audiogram has improved as a result of working with the computer simulation.

   1 2 3 4 5

4. My confidence to perform audiometric testing on real clients in the future has increased as a result of working with the computer simulation.

   1 2 3 4 5

5. I learned a new skill as a result of working with the computer simulation.

   1 2 3 4 5

# Appendix G: Speech Audiometry Marking Criteria

| Item | Potential Marks | | Awarded |
|---|---|---|---|
| | Better ear | Worse ear | |
| | | | |
| **Instructions to clients (-1 for missing instruction)** | 2 | | |
| | | | |
| **Reading audiogram and otsocopy** | | | |
| Start with better ear | 2 | | |
| Start with correct transducers | 2 | | |
| | | | |
| **Calibration** | 1 | | |
| Correctly done | 1 | | |
| | | | |
| **Presentations:** | | | |
| *PI Max* | | | |
| Start at appropriate presentation level | 2 | 2 | |
| Correct phonetic scoring (4 marks, -½ for each mistake) | 4 | 4 | |
| Record mistakes (If req., if not full marks) | 1 | 1 | |
| Correct sum (1 mark) | 1 | 1 | |
| Score recorded on audiogram (-1 if incorrect symbols, -1 if wrong placement) | 2 | 2 | |
| Masking required (excluded if unneeded) | 4 | 4 | |
| Correct masking level (excluded if unneeded) | 6 | 6 | |
| Correct stimuli (-1 for incorrect stimuli) | 2 | 2 | |
| | | | |
| *Slope (2nd point)* | | | |
| Obtain higher level if necessary, if not for appropriate slope level | 2 | 2 | |
| Correct phonetic scoring (4 marks, -½ for each mistake) | 4 | 4 | |
| Record mistakes (If req., if not full marks) (1 mark) | 1 | 1 | |
| Correct sum (1 mark) | 1 | 1 | |
| Score recorded on audiogram (-1 if incorrect symbols, -1 if wrong placement) | 2 | 2 | |
| Masking required (excluded if unneeded) | 4 | 4 | |
| Correct masking level (excluded if unneeded) | 6 | 6 | |
| Correct stimuli (-1 for incorrect stimuli) | 2 | 2 | |
| | | | |
| *Half-peak level* | | | |
| Appropriate level | 2 | 2 | |

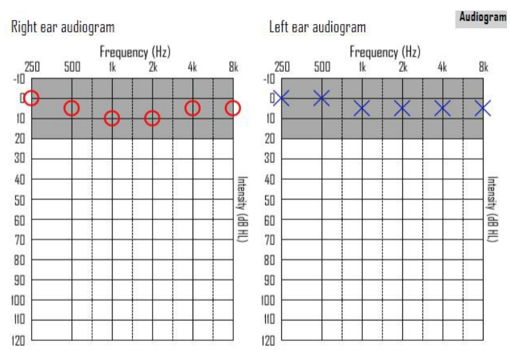
| | | | |
|---|---|---|---|
| Correct phonetic scoring (4 marks, -½ for each mistake) | 4 | 4 | |
| Record mistakes (If req., if not full marks) (1 mark) | 1 | 1 | |
| Correct sum (1 mark) | 1 | 1 | |
| Score recorded on audiogram (-1 if incorrect symbols, -1 if wrong placement) | 2 | 2 | |
| Masking required (excluded if unneeded) | 4 | 4 | |
| Correct masking level (excluded if unneeded) | 6 | 6 | |
| Correct stimuli (-1 for incorrect stimuli) | 2 | 2 | |
| | | | |
| **General** | | | |
| Pausing as required | 1 | 1 | |
| | | | |
| **Estimation** | | | |
| PI Max percentage | 2 | 2 | |
| Half-peak level | 2 | 2 | |
| | | | |
| **Consistency with audiogram** | 1 | 1 | |
| Reason given | 2 | 2 | |
| | | | |
| **Feeding back to client (-1 for not attempting to explain purpose of speech audiometry)** | 2 | | |
| | | | |
| **Efficiency** | | | |
| At least 3 points per ear | 1 | 1 | |
| 5 or less points per ear | 1 | 1 | |
| Less than or equal to10 minutes | 1 | | |
| Less than or equal to 15 minutes | 1 | | |
| Less than or equal to 20 minutes | 1 | | |
| **GRAND TOTAL (OUT OF 165)** | 89 | 76 | |

The items highlighted in Yellow have been excluded from the marking, as practicing with the CAS did not encompass them. The Items highlighted in green are items that do not appear in each case.

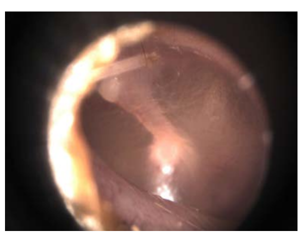
# Appendix H: Speech Audiometry Paper Assessment

### Sub-test 1

This 35 year-old woman has come in to have her hearing checked as part of a job interview. There are no complains of balance problems, tinnitus, or other health concerns. Here is her audiogram.



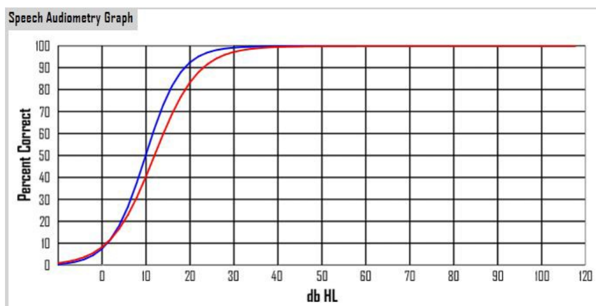

1. Which ear should you test first? (half mark to be given if corresponding to 1)
    a. Right ear
    b. Left ear
    c. Either ear
2. Here are the otoscopy images for her left and right ears respectively:



Considering them both together, which of the following transducers is most recommended?
   a. Supra-Aural headphones
   b. Bone-conductor headphones
   c. Insert headphones
   d. Soundfield

3. Is the hearing loss configuration for her right ear flat or sloping?
   a. Flat
   b. Sloping
   c. Rising
4. Based on 3), what is the initial presentation level for speech in her right ear?
   a. 50 dB
   b. 25 dB
   c. 45 dB
   d. 40 dB
5. Is the hearing loss configuration for her left ear flat or sloping?
   a. Sloping
   b. Rising
   c. Flat
6. Based on 5), what is the initial presentation level for speech in her left ear?
   a. 45 dB
   b. 40 dB
   c. 25 dB
   d. 35 dB
7. Based on your answers for questions 2–6, would you need to put masking into the better ear? (mark accordingly)
   a. Yes, because interaural attenuation is going to affect the other ear
   b. No, because sound will not cross over to the other ear
   c. Yes, because masking is required for flat hearing losses
   d. No, because there is no air bone gap
8. How much masking is required for the initial presentation levels from questions 2–6?
   a. No masking is required
   b. 0 dB HL
   c. –5 dB HL
   d. 5 dB HL
9. Based on your answers for questions 2–6, would you need to put masking into the worse ear? (mark accordingly)
   a. Yes, because interaural attenuation is going to affect the other ear
   b. No, because sound will not cross over to the other ear
   c. Yes, because masking is required for flat hearing losses
   d. No, because there is no air bone gap
10. How much masking is required for the initial presentation levels from questions 2–6?
   a. No masking is required
   b. 0 dB HL
   c. –5 dB HL
   d. 5 dB HL

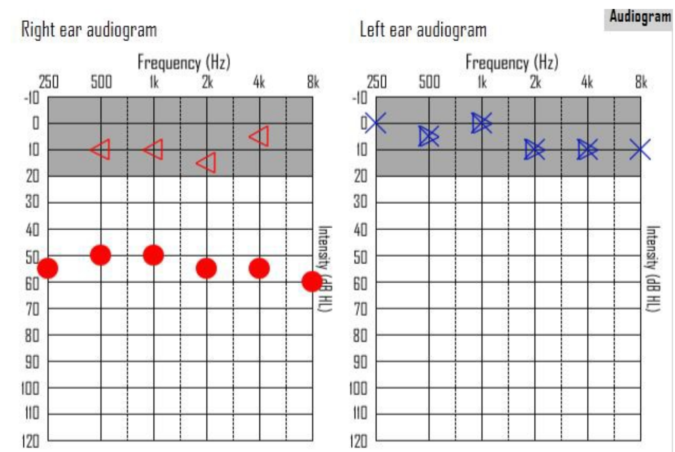
This is her speech audiogram for both ears (Fig C). Use this information to answer questions 11-16.



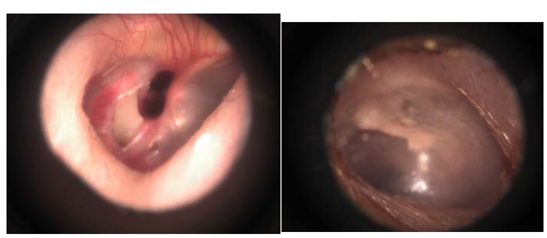

11. What is the PI Max of her Right
    a. 100%
    b. 93%
    c. 83%
    d. 95%
12. What is the Half-peak Level for her Right ear?
    a. 10 dB HL
    b. 12 dB HL
    c. 15 dB HL
    d. 50 dB HL
13. Are the results for her right ear consistent with the audiogram?
    a. Yes because the HPL is within 15 dB of the 1 kHz threshold
    b. Yes because the HPL is within 15 dB of the average loss between 2-4 kHz
    c. No because the HPL is more than 15 dB away from the 1 kHz threshold
    d. No because the HPL is more than 15 dB away from the average loss between 2-4 kHz
14. What is the PI Max of her Left ear?
    a. 93%
    b. 83%
    c. 97%
    d. 100%
15. What is the Half-peak Level for her Left ear?
    a. 8 dB HL
    b. 10 dB HL
    c. 12 dB HL
    d. 50 dB HL
16. Are the results for her left ear consistent with the audiogram?
    a. Yes because the HPL is within 15 dB of the 1 kHz threshold
    b. Yes because the HPL is within 15 dB of the average loss between 2-4 kHz
    c. No because the HPL is more than 15 dB away from the 1 kHz threshold
    d. No because the HPL is more than 15 dB away from the average loss between 2-4 kHz

**Sub-test 2**

This 27 year-old man received a traumatic injury to the right ear when playing around with a chopstick. The ear drum is healing but the hearing still "feels down" in the right ear. He wants to have it checked and see what can be done.

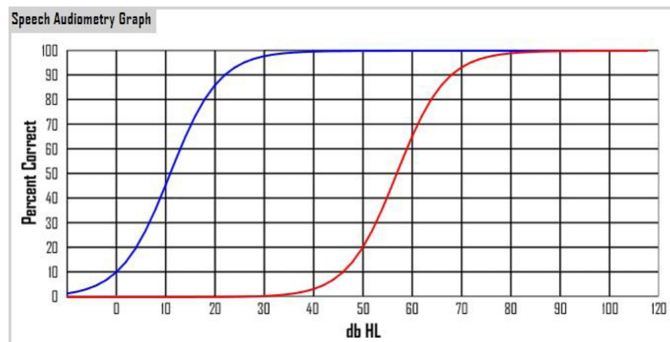



1. Which ear should you test first?
    a. Right ear
    b. Left ear
    c. Either ear
2. Here are the otoscopy images for his right and left ears respectively.



Considering them both together, which of the following transducers is most recommended?
  a. Supra-Aural headphones
  b. Bone-conductor headphones
  c. Insert headphones
  d. Soundfield
3. Is the hearing loss configuration for his right ear flat or sloping?
  a. Flat
  b. Sloping
  c. Rising
4. Based on 3), what is the initial presentation level for speech in his right ear?
  a. 50 dB
  b. 65 dB
  c. 80 dB
  d. 90 dB
5. Is the hearing loss configuration for his left ear flat or sloping?
  a. Sloping
  b. Rising
  c. Flat
6. Based on 5), what is the initial presentation level for speech in his left ear?
  a. 50 dB
  b. 30 dB
  c. 20 dB
  d. 45 dB
7. Based on your answers for questions 2-6, do you need to put masking into the better ear? (mark accordingly)
  a. Yes, because interaural attenuation is going to affect the other ear
  b. No, because sound will not cross over to the other ear
  c. Yes, because we have normal hearing in the better ear
  d. No, because there is no air bone gap
8. How much masking is required for the initial presentation levels from questions 2-6?
  a. No masking is required
  b. 60 dB HL
  c. 40 dB HL
  d. 50 dB HL
9. Based on your answers for questions 2-6, would you need to put masking into the worse ear? (mark accordingly)
  a. Yes, because interaural attenuation is going to affect the other ear
  b. No, because sound will not cross over to the other ear
  c. Yes, because masking is required for flat hearing losses
  d. No, because there is no air bone gap
10. How much masking is required for the initial presentation levels from questions 2-6?
  a. No masking is required
  b. 0 dB HL
  c. -5 dB HL
  d. 5 dB HL

This is his speech audiogram for both ears (Fig C). Use this information to answer questions 11–16.




11. What is the PI Max of his Right ear?
    a. 97%
    b. 93%
    c. 100%
    d. 95%
12. What is the Half-peak Level for his Right ear?
    a. 56 dB HL
    b. 50 dB HL
    c. 12 dB HL
    d. 80 dB HL
13. Are the results for his right ear consistent with the audiogram?
    a. No because the HPL is more than 15 dB away from the average loss between 2–4 kHz
    b. Yes because the HPL is within 15 dB of the 1 kHz threshold
    c. No because the HPL is more than 15 dB away from the 1 kHz threshold
    d. Yes because the HPL is within 15 dB of the average loss between 2–4 kHz
14. What is the PI Max of his Left ear?
    a. 93%
    b. 83%
    c. 97%
    d. 100%
15. What is the Half-peak Level for his Left ear?
    a. 15 dB HL
    b. 50 dB HL
    c. 12 dB HL
    d. 10 dB HL
16. Are the results for his left ear consistent with the audiogram?
    a. Yes because the HPL is within 15 dB of the average loss between 2–4 kHz
    b. No because the HPL is more than 15 dB away from the 1 kHz threshold
    c. Yes because the HPL is within 15 dB of the 1 kHz threshold
    d. No because the HPL is more than 15 dB away from the average loss between 2–4 kHz

## Sub-test 3

1. What is the interaural attenuation for supra-aural headphones for speech audiometry?
2. What is the formula for initial presentation level for a flat audiogram?
3. Here is a word list. Complete the scoring and total score.

| Word | Response | Score |
|---|---|---|
| Lean | Leap | |
| Hag | Heap | |
| Bed | Dance | |
| Sews | Ruse | |
| Cop | Cot | |
| Root | Root | |
| Pick | Pop | |
| Maim | Main | |
| Toss | Tom | |
| Dial | Dial | |
| **STIMULUS** | 50 dB HL in left ear | |
| **MASKING** | 10 dB HL in right ear | |
| **TOTAL** | | |

4. What symbol is used to mark this list on the audiogram?
5. If this was the first list for this ear, would you increase or decrease the presentation level?
6. What is the formula for deciding if you need to mask?
7. What is the formula for deciding how much masking to put in?
8. What type of masking noise is used for speech audiometry?
9. How do you check if the half-peak level is consistent if the audiogram is sloping?
10. If the PI Max is 90%, what is the percentage for its half-peak level?
11. How many points should we obtain per ear?
12. What are the three purposes of speech audiometry?

# Appendix I: Feedback Paper Assessment

**Case 1: A 53 yr old woman comes in for a hearing check. She is worried things sound a bit muffled, and she needs to turn the TV up louder than usual. Otoscopy reveals a wall of dry wax in both ears.**

1. What transducers do you select to test her hearing?

A Insert earphones
B Bone conductor
C Supra aural headphones
D None, use the sound-field speaker

2. What ear will you test first?

A Left
B Right
C Both as I'm testing in the sound-field
D I wouldn't test this woman's hearing until she's had the wax removed

3. What frequency do you begin testing at?

A 1000 Hz
B 250 Hz
C 10 kHz
D 2000 Hz
E 8000 Hz

4. At what level do you first present the tone?

A -10 dB
B 20 dB
C 30 dB
D 50 dB
E 90 dB

5. She doesn't respond to the initial tone presentation. What do you do?

A Discontinue testing, she's obviously not hearing the tone due to the wax

B Present the tone again at the same level for 1-2 seconds

C Present the tone again at the same level, holding the presentation button down until she responds

D Increase the presentation level by 15-20 dB and present the tone for 1-2 seconds

E Increase the presentation level by 15-20 dB, present the tone, holding the presentation button down until she responds

6. She responds: how do you find her hearing threshold?

A drop in 10 dB steps until she no longer responds, then return to the initial presentation level and repeat dropping in 10 dB steps until she no longer responds. Her threshold is the level at which she no longer responds on two descending runs.

B drop in 10 dB steps until she no longer responds, go up in 5 dB steps until she responds again. Repeat dropping 10 dB and increasing 5 dB. Her threshold is where she responds two out of three times on an ascending run.

C drop in 5 dB steps until she no longer responds, go up in 10 dB steps until she responds again. Repeat dropping 5 dB and increasing 10 dB. Her threshold is where she responds two out of three times on an ascending run.

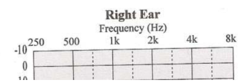
D start at -10 dB and increase in 5 dB steps until she responds. Repeat. Her threshold is where she responds on two ascending runs.

**Case 2. A 48yr old man has come in for a hearing check. He has been feeling dizzy and nauseous at times, with a sound like the roaring ocean in his right ear when he gets dizzy. He feels his hearing hasn't been that great on the right side since he's been having the dizzy spells. Otoscopy reveals clear ear canals and healthy looking eardrums on both sides.**

1. What transducers do you choose? ________________________________________

2. What ear do you begin with? ____________________________________________

3. Write the frequencies in the order you test them:
1) ______ 2) ______ 3) ______ 4) ______ 5) ______ 6) ______




4. At what level do you initially present the tone when testing his right ear? ______

5. At 2000Hz in the right ear, you obtain the following responses:

| Presentation level: | Response: |
|---|---|
| 50 dB HL | Yes |
| 40 dB HL | Yes |

What will the next presentation level be? _________

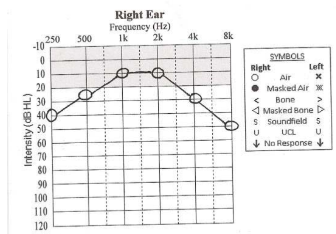
6. At 500 Hz in the left ear, you obtain the following responses:

| Presentation level: | Response |
|---|---|
| 30 dB HL | Yes |
| 20dB HL | Yes |
| 10 dB HL | Yes |
| 0 dB HL | No |

What will the next presentation level be? _________

7. His threshold at 4000Hz in the right ear is 45dB HL. At what level do you start presenting the tone at 8000Hz? ___________

**Case 3. This person is having their hearing tested with insert earphones. So far you have obtained the following results:**




1. At what frequencies do you need to obtain bone conduction thresholds?

A All of them

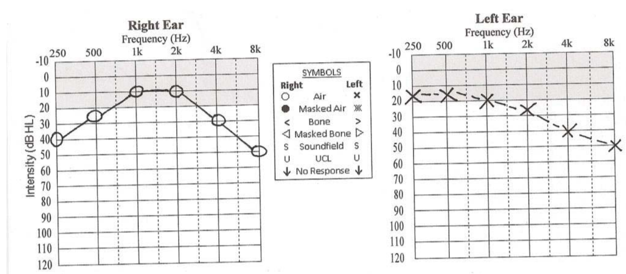
B Only those frequencies where there is a hearing loss (where AC thresholds are greater than 20 dB HL)

C At 500 Hz, 1 kHz, 2 kHz and 4 kHz

D At 500 Hz and 4 kHz

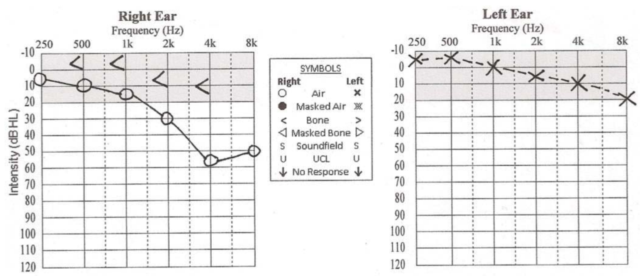
E None of them

2. Below are the results for the left ear. At what frequencies do you need to obtain bone conduction thresholds? ______________________________________________________




3. Is there a need to obtain air conduction thresholds for any of the inter-octave frequencies? If so, for which ear at which frequencies? _________________________

**Case 4. This person is having their hearing tested with supra-aural headphones. So far the following results have been obtained:**




1. Which bone conduction thresholds need to be masked?

A All of them
B 500 & 1000 Hz
C 1, 2 & 4 kHz
D 2 & 4 kHz
E None of them

2. You decide to mask the BC threshold at 2 kHz. What ear do you play the masking noise into?

A Both ears via the headphones
B Both ears via the sound-field speaker
C Both ears via the bone conductor
C Left ear only via the headphones
D Right ear only via the headphones

3. At what level do you start presenting the masking noise?

A 5 dB HL
B 15 dB HL
C 20 dB HL
D 30 dB HL
E 40 dB HL

4. At what level do you start presenting the test tone?

A 5 dB HL
B 15 dB HL
C 20 dB HL
D 30 dB HL
E 40 dB HL

5. The patient came back for a re-test the following week. You are part way through masking their BC threshold at 2 kHz. So far you have these responses using the plateau method:

| Masking noise level: | Test signal level: | Response: | Action: |
|---|---|---|---|
| 10 dB EM | 10 dB HL | Yes | masking noise up 10 dB |
| → 20 dB EM | 10 dB HL | No | test signal up 5 dB |
| → 20 dB EM | 15 dB HL | Yes | ??? |

What is your next move?

A Mark 15 dB HL as the masked BC threshold for 2 kHz

B Leave settings as they are and present the test tone a second time
C Increase the test tone by 5 dB
D Decrease the test tone by 5 dB
E Increase the masking level by 10 dB

6. There are some mistakes on the audiogram. What should also have been done?

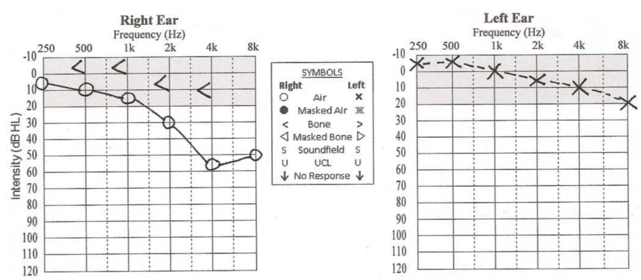
A Obtain a masked air conduction threshold for the right ear at 4 kHz and obtain unmasked bone conduction thresholds for the left ear

B Obtain an air conduction threshold at the 3 kHz inter-octave frequency on the right ear and obtain unmasked bone conduction thresholds for the left ear

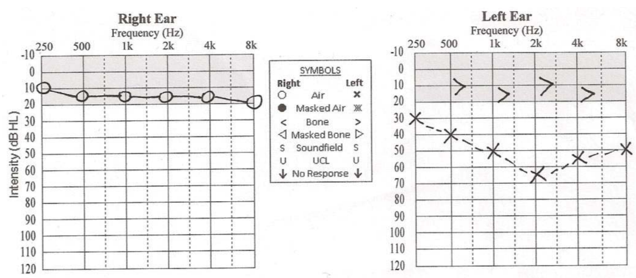
C Obtain a masked air conduction threshold for the right ear at 4 kHz and obtain an unmasked bone conduction threshold at 8000 Hz for the same ear

D Obtain an air conduction threshold at the 3 kHz inter-octave frequency on the right ear and obtain a masked air conduction threshold at 4 kHz for the same ear

E Obtain unmasked bone conduction thresholds for the left ear and obtain an unmasked bone conduction threshold at 8000 Hz for the right ear



**Case 5. This person is having their hearing tested with insert earphones. So far you have obtained the following results:**




1. Which bone conduction thresholds need to be masked? ___________________________

2. You decide to mask the bone conduction threshold at 1 kHz. What ear do you play the masking noise into? _____________________________

3. What would the initial masking level be at this frequency? _________________

4. What ear do you present the test tone to? ________________________

5. What level do you start presenting the test tone at? _____________________

6. At a re-check 2 weeks later, you obtain the following responses using the plateau method when masking their BC threshold at 500 Hz.

| Masking noise level: | Test signal level: | Response: | Action: |
|---|---|---|---|
| 40 dB EM | 15 dB HL | Yes | masking noise up 10 dB |
| → 50 dB EM | 15 dB HL | Yes | masking noise up 10 dB |
| → 60 dB EM | 15 dB HL | Yes | ??? |

What is your next move? __________________________________________________

7. Is there a need to obtain any masked air conduction thresholds? If so, for which ear at which frequencies? ______________________________________

8. Is there a need to obtain any air conduction thresholds for any of the inter-octave frequencies? If so, for which ear at which frequencies? ________________________