PHYSICAL-VIRTUAL PATIENT SIMULATORS: BRINGING TANGIBLE HUMANITY TO SIMULATED PATIENTS

by

SALAM DAHER
M.S. University of Central Florida, 2015
M.S. University of Florida, 2006
B.S. Lebanese American University, 2004

A dissertation submitted in partial fulfilment of the requirements
for the degree of Doctor of Philosophy
in the School of Modeling, Simulation, and Training
in the College of Engineering and Computer Science
at the University of Central Florida
Orlando, Florida

Fall Term
2018

Major Professor: Gregory F. Welch

ABSTRACT

In lieu of real patients, healthcare educators frequently use simulated patients. Simulated patients can be realized in physical form, such as mannequins and trained human actors, or virtual form, such as via computer graphics presented on two-dimensional screens or head-mounted displays. Each of these alone has its strengths and weaknesses. I introduce a new class of physical-virtual patient (PVP) simulators that combine strengths of both forms by combining the flexibility and richness of virtual patients with tangible characteristics of a human-shaped physical form that can also exhibit a range of multi-sensory cues, including visual cues (e.g., capillary refill and facial expressions), auditory cues (e.g., verbal responses and heart sounds), and tactile cues (e.g., localized temperature and pulse). This novel combination of integrated capabilities can improve patient simulation outcomes.

In my Ph.D. work I focus on three primary areas of related research. First, I describe the realization of the technology for PVPs and results from two user-studies to evaluate the importance of dynamic visuals and human-shaped physical form in terms of perception, behavior, cognition, emotions, and learning. Second, I present a general method to numerically evaluate the compatibility of any simulator-scenario pair in terms of importance and fidelity of cues. This method has the potential to make logistical, economic, and educational impacts on the choices of utilizing existing simulators. Finally, I describe a method for increasing human perception of simulated humans by exposing participants to the simulated human taking part in a short, engaging conversation prior to the simulation.

To my parents Amal and William for all their continuous sacrifices, care, and support that shaped the person I am today and helped me go higher and further during every step of the way.

To my brother Bassam for always being there for me, especially when it matters the most.

To my grandmother Teta Nabiha for your wisdom, love, and prayers. I am proud of being your first grand daughter to graduate with a PhD.

To my second family away-from-home, my Orlando friends for sharing my special moments and offering their physical and psychological support.

ACKNOWLEDGMENTS

This work would not have been possible without the financial support primarily by National Science Foundation (NSF) Award #1564065, C̈HS: Medium: Physical-Virtual Patient Bed for Healthcare Training and Assessment," Program Director Dr. Ephraim P. Glinert, and in part by the Office of Naval Research (ONR) Code 30 under Dr. Peter Squire, Program Officer (ONR awards N00014-14-1-0248 and N00014-12-1-1003).

I deeply appreciate the scholarships and fellowships from the National Center for Women and Technology (NCWIT), the I/ITSEC committee for the RADM Fred Lewis scholarship, the Link Foundation, and the UCF Modeling and Simulation graduate program. I also appreciate Florida Hospital for their support of my advisor Prof. Welch via their Endowed Chair in Healthcare Simulation.

I am especially indebted to my advisor and chair of my committee Prof. Gregory Welch for caring, advising, and supporting me and the team in every aspect of the research. I am inspired by your creativity, innovation, energy, collaboration, and strong work ethics. Thank you for taking the time to and having no reservations in directly or indirectly passing all of these positive qualities to people around you. As the proverb says: "He gives twice who gives quickly", and I double my thanks to you for always giving quickly!

I am grateful to my committee Prof. Laura Gonzalez, Prof. Juan Cendan, and Prof. Michael Proctor for your time, expertise, and advice. Similarly, I am also grateful for the feedback from Prof. Gerd Bruder, Prof. Andrew Raij, Prof. Arjun Nagendran, and Prof. Charlie Hughes from the SREAL lab, and for Prof. Mindi Andreson, and Prof. Desiree Diaz from College of Nursing at UCF, and for Prof. Jeremy Bailenson from Stanford University.

It takes a village to do the work described in this dissertation. First, I acknowledge my team-mates and co-authors on many papers (not in any specific order): Jason Hochreiter, Ryan Schubert, Nahal Norouzi, Kangsoo Kim, and Myungho Lee. I acknowledge behind the scenes work of Barbara Lee, Katie Ingraham, and Eric Imperiale at the SREAL lab, and the overall support from the Institute of Simulation and Training and the school of Modeling and Simulation faculty and staff at University of Central Florida. I would like to mention by name Prof. Randall Shumaker, Prof. Paul Wiegand, Dr. Sabrina Gordon, Kirsten Seitz, and Naya Ramirez. I also acknowledge the behind the scenes work of people from College of Nursing, I would like to mention by name: Syretta Spears, Jorge Nieves, and Chris Upchurch.

Furthermore, I recognize the efforts of my teachers, instructors, professors, and anyone who contributed in planting the seeds of knowledge throughout my life, this includes a very long list of people from College des Soeurs des Saints Coeurs, the Lebanese American University, University of Florida, and University of Central Florida.

A special thank you to a unique friend, family member, and researcher, Dr. Joseph Najem for the insightful research discussions, perspective, and motivation.

A huge thank you for my friends who have been there for me (in no specific order): Dr. Shainna Ali Borenstein, Dr. Rachel Eyma, Jo Bohn, Colleen O'Sullivan, Cindy Vincent, Rita Carnero, Veronica Lavenworth, Belda Stack, Dr. Linda Rosa-Lugo, Suspira Tiouat, Yana Maxwell, Kim Moore, Marie Hewitt, Shehan Sirigampola, and Ravi Melaram.

Last but not least, nobody has been more important to me than my family Amal, William, Bassam, and Teta Nabiha. You are my rock, my strength, and my inspiration. I love you.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

LIST OF ACRONYMS AND ABBREVIATIONS

COL = Co-located

HMD = Head Mounted Display

IPVA = Intelligent Physical-Virtual Agent

PV = Physical-Virtual

PVP = Physical-Virtual Patient

PVPBed = Physical-Virtual Patient Bed

PVPH or PVPHead = Physical-Virtual Patient Head

SAR = Spatial Augmented Reality

SEP = Separated

VH = Virtual Human

CHAPTER 1: INTRODUCTION

Historically, simulation is broadly defined as "an imitation of some real thing, state of affairs, or process for the practice of skills" [2]. Simulation is used in various domains including high-hazard ones to reduce cost, uncertainty, and the risk of undesirable outcomes such as the possibility of a loss or a catastrophe, and to increase safety by providing a learning and practice environment under controlled circumstances [3, 4]. Sometimes the world does not offer a replica for practice but a simulator could. The need to practice in a safe environment is common in fields where mistakes are costly and sometimes deadly, such as in aviation, military, and medical fields [3]. It takes about 10,000 hours of practice for people to become masters in their fields [5]. Simulation allows for virtually unlimited practice necessary for healthcare simulation.

Healthcare simulations can be said to have four main purposes – education, assessment, research, and health system integration in facilitating patient safety [6]. Miller's framework for medical assessment starts with "knows" (knowledge), then "knows how" (competence), then "shows how" (performance), and ends with "does" (action) [7]. When learning new skills, visual observation alone does not lead to proficiency. Repeated deliberate practice is indispensably necessary for skill development, improvement, and maintenance [8].

Using simulation for training improves healthcare provider's self-efficacy, competence, and actual operational performance in clinical settings [3].

1.1 Simulation Fidelity

Fidelity is the accurate representation through cues and stimuli from the perspective of the participant [9]. The term "fidelity" is overloaded and can have different meanings depending on the discipline. In healthcare "high fidelity" often refers to mannequins that provides a good control of physiology even though their appearance is static. High-fidelity simulation refers to structured student learning experiences with the use of a technologically advanced computerized mannequin [10]. In computer graphics, the term "high fidelity" refers to realistic visuals that are (or are very close to being) photo-realistic, starting with static graphics and extending to when they are animated regardless of where these graphics are displayed; it is usually assumed they are displayed on a flat screen or a head mounted display.

1.1.1 Importance of Simulation Fidelity

There is evidence that high-fidelity human simulation might positively impact a high level of cognitive and clinical skills acquisition [11], and has a potential to support and affect the development of clinical judgment in nursing students [12]. It is desired to increase fidelity and realism in simulation training as high-fidelity simulation is associated with effective learning. The two most highlighted features of high-fidelity simulation in journal articles are providing feedback, and repetitive practice [13]. In game-based simulations, more realism increases attention and retention [14]. According to the National Council State Boards of Nursing realism during simulation is required [10]. The closer the realism is to clinical reality, the easier it is for participants to engage in the simulation scenario [15].

1.1.2 Patient Simulation Fidelity Space

The healthcare simulation fidelity space is divided into facilities, clinical, and patient fidelity. The patient fidelity dimension encompasses representations of interactions with all or part of a patient, such as communicating or performing a procedure, and takes into account fidelity of anatomy and physiology [9]. In this work, I focus on the the simulated patient fidelity and realism in terms of shape (i.e. physical manifestation) and appearance (i.e. visual manifestation) in a healthcare simulation setting.

1.1.3 Patient Simulation Fidelity Challenges

Compared to simulating machines, simulating humans seems much more challenging since humans are very complex not only in terms of physiology, but also in terms of appearance, behavior, communication, social interactions, culture, and psychology. Another challenge is people's expectations of what is considered "more realistic" when it comes to simulating humans, and how they feel about that increase in realism when the simulator gets closer to a human [16]. There is a need to create more effective simulated patients. One challenge is technological, and another is psychological. From a technology perspective, simulated patients should be able to realistically represent subtle and not so subtle signs and symptoms in terms of visuals, sounds, and haptics. Real humans have these channels dynamic, interactive and all in one location, for example if a real human raises their arm, the shape and appearance of their arm move together. In 3D virtual environments, co-locating haptic cues with visuals cues greatly improves performance [17]. Co-location between visual and haptic sensory modalities are linked to eliciting sense of presence [18].

From a psychological perspective, how you perceive the human matters for training. For example in a simulation setting, participants are well aware that this is a training and they are interacting with a simulated human, therefore they may already be biased into behaving differently. It is possible to counter the initial bias of participants by exposing them to a short engaging conversation prior to interacting with simulated humans [19, 20]. A realistic simulator that looks like, feels like, and behaves like a human may affect the perception. People's expectations increase the more the simulator is similar to humans [21]

1.2 Current Healthcare Education

In a clinical setting, the patient thought of as having a 100% fidelity (e.g. they are actually sick) and healthcare providers respond back with what would correspond to a 100% natural behavior (e.g. they do not pretend to do a procedure), in other words they do not need to modify their natural behavior. In a simulation setting, a simulated patient has a fidelity of less than or equal to 100%. When the Simulated Patient's inherent capability to show essential medical cues is limited, a mitigation (i.e. change to compensate for a simulator deficiency) for these capabilities is necessary. Depending on the mitigation the behavior of the participant in the simulation may be modified which could affect their performance.

1.2.1 Didactic vs. Experiential Learning

For the purpose of this dissertation, I distinguish between didactic learning and experiential training for nurses and doctors. My focus is on the experiential training. In didactic learning, students learn about a human without directly experiencing a human (e.g. lecture, books, images, case studies). In a non-experiential classroom-based learning setting, what represents a human can be effective without having to look like and act like a full human (e.g., illustrated image, orange to practice injections, part task trainer), the interface can be free (e.g. pinch zoom to certain areas) allowing the content can be more forgiving. This type of learning is common and appropriate early on when learners are novice as "experiencing" a whole human may be overwhelming at that stage. As learners become more experienced, they get more exposed to "experiencing" a human (using simulated patients) before before actually practicing on real patients.

When experiencing a human, learners' sensitivity to what the simulated human looks like and how the they behave increases. The (simulated) human *is* the interface (e.g. speak directly to the patient as opposed to pressing a button, move around the patient as opposed to moving the mouse...etc). If experiencing a human is done right, we can expect more natural behaviors from the learners. I am interested in the area where learners experience a human as opposed to learning about a human. Below is brief review of potential existing simulators in my domain of interest:

1.2.2 Current Types of Patient Simulators

Currently healthcare educators use standardized patients (real humans in an acting role), mannequins (e.g. SimMan3G, pediatric HAL, CAE simulators), computer-based simulations (e.g. Shadow Health, i-Human, Second Life), and virtual and augmented reality for patient simulation (e.g. CAE Vimedix) [2, 22]. Each of these simulators has its strengths and weaknesses. The next section briefly describes each type and examines their pros and cons. I used the following tags in brackets to structure the presentation of different types patient simulators in terms of their ability to **[#Physiology]** portray physiology (e.g., temperature, pulse, heart rate...etc), **[#Shape]** physically occupy volume (e.g., have a tangible material), **[#Appearance]** display dynamic visuals (e.g., change appearance, facial expressions), **[#Intelligence]** intelligence (e.g., how the responses are controlled), and **[#Logistics]** logistics (e.g., availability, cost),

1.2.2.1 Standardized Patients (real humans)

Standardized patients (SPs) (i.e. living humans serving as patients), are trained to act like a patient in a consistent and standardized manner [23]. They are frequently used for to develop history-taking skills and perform physical exam [24]. **[#Physiology]** SPs cannot exhibit certain symptoms such as changing physiology or dynamic appearance at will. For example, a healthy SP may not always be able to portray useful diagnostic cues for example they cannot manually control their blood pressure or temperature of specific body parts or change how their pupils react to light to simulate abnormal findings. **[#Shape]** SPs are real humans, so they naturally occupy volume and can be physically touched. **[#Appearance]** SPs can dynamically display certain visuals that are naturally associated with their inherent appearance. For example, SPs can show facial expressions and move their body, but they may not be able to portray other appearances such as showing abnormal pupil dilation, or ptosis. **[#Intelligence]** SPs are real humans who are trained to respond in a manner authentic to the scenario, and at the same time they can seamlessly adapt to any unexpected questions. **[#Logistics]** It is not always easy to schedule SPs [25], and even more difficult to recruit children and infants as SPs due to child labor laws [26, 27]. Using mannequins or software-based simulations is more feasible compared to recruiting children.

1.2.2.2 Mannequins (physical simulators)

The use of mannequins originated in the 1960s with Resusci Anne for cardiopulmonary resuscitation (CPR) mannequin [28, 29]. In the following years, mannequin-based simulators, such as SimOne, CASE, GAS, and others, were developed to train anesthesiologists and increase patient safety [30, 31]. There was no need for Mannequins to represent dynamic visuals (such as facial expressions) for proper training of anesthesiologists; in general, patients under anesthesia do not converse with their healthcare providers. **[#Physiology]** Modern Mannequins can exhibit certain physiology changes that are controllable to accommodate different medical conditions. Changes such as temperature of localized skin are more challenging and may require mitigation before the simulation (e.g., place an ice bag on the mannequin's hands). **[#Shape]** Modern Mannequins occupy physical volume that is comparable to a real human and can be touched. **[#Appearance]** Modern Mannequins with a static appearance, much like those first used by anesthesiologists, are still commonly used in simulation across multiple healthcare domains (e.g. medical-surgical, trauma, critical care, and obstetrics), even for scenarios that inherently include dynamic appearance changes with real human patients, such as facial expressions, gestures, and abnormal visual findings (e.g., ptosis, capillary refill), which cannot be easily portrayed on a mannequin [32, 33]. Moulage is a technique used to increase the fidelity of the appearance, but it is typically static (e.g., stickers or silicone wounds), and can be time consuming to create and setup. **[#Intelligence]** While the physiology can be pre-programmed and controlled by a computer, usually the verbal responses are controlled using a human in the loop (typically a trainer or facilitator). Similarly to an SP, the set of responses can be standardized and can allow for deviations from a script if needed in order to adapt to any unexpected questions or actions. **[#Logistics]** Mannequins can be expensive, especially the high-fidelity ones. They could offer an improved availability compared to SPs, but they are still typically operated by a human that also requires scheduling.

1.2.2.3 Computer-Based (Virtual Simulators)

Computer-based patient simulation involving a Virtual Humans (VH) such as Shadow Health [34], CliniSpace [35], and iHuman [36] are typically displayed on devices such as monitors, television (TV) screens, and projectors. **[#Physiology]** A virtual human can portray physiology that can be controlled by a computer such as heart rate, respiration rate, O2 saturation. It is more difficult to displays scenario-driven pulse, and localized temperature cues on a computer monitor or a TV screen. **[#Shape]** The flat displays where computer based patients are represented do not occupy a physical three-dimensional (3D) volume besides what the display medium occupies. **[#Appearance]** Virtual humans in computer-based simulation can visually represent changes in appearance and behavior in a compelling rich manner allowing them to simulate symptoms that are otherwise hard to simulate. Virtual humans can be animated to change their body language, show dynamic signs and symptoms (e.g., pupil reaction to light). **[#Intelligence]** Virtual humans' control ranges from being fully automated (e.g., artificial intelligence) to having a human in the loop triggering live answers or pre-recorded answers. **[#Logistics]** Computer-based simulators offer the potential of flexible availability where different people can use the software at the same time minimizing scheduling concerns.

1.2.2.4 *Augmented Reality (Physical-Virtual Simulators)*

Mannequins can be augmented using Augmented Reality (AR) technology. AR combines dynamic visuals with the user's physical environment, for example a physical mannequin can be supplemented with dynamic imagery. AR can combine features from both physical and virtual simulators **[#Physiology]** AR can portray physiology through the physical and/or the virtual component. **[#Shape]** The physical component of AR can occupy volume and have a tangible material. **[#Appearance]** as with virtual humans in computer simulations, the virtual component of AR can display realistic dynamic visuals (e.g., change appearance, facial expressions) **[#Intelligence]** as with computer simulations, the simulation can be controlled in an automated way or using a human in the loop. **[#Logistics]** The logistics depend on the type of choice of control and hardware for scheduling and cost.

One way to achieve AR is through head-mounted displays (HMDs) [37–39], another way is through projected imagery [32,40,41]. Modern HMDs (e.g. HoloLens [42], Meta [43]) are limited in their field of view, are heavy [44], and it can be hard to have multiple users for the same scenario at the same time due to difficulty in synchronizing graphics across various devices. In addition, the HMD's augmented graphics can occlude the hands of users unless significant hand tracking development is implemented [45–47]. Front-projection AR has a different issue in the opposite direction. The shadow of the user's hands can occlude the projection [48]. Rear-projection AR can solve both multi-user problem and the occlusion problem, but requires enough physical space for the projectors placed behind any augmented objects.

1.2.3 Challenge

In healthcare simulation, it is often a challenge to represent certain symptoms “just right,” making them easier or harder to detect compared to a real human. Presenting symptoms inaccurately in the simulation environment can reinforce incorrect behaviors and possibly lead to diagnostic and treatment mistakes [49]. For example, participants may not recognize a symptom (e.g. mottled skin) on their own if the simulator cannot naturally represent it. To mitigate such simulator limitations, the simulation facilitator may have to cue the participant [50], e.g., explicitly say “mottled skin.” However, this can provide participants with a hint (or cue) towards the correct diagnosis, which could result in reaching the correct diagnosis faster and more easily than it would have been otherwise. If the facilitator chooses not provide a hint, then the simulation lacks needed information that is supposed to be available. This might unintentionally make reaching the correct diagnosis more difficult.

1.3 Research Overview

The first objective is to present a new type of patient simulator, a Physical-Virtual Patient (PVP) simulator that combines strengths of other patient simulators to support realistic multi-sensory cues all in one location for participants to directly experience a human (simulated patient). I used rear-projection AR for creating a Physical-Virtual Patient, a new type of patient simulator that occupies volume and can dynamically change its appearance, also it integrates other sensory cues such as localized temperature and haptic feedback. The PVP can display subtle symptoms in different senses (visual, haptic, sounds), a variety of scenarios (e.g., burns, sepsis, stroke), and a diversity of patients shapes (e.g. child, adult, male, female, obese, amputee) and appearances (e.g. eyes, varied skin, and hair color and other attributes), with the ability to quickly change clinical cases to expand the users’ exposure to cases. The term "Physical" is used because the patient has a physical human-shaped shell. The term "Virtual" is used because the dynamic imagery can be projected on the human-shaped shell. The shell can present pulse and localized temperature. Touching the shell can trigger automated and non-automated responses (e.g capillary refill, show patient's teeth and eyes, speech, and facial expressions.) The PVP allows the direct experience of a patient interaction so healthcare providers can directly recognize certain signs and symptoms first-hand (e.g facial expressions, subtle changes in appearance, and localized temperature.). This is covered in chapter 3.

The Second objective is to use the PVP to manipulate certain cues (i.e. dynamic visuals and physical shape) during the simulation in order to learn how these cues affect the simulation outcomes, and to provide an insight about their importance for healthcare patient simulation. This is covered in chapters 4, 5,and 6.

The third objective is to propose a method that can be applied to any simulator (not limited to PVPs) to objectively evaluate the compatibility of a simulation by matching the importance of scenario cues to simulator's capabilities in representing these cues. This method applies to any and every healthcare patient simulator. The method produces scores for matching, utilization, and underutilization (waste) of simulator resources. Being able to objectively evaluate the compatibility of scenario-simulator pairs which can have logistical, economic, and educational impacts.This is covered in chapter 7.

The Fourth is an even more general method that applies to simulating humans (not limited to healthcare) where participants can be primed before the simulation by witnessing a short engaging conversation to improve their perception of social presence. This is covered in chapter 8.

1.3.1 Thesis Statements

the following section list the thesis statements [TS] and briefly describe the experiments that support them.

[TS 1] Combining Physical and Virtual Patient Simulation Can Improve Training

[TS 1A] Replacing a static mannequin head with realistic dynamic visuals on a matching physical form can increase social presence, increase engagement, and improve learning.

[TS 1B] Separating the realistic dynamic visuals from the matching physical form decreases the perception of realism.

[TS 2] Fidelity of the Physical Shape Can Affect the Cognitive Load

[TS 3] Cue Importance and Fidelity Ratings Can be Used to Quantitatively Compare Scenario-Simulator Pairs

[TS 4] Social Presence Priming Can be Used to Improve a Participant's Perception of a Simulated Human

1.3.2 Support for Thesis Statements

[TS 1A] is supported in Chapter 4 by comparing a PVP head and a mannequin in a between-participant design study where nursing students conducted a neurological assessment on a patient with stroke. **[TS 1B]** is supported in Chapter 5 by using the PVP to compare co-located visuals (COL) with a separated visuals (SEP) in a mixed design study. The Co-location portion had a within-subject design where graduate nurse practitioner students assessed 2 simulated children, one with sepsis and one with signs of child abuse. **[TS 2]** is supported in Chapter 5 by using the PVP to compare a patient with physical human-shaped form (PVP) with a patient that has a flat shaped form (FVP) as part of a mixed design study. The form of the simulator was a between-subject design part where graduate nurse practitioner students assessed 2 children, one with sepsis and one with signs of child abuse. It is also supported in Chapter 6 where we evaluate the cognitive load during a task that requires touching dots on a head while counting backward by seven. The between-subject design has 2 independent variables: Shape and display method. The shape can have a physical form, flat form, or no shape. The imagery can be projected with Spatial Augmented Reality (SAR) or displayed on a Head-Mounted Display (HMD). **[TS 3]** supported by chapter 7 where a method is described to evaluate scenario-simulator pairs for any simulator based on the importance of the scenario cues, and the fidelity of these cues as portrayed by the simulator. The method finds a match score, a utilization score, and a waste score. **[TS 4]** is supported in chapter 8 in a between-subject design study where, prior to interacting with a virtual human, participants were primed in one condition with a short engaging conversation with a virtual human compared to not being primed.

1.3.3 Notable Findings

This sections contains noteworthy results [NR] form the experiments.

[NR 1] When the realistic dynamic visuals are separated from the haptic cues, the physical shape where the haptic cues are displayed affects the perception what represents the patient. Compared to flat haptic cues, separate realistic dynamic visuals are perceived more representative of the patient. Compared to Human shaped haptic cues, separate realistic dynamic visuals are perceived as equally representative of the patient. These are supported in Chapter5 .

[NR 2] People stand by the side of the patient when it has a human physical form, than when the patient is flat . This is supported in Chapter 5.

[NR 3] People's sensitivity to the realism of realistic dynamic visuals increases when displayed on a Human Physical Form compared to when the realistic dynamic visuals are displayed on a flat form. This is supported in Chapter 5.

[NR 4] When the task requires touch, representing the visuals using Spatial Augments reality is preferred and easier to use than using a Head Mounted Displays. . This is supported in Chapter 6

[NR 5] A physical surface with geometry that matches the 3D model is preferred over a mismatching one. This is supported in Chapter 6.

1.4 Terminology

A *Simulation Environment* is the physical setting where the activities may take place, inclusive of the people and equipment that form part of the simulation experience [51]. The Simulation Environment can be real (clinical), simulated, or mixed.

A *Simulated Patient* is an individual who is trained to portray a real patient in order to simulate a set of symptoms or problems used for healthcare education, evaluation, and research [51]. The "individual" is not limited to a real human, in fact an individual can also be a mannequin, or a computer based patient. Simulated patients can be real humans, simulated, or mixed. A Simulated Patient has an inherent capability to portray certain symptoms of medical cues. When the Simulated Patient's inherent capability to show essential medical cues is limited, a mitigation for these capabilities is necessary. In a simulated setting, a simulated patient has a fidelity of less than or equal to 100%.

Simulated Equipment or *Tools* (such as a bed, stretcher, IV pump, monitor, and medication...etc) can be real, simulated, or mixed. Also, these tools can offer an inherent capability or a mitigated capability to support the simulation.

Mitigation refers to an interference from the Trainer or Facilitator before, during, or after the simulation in order to compensate for sub-optimal or missing simulator inherent capability.

A *Participant* in healthcare simulation is a person who engages in a simulation activity for the purpose of gaining or demonstrating mastery of knowledge, skills, and/or attitudes of professional practice [52]. When the Participant interacts with the simulated patient with less than 100% fidelity, they may have to change their natural behavior to accommodate for the simulate patient's technical capabilities. In this case, the Participant may exhibit a Natural Behavior and a Modified Behavior.

A *Scenario* is a deliberately designed simulation experience that provides Participants with an opportunity to meet identified objectives [53]. Each Scenario can include Participants (such as medical students, nursing students...etc), Simulated Patients, and tools.

A *Trainer* or *Facilitator* or *Controller* observes what happens during the simulation Scenario and possibly provides feedback. The Trainer can have a role before the simulation (e.g. setup, mitigation to make up for a lack of capability of the simulator), during the simulation (e.g. real time changes, mitigation for a limited simulated patient capability, live comments...etc) and/or after the simulation (e.g. debriefing or after action review).

Facilities Fidelity encompasses representations of the clinical equipment and environment, such as the instruments, the monitors, and the environment in which clinical activities or patient encounters take place [9].

Clinical Scenario Fidelity encompasses representations related to the script and progression of a scenario, and situational complexity such as team and family dynamics. It includes the educators' involvement and the debriefing [9].

Patient Fidelity encompasses representations of interactions with all or part of a patient's body, such as communicating or performing a procedure, and takes into account fidelity of anatomy and physiology [9], in this work the focus is on the shape and overall appearance of the patient.

Co-location of cues is the simultaneous temporal and spatial integration of visual cues to the physical form. It is not enough to have visual and haptic cues near each other, they have to be registered to each other.

Physical-Virtual Patient is a mixed reality Simulated Patient that supports high fidelity visuals cues co-located with physical form and haptic feedback such as pulse and localized temperature.

Social Presence is one's sense of being socially connected with the other. *Co-Presence* is one of several dimensions that make up social presence, it is one's sense of the other person's presence [54].

1.5 Dissertation Overview

Below is a quick guide for chapters that are more relevant to readers with healthcare background, and readers with engineering background. Readers with an Engineering background may be interested in chapters 3, 6, 7, and 9. Readers with a Healthcare background may be interested in chapters 4, 5, 8, 9, and the scenarios in Appendix B, C, and D. Readers interested in related works and history may appreciate Chapter 2 and Appendix A

Chapter 2 presents related work for the dissertation. Details about the history of human simulators are in Appendix A.

Chapter 3 describes the development of the Physical-Virtual Patient Head and Physical-Virtual Patient Bed and the scenarios developed to support the experiments. Details about scenario development are in Appendix B, C, and D.

Chapter 4 describes a study in support of **[TS 1A]**. In this between-participants study, undergraduate nursing students were asked to perform a neurological exam on a simulated patient that has a stroke. We compared two experimental conditions: (1) adding a PVP head near mannequin's body vs. (2) using a full body mannequin only without adding the PVP head. We measured perceived realism, social presence, communication, urgency, engagement, and learning

Chapter 5 describes a study in support of **[TS 1B]** and **[TS 2]**. In this mixed design study, graduate nurse practitioner students interacted with 2 patients (Child with sepsis, and child with signs of abuse). We used a Physical-Virtual Patient where the shape of the shell changes between human shaped and flat. The location of the dynamic visual cues changes between co-located and separated.We measured perception of realism, preferences, cognitive load, and behaviors.

Chapter 6 describes a study in support of **[TS 2]**. In this within-participant design, participants from the UCF community were asked to perform a task that requires touching a Physical-Virtual Head while counting backward by seven. The physical shape (Physical shape vs Flat shape) and the display methods (Projected Visuals vs. Head Mounted Display) changed between conditions. We measured the cognitive load, preference, and performance.

Chapter 7 in support of **[TS 3]**. This chapter present a method to objectively evaluate the compatibility of any scenario-simulator pair. Every Simulation has a Scenario and a Simulator. Every Simulator provides a set of Cues with varying fidelity depending on its capabilities. Every Scenario depends on certain Cues. Some cues are more important than others. For every scenario one can develop a cue importance vector comprising a normalized importance score for each cue. For every simulator one can develop a cue fidelity vector comprising a normalized fidelity score for each cue. For every scenario-simulator combination one can develop a simulation fitness score that is formed by dot product of cue importance vector and cue fidelity vector, and used to compare simulator-scenario pairs in an objective way.

Chapter 8 in support of **[TS 4]**. In this between participants study, participants from the UCF community played a 20-questions game with the Virtual Human. One group was exposed to a brief social priming before the simulation (i.e. short conversation between two virtual humans). For the priming we used a human size monitor with a virtual human talking to another virtual human on a separate screen. The second group was not exposed to the priming. The simulation consisted of one virtual human without adding any social priming that involves another virtual human. We measured social presence.

Chapter 9 presents a summary of the results.

CHAPTER 2: RELATED WORK

2.1 Healthcare Setting

This section compares the real world clinical setting with the healthcare simulation setting.

2.1.1 Real World Clinical Setting

In the real world clinical setting such as a clinic, a hospital, or an emergency department the environment includes real equipment or tools (such as a bed, stretcher, IV pump, monitor, and medication...etc), real healthcare providers (such as physicians, nurses, techs...etc), and real patients. Scenarios or medical cases include interaction between a healthcare provider, real patients, and tools. In the real world clinical setting, patients and the tools have a 100% fidelity and healthcare providers behave 100% naturally. Optionally, you can have an observer such as an attending physician or a preceptor that observes inside or outside the environment observing the scenario and possibly providing feedback (See Figure 2.1)

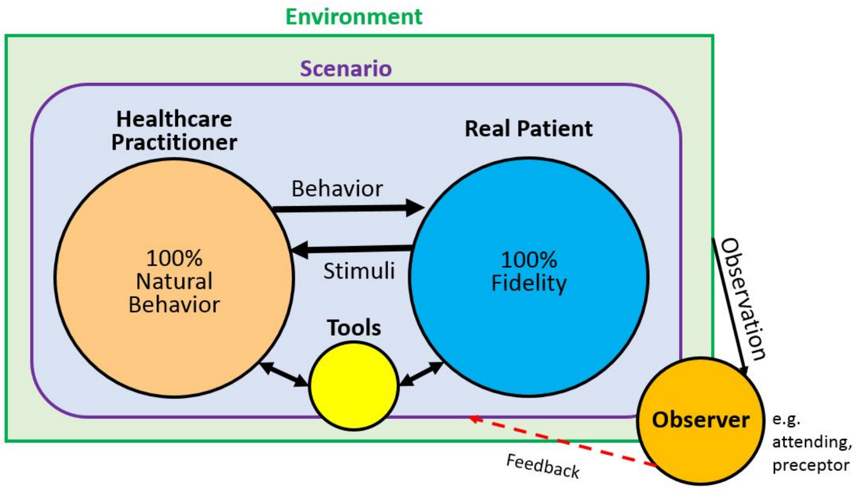


Figure 2.1: Real World Clinical Setting

2.1.2 Healthcare Simulation Setting

A *Simulation Environment* is the setting where the activities may take place, inclusive of the people and equipment that form part of the simulation experience [51]. The Simulation Environment can be clinical, simulated, or mixed. The Simulation Environment includes Participants(s), Simulated Patient(s), Tool(s), Scenario(s), and Trainer(s) (See definitions in chapter1) (See Figure 2.2).

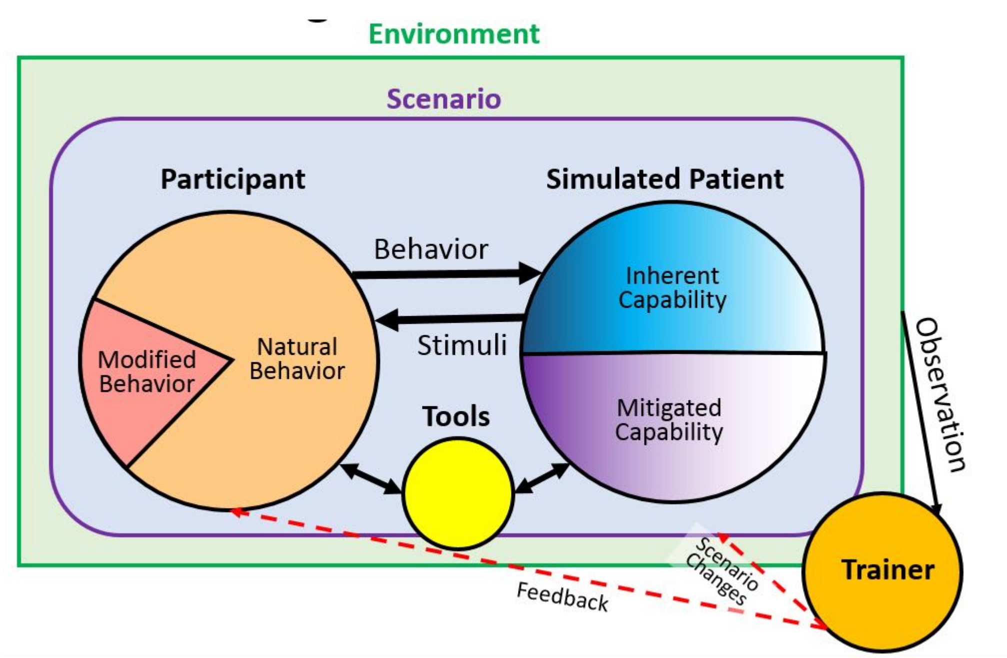


Figure 2.2: Healthcare Simulation Setting

2.2 Patient Simulators

Healthcare simulators range from part task trainers to full body simulators. Part task trainers are simple anatomical models of body parts in a normal or diseased state [2]. Full body simulators represent a full patient. Typically, healthcare educators use standardized patients, mannequins, and computer-based simulation [55]. Mannequins and computer-based patient simulators are more readily available and versatile, making up the largest portion of practically used patient simulators. This section explores different types of patient simulators

2.2.1 Standardized Patients

Standardized patients are individuals trained to simulate a set of symptoms to portray a real patient [51].

2.2.1.1 Description and History

Standardized Patients is the umbrella term for both simulated patients who are well trained people to simulate a patient's illness in a standardized way, and actual patients trained to present his/her illness in a standardized way [56]. In 1964 Dr. Howard Barrows originally conceived and developed the concept of Standardized Patients, at first it was not widely accepted but became increasingly popular and now recognized across the continuum of medicine [57]. Dr. Barrows was responsible for acquiring patients for the Board Examinations in Neurology and Psychiatry; he realized that the use of real patients was physically straining and harmful to the nature of examinations due to patients getting tired or inconsistent in their responses. The terminology changed from what was once called Programmed Patients, to Simulated Patients, to Standardized Patients. There are other subtle differences in terminology for simulated patients and roles between Role Playing, Pseudo Patients, Practical Instructors, Patient Instructors, and Subjects but scripted actors in educational films are wrongly referred to as simulated patients. The term Standardized Patient was coined by Geoff Norman as a replacement to Simulated Patients to highlight the importance of providing a standardized patient problem that does not vary from student to student, also the term "standardized" does not reveal to the students if the patient is simulated or a real patient [56]. The importance of the fact that the patient is standardized is very important for the certification process of healthcare practitioners (See Figure 2.3.)

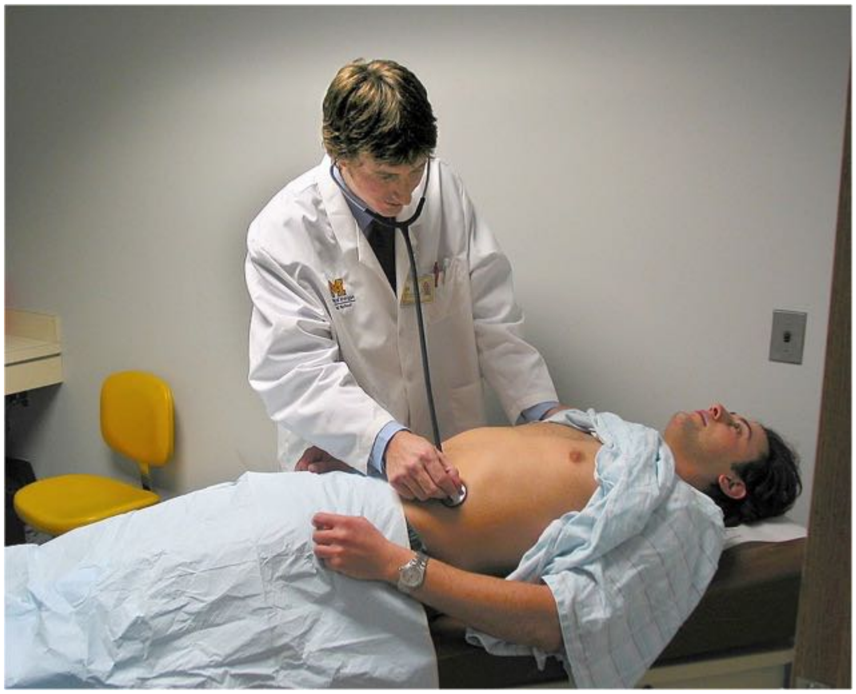


Figure 2.3: Standardized Patient.
Source: By University of Michigan Medical School Information Services (Standardized-Patient-Program-examining-the-abdomen) [CC BY 2.0 (https://creativecommons.org/licenses/by/2.0)], via Wikimedia Commons

2.2.1.2 Pros and Cons

There are advantages to use Standardized Patients (SP) over real patients beyond the availability of real patients since unlike real patients, SP can be available at any setting and time where they can present the same problem for students. The use of SP for educational settings avoids mistreating real patients as SPs are prepared when students perform inadequately and are ready to be used as teaching tool. There are situations such as emergencies and sensitive medical conditions where the students are not allowed to work with real patients, SPs allow students to get practice in these situations. The SP can be manipulated for educational purposes, one powerful technique is called “time in - time out” where the simulation pauses to give live feedback for students that they would not get with a real patient due to the nature of what is being discussed, then the SP resumes the scenario after the feedback. Also time for progression of illness or improvement can be manipulated with a SP and not with a real patient. The process to train SPs can be as short as 2-3 hours for people who have never done it before [56]. In addition to the physical findings that can be simulated, and compared to Mannequins and Computer-Based simulators (in the next sections), human actors including SPs have the ability to feel and react to human touch, to react physically, verbally and non-verbally to stimuli and to improvise and adapt to new situations outside the domain when needed. Compared to Computer-Based Simulators, people occupy physical space and provide a capability for interaction without giving the students hints to what is on the next menu of choices. While SP are very valuable and are still frequently used for training today [23], they are paid per hour and scheduling them is a factor that has to be taken into consideration (when compared to a computerized simulator that is always available). SPs can be hard to recruit for certain population such as babies, or children [26, 27, 58]. Also there can be issues such as people’s unwillingness to simulate certain uncomfortable procedures. Despite all good intentions, SPs cannot immediately and voluntarily change certain appearance aspects or physiology at will (e.g. temperature, blood pressure, facial droop), which can make certain medical conditions hard to simulate (e.g. stroke, sepsis). Needless to say, it is not ethical to inflict harm on a SP to simulate and illness.

2.2.2 Mannequins

Mannequins are electro-mechanical robotic life-sized human-like patient simulators [51] that physically occupy volume and can change physiology to simulate a wide range of medical conditions, but they have a static appearance (e.g., facial expressions, skin color, inability to move) [59, 60]. When a full body is not required, task trainers—models that represent a part or region of the human body, such as an arm or an abdomen—can be effective training tools.

2.2.2.1 Description

Mannequin Based Simulators are sometimes referred to as low, mid, and high fidelity simulators. The term Robotic Mannequin can be used for physical mannequins that have mechanical and or electronic parts to simulate physiology of a patient, typically those are mid-fidelity and high fidelity mannequins, also known as “High Fidelity Simulators.” The High Fidelity refers to the ability to customize scenarios by varying the simulator’s physiology (See Figure 2.4.)

Figure 2.4: Mannequin.
Source: By U.S. Navy photo by Photographer’s Mate 3rd Class Rebecca J. Moat [Public domain], via Wikimedia Commons

2.2.2.2 *Examples and History*

In the 1960s Laerdal developed Resusci Anne, the first doll designed for mouth to mouth resuscitation after a young girl drowned in the Seine. In 1968 was the beginning of Harvey, a cardiac training mannequin at University of Miami. Dr. Michael Gordon named the mannequin after his mentor Dr. Proctor Harvey. Students can assess heart beats, pulses and respiration. Harvey was progressively improved over the next three decades and is currently available from Laerdal. In the 2000s Laerdal purchased Medical Plastics Laboratories and entered the arena of full-scale computerized simulation with SimMan3G [57]. In 1995 Gaumard started with CodeBlue III and in 2004 they developed HAL the 1st tether-less human patient simulator [61]. In 2001 METI developed ECS (Emergency Care Simulator), followed by iStan in 2007 and by METIman in 2009 [57, 62]. CAE Healthcare, formerly METI, developed the trauma simulator Caesar in 2014 and Fidelis Lucina, a birthing simulator in 2015 [63, 64].(See Appendix A for a more detailed history of patient simulators.)

2.2.2.3 *Pros and Cons*

Some of the advantages of using a "high fidelity simulator" (high fidelity referring to a mannequin with physiology) include no direct risk to patients [65], and standardization of learning by guaranteeing the same experience for all students and skills can be tailored and repeated. Mannequins have what is sometimes referred to as *cosmetic reality* [12] which means the same size as a human being, and occupy physical space (e.g proportionate limbs, scale). Compared to a computer based simulation (with a user interface), the students don't need to learn a new computer user interface (e.g. menu, buttons, mouse click...etc) to navigate the software, in other words this reduces the chance of accidentally providing the students with cues or hints (i.e. multiple choice in a menu item) for actions they were supposed to recognize on their own. This is related to Miller's pyramid for medical assessment [7] where we want the learners to "show how" as opposed to "know how". High Fidelity Mannequins can easily change their physiology such as pulse, temperature (with pre-simulation mitigation), blood pressure... etc and some can even sweat due to their electro-mechanical components. Repeat-ability of pre-programmed scenarios is important to keep consistency between scenarios.

While Mannequins occupy the same space and volume as real humans, their appearance is static making them incapable of changing skin color, facial expressions, eye gaze and other non-verbal communication [66]. Some people use medical moulage to change the appearance but that is also static and needs preparation ahead of time, for example you cannot see animations such as dynamic wound progression through time or movement. Also, 94% of Mannequins (including body parts) represent the average white patient, and 6% represent a black patient [67]. There is not much diversity in existing mannequins and body parts based on skin color, ethnicity, gender, body form and other deviation from the majority normal (See Appendix A), for example physical mannequins do not have a patient with tattoos and piercing which may make it harder to assess certain body areas, resulting in less practice and less readiness when faced with similar presentations in real life. Similarly, Mannequins have a static face which means participants do not see any facial expressions whether regarding pain or emotional expressions such as smile, frown, surprise ... etc [68].

2.2.3 Computer Based Patients

Computer-based simulation, typically presented on interactive flat displays with graphical and text output, can include virtual humans (e.g., patients, nurses, doctors), virtual reality task trainers, and immersive virtual reality simulation [69].

2.2.3.1 Description

Computer based patients are also known as web-based patients, or virtual patient. They include interactive dynamic computer graphics rendered visuals, but they are typically displayed on a flat computer screen without a physical shape or haptic physiology cues [34, 58, 70].(See Figure **??**.) While virtual humans can be used as a replacement for real humans in certain situations, people usually do not treat an virtual humans exactly as they would treat a real human. For instance, in studies where medical students interacted with either an virtual humans or a real human pretending to have the same symptoms, participants appeared less engaged, sincere, and interested, and had a poorer attitude towards the virtual [71]. In an experiment with a computer graphics representation of a virtual human, its advice was more rarely sought out compared to a physically present robot [72]. One explanation for this phenomenon is the low sense of presence, social presence, and co-presence induced by the virtual human.

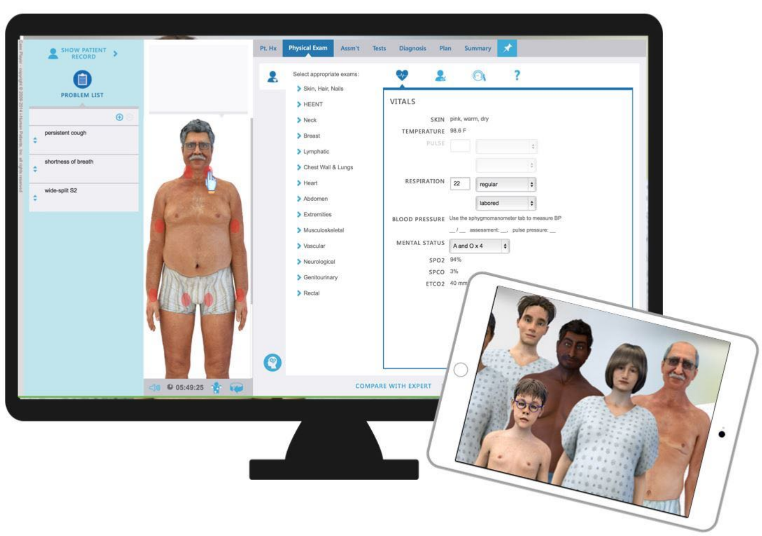


Figure 2.5: Computer Based Patient.
Source: The sample image on the screen is taken from iHuman software.

2.2.3.2 Examples

Some examples of virtual patient simulators include Clinispace [73] which is a 3D immersive computer application for practicing patient care with interactive devices and virtual patients. Learners can create their own patients and scenarios while their performance is tracked. Similarly Shadow Health [74] allows students to practice communicating with and examining the 3D virtual patients. iHuman provides virtual patients that simulate a complete medical patient encounter [75] . Anatomage is a virtual table that allows students to explore the human body in 3D and to slice it at any angle and to filter by type such as skeleton, organs, circulation system [76].

2.2.3.3 Pros and Cons

Virtual patient simulations are consistently associated with higher learning outcomes compared to other educational methods [77], and virtual patient assessment shows high acceptance among medical students [78]. For example, learners can improve their pain observation skills by training with virtual patients [79], and it can be easier for learners to detect pain in a virtual human than in a high-fidelity, facially expressive humanoid robot patient simulator [80]. Computer-based virtual patient can have substantial emotional effects on medical students [81]. Virtual patients can have the capacity for exhibiting a high level of interactivity and realism [82]. Computer-based patients allow the instructors to repeat pre-programmed scenarios where physiology changes can be simulated (pulse, temperature, BP, sweat...etc) and the characters are capable of changing their appearance instantaneously. The changes can be as animation (for example dynamic wound bleeding or healing) or as general appearance to represent different people with different skin color, age, gender, and ethnicity. Virtual patients can easily support showing non-verbal communication such as eye gaze, facial expressions, and gestures. Also, multiple instances of the virtual patients can run independently in parallel on multiple machines to allow for more people to practice simultaneously.

Students may need to learn a new interface to navigate the software and sometimes the interface may provide cues or hints for what to do next through multiple choice question, menu choices, information that is explicitly visible. These “hints” are not available in the real world which makes it hard to determine how much of the success using the simulator comes from the student knowing what to do next on their own and how much comes from the student receiving a hint from the system. Display size affects the sense of being together [83]. Typically virtual simulators do NOT occupy volume on their own (besides the flat display they are on). While this can be considered good, it can also be bad as occupying space may affect the student’s behavior around the patient.

2.2.4 Augmented and Mixed Reality Simulators

Mixed Reality simulators are hybrid simulators that involve physical and virtual simulated components, in this case full body patients or parts a patient’s body. Depending on the need and the design, they can have the pros (and sometimes cons) of both physical and virtual worlds. Augmented reality (AR) supplements the real world with virtual objects, such that virtual objects appear to coexist in the same space as the real world [84, 85]. It combines the physical world and dynamic visuals, for example a physical mannequin dynamic imagery can be added on top of a mannequin. There are multiple ways to achieve AR, one way is through head-mounted displays (HMDs)(e.g. HoloLens [42], Meta [43]), and another way is through projected imagery [86, 87].

2.2.4.1 Description

This is a relatively new area in medical simulation where the physical and virtual modalities are combined in the same space. A new class of *physical-virtual* simulators is capable of presenting many of these physical and haptic cues *co-located* together and that can quickly change scenarios to convey a diversity of patients and variations of medical conditions. One type of Physical-Virtual (PV) patient simulators is Shader-Lamps based virtual patients that combines a front- or back-projected human-shaped surface with a projector to provide computer-generated visual feedback to observers, thus allowing medical students to conduct ophthalmic exams in an interactive training experience [88].

2.2.4.2 Examples

Researchers at University of Florida developed a clinical breast exam virtual human patient where the student receives feedback on their palpation pressure of the breast from the virtual patient [89]. Similarly Drexel developed a robotic bottom to facilitate intimate exams [90]. Other researchers explored haptic palpation for medical students in virtual environments [91]. Some researchers have supplemented static mannequins with videos and audio on a nearby screen to make up for the lack of dynamic visuals [92] One example of commercial AR simulator that uses HMDs is CAE Vimedix [37]. The *shader-lamps virtual patient* ophthalmic examination is an example of PV simulator where the patient is manifested physically with a front projection of his head on a Styrofoam head [93–96], including for ophthalmic [88, 97, 98] or general neurological assessment [99, 100], offering medical students standardized experiences to interview, examine, and diagnose virtual patients [77]. Physical-virtual agents can support automated touch sensing with integrated graphical response through rear-mounted IR lights and cameras [44, 99, 100] (See Chapter 3).

2.2.4.3 Pros and Cons

Modern HMDs are limited in their field of view, are heavy [44], and it can be hard to synchronize multiple users for the same experience. The HMD's augmented graphics can also occlude the hands of users unless significant hand tracking development is implemented [48]. Imagery from front projected AR can be occluded with objects between the projector than the target (e.g. shadow of one's own hand). Rear-projection AR easily allows for multiple users to participate in the same simulation without the occlusion problem or the field of view limitation, but it requires enough physical space for the projectors placed behind any augmented objects.

2.3 Fidelity and Realism for Healthcare Simulation

In healthcare simulation, *fidelity* is the degree of accuracy to which a simulation represents a given frame of reality in terms of cues, stimuli, and permissible interactions [9]. This can refer to the realism of different aspects of the facilities, clinical scenario, or patient. In particular, each simulated patient has a physical shape, appearance, and physiology of a different fidelity.

2.3.1 Usage of the Term

The term *fidelity* is often used inconsistently in the literature [50], for example in the medical field the term "high fidelity simulator" is used to represent a physical simulator with a physiology model that can be programmed for different conditions and scenarios without taking into consideration other factors such as appearance of the patient, or possibly newer "higher" fidelity simulated patients. A "high-fidelity" mannequin in healthcare simulation typically provides good control over the physiology, even if its visuals are static. In healthcare, the terms low-, medium-, and high-fidelity are often used to denote the physiology of the patient and possibly its physical shape, while in computer graphics fidelity generally refers to visual appearance.

In 1992, Lane and Alluisi identified over 22 different definitions for simulation fidelity [101]. In some definitions, fidelity is the extent to which the appearance and behavior of the simulator simulation match the appearance and behavior of the simulated system [102]. There is a distinction between engineering fidelity and psychological fidelity which was addressed first by Miller [103]. Psychological fidelity or functional fidelity is the degree to which the skills in the real task are captured in the simulated task [102]. Engineering fidelity is the degree to which the device replicates the physical characteristics of the real task [102]. In engineering, specifically in flight simulation, the fidelity is divided into equipment fidelity, environment fidelity, and perceptual fidelity; while these dimensions are suitable for aviation they are not suitable for representing patients [9, 104]. Previously engineering fidelity has been used to describe healthcare simulation fidelity, but maybe it should be according to anatomy and physiology [102, 105].

Curtis et. al looked into fidelity in medical education and classified it into physical, functional, and psychological fidelity. Physical fidelity involves the degree in which physical characteristics are represented. Functional Fidelity is related to the action the simulation system initiates and how much it matches its real world counterpart. Psychological Fidelity involves the degree in which the trainee is engaged in tasks that generate experience and actions [106]. This classification does not distinguish between simulated patient, scenario, and environment.

Tun et. al. define fidelity for healthcare as the degree of accuracy to which a simulation represents a given frame of reality in terms of cues and stimuli, and permissible interactions, it does not necessarily require faithful replication of reality, but the accurate representation through **cues and stimuli** from the **perspective of the participant**. They define the healthcare fidelity space into *Facilities Fidelity*, *Clincal Scenario Fidelity*, and *Patient Fidelity*. The Facilities Fidelity dimension encompasses representations of the clinical equipment and environment, such as the instruments, the monitors, and the environment in which clinical activities or patient encounters take place. The Clinical Scenario Fidelity dimension encompasses representations relating to the script and progression of a scenario, and situational complexity such as team and family dynamics. It includes the educators' involvement and the debriefing. The Patient Fidelity dimension encompasses representations of interactions with all or part of a patient, such as communicating or performing a procedure, and takes into account fidelity of anatomy and physiology [9]. Also, in addition to these dimensions, trainers can affect the fidelity of the simulation by using "deception" for example they may prime the participants by giving the illusion that the standardized patient is real

2.3.2 Importance of Fidelity and Realism

Why should we care about high-fidelity? High-fidelity human simulation might positively impact a high level of cognitive and clinical skills acquisition [11], and has a potential to support and affect the development of clinical judgment in nursing students [12] and improve performance [107].

One way to increase fidelity is to increase *realism*, which could include the patient's appearance, physiology, shape, or other cues.

There is evidence that more realistic video games increased attention and retention [14]. According to the International Nursing Association for Clinical Simulation and Learning (INACSL) standards, creating a perception of realism using various types of fidelity is recommended [108]. According to the National Council State Boards of Nursing, equipment fidelity, psychological fidelity, and environment realism where the simulation takes place are all required, for example, mannequins cannot show emotional stress and standardized patients are not actually sick [10]. The closer the realism is to clinical reality, the easier it is for participants to engage in the simulation scenario [15]. Also, in most circumstances, it is a good simulation practice to refrain from disrupting the frame of realism to stay true to the realism the simulation is supposed to deliver [109].

2.3.2.1 Multi-sensory Realism

Our perception of realism starts from the basic sensory stimuli: visual, auditory, tactile, olfactory, and gustatory. The brain combines information from multiple senses and generates what looks like a seamless perception of the external world. Ding et. al. found that adding tactile, olfactory, and auditory cues to the visuals of a virtual environment increase the sense of presence and enhance memory for objects in that environment. Increasing only the level of visual detail did not result in an increase in the user's sense of presence or memory of the environment [110].

Before doctors used to rely on CT scans, PET scans, MRI and laboratory tests, they only had their eyes, ears, nose, and touch. Touch is used for both diagnosis and for its therapeutic effects as patients "need to feel cared for, and touch is a ritual of caring" [111].

A simulated patient's *physicality* includes the fidelity with which it occupies a physical space in terms of its size, volume, position, and other factors along with its ability to sense and cause changes in that environment [112]. Previous research has shown that it is possible to prompt users to behave more realistically by increasing the physicality of virtual humans, which can be achieved in a variety of ways [113]. In particular, interpersonal touch is of vital importance in healthcare simulation, both in terms of allowing a healthcare providers to assess or comfort a patient by touch and in terms of the added realism and presence benefits provided by simulators with physical affordances [95, 114, 115].

2.3.2.2 Separation of Sensory Cues

When cues from different senses are incoherent that can result in illusions such as "ventriloquism effect" [116]. The ventriloquist effect refers to the perception that the sound comes from a different direction than the true direction due to the influence of the visual stimuli. Also, active touch feedback can change the visual perception of a slanted surface [117]. Congedo et. al. found that spatial separation of visual and touch cues promote a visual dominance when judging rotation angles of operating a handle [118].

2.3.2.3 Co-location of Sensory Cues

Swapp, Pawar and Loscos performed experiments co-locating haptic cues with visuals cues in a desktop setting, they found that co-located haptics with force feedback in 3D virtual environments greatly improved performance [17]. Viviana and Reyes analyzed the effects of co-location between visual and haptic sensory modalities and found that haptic cues are essential to elicit sense of presence, but adding visuals and audio had a slight additional increase in sense of presence. They note that the more the virtual experience fits reality the higher the expectations of participants are [18]. Researchers explored co-locating virtual reality over a mannequin for select surgical procedures to assess effectiveness, usability, and acceptability [119]. Some high-fidelity mannequins can accurately portray co-located cues such as heart and lung sounds, pupil response, and limited emotional expressions, leading to more realistic healthcare simulations [120]; however, they are generally limited in their ability to show dynamic imagery and thus may not be suitable for portraying certain abnormal visual symptoms. Similarly, some mannequins lack realism in the area of skin temperature [121].

2.4 Human Senses and Technology

2.4.1 Sight: Visual Displays

Computer graphics deals with the creation and manipulation of image content, it can refer to 3D representations of a scene in real time and as offline generation of images. The building blocks of computer graphics involve domains such as photography, mathematics, geometry, mechanics, psychology, and design. Computer graphics is used in areas such as entertainment, data visualization, graphics design, marketing, and virtual reality applications [122]. The Computer graphics can be sent to any type of display to help in creating an immersive virtual, mixed, or augmented reality.

In Virtual Reality (VR) systems display hardware can be a desktop VR, projective VR, or immersive VR. Desktop VR is the simplest solution for a VR setup that includes a standard computer screen with additional components such as stereoscopic rendering or tracking. Such setups are usually referred to as fish-tankVR [123]. Projective VR is an extension of desktop VR that involves projection on flat, parametric, or non-parametric surfaces [86, 122, 124]. The goal of Immersive VR is to place the user in a virtual world by blocking out cues from the real environment (e.g. CAVE [125], head mounted display) [122]. In augmented reality (AR), 3D virtual objects are integrated into the real environment [126]. *Spatial Augmented Reality* is a type of projective virtual (or augmented) reality where the images are integrated directly to augment the user's environment as opposed to simply being in the visual field (e.g. no need to wear a head mounted display to see the imagery) [87]. *See-Through Augmented Reality* is when the user wears a head mounted display to get 3D objects superimposed over the real world. It can be optical or video-see through [127]

2.4.2 Touch: Haptic Displays

Haptic displays can be passive or active, they can be classified according to the kinematic design (serial kinematics, or parallel kinematics), or according to the actuation principle, in other words the generated force can be electromagnetic, pneumatic or hydraulic, electrorheological, magnetorheological, shape memory alloys, electroactive polymers, or piezoelectric. Haptic Displays can also be classified based on which sensory pathway is addressed, either tactile (thermal), kinesthetic, or pain receptors [122]. Examples of haptic displays include desktop systems, ground and wall mounted displays, portable systems, and tactile displays such as pressure displays, vibration interfaces, pin and bar arrays, electro-tactile displays, and thermal displays.

2.4.3 Hearing: Auditory Displays

In human factors field the auditory displays are classified as warning and alarm signals, three dimensional displays, and speech displays [128]. Audio can be displayed using headphones, mono,stereo, and loud speaker system, wave field synthesis, and sound focusing [122]. In a VR environment, auditory rendering can use pre-recorded sound samples, real-time synthesizing of sound, or triggering of sound based on time, position or collision [122]

2.5 Human Information Processing

The human information processing model starts with stimuli as input and ends with a response as output [128]. The information enters auditory, visual, tactile, and other processors, then it is passed to the working memory [129]. The information or stimuli pass through three stages: perceptual stage, cognitive stage, and action stage [128] (See Figures 2.6, 2.7, and 2.8.) In the perceptual stage participants make sense of the the stimuli, that's where they perceive them as realistic or not, that information is passed to the cognitive stage where they pass by the working memory before initiating an action which results in a response that we use to measure performance. A deterioration in the quality of stimuli leads to reduction in performance. The cognitive load theory suggests that instruction should be designed in a way where the material matches the learner's previous knowledge, should minimize the extraneous cognitive load (the unnecessary type of load on the memory), and the learners should engage in a process that evokes germane cognitive load (the type of memory that leads to learning) [130–132].

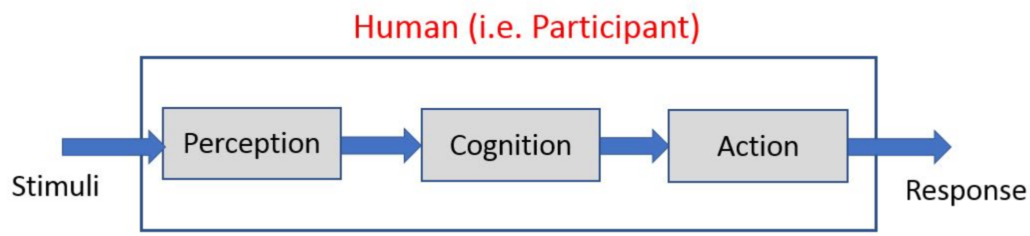


Figure 2.6: Human Information Processing

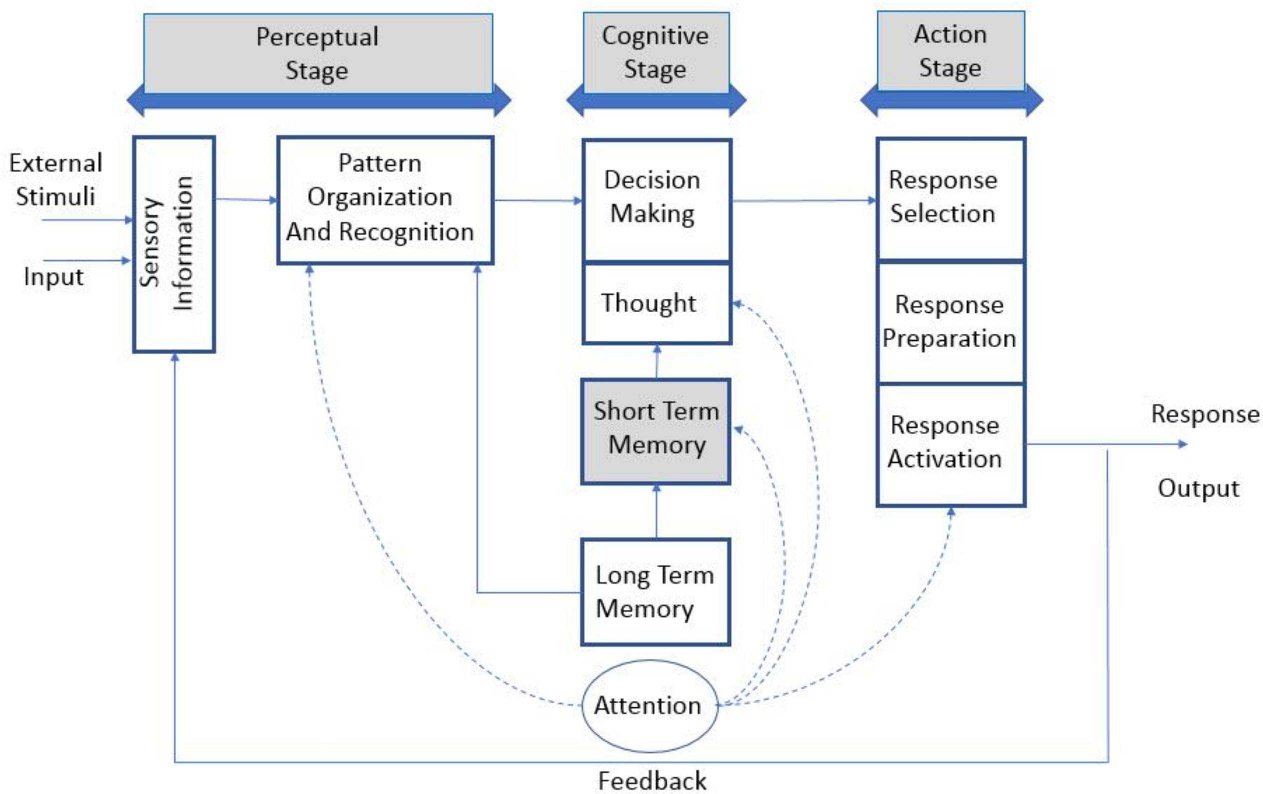


Figure 2.7: Human Information Processing in More Details

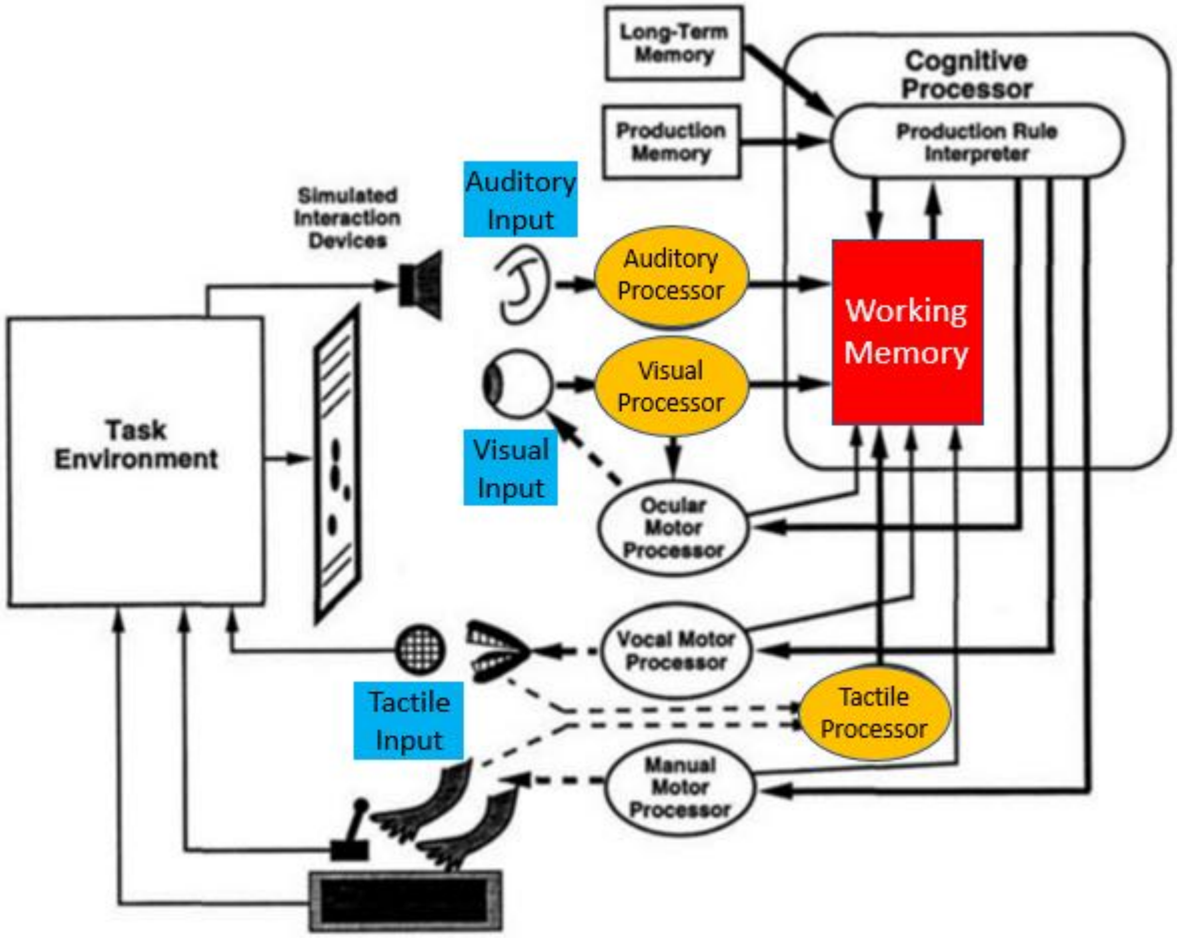


Figure 2.8: Executive Process Interactive Control: Information enters auditory, visual, and tactile processors, and then passed on to working memory

2.5.1 Perceptual Stage

During the perceptual stage, stimuli pass through the sensory organs and the information is recognized and can be organized into patterns. The ability of the brain to extract information from the signal depends on the quality of the sensory input. In other words, a degraded input can lead to restricted information which can lead to restricted performance [128, 133].

2.5.1.1 Visual Sense and Perception

The eyes have receptors sensitive to a portion of the electromagnetic spectrum which turn the light into an electrochemical signals sent through the optical nerve to the primary visual cortex. The cones in the retina respond to light at different wavelengths which leads to the perception of color [122, 134]. Most display hardware (e.g. monitors, projectors) mix the three primary colors (red, green, and blue) to create a wide spectrum of colors. Depth perception is the ability to perceive distances and relationships in space, it is critical in regards to immersion and interaction in a virtual environment. Three major categories contribute to the depth perception: monocular, oculomotor, and binocular cues. Monocular cues are obtained from one eye, they can be static or dynamic. Static monocular cues include the size of the image on the retina, the linear perspective, texture gradients, aerial perspective, occlusion, and shadows. Dynamic monocular cues originate from motion (e.g. motion parallax). The oculomotor cues are related to the function of the eye, the related key processes are accommodation and convergence. The binocular cues require information from both eyes. Our eyes are located at slightly different positions on the horizontal plane, the brain used information from both eyes an fuses then into a single visual precept [122].

2.5.1.2 Haptic Sense and Perception

The origin of the word haptic comes from the Greek word "haptikos" which means able to touch/grasp. Today we use the word haptic to describe all tactile sensations (deformations of skin), kinesthetic sensations (muscle forces), and proprioceptive sensations (joint positions) of the body. Haptics can be passive or active [122]. The skin has different receptors to sense touch (mechanoreceptors, thermoreceptors, nociceptors, and proprioceptors) which are used to sense positive or negative pressure, vibration, texture, normal and tangential forces, temperature, electric voltage and current [135]. The Merkel cells, the Meissner corpuscles, and the Pacinian corpuscles are the most commonly used in tactile display applications. Depending on the region of the body, Merkel cells have the highest spatial resolution and a temporal resolution of around 10Hz, they are most sensitive to surface texture of objects, especially for moving mechanical stimuli. Meissner corpuscles are used to detect relative movements of objects on the skin and are needed to regulate grip forces, e.g. for holding an object with the fingers. They are most sensitive to low-frequency vibrations around 30 Hz. Pacinian corpuscles are most sensitive to high-frequency vibrations around 250 Hz [136]. The threshold of perception of vibration depends on the frequency of vibration, the location on the body, surface area of contact, and the type of probe used in the experiments. Fingertips have the lowest vibro-tactile thresholds, compared to the large toe, the heel, and the forearm [137]. Also, the skin has free nerve endings that sense temperature and pain which contribute to the tactile sense. Our receptors cannot detect the exact temperature of the surface but rather sense the thermal energy flow [135]. When haptic stimuli get detected by the tactile or kinesthetic receptors a signal is sent to the brain for interpretation.

2.5.1.3 Auditory Sense and Perception

The basic auditory properties for perception are loudness (psychological), pitch (qualitative), timbre (qualitative), and consonance and dissonance (qualitative). On a higher level, temporal proximity of tones is more important than spatial proximity. Humans depend on the intensity between both ears to localize sound, and must be able to recognize and identify complex auditory patterns to to perceive speech. The phoneme is the basic unit of speech [128]

2.5.1.4 Priming

Mood and even racial biases can be "contagious", i.e., transferred to other humans via implicit non-verbal behaviors [138, 139]. Cahrtrand and Bargh introduced the *chameleon effect*, which explains one human's subconscious mimicry of another human's behaviors [140]. They found that "the mere perception of another person's behavior automatically increases the likelihood of engaging in that behavior oneself" [141].

In general, *priming* can be seen as the incidental activation of a person's knowledge structure, which can lead the person to specific behaviors and attitudes [141]. It can affect social judgment [142], as well as goal-driven tasks, as Bargh et al. demonstrated by showing that primed participants performed comparatively better in an intellectual task [141]. Dijksterhuis and Bargh indicate that perception itself can prime or activate a behavioral tendency. Apart from perceiving observables of what is literally present, people make trait inferences and activate social stereotypes as forms of social perception that elicit the tendency to imitate in the social perceiver [143]. Qu identified three main elements in a conversation between a real human and a virtual human: the surrounding environment, the virtual conversation partner, and the virtual bystanders [144]. Qu showed that priming with surrounding media content had a guidance effect in both the real world and the virtual world [145]. Various studies have examined the concept of priming, some related to virtual reality [146, 147], but most of them explore the theory underlying the priming phenomenon. Researchers explored racial biases, gender, and virtual human personality in virtual environments [148–150].

2.5.2 Cognitive Stage

The cognitive stage is very complex, it might include retrieval of information from short term or long term memory, comparisons, arithmetic operations, decision making, and thoughts. This stage imposes its own constraints on the the performance. The fewer resources there are the more performance suffers. [128, 151] (See Figure 2.9.).

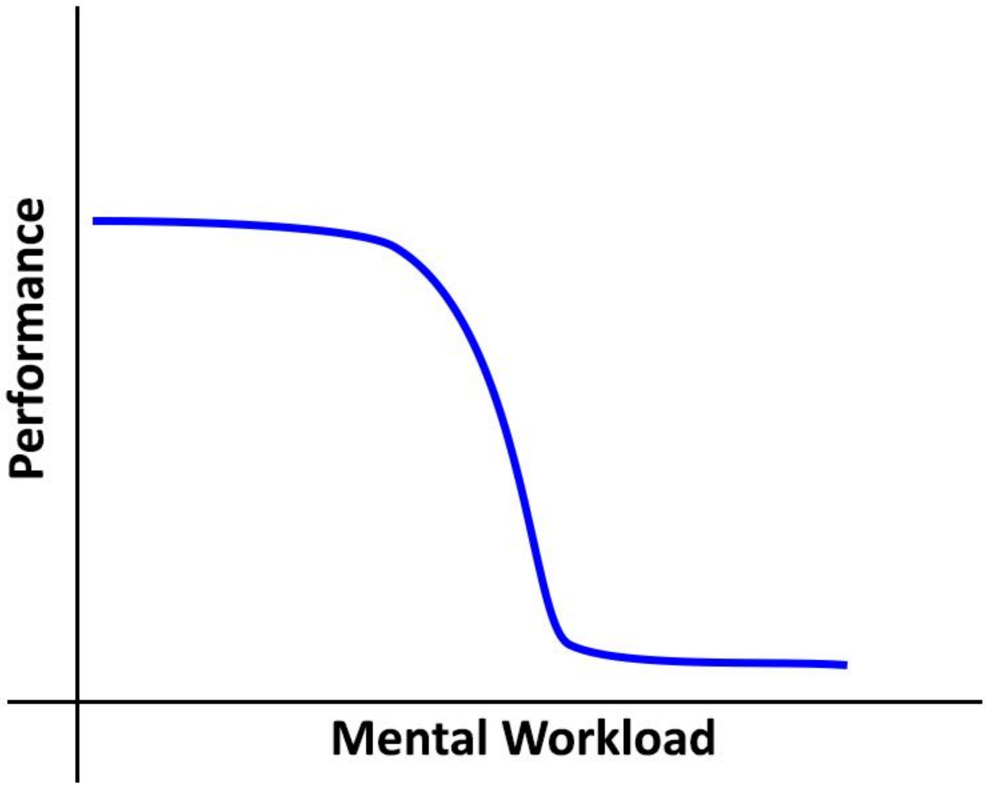


Figure 2.9: Mental Workload and Task Performance

2.5.2.1 Working Memory

The working memory holds a seven (plus or minus two) items or chunks of information at a time when merely holding information [152] and even fewer when processing information [153]. The number of information that needs to be processed at the same time determine the cognitive load [154]

2.5.2.2 *Cognitive Load Theory*

The cognitive load theory distinguishes between three sources of cognitive loads: *intrinsic cognitive load*, *extraneous cognitive load*, and *germane cognitive load* [130, 155, 156] (See Figure 2.10.)

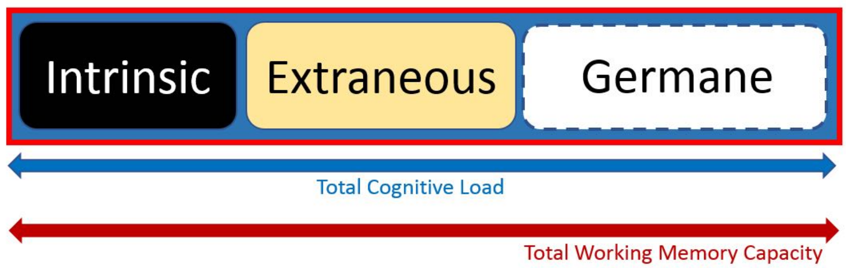


Figure 2.10: Working Memory: Intrinsic Load, Extraneous Load, Germane Load

The intrinsic cognitive load is essential and unavoidable. The intrinsic nature of the information such as level of the learner, number or elements, or learning material determines this cognitive load. Novice have a much higher intrinsic cognitive load than experts [156, 157] (See Figure 2.11.)

Figure 2.11: Working Memory: Intrinsic Load Higher for Novice.

The extraneous cognitive load is the "noise" or unnecessary to the learning goals, it comes from sub-optimal instructional methods. When this load is high, it can impede learning [158, 159].

The germane cognitive load comes from relating new information to relevant information from long term memory which leads to learning [155, 156].

The cognitive load theory suggests that instruction should be designed in a way where the material matches the learner's previous knowledge, should minimize the extraneous load, and the learners should engage in a process that evokes the germane cognitive load [130–132] (See Figure 2.12.)

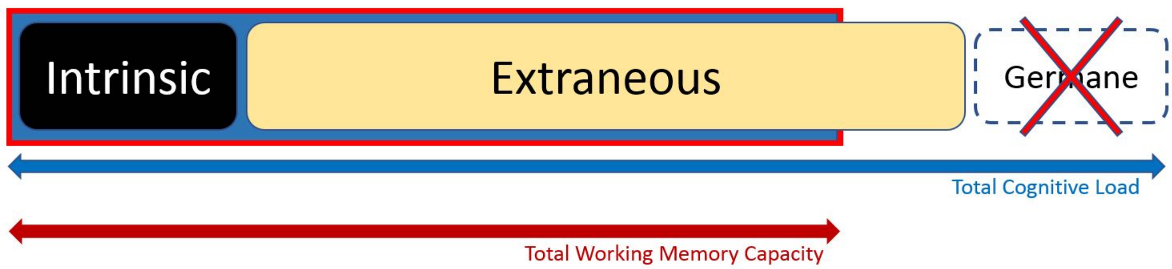


Figure 2.12: Working Memory: Extraneous Load Exeeds Memroy Capacity.

2.5.3 Action Stage

The action stage follows the perceptual and cognitive stages of information processing where if a response is required then it is selected from a set of alternatives under the circumstance, translated to a set of neuro-muscular commands, and executed. Limitations in the action stage can result in reduced performance [128].

2.6 Select Metrics of Interest for Simulated Patients

2.6.1 Perceptual

This section is related to perceived realism, presence, and social presence:

An important concept is the user's sense of *being* with a real person, often measured by presence, co-presence, and social presence. *Presence* can be defined as "the sense of non-mediation": one who is oblivious to the existence of a technological medium can perceive presence via that medium [160]. The sense of another person's presence, called *co-presence*, exists when people mutually perceive one another [161]; it can also be defined as the degree to which one believes himself or herself to be in the presence of and interacting with other veritable human beings [162, 163]. *Social presence* is sometimes defined as one's sense of being socially connected with another person [19]. Harms and Biocca explain co-presence as one of several dimensions of social presence, confirming the validity of their measures with questionnaires [164]. Chuah et al. defined an virtual human's *physicality* to include the physical size, volume, and position it occupies and its ability to interact with its surrounding environment [112]. Increasing a virtual human's physicality has been observed to increase social presence [165], communication, and pro-active and re-active behavior [166]. Using agents with a physical component, Lok et al. showed that the physicality of an agent can have benefits for social presence and the training of communication skills in the scope of medical team training [167].

Bailenson et al. studied participants' sense of co-presence in a multi-user shared immersive virtual environment while manipulating the non-verbal behavior of their virtual self-representations. The participants reported a higher sense of co-presence in a condition with head movement compared to the other conditions [168]. Bailenson et al. also compared different forms and behaviors of realism using a virtual human. In their study, they used video-conference (high behavioral and form realism), voice-only (low behavioral and form realism), and "emotibox" avatar (high behavioral realism and low form realism), and presented that people rated "emotibox" lowest in terms of the self-reported co-presence score [169]. Garau et al. evaluated participants' responses, including presence, co-presence, and physiological signals, with respect to a virtual human's degree of responsiveness. Their results did not show a significant relationship between perceived co-presence and the virtual human's degree of responsiveness. However, they did suggest a link between higher levels of co-presence and participants who reported using computers less [170]. In an educational context, Bulu compared learners satisfaction with their sense of presence (including social and co-presence), and found that presence is a predictor of a students' satisfaction in a virtual environment [171]. In a medical context, Robb et al. explored the effects of real versus virtual surgical team members on social presence and related constructs, and found that participants experienced less social presence with a virtual anesthesiologist [172]. There are multiple possibilities for how the sense of social presence or co-presence of a virtual human can be improved through modifications to its behavior *during* an interaction [170, 173, 174], and *prior to* a direct interaction.

2.6.2 Cognitive

This section focuses on cognitive load and learning: Under optimal conditions, the mental effort required to complete a task with a simulated patient would match that induced by a real patient. This mental effort is referred to as *cognitive load*, which can be broken down into three primary categories [175]. When cognitive resource demands are inherent to a particular topic or task, they are referred to as *intrinsic* cognitive load; they cannot be reduced or omitted without affecting the learning task. In contrast, *extraneous* cognitive load is induced by the way in which learners or task performers receive information, which can potentially be reduced by modifying the means of instruction. The category of cognitive load that promotes learning is called *germane* cognitive load. All types of cognitive load are bound by working memory, which has a limited capacity for cognitive processing [176]. Because of this balance, it is important to manage extraneous cognitive load so that learning is not adversely affected. Many researchers have considered how cognitive load and learning relate in the context of healthcare simulation [177–179]. For example, students who received training to identify a heart murmur only one hour prior to simulation training were less likely to be successful in the face of increased cognitive load [180]. Fraser et al. describe the increased cognitive load due to a “split-attention effect” that occurs when multiple information sources are separated in space or time [176], pointing to a potential advantage of physical-virtual simulators that can present multi-sensory cues in a co-located manner. Reducing cognitive load and increasing performance in medical simulation training depends to a large degree on learning certain behaviors when interacting with a patient.

2.6.3 Emotions and Feelings

Learning through simulation generates emotional experiences before, during, and after the simulation. There are not many studies that focus on learning and emotions [181]. The effects of simulated clinical immersion and simulated patients on emotion impact how students learn [178]. The Cognitive-Affective theory of learning with media indicates that the multimedia learning process is mediated by the learner’s mood, and positive mood has a facilitating effect on multimedia learning [182]. During simulation practice, nursing students can feel unready, fearful, anxious, and worried that could affect learning outcomes [183]. According to The Positive and Negative Affect Schedule (PANAS), “Excited” and “Alert” are classified as positive affects [184]. Emotional engagement enhances interaction with mannequins [185]. When nurses feel a sense of urgency it usually comes from recognizing that a situation needs quick action. The fidelity of the simulation can affect the sense of immersion and sense of urgency [186]

In psychology, there is evidence that an person’s apparent mood can be affect another person via implicit nonverbal behaviors [138, 140, 141]. Researchers have used questionnaires to measure affective attraction [187], and Mood Rating Questionnaires to measure the subjective emotional state [188].

2.6.4 Behavioral

This section focuses on behavioral measures such as communication, standing positions and eyetracking.

When employed in simulation facilities, motion-tracking systems, such as head-tracking devices, can provide information about the movement patterns of the users in their environment [189], while eye-tracking systems are a means of measuring the visual focus of attention of users [190]. Researchers have used head pose and orientation as a reasonable substitute for a user's gaze to estimate interest, visual focus of attention, and movement patterns [190–192]. Many researchers in the field of healthcare have utilized eye-tracking systems, both as measures of behavior and as student feedback for performance and learning goals [193, 194]. Suetsugu et al. compared the attention behavior of nursing professionals and students by their gaze duration and gaze locations [195]. In an experiment to interpret physiological vital signs, Currie et al. predicted the performance of nursing professionals and students using some metrics collected from their eye-tracking data, such as fixation frequency [196]. The promising body of eye-tracking research inspired us to investigate similar relationships in our experiment.

CHAPTER 3: PHYSICAL-VIRTUAL PATIENT DEVELOPMENT AND TESTING

This chapter covers the development of the Physical-Virtual Patient Head (PVHead) and the Physical-Virtual Patient Bed (PVPBed) for a child. Together with another PhD student, Jason Hochreiter, we developed the PVHead that has touch sensing capability. Jason focused on the touch sensing part, and I focused on other aspects such as creating the interactive graphics, providing real-time responses to touch by updating the visuals and audio, creating medical scenarios to use in experiments. While the PVHead is designed to be able to support automated touch sensing, the touch sensing part is outside the scope of what I am interested in for the experiments in this dissertation. This chapter does not go into the details of touch sensing, those can be found in published papers [44, 100, 197] . After the development of the PVHead I lead the effort in the development of a patient with a body that can be transported for experiments.

Papers Published related to the development of the Physical-Virtual Patient Head:

Title: Touch Sensing on Non-Parametric Rear-Projection Surfaces: A Physical-Virtual Head for Hands-On Healthcare Training,

Authors: Jason Hochreiter, Salam Daher, Arjun Nagendran, Laura Gonzalez, Gregory Welch.

Published in: IEEE Virtual Reality, 2015

Title: Optical Touch Sensing on Non-Parametric Rear-Projection Surfaces for Interactive Physical-Virtual Experiences.

Authors: Jason Hochreiter, Salam Daher, Arjun Nagendran, Laura Gonzalez, Gregory Welch.

Published in: Presence: Teleoperators and Virtual Environments, 2016.

3.1 Physical-Virtual Patient Head

This section describes the development and testing of the PVPBed.

3.1.1 Development

This section focuses the hardware and software development, and how the PVHead was evaluated and used for experiments.

3.1.1.1 Hardware

An 8020 rig supports the electronics (i.e., projector, infrared lights, infrared cameras). A human shaped plastic shell sits on top of the rig. Four monochrome Cameras (Point Grey Blackfly) with infrared filters, two infrared lights, and one AAXA P300 pico projector [198] are positioned on the rig under the shell. The cameras detect touch events in the infrared spectrum. The projector sends imagery to the shell (Figure 3.1)

3.1.1.2 Software

I used 123D Catch to scan the shell. The scan results in a very dense mesh that is hard to texture and animate. I used the dense mesh as a guide to create a quad-based low-density mesh with a quad-based edge-flow topology (instead of triangles) suitable for animation. Quads are the polygon of choice to design a topology that follows the natural movement of a human body and that reduces artifacts which sometimes get generated by subdividing the polygons during the smoothing process. Edge-loop modeling relies on quads as they naturally support this type of modeling [199].

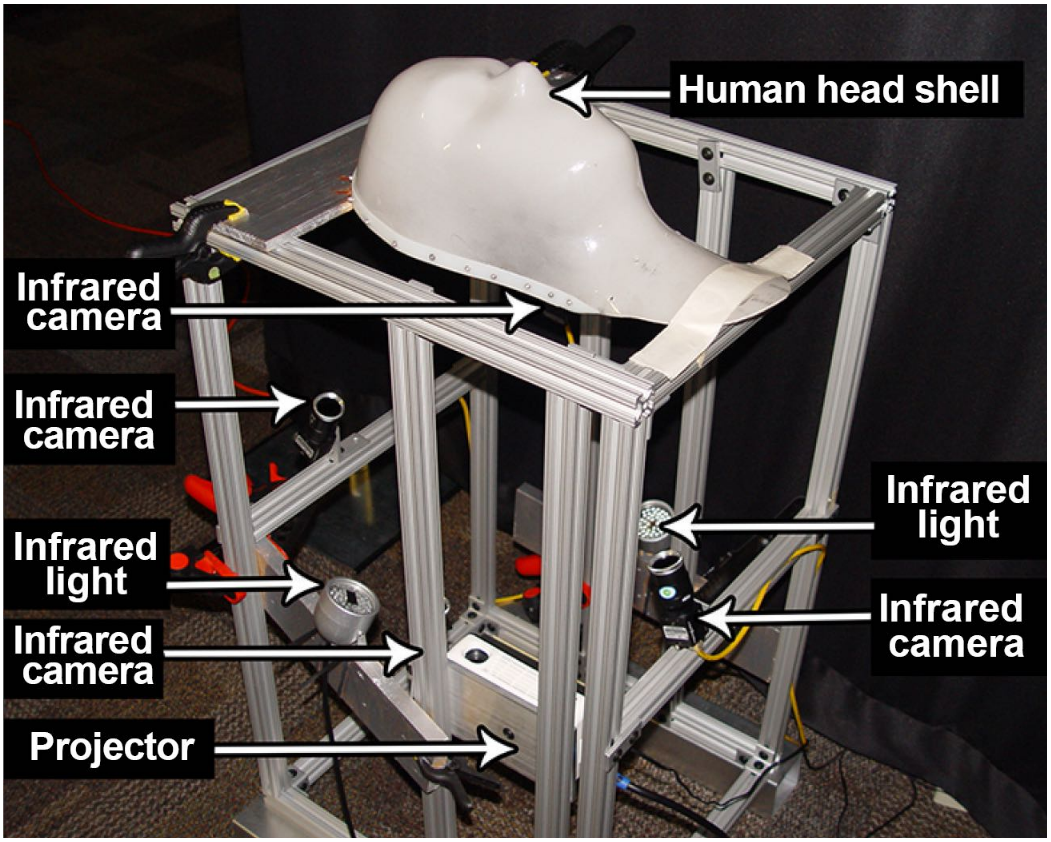


Figure 3.1: Rig for PVHead

Where a vertex has three or five or more edges it is called a Pole. During animation the chances to see artifacts are higher. Poles were positioned at specific locations to minimize those artifacts. For our head shell, I used Maya [200] to create eyeballs and inner mouth geometry. I attached the inner mouth geometry to the outside mesh by merging vertices. I UV-mapped the mesh to a texture created in Photoshop. The texture is an unwrapped image of a face. The mesh was rigged in Maya using a combination of joints and blend-shapes for various animations. Following the anatomy of the head, a joint in Maya controls the jaw bone for opening and closing the mouth. Also the eyeballs were attached to joints so the joints can control the rotations (up, down, right, and left). Each blend-shape controls a facial muscle groups or Action Unit (AU) in the Facial Action Coding System (FACS). By combining different blend-shapes the face is capable of supporting animation for any facial expression (Figure 3.2).

I exported the 3D model to Unity3D [201] and positioned it in a scene. The size of the scene matches the projector's resolution (1920×1080 pixels).

We modeled a virtual camera that matches the field of view, position, and rotation or the rear projector's intrinsic and extrinsic parameters. The virtual camera is located under the mesh to match the projector's position in the physical rig. The camera renders the scene and sends the resulting imagery from that view to the projector. This camera/projector calibration process can be automated. (Jason worked on the automation of the calibration. This part of the development is not an area of focus for my dissertation.)

When a touch is detected, the data is sent to unity. A region in the face is detected (e.g., upper lip, lower lip) and the corresponding blend-shape is activated. As the touch location changes I compute the distance and update the blend-shapes percentage to follow the location of the touch.

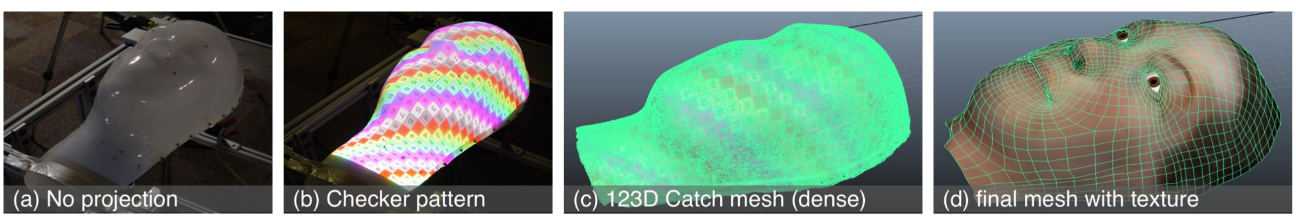


Figure 3.2: stages of development for the 3D mesh starting from scanning the physical shell (a) by projecting a pattern on it(b), then converting the dense 3D mesh from the scan (c) to a simplified animatable 3D mesh (d)

3.1.2 Evaluation of the PVHead

The evaluation of the PVHead was technical at first. Jason and I tested the accuracy of automated touch sensing on the head. More details about the touch sensing testing can be found in [100].

After the technical test, I developed a scenario for a neurological assessment (stroke). Prof. Laura Gonzalez from College of Nursing provided feedback regarding the behavior of the patient, for example Figures 3.3, and 3.4 show the interaction regions for audio and visual responses to touch. The head was used in 2 pilot studies where we received feedback. Some of the feedback from the pilots was incorporated into the next study that used the PVHead (e.g., character responses and behaviors), and some was incorporated in the design of the PVPBed in the next section of this chapter (e.g., full body, position patient as a slight angle instead of 100% parallel to the floor). In an iterative process, and using feedback from the pilots the PVHead was used for the study described in Chapter 4. The Scenario used in the Stroke Study is detailed in Appendix B. Figures 3.5, 3.6, and 3.7 show a sample of what the graphics looks like. Figure 3.8 shows the interface of the software.

The head was also used in another study that focuses on the performance and cognitive load when given a touching task. This study is covered in chapter 6.

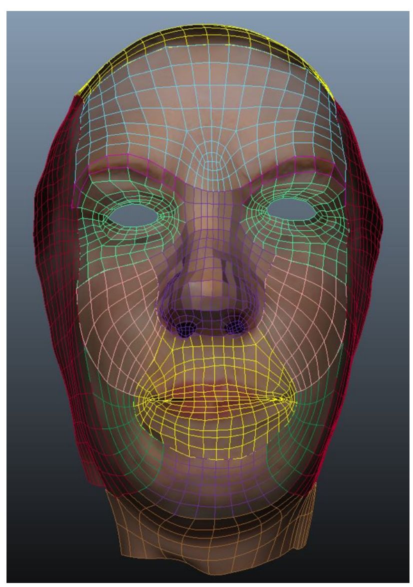


Figure 3.3: Regions on the mesh to trigger different audio

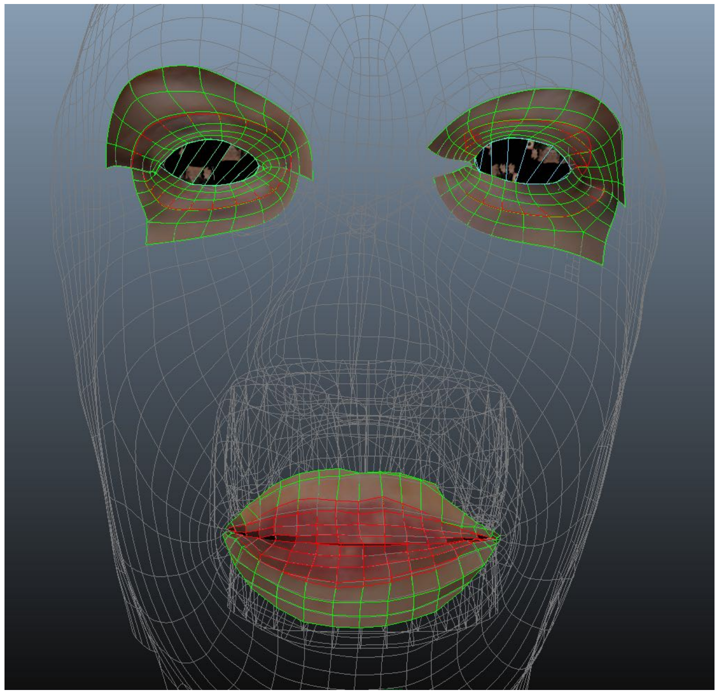


Figure 3.4: Regions on the mesh to trigger visual changes when a touch is detected in a start region (in red) and the touch moves to its corresponding destination region (in green)

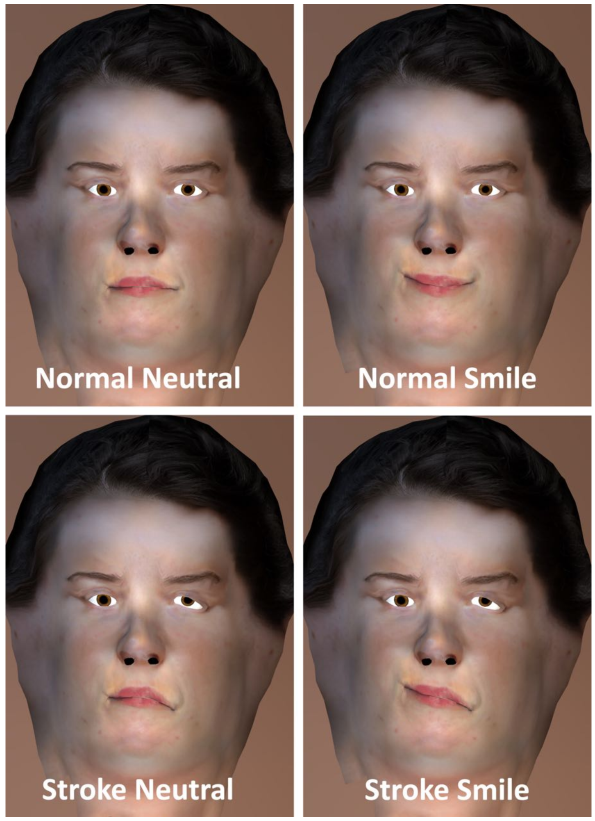


Figure 3.5: Head images sent to the Projector featuring a normal patient and a stroke patient with a neutral and smile expressions

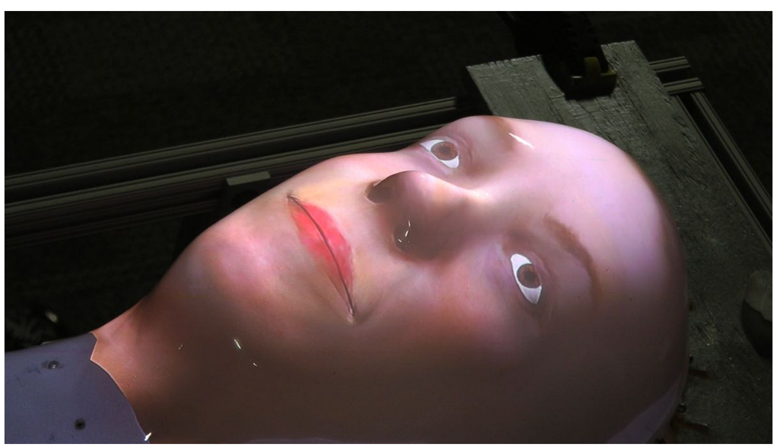


Figure 3.6: Projection on the head shell with neutral facial expressions

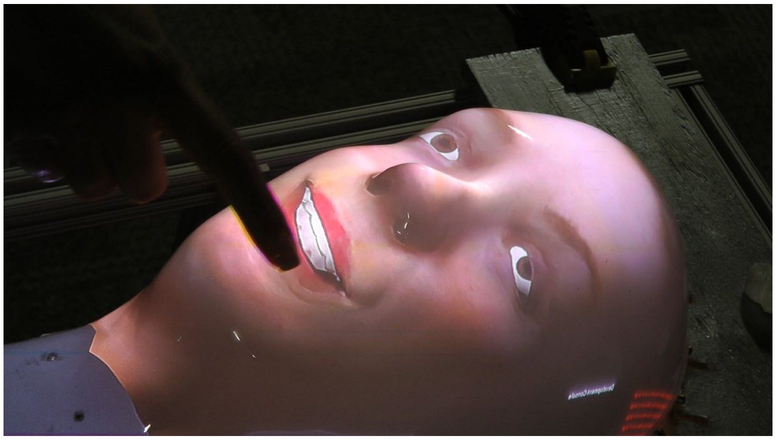


Figure 3.7: Lip tug on the PVHead

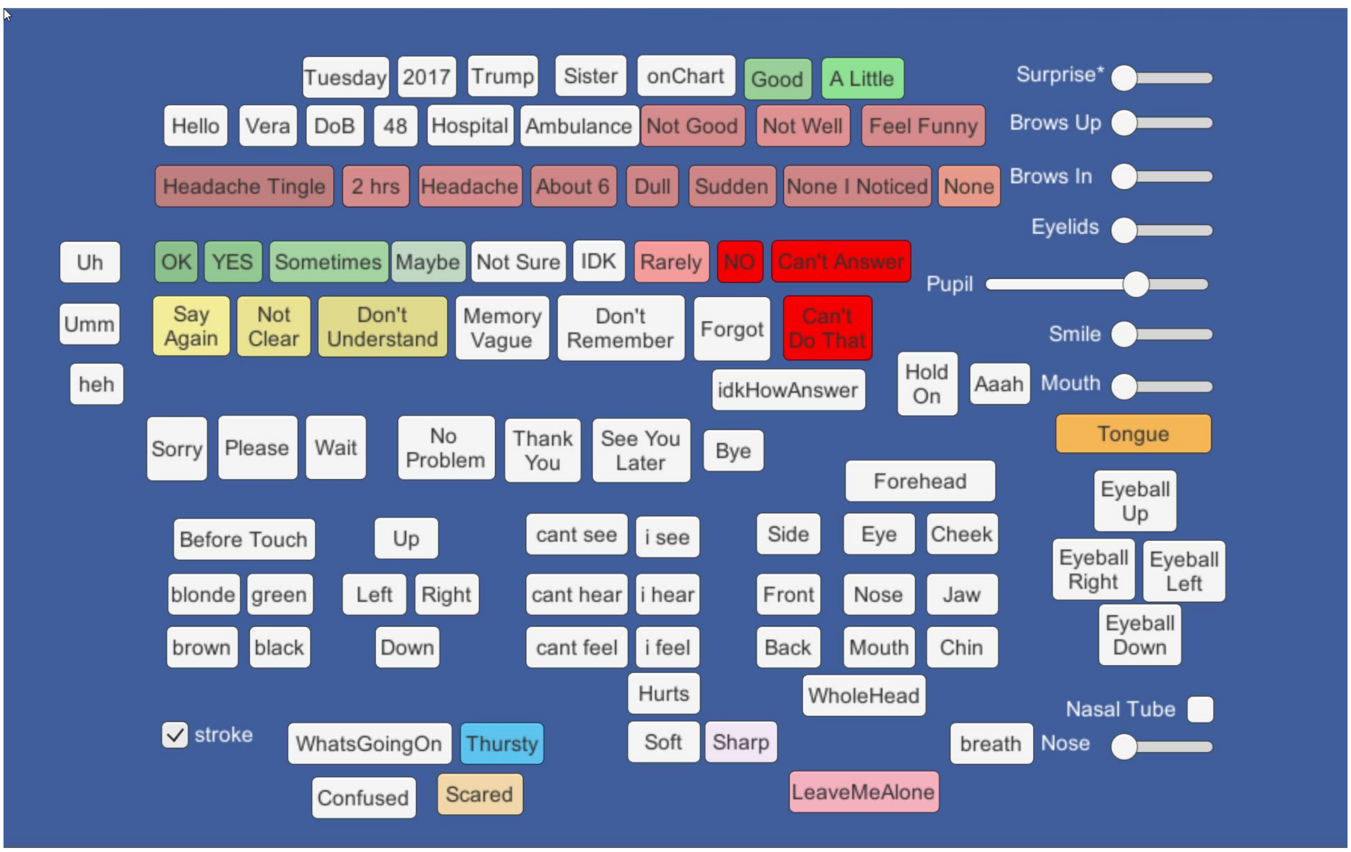


Figure 3.8: Graphical User Interface to control the verbal and non verbal responses

3.2 Physical-Virtual Patient Bed

This section describes the hardware and software development, and how the PVPBed was evaluated and used for experiments. This is submitted to Simulation in Healthcare Journal as a technical report and it is accepted pending modifications

Paper accepted related to the Physical-Virtual Patient Bed [202]:

Title: The Physical-Virtual Patient Simulator: A Physical Human Form with Virtual Appearance and Behavior.

Authors: Salam Daher, Jason Hochreiter, Ryan Shubert, Laura Gonzalez, Juan Cendan, Mindi Anderson, Desiree Diaz, Gregory Welch

Accepted in: Simulation in Healthcare Journal (to appear 2018-2019)

3.2.1 Development

3.2.1.1 Hardware

The PVPBed consists of a 51cm x 76cm x 76cm metal frame that houses electronic equipment inside and an interchangeable translucent plastic "shell" on top (Figure 3.9.) While the shape of the shell could represent virtually any human or condition, the shell in our initial prototype is shaped like a small child, spanning from the head to just above the knees. The PVPBed is designed to be transportable, so it can be moved to an appropriate healthcare environment for in situ training or experiments. The imagery is projected onto the surface of the shell from below using 2 AAXA P300 Pico projectors [198] mounted inside the frame and pointing upwards, each with a resolution of 1920 x 1080 pixels. Presently, one projector is positioned to provide the imagery for the head of the patient, while the other covers the rest of the body on the shell. The computer graphics application uses Unity [201] to send imagery to the projectors. In Unity, I modeled a virtual representation of the physical setup, with two virtual cameras pointing toward a 3D model of the patient. Each virtual camera matches the position, orientation, and field of view of its corresponding physical projector. The imagery rendered by these virtual cameras is sent directly to the projectors for display on the shell. To form the shell, I modeled a 3D computer-aided design (CAD) model of a child patient with appropriate proportions [203–206], had it milled to create a "positive" mold, and had the shell material vacuum-formed over the mold. We researched several types of plastic material that simultaneously allow for projected imagery to form clearly on the surface of the shell and infrared light to pass through the shell. The shell was vacuum-formed using 1/16-inch Optix 2447 plastic sheet material [207]. The CAD model, mold, and shell were designed to comprise a relatively smooth surface shape so that the virtual patient could exhibit some degree of animated movement in areas such as the nose and fingers, without creating the disturbing visual distortions that would otherwise occur if the imagery moved over areas of sharp shape changes (Figure 3.10). Together with PhD student Ryan Schubert, we simulated the projection of the 3D virtual character onto the 3D CAD model of the shell before manufacturing the physical shell to find a shape suitable for projection (Figures 3.11, and Figure 3.12). Speakers were added near the head to allow the patient so speak and breathe. The speech and breath sounds were pre-recorded, with the inhale/exhale sounds created to match the respiratory rate and animations of each scenario. Five Honeywell HCE100B Heat Bud Ceramic Heaters [208] were installed below the head of the shell, the left and right sides, and the bottom. Each heater can be independently set to high, medium, or off to provide different degrees of warming of different parts of the body. To simulate pulse, audio signals were sent to two acoustic haptic Techtile Toolkit [209] devices located under the shell at the right and left arms.

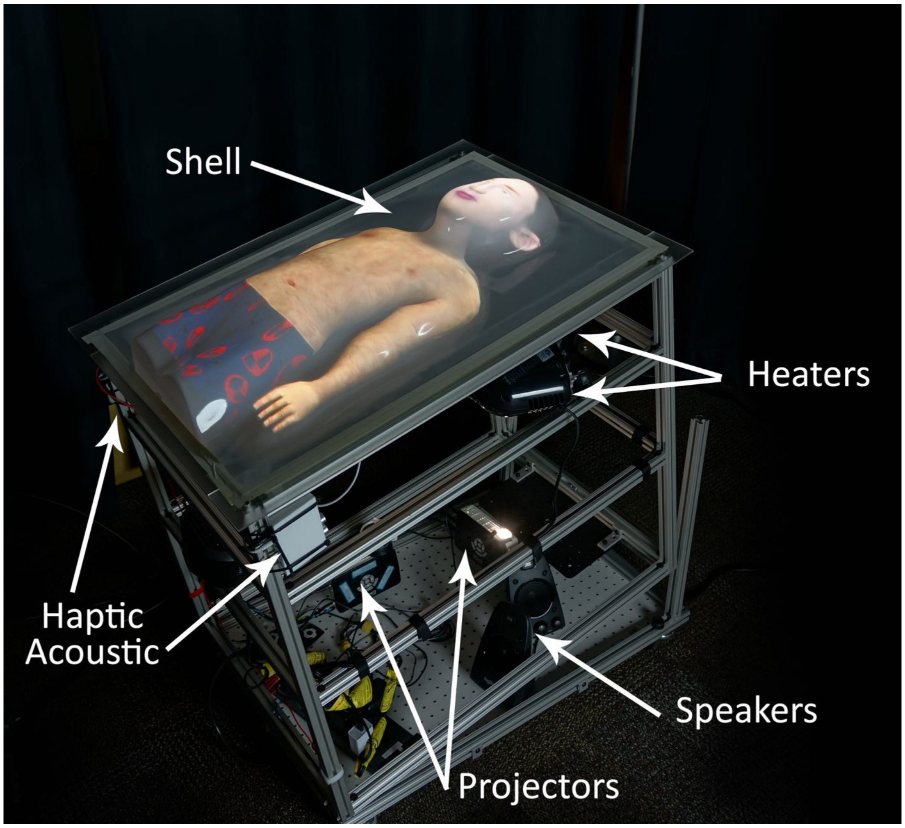


Figure 3.9: PVPBed rig featuring the shell with projection, projectors, heaters, and haptic acoustic devices.

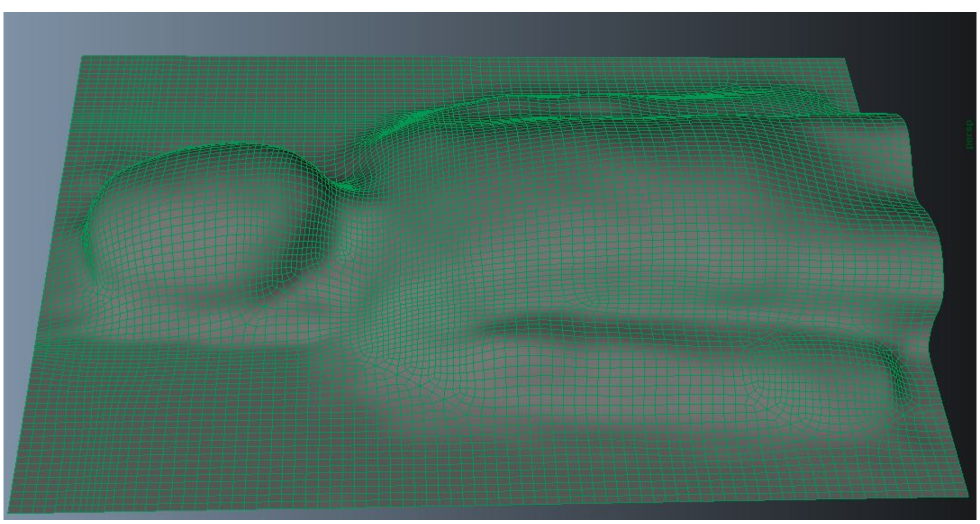


Figure 3.10: 3D shell modeled in Maya in preparation for manufacturing the plastic shell.

Figure 3.11: Mold milled using the 3D design

Figure 3.12: Vacuum formed acrylic shell

3.2.1.2 Software

After researching children's proportions [203–206], I modeled a 3D child character using the 3D modeling software package Maya [200]. The full body 3D model has a low number of polygons to ensure fast real-time interaction. I created textures in Adobe Photoshop [210] and used a UV mapping technique to map the 2D textures, which use (U, V) coordinates in a 2D plane, onto the 3D mesh, which uses (X, Y, Z) coordinates in 3D space. Various body parts in the 3D character were rigged for animation using joints and blend-shapes to allow for animating different body parts. The eyeballs, jaw, neck, torso, breathing, arms, hands, fingers, and legs were rigged using joints to smoothly control the mesh vertex positions during animations. The rest of the facial muscles besides the eyeballs and jaw were controlled using blend-shapes (Figure 3.14 and Figure 3.13). Twenty-two blendshapes were created for facial expressions and for phonemes (e.g. blink, open/close mouth, open/close lips, open/close eyelids, smile, frown, nose wrinkle, disgust, fear, sadness, pupil changes, etc.) that can be further combined to create more complex variations with different intensities. The blendshapes for phonemes allow the character to appropriately move his/her lips regardless of the scenario (Figure 3.15 and Figure 3.16). In Unity, two "virtual cameras" were positioned under the 3D character, and a custom shader renders the graphics in reverse depth. The shader was needed so that geometry located inside the 3D character, such as the internally modelled mouth and eyeballs, are rendered appropriately on the shell. The imagery from each virtual camera was sent to the corresponding real projector. The control screen presented to the simulation operator provided an interface to trigger speech, animations (e.g. open eyes, change pupil size, move eyeballs, open mouth, move head, facial expressions, move arms, move fingers, etc.), dynamic visuals (e.g. capillary refill, removal of the patient's shirt), and toggle vitals (temperature strip, blood pressure, O2 saturation).

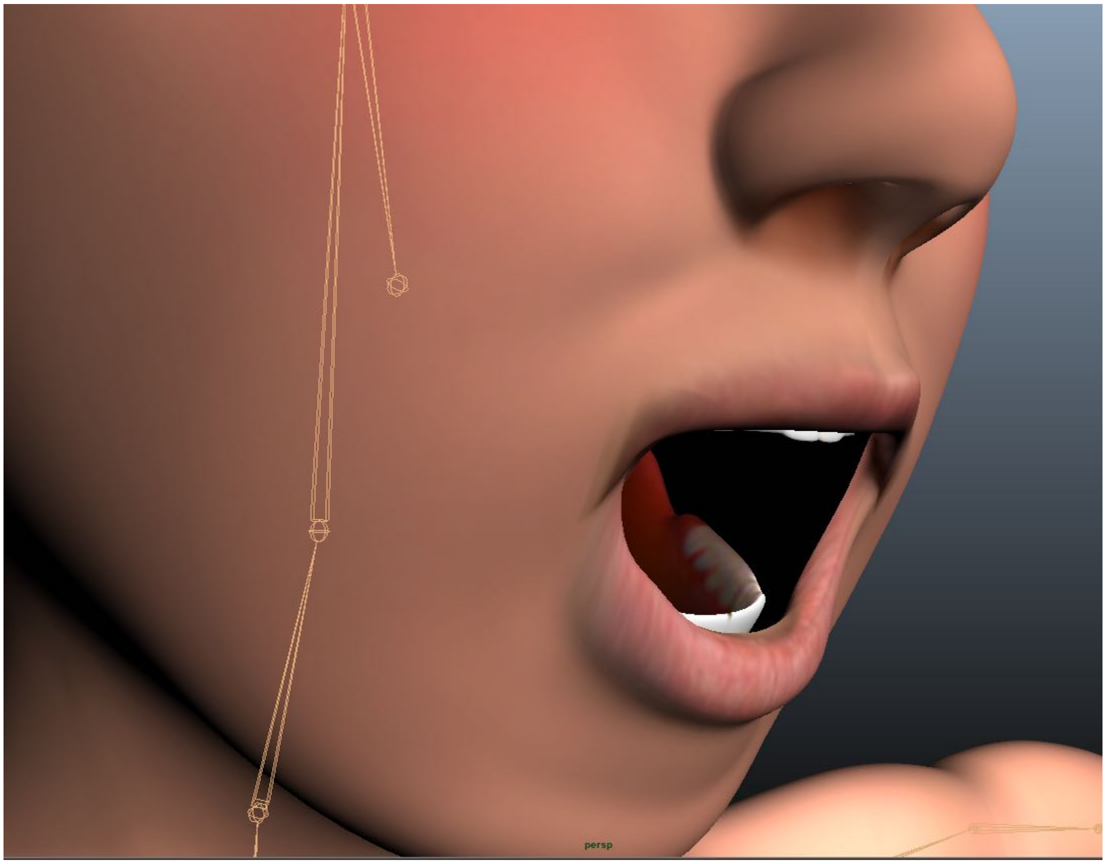


Figure 3.13: Joints in the head control the jaw movement

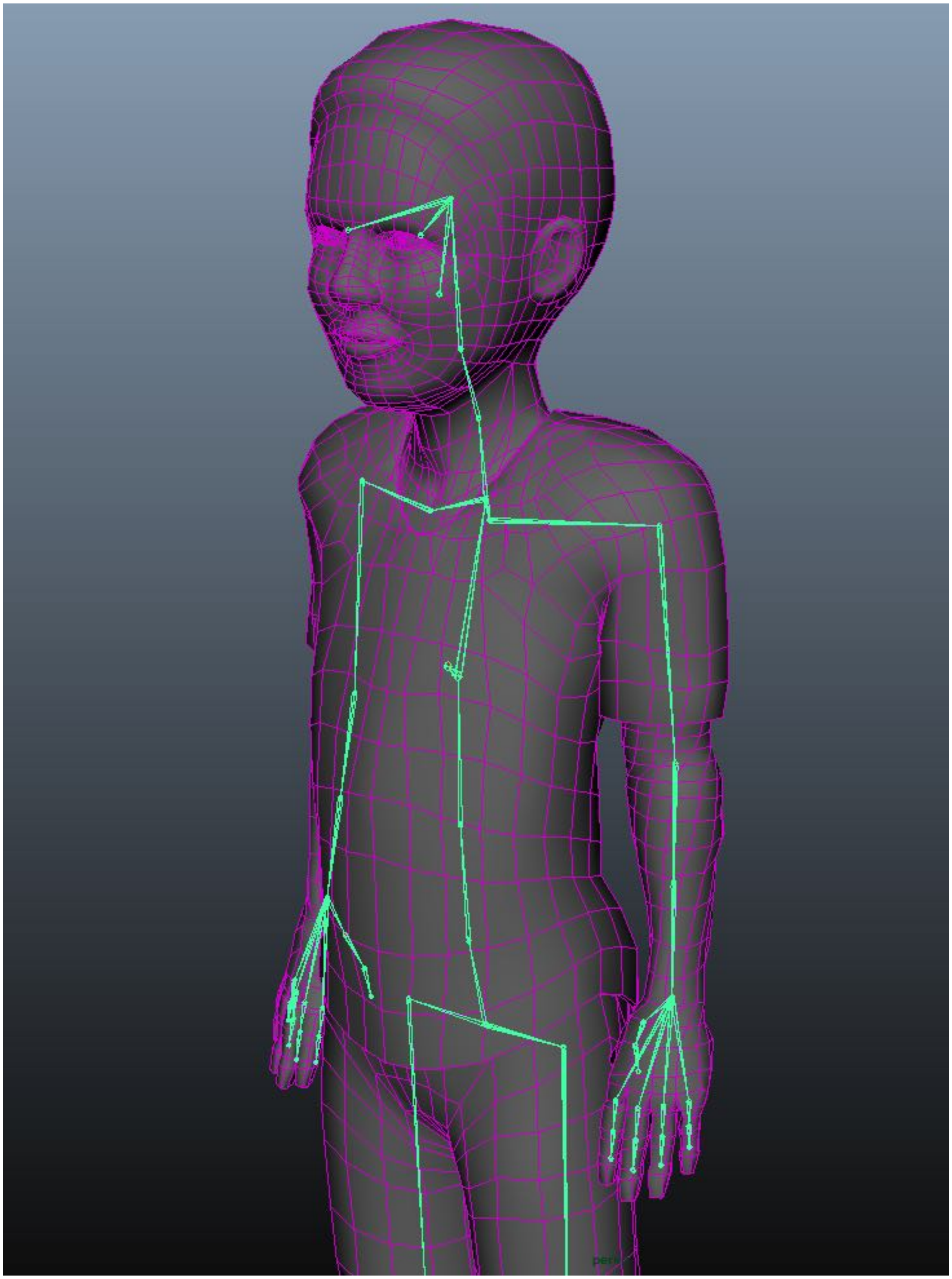


Figure 3.14: character in Maya showing the mesh and joints

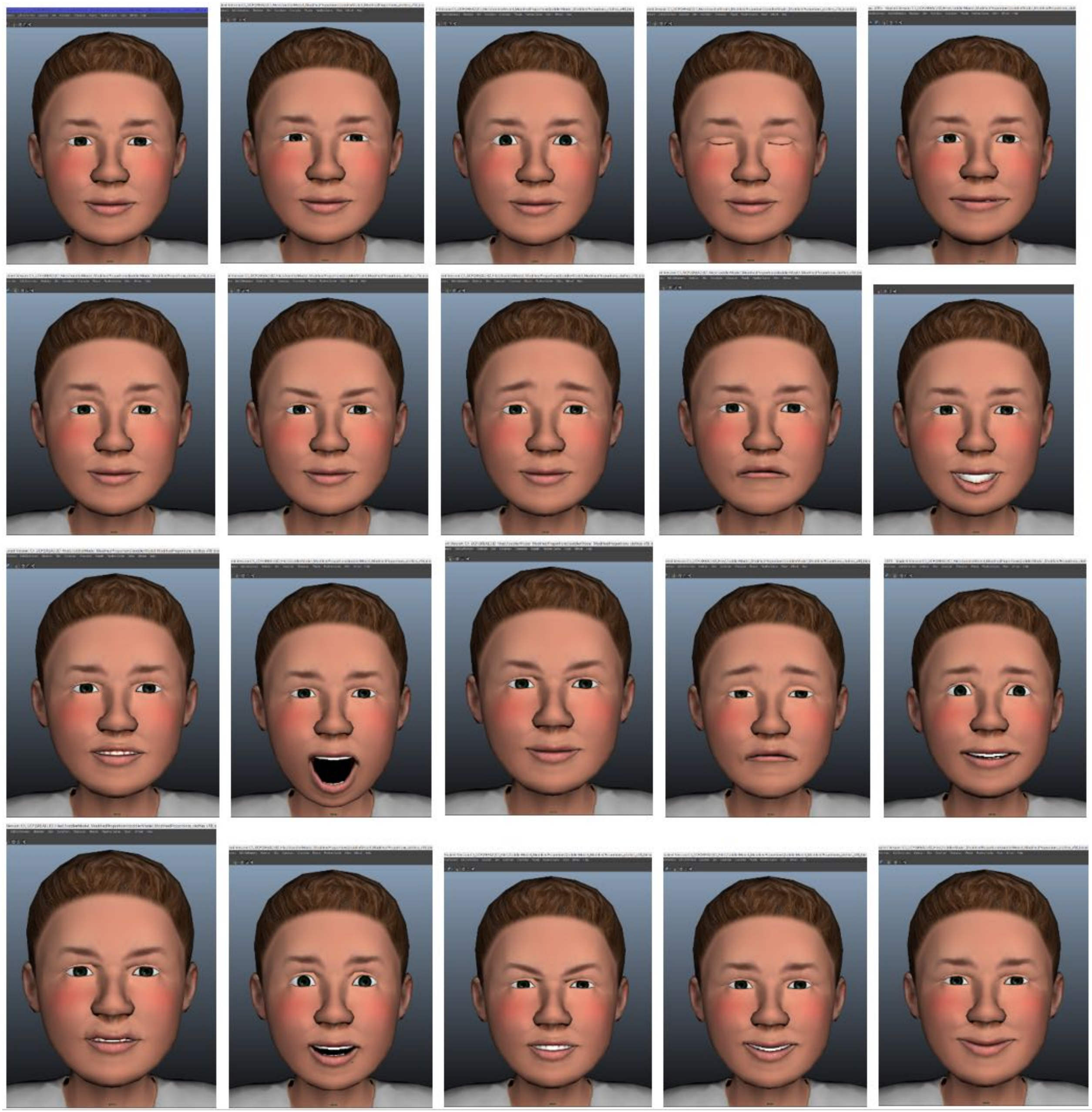


Figure 3.15: Facial Expressions using a combination of blendshapes and joints

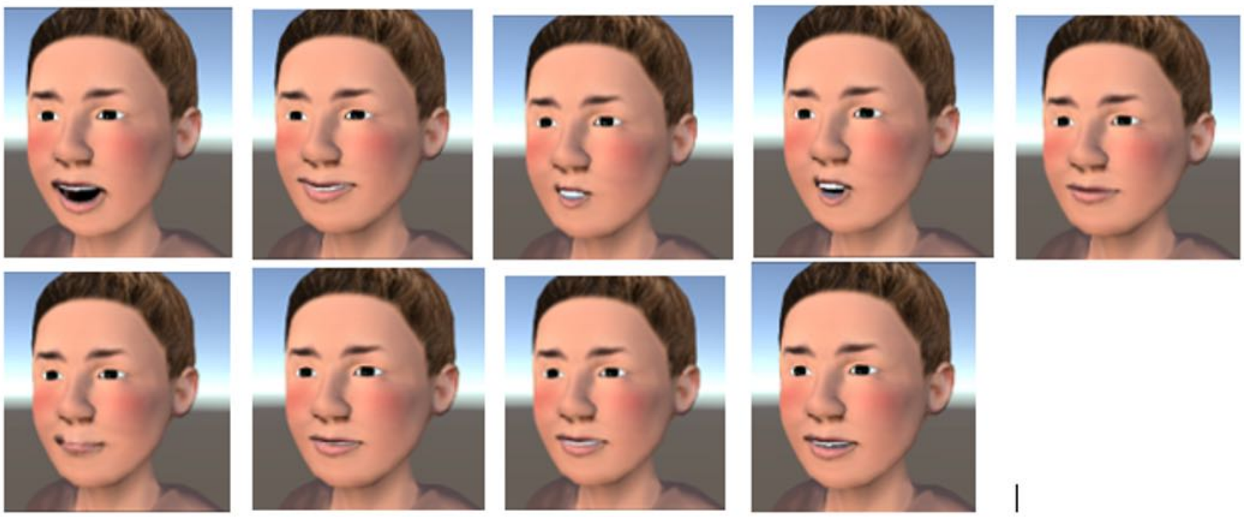


Figure 3.16: Visemes to animate speech

3.2.1.3 Scenario-Driven Content

After collaborating with subject matter experts from College of Nursing (Prof. Laura Gonzalez, Prof Mindi Anderson, and Prof Desiree Diaz) and from College of Medicine (Prof. Juan Cendan), I developed content for a normal healthy child to serve as a baseline (Figure 3.17, Figure 3.18, Figure 3.19, Figure 3.20, and Figure 3.21.)

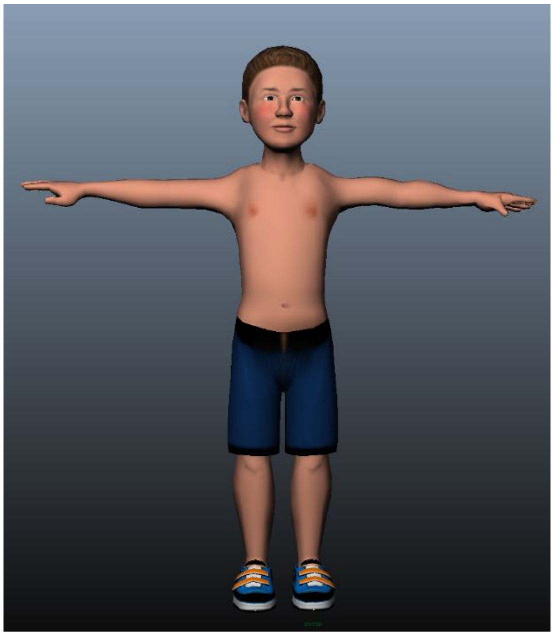


Figure 3.17: 3D character in a t-pose

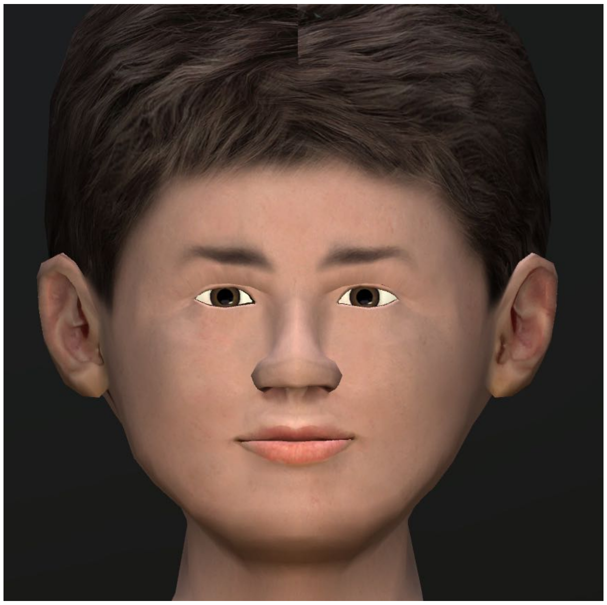


Figure 3.18: Close up to the 3D character of a healthy child's head

Figure 3.19: Close up to the 3D character of a healthy child's body with shirt

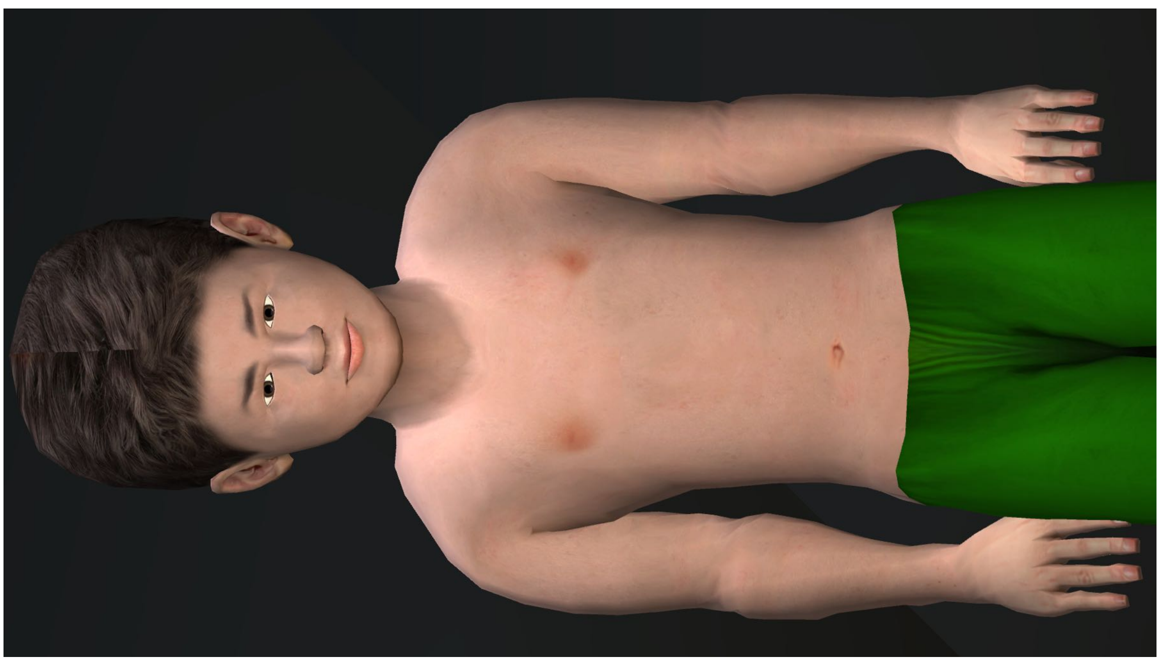


Figure 3.20: Close up to the 3D character of a healthy child's body without a shirt

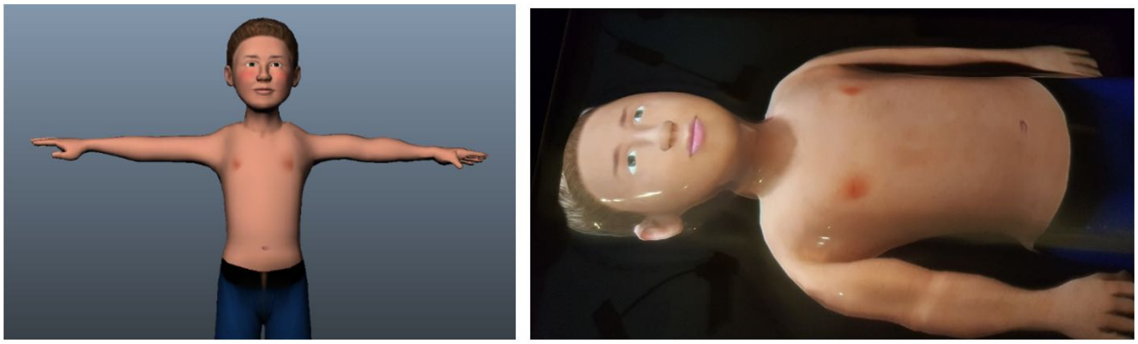


Figure 3.21: Projection of a child on the shell

We adapted a previously validated checklist of a pediatric patient demonstrating early signs/symptoms of sepsis to the PVPBed. We developed software to illustrate that the simulator can convey multiple subtle symptoms, such as skin mottling, cyanosis, ptosis, delayed capillary refill, tachypnea, fever, hypotension, and tachycardia (See Appendix C). Additionally, we developed content for other child patient scenarios with associated signs/symptoms that could be represented using the same PVPBed. For example, we developed a model for a child subjected to physical abuse, showing symptoms such as ecchymosis, cigarette burns, and bites (See Appendix D); and a burn patient showing burns, blisters, and swelling (See Appendix E.) A total of 446 audio clips for patient speech were recorded in three different tones: 148 for a healthy patient, 149 for a patient in pain, and 149 for a low-energy (lethargic) patient. As it is challenging to get a child to act and record in a recording studio setting, we initially recorded an adult. Using MorphVOX software [211], I recorded the voice of another PhD student, Ryan Schubert, then applied a set of modifications to these recordings to simulate a child's voice. The resulting audio was cleaned and imported in Unity along with the 3D character detailed below. The Rogo Digital LipSync [212] plugin was used to automate the visual lip motion of the 3D character to match the audio clips, and I developed an accompanying graphical user interface to trigger the clips and other animations. The control for the patient could be fully automated (i.e., via artificial intelligence), a real person could speak "live" into a microphone, or a hybrid "Wizard of Oz" [213] approach can be used in which an operator triggers pre-recorded responses. Because our medical scenarios are relatively specific, pre-recorded responses were chosen to ensure consistency (Figure 3.22.)

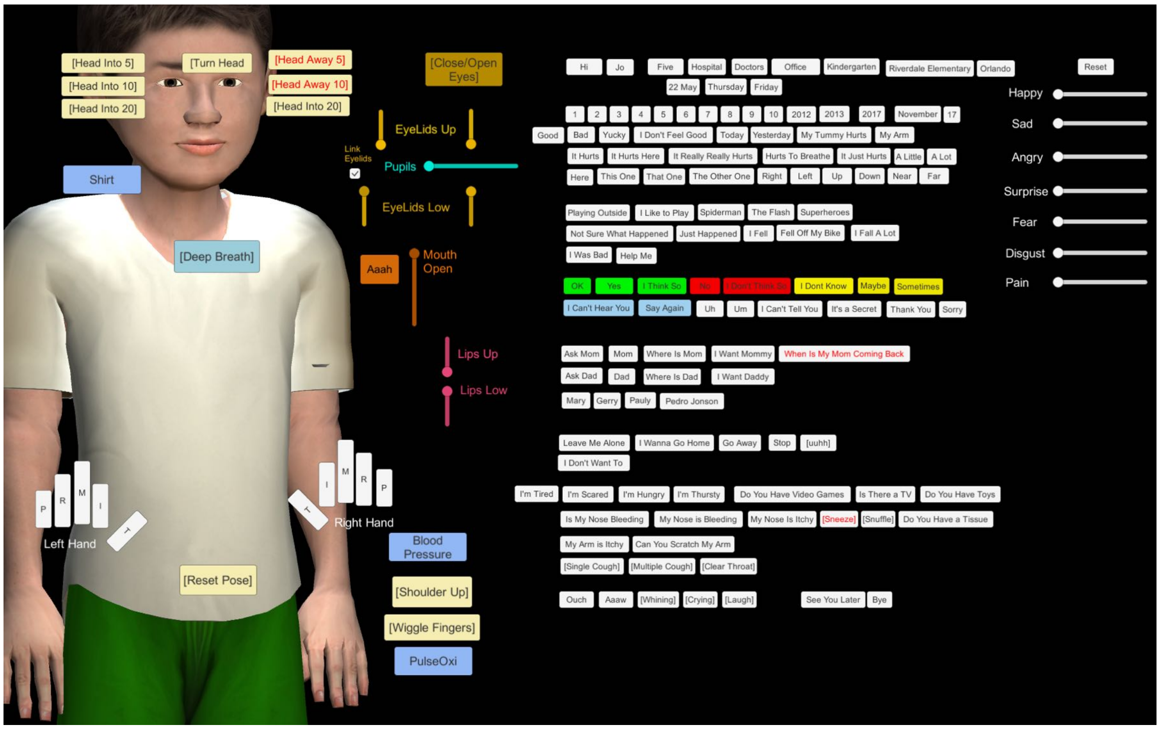


Figure 3.22: Graphical user interface to control the patient's responses

3.2.2 Evaluation

3.2.2.1 Technical Evaluation

First, we tested the hardware components of the PVPBed, comparing imagery projection quality and infrared transmission on different plastic materials with various transparencies and thicknesses (Figure 3.23).

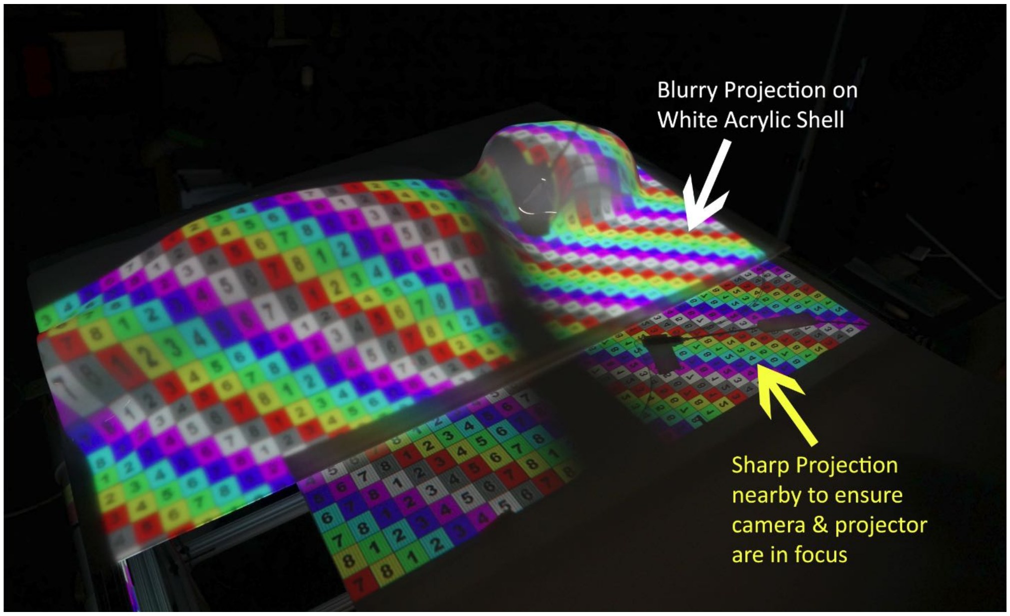


Figure 3.23: Material testing for sharpness of images

Different configurations (orientation and distance from the shell) for the heaters were tested and various locations (forehead, chest, arms, hands, groin, etc.) were spot-checked with a surface thermometer. Similarly, we simulated pulse using pre-created audio files that we sent as an input to the haptic-acoustic device which produced a tactile pulse. The audio files were created from an existing sample sound of a pulse and replicated using the intended rate (e.g. 80 bpm for the child abuse scenario, and 100 bpm for the sepsis scenario). The tactile pulse was measured and verified by three nurses on the child surface to verify they matched the intended pulse. We iteratively developed the simulation software using feedback from professors in nursing and medicine. We recorded and analyzed the actions of our medical team members performing a mock simulation on an ordinary mannequin to assist in further development of audio responses, graphical reactions, and other simulator capabilities (Figure 3.24.)

Figure 3.24: Mock simulation with a mannequin to help determine what is needed for the development

After we integrated the software and hardware, nursing professors conducted simulation sessions using the PVPBed and provided formative feedback (Figure 3.24.)

Figure 3.25: Mock simulation with a version of the PVP.

3.2.2.2 Human-Subject Experiment

After obtaining Institutional Review Board (IRB) approval, we conducted a formative human-subject study with twenty-two Nurse Practitioner (NP) students in an advanced health assessment class where they interacted with simulated child patients using the PVPBed. Twelve participants interacted with a sepsis patient, and ten participants interacted with a child showing signs of abuse (Figure 3.26, and Figure 3.27.)

Figure 3.26: Projection of a child with signs of abuse.

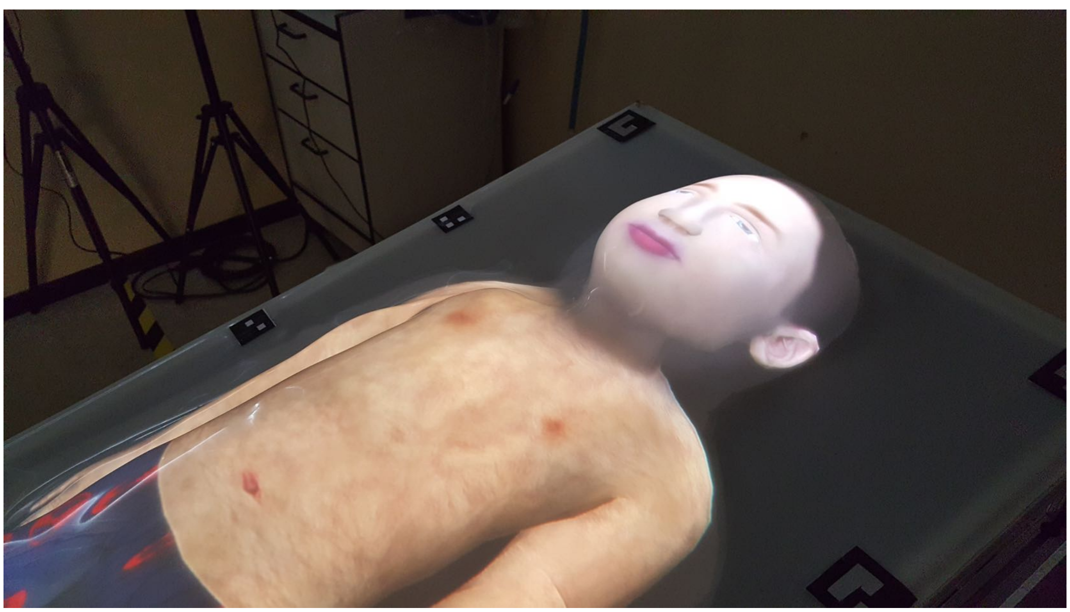


Figure 3.27: Projection of a child with sepsis.

3.2.2.2.1 Protocol

Students were first familiarized with the PVPBed via a video showing the PVPBed simulating a healthy child and providing an overview of the capabilities of the simulator. Next, pairs of students collaboratively assessed a child patient using the PVPBed. By design, a slightly inconsistent story was provided for the child abuse scenario between the patient report (patient fell off the sofa) and the patient's responses when participants asked the patient probing questions about his condition (patient fell off his bike), to see if the participants would notice the discrepancies. The participants were observed interacting with the patients from the laboratory's control room using a video system that had a (roughly) two-second delay. The delay was from the video recording system used in the university simulation lab, which streams the video feed to the controller in a remote room; there was no delay caused by the simulator itself. We did not want the controller to be in the same room as the participants so that they would not be affected by the presence of someone other than the simulated patient. The patients' animated behaviors and audio clips were initiated pro-actively in some cases and in response to participant behaviors or questions other times.

3.2.2.2.2 Instrument

After the simulation, participants were asked to provide an open response to "How easy was it to interact with the patient?"; "Did the patient seem real? Why or why not?"; and "List the findings you identified during your assessment that led you to your diagnosis." We were interested to see which cues they noticed on their own without intervention from the observer or researchers. We used words they mentioned in their qualitative answers to come up with common categories and then aggregated those answers in the results below.

3.2.2.2.3 results

Participants open responses were classified as "easy" vs. "not easy". There were 19 participants who thought it was easy to interact with the patient, and 3 that thought it was not easy to interact with the patient. We used a Chi Square test to compare these proportions resulting in a significant difference (z=11.636, p=0.0006). More participants thought it was easy to interact with the patient compared to those who did not. Note that two participants that interacted with the sepsis child stating it was not easy and one mentioning it was difficult to hear the patient.

The sepsis patient's speech was intentionally recorded to sound lethargic. When asked if the patient seemed real, participants provided mixed responses. Of the 12 participants who interacted with the sepsis child, two said that it was "the most real I have ever seen in simulation" and indicated that it was a "great use of technology." Four particularly liked the patient's speech and answers to questions, one person indicated that the patient "had feelings," and another person mentioned the patient "felt warm [and had] mottled skin." One participant remarked on the delay in patient responses which was due to the camera/audio of the video communication system. Of the ten participants who interacted with the child in the abuse case, seven indicated that the patient's responses, behavior, and reactions seemed real, while two felt that it was not real due to the lack of a lower body. Participants were asked to list the cues they noticed that led to their diagnosis. Participants noticed multiple cues. For the 12 participants who interacted with the sepsis patient, eight mentioned lethargy or weakness (voice/attitude), five mentioned temperature, four mentioned mottled skin, three mentioned cough, two mentioned vital signs (oxygen saturation), two mentioned blue/red lips, two mentioned facial expressions, one mentioned audible wheezes, and one mentioned the respiratory rate which was provided visually and by breath sounds. For the 10 participants in the child abuse scenario, six mentioned physical trauma/wounds, six mentioned cigarette burns, five mentioned abrasions/skin lacerations/scrapes, six mentioned bruises/contusions, six mentioned fearful attitude, three mentioned mismatching story, two mentioned facial expressions, one mentioned distress/anxiety, one mentioned swelling, and one mentioned bite marks.

CHAPTER 4: PHYSICAL-VIRTUAL AGENTS FOR HEALTHCARE SIMULATION (STROKE STUDY)

The work in this chapter has been published with ACM Intelligent Virtual Agent 2018 [32] indicated by the tag **[IVA2018]** or currently in the process of being submitted to Journal of Nursing Education indicated by the tag **[JNE2019]**.

Paper Published:

Title: Physical-Virtual Agents for Healthcare Simulation

Authors: Salam Daher, Jason Hochreiter, Nahal Norouzi, Laura Gonzalez, Gerd Bruder, Gregory Welch.

Published in: Intelligent Virtual Agents, 2018.

This chapter supports **[TS 1A]** in Chapter 1.3.1: Replacing a static mannequin head with realistic dynamic visuals on a matching physical form can increase social presence, increase engagement, and improve learning.

I lead the design, development, execution, and analysis of the work described below. I acknowledge the effort of co-authors and collaborators.

4.1 Abstract

[IVA2018] Conventional Intelligent Virtual Agents (IVAs) focus primarily on the visual and auditory channels for both the agent and the interacting human: the agent displays a visual appearance and speech as output, while processing the human's verbal and non-verbal behavior as input. However, some interactions, particularly those between a patient and healthcare provider, inherently include tactile components. We introduce an Intelligent Physical-Virtual Agent (IPVA) head

that occupies an appropriate physical volume; can be touched; and via human-in-the-loop control can change appearance, listen, speak, and react physiologically in response to human behavior. Compared to a traditional IVA, it provides a physical affordance, allowing for more realistic and compelling human-agent interactions. In a user study focusing on neurological assessment of a simulated patient showing stroke symptoms, we compared the IPVA head with a high-fidelity touch-aware mannequin that has a static appearance. Various measures of the human subjects indicated greater attention, affinity for, and presence with the IPVA patient, all factors that can improve healthcare training.

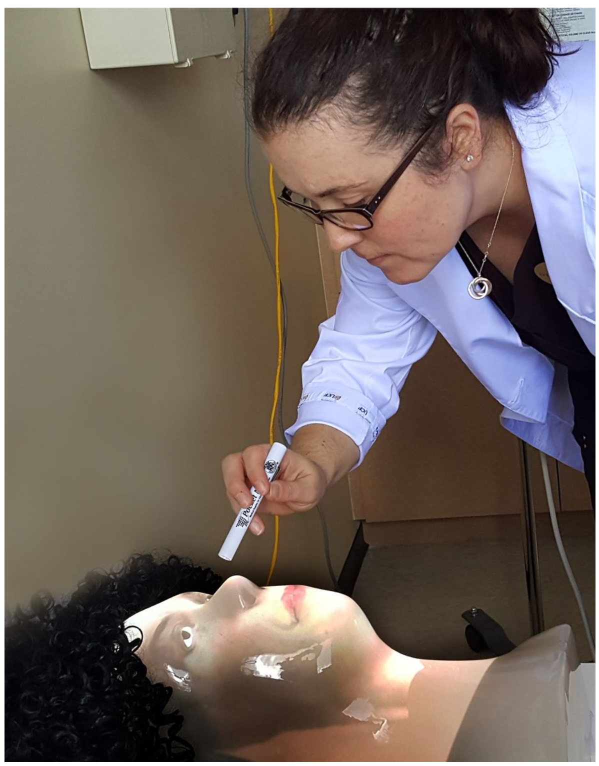


Figure 4.1: Nursing student interacting with the Intelligent Physical-Virtual Agent during a neurological assessment.

4.2 Introduction

[IVA2018] Intelligent Virtual Agents (IVAs) are commonly used in training, simulation, and education across different fields and applications, such as healthcare, military and police training, serious games, entertainment, interview training, and educator training [214]Most previous research and development efforts focused on the visual and auditory channels of communication between IVAs and users, which are arguably the most important modalities for most real-world applications. However, adding touch input and tactile feedback capabilities to an IVA has shown much potential to additionally improve the user's perception of the IVA [215–217]. Facilitating a high sense of *being* with a real person can make training with IVAs more effective, engaging, and relevant [218, 219].

In this chapter, I introduce an Intelligent Physical-Virtual Agent (IPVA) in the shape of a life-size physical head supporting dynamic imagery (see Figure 4.1). An interactive computer-generated virtual agent is projected onto the physical shell of a head. The IPVA is capable of displaying a wide variety of symptoms related to its intended use as a simulated patient in healthcare applications. In particular, for the considered simulated stroke scenario we developed and integrated appropriate visual and behavioral content. In a user study, participants performed a simulated patient stroke assessment using our IPVA, which we compared to assessment of a high-fidelity mannequin in a baseline condition. Both simulators were able to respond verbally to participants, drawing from the same finite set of responses. Though both simulators were aware of participant touch, the mannequin's reactions were limited to verbal responses, while the IPVA was additionally capable of visual feedback (Table 4.1).

In this chapter, I address the following research questions in support of thesis statement **[TS 1A]**:

Q1 Will users experience greater *social presence* with an IPVA than when interacting with a mannequin? **[IVA2018]**

Q2 Will users have a more positive *user experience* with an IPVA than when interacting with a mannequin? **[IVA2018]**

Q3 Will users rate the *communication and interaction abilities* higher for an IPVA than a mannequin? **[IVA2018]**

Q4 Will users focus their *attention* more often on an IPVA than a mannequin? **[IVA2018]**

Q5 Will users demonstrate higher *engagement* with an IPVA than when interacting with a mannequin? **[JNE2019]**

Q6 Will users indicate a higher sense of *urgency* with an IPVA than when interacting with a mannequin? **[JNE2019]**

Q7 Will users demonstrate higher *learning* with an IPVA than when interacting with a mannequin? **[JNE2019]**

In the sections below I describe the experiment, results and discussion.

4.3 Experiment

[IVA2018] In this section, I describe the development of an IPVA head as a Shader Lamps based healthcare simulator, the development of a medical training scenario with a simulated patient showing signs of a stroke, and a human-subject study that we performed to evaluate and compare this IPVA to a high-fidelity mannequin. In a between-subject study design, nursing students assessed the two simulators, and we evaluated their sense of social presence, mood, and other attributes using subjective responses and head tracking.

4.3.1 Participants

[IVA2018] Overall, 59 undergraduate nursing students (51 females, 8 males) from health assessment participated in this human-subject study. All 59 participants had been exposed to physical mannequins, 57 participants had been exposed to standardized patients (real actors), and 32 participants had been exposed to virtual simulations such as Shadow Health [34] or Second Life [70].

The study was performed as part of a nursing class where it is common practice to split students into pairs or triplets depending on the number of students and available simulators. Typically, two students interact with a patient, and the third person observes the interaction from inside the room. In this study, 44 students interacted with the simulator and 15 observed the simulation due to limited space around the patient. All students (interactors and observers) participated in the study. There is evidence that learning outcomes are not strongly affected based on whether students interacted or observed the interaction [220].

4.3.2 Scenario

[IVA2018] Prof. Gonzalez and I developed a training scenario in which a patient, called *Vera Real*, who is a woman in her 40s, shows up to the emergency department with one-sided upper and lower hemiplegia (paralysis of one side of the body). Upon examination, nurses would discover that the patient has visual loss and is complaining of a headache. The patient's speech is slurred with a one-sided facial droop. The patient's character is polite and neutral in general (neither positive nor negative). Vera is capable of showing different one-sided facial expressions when asked to do so (e.g., smile, frown, raise eyebrow). In neurological assessments, it is important to determine whether a patient can perceive touch; in particular, as Vera is experiencing a stroke, she is unable to feel and respond to touch on the affected half of her face. As part of the medical scenario, Vera starts alert and responsive, but as the examination continues, her condition deteriorates and she becomes forgetful, confused, and scared. Nursing students are expected to perform a neurological exam, recognize the stroke symptoms, and call the doctor for further instructions.

4.3.3 Apparatus Development

[IVA2018] For this experiment, we developed and adapted two setups representing this healthcare training scenario: one with a physical-virtual head described in details in chapter 3 and one with a physical mannequin head. Both conditions used the same commercial Laerdal full-body Sim-Mom mannequin and its physiology capabilities (i.e. breathing, heart sounds, and pulse) [221], whereas they differed in the type of head used (Figure 4.2). Typically in mannequin simulation, a person controls the patient's verbal responses from a control room by speaking into a microphone, and the sound comes out of the mannequin. The patient's physiology is often controlled by this same operator using a computer interface. For this study, we created a finite set of garbled verbal responses common to both conditions. The verbal responses were limited to specific information regarding the patient (e.g. name, age), her condition (e.g. "I have a headache,") and basic responses (e.g. "yes," "no"); they did not include backchannels. Using a graphical interface, the simulator controller triggered patient responses that come from speakers located below the mannequin. The same software was used in both experimental conditions to control the verbal responses of the patient. The IPVA's verbal responses were lip-synced and contained facial expressions commonly used in speech, such as occasional eyebrow raising and blinking. Table 4.1 shows a comparison between the properties and capabilities of the simulated patient head for the two study conditions.

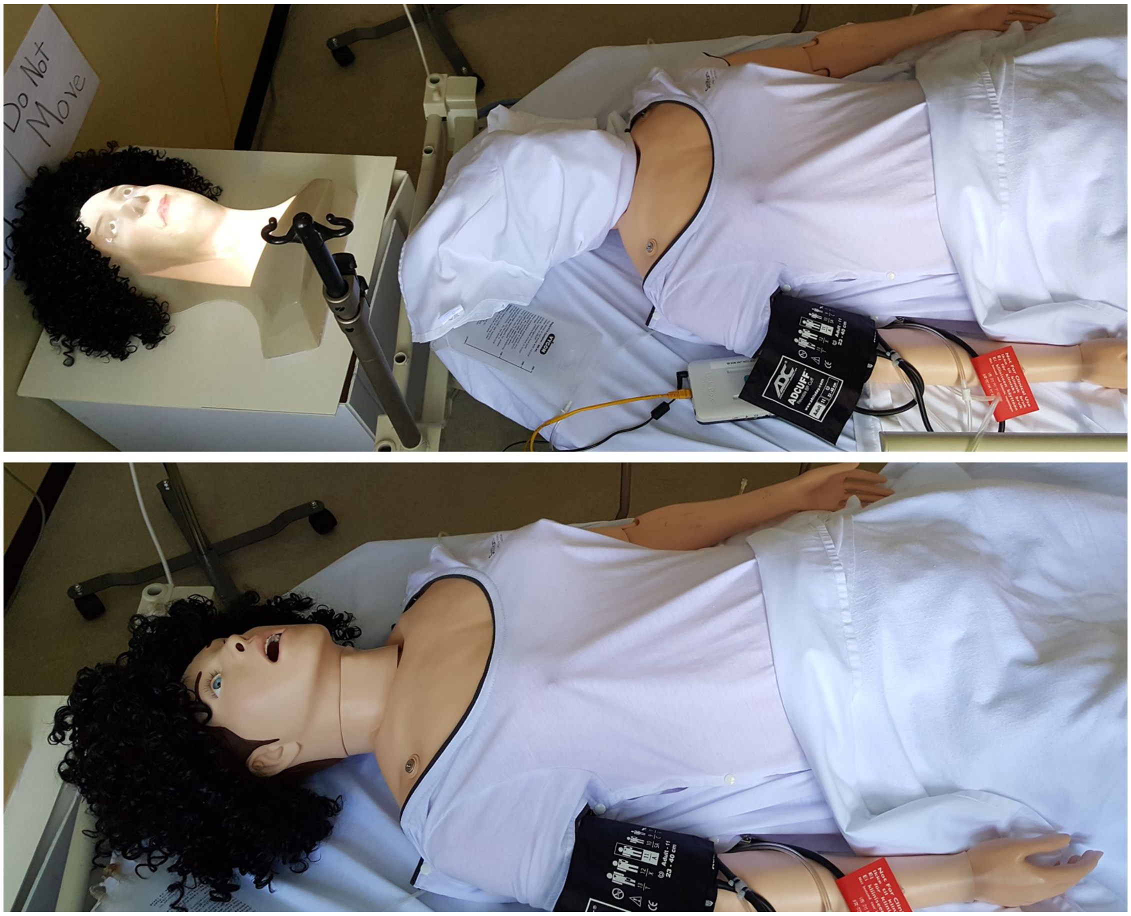


Figure 4.2: The two study conditions: IPVA (top) and Mannequin (bottom).

Table 4.1: Properties and capabilities of the simulated patient head in the IPVA and Mannequin conditions.

Property/Capability	**IPVA Head**	**Mannequin Head**
Realistic physical shape	Yes	Yes
Human-in-the-loop control	Yes	Yes
Additional operator responses	Available	Required
Touch-aware	Yes	Yes
Verbal responses	Yes	Yes
to touch	Yes	Yes
to questions	Yes	Yes
Facial appearance	Dynamic	Static
Facial expressions	Yes	No
Lip syncing	Yes	No
Eye/pupil movement	Yes	No
Visual responses	Yes	No
to touch	Yes	No
to questions	Yes	No
to light and motion	Yes	No

4.3.3.1 Physical-Virtual Agent Head

[IVA2018] To develop the IPVA head, I built a wooden rig that supports a semi-transparent plastic head-shaped shell and a projector (Figure 4.3 right). I designed an interactive 3D graphical face that matches the shape and size of the head shell. An AAXA P300 pico projector (resolution

1920x1080) projects imagery of the animated face onto the shell from below.

I scanned the shell using the photogrammetry software Autodesk ReCap. The geometry from the scan is dense with a topology that is difficult to texture and animate. In Maya, I created a 3D mesh of the head that matched the form of the shell and designed a topology appropriate for facial animation. The head is made out of one mesh for the face (2657 vertices, 2541 faces) and two meshes for the eyeballs (72 vertices, 71 faces each). Each eyeball is connected to a joint to control its movement. The vertices of the face mesh are weighted to the head and jaw joints to allow for jaw opening. Combinations of blendshapes (e.g. upper eyelid up, lower eyelid down, upper lip up, lower lip down, blink, one-sided facial droop, right smile, left smile, nose up, right eyebrow inward, both eyebrows up, right eyebrow up) are used to create facial expressions and visemes (E, A, O, FV, MBP, and variations). I imported the 3D model to Unity and designed a graphical user interface to trigger the 129 pre-recorded audio responses, which were lip-synced to the patient's lips using Rogo Digital Lipsync. Verbal responses were played on speakers below the head. The simulator controller observed the assessment from the control room and triggered the patient's verbal and facial responses with a graphical interface, whether prompted by participant speech or touch. The healthy and stroke patient models I designed for the IPVA are shown in Figure 4.3.

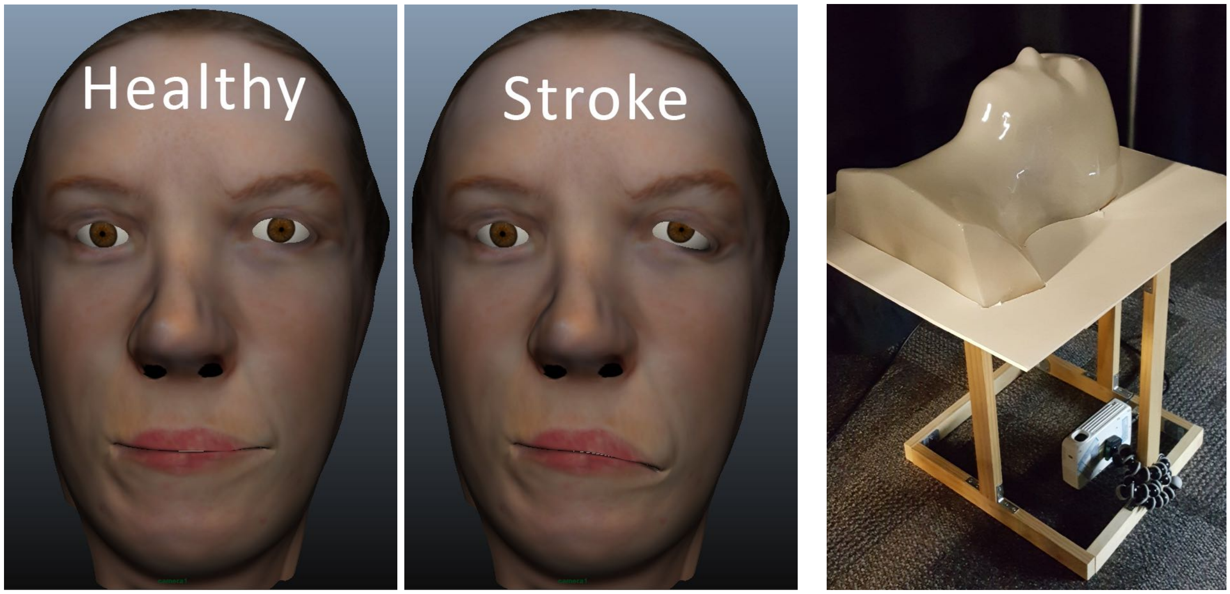


Figure 4.3: Left: virtual patients projected onto the IPVA. Right: wooden rig with projector and head-shaped shell.

4.3.3.2 Physical Mannequin Head

[IVA2018] The control condition consisted only of the SimMom mannequin, including the full body and the head. As in the experimental condition, the simulator controller observed the participants and triggered appropriate verbal responses from the patient whether prompted by participant speech or touch interactions. The mannequin was not capable of portraying certain visual symptoms, such as facial asymmetry; to obtain information related to such symptoms, participants would verbally state their question and receive a verbal response from the simulator controller.

4.3.4 Study Design and Procedure

[IVA2018] In a between-subject design, 26 participants interacted and 10 observed the patient in the control condition (Mannequin), and 18 interacted and 5 observed the patient in the experimental condition (IPVA). The Institutional Review Board approved this human-subject study. The study was conducted over 4 class sessions, with one condition operational per session due to setup time; participants were assigned to conditions based on their date of attendance. First, all participants were asked to fill out a demographics pre-questionnaire. All participants were already familiar with the capabilities of the mannequin, and those who interacted with the IPVA watched a two-minute video to familiarize them with its capabilities. The video featured a healthy agent with a non-slurred voice in an interaction with a healthcare provider that highlighted the capabilities of the IPVA, including verbal behavior, non-verbal behavior (such as facial expressions, eyeball movement following an object, pupillary reaction to light), and response to touch. Participants were given the patient's history and asked to perform a physical assessment of the patient. The simulation started when the controller said "begin simulation" and had a time limit of 15 minutes. After the simulation, participants were asked to answer a post-questionnaire with qualitative and quantitative sections.

For both conditions, the controller observed the simulation from a separate control room and used the speakers in the simulation room to inform participants about the start of the simulation and any patient actions (or inactions) that the simulator was (or was not) capable of representing (e.g., "begin simulation," "patient cannot move this arm," "one-sided smile"). When the participant touched the patient and asked if she could feel the touch, she would respond accordingly depending on the side (stroke vs. non-stroke). Participants had to actively inquire to receive certain cues; for example, in the Mannequin condition, the participant had to ask the patient to smile in order to receive the answer "asymmetric smile." In the IPVA condition, if the participant asked the patient to smile, they received their response directly as visual feedback by observing the patient; if the participant asked the patient to move her arm (part of the mannequin's body) they received a response from the controller (e.g., "patient cannot move right arm").

4.3.5 Measures

[JNE2019] Before the simulation, a pre-questionnaire containing demographics questions was asked, as well as a pre-simulation test: "When performing a neurological assessment what are all the potential findings you can remember?" to capture what they remember from class before the simulation. **[IVA2018]** During the simulation, participants were video recorded and tracked using a Kinect. After the simulation, participants answered a questionnaire with qualitative and quantitative questions. Observers completed the same questionnaires but indicated that they did not actually interact with the patient. **[JNE2019]** At the end of the semester they were asked the same pre-simulation question in an optional extra credit question for their Final Exam. Considering only participants who answered both the pre-simulation test and post-simulation test, there were 13 people in the IPVA group and 12 people in the Mannequin group. For each participant, both the pre-simulation and post-simulation answers were parsed for answers that indicate knowledge in the following categories: Facial (droop, ptosis, asymmetry, facial expressions), Tongue, Pupils, Touch, Slurred Speech, and all the categories combined to represent an overall score. The categories focused on the head since that was the part of the simulator that was different between the experimental conditions. If the participant mentioned one of more words that corresponds to one category, they get a score of 1 for that category. If they do not mention any word related to that category they get a score of zero. The scores of participants in the same pair or triplet were aggregated to get a team score. The difference between the pre-simulation and post-simulation test was computed for each category keeping track if that score shows an overall "Progress", "No Change", or "Regress" for that team. The scores between the teams are compared between the IPVA and the

Mannequin for each of “Progress”, “No Change” and “Regress”. **[IVA2018]** We measured realism by asking specific questions about the avatar, such as facial expressions and animations, and we measured social presence using a modified questionnaire from Harms and Biocca [164] shown in Table 4.2. Two questions from the original Harms and Biocca’s perceived behavioral interdependence (Bhv) questionnaires were omitted as they did not fit in this patient-provider case. In addition, participants were asked to complete the affective attraction (AffAtt) questionnaire [222], shown in Table 4.3. As a gauge of their mood during the simulation, participants were asked “How did your interaction with the patient (Vera) make you feel?” with respect to feeling anxious, excited, tense, alert, in control, and having a desire to leave the situation, each as a Likert scale from “not at all” (0) to “extremely strong” (10). **[JNE2019]** Also, participants were asked to complete the Virtual Patient Design Evaluation (VPEval) questionnaire [223]. Two of the 4 domains identified by Huwendiek were used: (1) authenticity of patient encounter and (2) cognitive strategies on the consultation. The other two domains did not apply in our case. This questionnaire is used to evaluate virtual patients and to ensure designs enable critical thinking of learners is shown in Table 4.5. **[IVA2018]** Finally, participants were also asked miscellaneous questions about the simulator’s touch/response interaction, communication abilities, sense of urgency, and closeness to a real patient shown in Table 4.4. The touch response interactions were achieved by the controller using a GUI to trigger pre-recorded responses.

Table 4.2: Social presence questions (scale from 1 to 7).

	Co-Presence (CoP)
1	I noticed the patient (Vera).
2	The patient (Vera) noticed me.
3	The patient (Vera)'s presence was obvious to me.
4	My presence was obvious to the patient (Vera).
5	The patient (Vera) caught my attention.
6	I caught the patient (Vera)'s attention.
	Attentional Allocation (Att)
1	I was easily distracted from the patient (Vera) when other things were going on outside this room.
2	The patient (Vera) was easily distracted from me when other things were going on outside this room.
3	I remained focused on the patient (Vera) throughout our interaction.
4	The patient (Vera) remained focused on me throughout our interaction.
5	The patient (Vera) did not receive my full attention.
6	I did not receive the patient (Vera)'s full attention.
	Perceived Message Understanding (Msg)
1	My thoughts were clear to the patient (Vera).
2	The patient (Vera)'s thoughts were clear to me.
3	It was easy to understand the patient (Vera).
4	The patient (Vera) found it easy to understand me.
5	Understanding the patient (Vera) was difficult.
6	The patient (Vera) had difficulty understanding me.
	Perceived Affective Understanding (Aff)
1	I could tell how the patient (Vera) felt.
2	The patient (Vera) could tell how I felt.
3	The patient (Vera)'s emotions were not clear to me.
4	My emotions were not clear to the patient (Vera).
5	I could describe the patient (Vera)'s feelings accurately.
6	The patient (Vera) could describe my feelings accurately.
	Perceived Emotional Interdependence (Emo)
1	I was sometimes influenced by the patient (Vera)'s moods.
2	The patient (Vera) was sometimes influenced by my moods.
3	The patient (Vera)'s feelings influenced the mood of our interaction.
4	My feelings influenced the mood of our interaction.
5	The patient (Vera)'s attitudes influenced how I felt.
6	My attitudes influenced how the patient (Vera) felt.
	Perceived Behavioral Interdependence (Bhv)
1	My behavior was often in direct response to the patient (Vera)'s behavior.
2	The behavior of the patient (Vera) was often in direct response to my behavior.
3	The patient (Vera)'s behavior was closely tied to my behavior.
4	My behavior was closely tied to the patient (Vera)'s behavior.

Table 4.3: Affective attraction questions (scale from 1 to 7).

	Affective Attraction (AffAttr)
1	How unpleasant/pleasant do you feel about the patient (Vera)?
2	How cold/warm do you feel about the patient (Vera)?
3	How negative/positive do you feel about the patient (Vera)?
4	How unfriendly/friendly do you feel toward the patient (Vera)?
5	How distant/close do you feel to the patient (Vera)?

Table 4.4: Miscellaneous questions (scale from 1 to 7).

	Miscellaneous
1	Rate the simulator's touch/response interaction.
2	From this interaction, rate the simulated patient's ability to communicate with you.
3	From this interaction, rate your ability to communicate with the simulated patient.
4	When assessing the head, did it provoke a sense of urgency?.
5	How close to a real patient did the patient feel?

Table 4.5: Virtual Patient Design Evaluation questions (scale from 1 to 7).

VPEval	
VPEval 1	While working on this case, I felt I had to make the same decisions a doctor/nurse would make in real life.
VPEval 2	While working on this case, I felt as if I were the doctor/nurse caring for this patient.
VPEval 3	While working through this case, I was actively engaged in gathering the information (e.g. history questions, physical exams, lab tests) I needed to characterize the patient's problem.
VPEval 4	While working through this case, I was actively engaged in revising my initial image of the patient's problem as new information became available.
VPEval 5	While working through this case, I was actively engaged in creating a short summary of the patient's problem using medical terms.
VPEval 6	While working through this case, I was actively engaged in thinking about which findings supported or refuted each diagnosis in my differential diagnosis.

4.4 Results

[IVA2018 and JNE2019] We used non-parametric statistical tests (Mann-Whitney U) to analyze the Likert-scale ordinal data from the questionnaires comparing the Mannequin condition with the IPVA condition. In some scientific disciplines it is common practice to treat Likert-type scales as interval-level measurements [224]. We avoid the discussion on whether parametric statistics can be a valid method for the analysis of non-parametric data [225, 226] by using non-parametric tests.

4.4.1 Subjective Questionnaires

[IVA2018]

4.4.1.1 *Social Presence*

Results show a significantly higher or a trend for higher social presence for the IPVA group in multiple dimensions (Figure 4.4), supporting our research question Q1.

Participants were more likely to have a higher perceived message understanding (Msg) in the IPVA group ($M = 5.342$, $SD = 0.861$) than in the Mannequin group ($M = 4.514$, $SD = 1.094$). The difference is statistically significant ($W = 117.0$, $p = 0.006$).

Participants were more likely to have a higher perceived behavioral interdependence (Bhv) in the IPVA group ($M = 4.150$, $SD = 1.375$) than in the Mannequin group ($M = 3.250$, $SD = 1.103$). The difference is statistically significant ($W = 143.0$, $p = 0.032$).

There is a trend suggesting participants could be more likely to have a higher co-presence (CoP) in the IPVA group ($M = 5.192$, $SD = 1.090$) than in the Mannequin group ($M = 4.420$, $SD = 1.334$). The difference is very close to being significant ($W = 150.0$, $p = 0.052$). Similarly, the attentional allocation (Att) shows a higher trend for the IPVA group. These results could become significant with a larger sample.

The remaining dimensions of social presence, perceived affective understanding (Aff) and perceived emotional interdependence (Emo), did not show a statistically significant difference ($p > 0.05$) between the groups.

4.4.1.2 *Affective Attraction*

[IVA2018] Participants felt a higher affective attraction (AffAtt) towards the patient in the IPVA group ($M = 4.780$, $SD = 1.180$) than in the Mannequin group ($M = 4.157$, $SD = 1.379$) (Figure 4.4). The difference is statistically significant ($W = 146.0$, $p = 0.043$).

4.4.1.3 *Mood Rating*

[IVA2018] Results for the mood questionnaires are shown in Figure 4.5. Participants felt more "in control of the situation" in the IPVA group ($M = 5.650$, $SD = 2.323$) than in the Mannequin group ($M = 4.261$, $SD = 2.200$). The difference is statistically significant ($W = 142.0$, $p = 0.031$), supporting our research question Q2.

There is a trend suggesting participants in the IPVA group could more likely feel more excited and more alert than those in the Mannequin group. The difference is very close to being significant (excited: $W = 154.5$, $p = 0.064$; alert: $W = 155.0$, $p = 0.065$) and could actually be significant with a larger sample.

The rest of the mood rating measures did not show a statistically significant difference ($p > 0.05$) for feeling anxious, tense, or a desire to leave the situation.

4.4.1.4 *Touch Response and Communication*

[IVA2018] Results for the touch response and communication questions are shown in Figure 4.6. Together, they support our research question Q3.

Participants were more likely to rate the IPVA's touch/response higher in the IPVA group ($M = 5.054$, $SD = 1.261$) than in the Mannequin group ($M = 3.294$, $SD = 1.750$). The difference is statistically significant ($W = 169.5$, $p < 0.001$). This is in spite of the fact that the touch response was achieved identically for both conditions.

Participants were more likely to rate the simulated patient's ability to communicate with them higher in the IPVA group ($M = 5.565$, $SD = 0.945$) than in the Mannequin group ($M = 4.559$, $SD = 1.599$). The difference is statistically significant ($W = 239.0$, $p = 0.011$).

Participants were more likely to rate their ability to communicate with the simulated patient higher in the IPVA group ($M = 5.652$, $SD = 0.647$) than in the Mannequin group ($M = 4.824$, $SD = 1.527$). The difference is statistically significant ($W = 258.0$, $p = 0.025$).

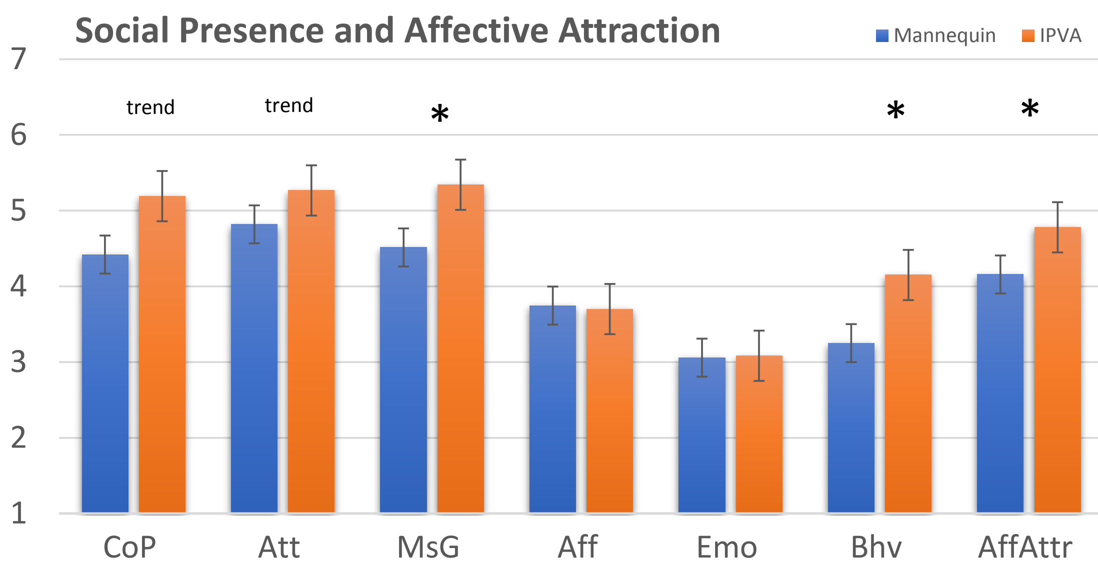


Figure 4.4: Social presence and affective attraction results.

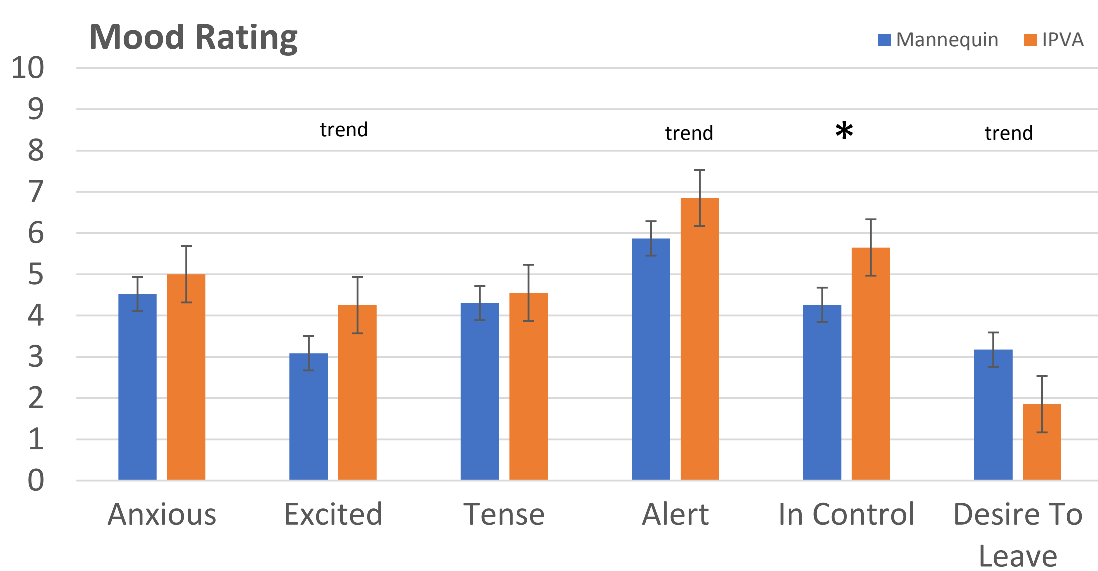


Figure 4.5: Mood rating results.

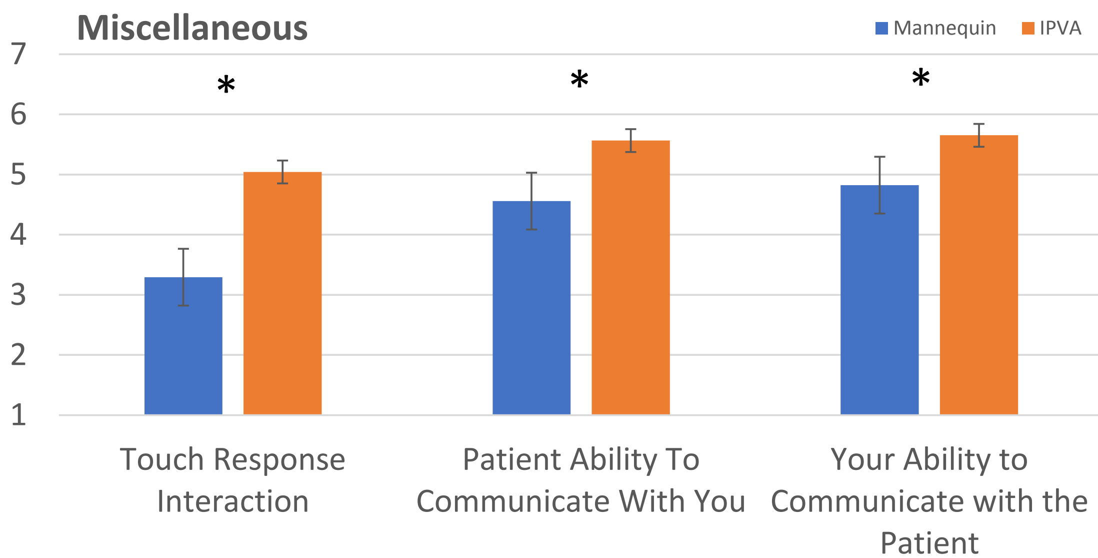


Figure 4.6: Miscellaneous results.

4.4.1.5 Realism

[IVA2018] Participants were asked to rate different aspects of the IPVA's realism on a 1 to 6 scale, where 1 represents the most "inexpressive" and 6 represents the most "expressive." Participants also had the option of "N/A" for questions that were not applicable for the Mannequin condition. Realism questionnaire results were all statistically significant ($p < 0.01$) and are shown in Figure 4.7.

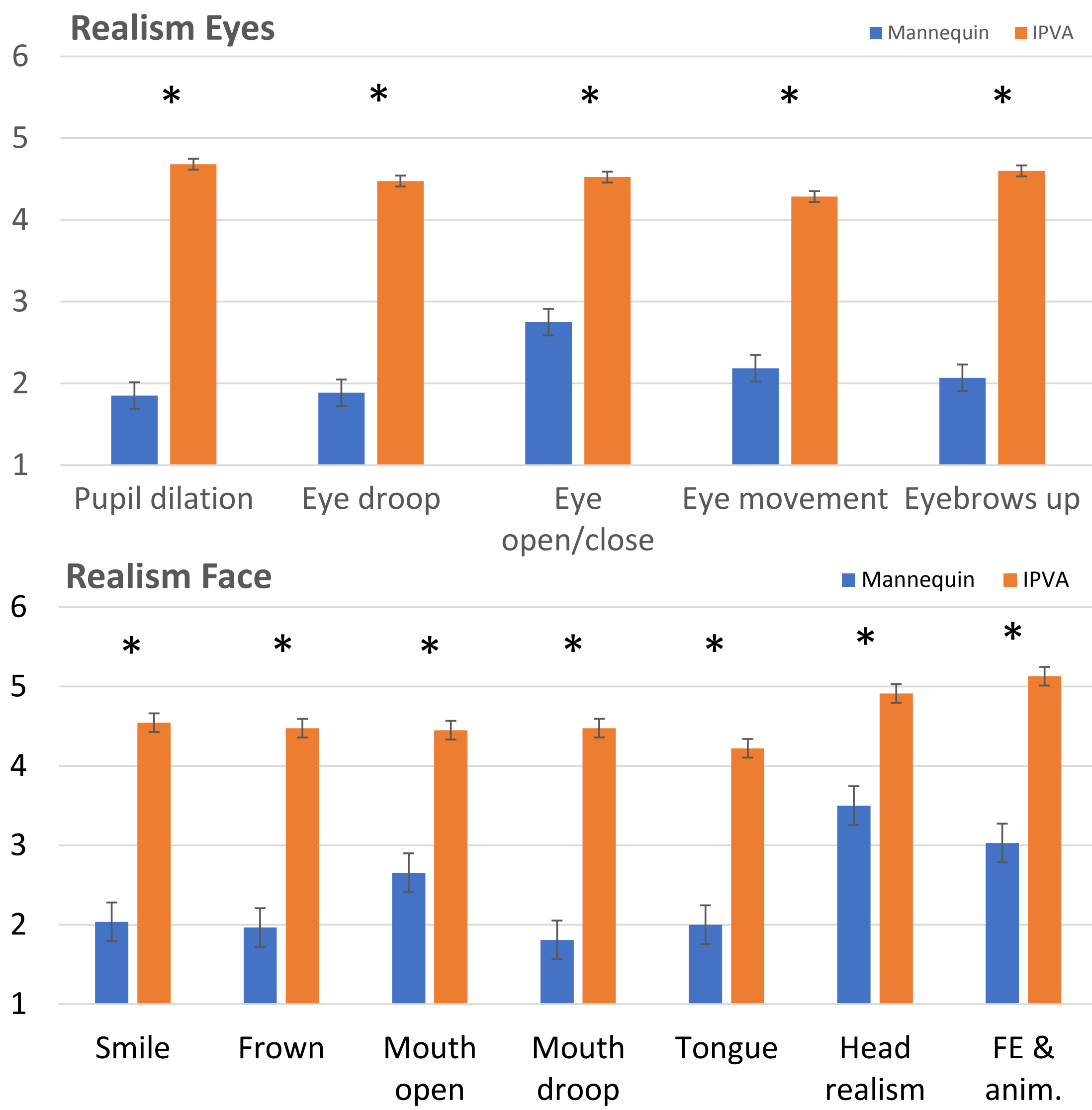


Figure 4.7: Realism results.

4.4.1.6 Urgency

[JNE2019] Results from the Mann-Whitney U test indicate a higher sense of urgency ($W = 207.50$, $p = 0.002$) for participants in the IPVA condition ($M = 4.696$,$SD = 1.636$) compared to participants in the Mannequin condition ($M = 3.294$,$SD = 1.360$).

4.4.1.7 *Closeness to Patient*

[JNE2019] Results from the Mann-Whitney U test indicate the simulated patient felt closer to a real patient ($W = 271.0$, $p = 0.046$) for participants in the IPVA condition ($M = 3.957$,$SD = 1.186$) compared to participants in the Mannequin condition ($M = 3.265$,$SD = 1.355$). (See Figure 4.8).

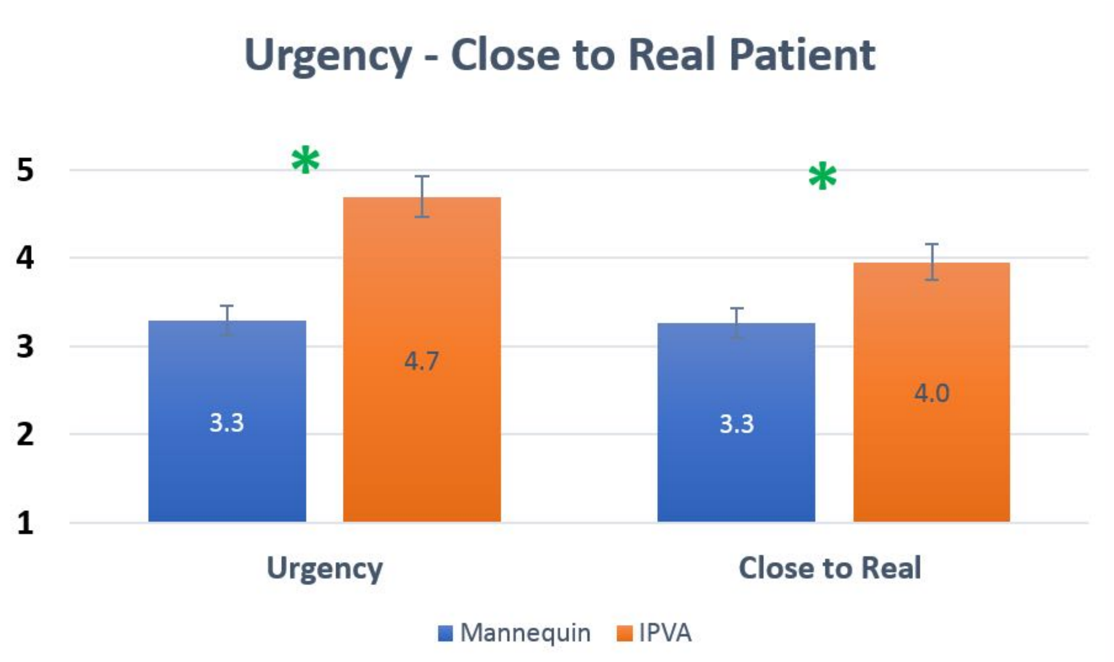


Figure 4.8: Urgency and feeling the patient is close to a real patient results.

4.4.1.8 *Engagement in Learning*

[JNE2019] In general participants in the IPVA condition had significantly or trend close to being significant for higher VPEval scores compared to participants in the Mannequin condition (See Figure 4.9). Mann-Whitney U test shows: There was a significant difference ($W = 228.00$, $p = 0.038$) where the IPVA participants felt they had to make the same decisions a doctor/nurse would make in real life ($M = 5.750$, $SD = 1.251$) more than the participants with the Mannequin (M

=5.059 , SD = 1.347).

There was trend close to being significant ($W = 239.50$, $p = 0.067$) where participants felt more as if they were the doctor/nurse caring for the IPVA patient ($M = 5.600$, $SD = 1.569$) than with caring for Mannequin ($M = 4.941$, $SD = 1.413$)

There was a significant difference ($W = 232.50$, $p = 0.049$) where the IPVA participants were more actively engaged in gathering the information (e.g. history questions, physical exams, lab tests) they needed to characterize the patient's problem ($M = 5.700$, $SD = 1.658$) than the participants with the Mannequin ($M = 4.853$, $SD = 1.708$)

There was a significant difference ($W = 217$, $p = 0.025$) where the IPVA participants were more actively engaged in revising their initial image of the patient's problem as new information became available ($M = 6.000$, $SD = 1.214$) than the participants with the Mannequin ($M = 5.059$, $SD = 1.516$)

There was a trend close to being significant ($W = 233.50$, $p = 0.052$) where the IPVA participants were more actively engaged in creating a short summary of the patient's problem using medical terms ($M = 5.700$, $SD = 1.261$) than the participants with the Mannequin ($M = 4.912$, $SD = 1.443$)

There was a significant difference ($W = 197$, $p = 0.009$) where the IPVA participants were more actively engaged in thinking about which findings supported or refuted each diagnosis in their differential diagnosis ($M = 5.900$, $SD = 1.165$) than the participants with the Mannequin ($M = 4.971$, $SD = 1.243$).

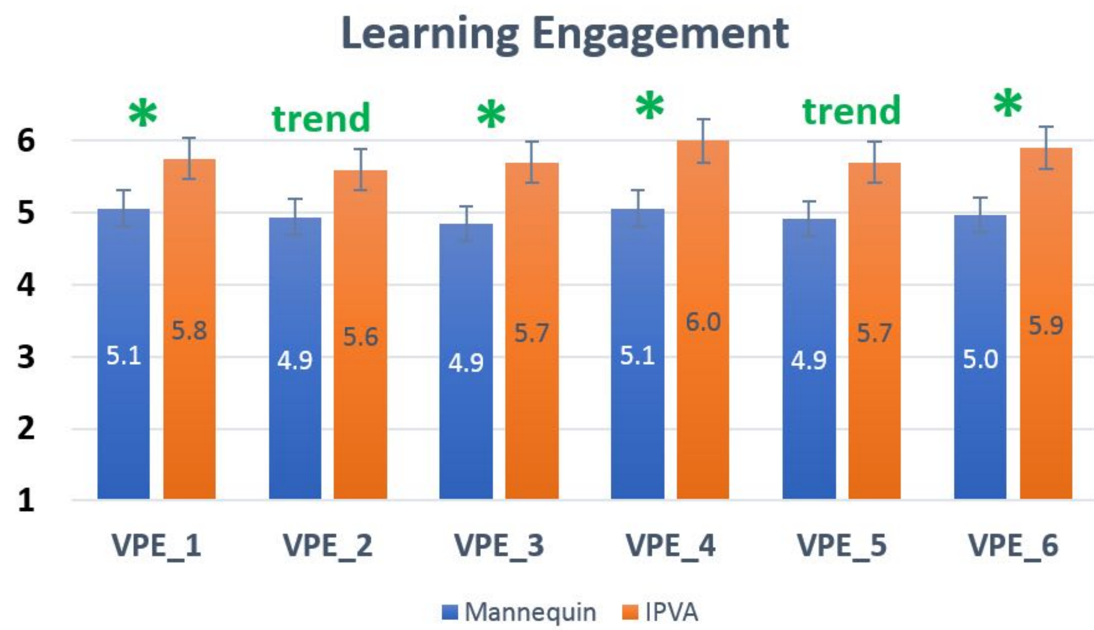


Figure 4.9: Engagement in learning results.

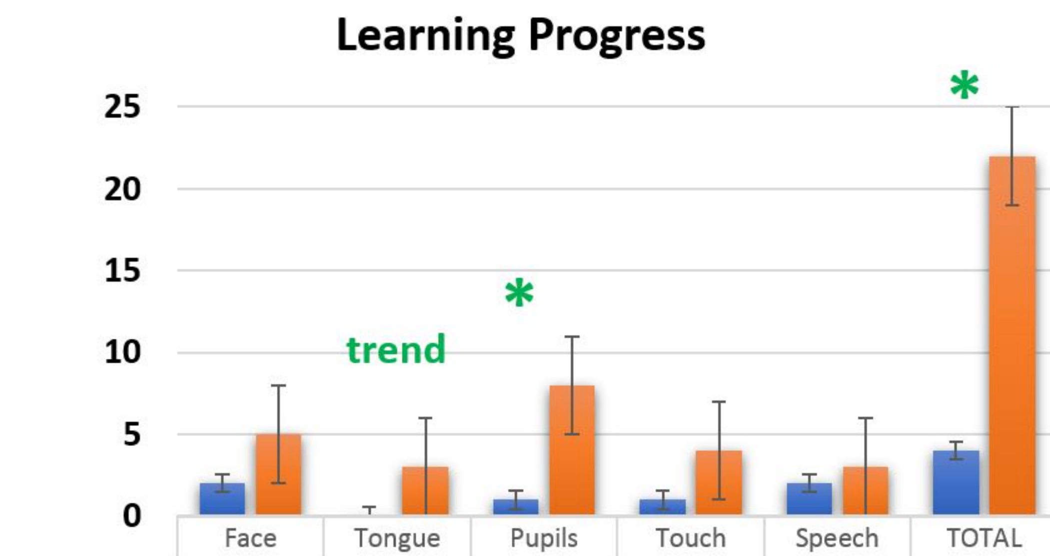
4.4.2 *Objective Learning*

[JNE2019] When using aggregate data for information retained, we compared the results of pre-simulation answers and post-simulation answers using the Chi-Square statistical test.

There was more overall knowledge acquired and retained ($z = 11.081$, $p = 0.0009$) when the participants were in the the IPVA conditions (score of 22 for 13 participants) compared to participants in the Mannequin condition (score of 4 for 12 participants).

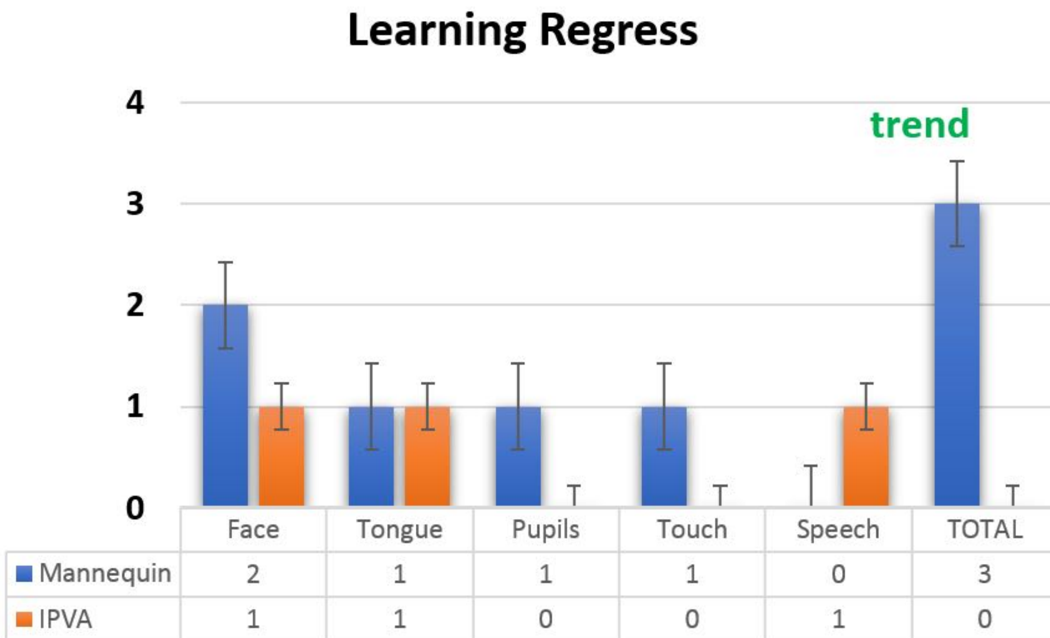
Looking further into the sub-categories for knowledge retained, participants retained and acquired more knowledge in the Pupils category ($z = 4.907$, $p = 0.027$) in the IPVA condition (score of 8 for 13 participants) compared to the Mannequin condition (score of 1 for 12 participants). We observed a trend close to being significant for the Tongue category where participants retained and acquired more knowledge ($z = 2.768$, $p = 0.096$) in the IPVA condition (score of 3 for 13 participants) compared to the Mannequin condition (score of 0 for 12 participants). There was no significant difference in the progress for the individuals sub-categories: Face, Touch, and Slurred Speech (See Figure 4.10).

There was a trend close to being significant ($z = 3.25$, $p = 0.0714$) with participants in the Mannequin group are more likely to regress (i.e. forget something they once knew) (3 out of 12) compared to participants in the IPVA group (0 out of 13). There was no significant difference in the other subgroups (See Figure 4.11).



	Face	Tongue	Pupils	Touch	Speech	TOTAL
Mannequin	2	0	1	1	2	4
IPVA	5	3	8	4	3	22

Figure 4.10: Objective learning progress results.



	Face	Tongue	Pupils	Touch	Speech	TOTAL
Mannequin	2	1	1	1	0	3
IPVA	1	1	0	0	1	0

Figure 4.11: Objective learning regress results.

4.4.3 Head Tracking

[IVA2018] Human gaze behavior has been extensively studied as an objective measure of a person's focus of attention [227–229]. Head orientation provides a less obtrusive but less accurate measurement of attention; several researchers have investigated estimating head pose and orientation from video recordings, including in the context of social interactions during group meetings [190, 191, 230–233]. As such, we tracked the participants' head positions and orientations using a Microsoft Kinect sensor to understand their head movement behavior and visual attention. From this data, we measured the amount of time participants spent facing Vera to compare how their attention varied between the two conditions. We only tracked the heads of interacting participants, not observers.

We collected 3D reference points in the simulation environment corresponding to Vera's body and head. Using these points, we constructed cubic volumes representing various regions of interest: Vera's head, Vera's body, and everywhere else. Intersections of head orientation vectors and these regions represent participant attention throughout the simulation. Using the Kruskal-Wallis test, we observed significant differences indicating that participants spent a greater portion of simulation time focusing on the head of the IPVA than on the head of the mannequin ($H = 6.0208, p = 0.014$). There were no significant differences between the two conditions regarding participant attention to the mannequin's body ($H = 0.3333$, $p = 0.56$) or to everywhere else ($H = 2.5208$, $p = 0.11$).

4.4.4 Qualitative Feedback

[IVA2018] Participants (both interactors and observers) were asked to answer the following question by free writing: "Based on your assessment session, what do you think about this simulated patient (specifically the head)?"

4.4.4.1 Mannequin

[IVA2018] Many of the participants (17 of 36) who assessed the mannequin felt it was difficult to perform a neurological assessment due its static appearance. Six of the qualitative responses mentioned the absence of facial expressions, and 9 mentioned the lack of eye and pupil movement. In terms of realism, 4 participants commented on the mannequin's general inability to "reenact neurological signs." Other participant responses indicated that mannequin's static appearance hindered the "realistic aspect of the scenario," that the lack of realism "[took me] out of my element of treating a real patient," and that the limited interaction capabilities prevented them from "physically [seeing] the abnormal." Three participants complained about the delayed responses. Seven of the participants had positive experiences with the mannequin, with several finding the patient and its responses to be realistic. One felt it was “as realistic as it could be,” and another suggested it was “fine for this test.”

4.4.4.2 IPVA

[IVA2018] Of the participants in the IPVA condition, 9 out of 23 “liked it,” and 3 indicated that it was “easy” or “helpful for neuro,” with 2 highlighting the facial expressions and pupils. Eight commended the realism, and 5 appreciated the head’s interactivity and ability to answer questions (e.g. “more interactive and reactive to prompts,” “more lifelike,” “it could sense touch,” “much better interaction with a face that can do things”). Eleven specifically mentioned that it was more realistic or more helpful than traditional mannequins. Three noted it was “slightly creepy” or “a little scary.” Five complained that the responses were “a bit laggy, slowing down assessment time.” Three did not like that the head was separate from the body, expressing a preference to see a full-body version. Three commented negatively on realism: the “quality of [the] face is poor. Should reflect [an] actual person,” the “tongue was a bit awkward,” and it is “not very lifelike but an improvement from the mannequins.”

4.5 Discussion

[IVA2018] Life-size touch-aware interactive virtual patients that occupy volume are desirable and can be effective healthcare training simulators, such as for assessing patients with stroke. Unsurprisingly, the IPVA was significantly more realistic than the mannequin; we measured realism to have a benchmark compared to the mannequin. Nursing students who interacted with the IPVA showed significantly higher Message Understanding, rating the communication abilities of the IPVA higher than the mannequin in terms of both of the patient’s ability to communicate with them and their ability to communicate with the patient. The perceived behavioral interdependence was also significantly higher for the IPVA. We think these results are due to the IPVA's ability to indicate understanding and respond to participant behavior; for example, the IPVA can respond verbally and visually to participant actions, including verbal requests and physical touches, whereas the mannequin was only capable of providing verbal responses.

The trend towards higher co-presence for the IPVA group could be due to the simulator's greater ability to show environmental awareness and react to participant behavior compared to the mannequin. For example, the IPVA can react to a participant's touch, exhibit pupillary movement in response to light, and display facial expressions in response to participant behavior. While the attentional allocation was not significantly different between groups, it was close enough to support head tracking results. The fact that the IPVA did not rotate towards the participants could have affected the results of the attentional allocation. With a larger sample, the co-presence and attentional allocation could show significance. Interestingly, touch interaction was rated higher for the IPVA than the mannequin even though the responses were controlled equivalently by an observer pressing buttons. It seems that the addition of visuals during the lip-sync of the verbal responses might have made it appear that the IPVA was more responsive to touch. We were not surprised that there was no difference in the Perceived Emotional Interdependence and Perceived Affective Understanding between conditions as the patient was designed to be neutral. The Affective Attraction questionnaire measured how participants felt toward the patient. The IPVA was more likely to be considered "pleasant, warm, positive, friendly, close" than the mannequin.

By design, the patient deteriorated cognitively during the simulation, starting from being alert and oriented to becoming confused, forgetful, and disoriented, which was demonstrated verbally and was equivalent in both conditions. While the patient's outcome was out of the control of participants, those who assessed the IPVA felt more "in control of the situation" than those who assessed the mannequin. The IPVA provided participants with more direct control over visual and tactile assessment through integrated graphics, while assessment of the mannequin was indirect, requiring participants to explicitly request information and wait for a verbal response from the simulation operator. The increased perception of being in control might have increased participants' social presence in the perceived behavioral interdependence dimension, which measures the degree to which patient actions were directly affected by participant actions.

There is a trend suggesting participants in the IPVA condition felt more alert and more excited. The excitement could be attributed to the novelty of the simulator, while the alertness could be attributed to the fact that the IPVA eyes could blink and look at them.

4.5.1 Limitations

[IVA2018] We designed this study with a consistent finite set of verbal responses between the Mannequin and IPVA conditions. Performing patient eye movement in response to medical tests in real time while observing participants was challenging; automating this capability of the patient is recommended. The IPVA's eyes and pupils could move, but the head could not physically rotate to face the participants.

4.6 Conclusion

[IVA2018] In this chapter, I described the development and evaluation of an Intelligent Physical-Virtual Agent (IPVA) head for neurological simulation. Participants were split into a group that assessed a high-fidelity healthcare mannequin and a group that assessed the IPVA with the mannequin body. In both groups, the simulator occupied space, allowing participants to touch the patient. The addition of interactive realistic visuals on the IPVA head resulted in higher social presence and Affective Attraction compared to the mannequin, supporting our research question

Q1. Participants felt more in control of the situation when interacting with the IPVA, supporting our research question Q2, and they rated the IPVA's touch response and communication abilities higher than those of the mannequin, supporting our research question Q3. Head tracking results indicated that participants in the IPVA condition spent a greater percentage of assessment time looking at the patient's head than those in the Mannequin condition, supporting our research question Q4. **[JNE2019]** Questionnaires results show that the participants were more engaged when they interacted with the IPVA compared to interacting with the mannequin supporting our research question Q5. The participants felt the patient was more like a real patient, and had higher sense of urgency when they interacted with the IPVA compared to interacting with the mannequin supporting our research question Q6. Objective results show that the participants learned more when they interacted with the IPVA compared to interacting with the mannequin supporting our research question Q7. **[IVA2018]** Adding a touch sensation to IVAs could enhance the user's perception towards them, which could lead to better training outcomes. Research questions Q1 to Q6 support **[TS 1A]**

Future work involves allowing for more head and neck rotations for the patient, whether virtual or physical. Also, many participants noted that the head in the IPVA condition was separate from the rest of the patient's body and expressed interest in seeing the same technology extended to a full body. We plan to apply these same principles to a full body physical-virtual simulator. We are also interested in incorporating low-latency touch input through automated touch sensing.

CHAPTER 5: THE IMPORTANCE OF MATCHED IMAGERY AND SHAPE FOR EMBODIED VIRTUAL AGENTS IN HEALTHCARE EDUCATION (CUE MANIPULATION STUDY)

The work in this chapter is under review:

Title: The Importance of Matched Imagery and Shape for Embodied Virtual Agents in Healthcare Education

Authors: Salam Daher, Jason Hochreiter, Nahal Norouzi, Ryan Schubert, Gerd Bruder, Laura Gonzalez, Mindi Anderson, Desiree Diaz, Juan Cendan, Greg Welch. **Under Review:** IEEE Virtual Reality

It supports:

[TS 1B]: Separating the realistic dynamic visuals from the matching physical form decreases the perception of realism.

[TS 2]: Fidelity of the Physical Shape Can Affect the Cognitive Load

I lead the design, development, execution, and analysis of the work described below. I acknowledge the effort of collaborators from College of Nursing and College of Medicine, professors in the lab, and my lab-mates for all their efforts without which this experiment would not have been possible.

5.1 Abstract

Embodied virtual agents serving as patient simulators are widely used in medical training scenarios, ranging from physical patients, such as mannequins and trained human actors, to virtual patients presented via virtual and augmented reality technologies. Physical-virtual patients are

a hybrid solution that combines the benefits of dynamic visuals integrated into a human-shaped physical form that can also present other cues, such as pulse, breathing sounds, and temperature. More generally, the shape of the physical form does not need to match that of the human, and the dynamic visuals do not need to match the physical form. In fact patient simulators are sometimes human-shaped and sometimes flat, while the visuals are sometimes co-located and sometimes displayed separately. To asses the impacts on user perception, cognitive load, and behavior, we carried out a human-subject study employing graduate nursing students in pediatric patient simulations comprising the four conditions associated with matching/not of the visuals and shape. Our results show that participants preferred co-located dynamic visuals (matched visuals) compared to separated cues. When the dynamic visuals were displayed on a flat TV screen (not matched shape) separate from the simulator (not matched visuals), participants indicated that the TV was perceived as the locus of the patient rather than the simulated child lying in front of them. This effect was reduced when the simulator had a physical human form (matched shape). Finally, when participants assessed a simulator with human form (matched shape) as opposed to a flat TV, they reported a higher cognitive load as expected, and behaved more realistically in terms of standing location.

Figure 5.1: A human subject examining a virtual child in a medical pediatrics training scenario associated with a controlled study. The face of the user, who is wearing eye-tracking gear, is blurred to hide their identity. Here, the physical-virtual patient simulator comprises separate dynamic visuals on a TV screen (left), with heat and pulse sensations integrated into a three-dimensional mannequin-like shell with static visuals (below). In this condition the shape is matched—the shell is the same shape as the embodied virtual agent (a child), and yet the imagery is not matched—the dynamic visuals of the child are separated from the child-shaped shell.

5.2 Introduction

An *embodied virtual agent* is a computer agent with a human face, hands, and voice affecting verbal and non-verbal communication [234]. Embodied virtual agents can include human-in-the-loop controls to augment or simulate the computer agent. For example, human patient simulators in the form of robotic physical mannequins play a crucial role in the hands-on training of nurses and physicians. Much progress has been made over the last few years in improving the physical realism of these simulators, incorporating robotics, moulage, blood, body heat, and pulse, depending on the simulator's training purposes [102, 235–237]. However, the visual realism of these simulators remains comparatively low, in particular related to dynamic visuals, e.g., when a nurse is trained on how to talk to and assess a patient's facial and bodily appearance and behavior. For this purpose, augmented and mixed reality technologies have much potential in improving the combined visual and physical training realism and, as a result, the effectiveness of the medical training.

Related research explored augmented reality (AR) head-mounted displays (HMDs), such as the Microsoft HoloLens, but found multiple limitations, including restricted fields of view and potential ergonomic or usability issues due to instrumentation. Furthermore, compared to projector-based spatial augmented reality (SAR [124]), AR HMDs are less suitable for multi-user scenarios. An important observation is that it is essential for medical training scenarios to maintain a haptic physical platform, an inherent benefit of physical-virtual patient simulators, which integrate visual displays and haptic feedback with a physical human form [44, 88, 95, 99]. For practical applications with such patient simulators in medical training and education, it is still an open question how to best incorporate dynamic visuals into the simulation to elicit realistic perceptions, cognitive load, and behavior, and to foster natural responses in trainees. For instance, some researchers have supplemented static mannequins with videos and audio on a nearby screen [92], but this separated presentation of physical human form and dynamic visuals may reduce the effectiveness of the simulator.

In this chapter, I examine the importance of matched imagery and shape of a physical-virtual patient simulator for pediatric medical training purposes based on two training scenarios, and a human-subject study with nursing students to assess the effects of the location of dynamic visuals (co-located with the physical surface or separated from the surface on a screen) and the physical shape (human shape or flat). I analyze and discuss the findings with respect to the considered medical training scenarios.

This research is guided by the following research questions on the location of dynamic visuals:

Q1 Will trainees perceive a simulator with co-located dynamic visuals as **more realistic** than one with separated dynamic visuals presented on a TV screen (as commonly used)? (to support **[TS 1B]**)

Q2 Will trainees experience **increased cognitive load** when dynamic visuals are separated from the physical patient?

Q3 Will trainees feel that the dynamic visuals **represent the patient** when they are separated from the other cues?

We are further guided by research questions on the shape of the patient simulator:

Q4 Will trainees perceive a child-shaped simulator to be **more realistic** than a flat simulator?

Q5 Will trainees experience **increased cognitive load** when assessing a child-shaped simulator compared to a flat one? (to support **[TS 2]**)

Q6 Will trainees naturally **stand by the side** of a child-shaped patient lying on a bed and will they show less natural behavior with a flat patient simulator?

5.3 Experiment

This section details our patient simulators and experimental design leveraging an advanced health assessment class involving graduate nursing students.

5.3.1 Apparatus

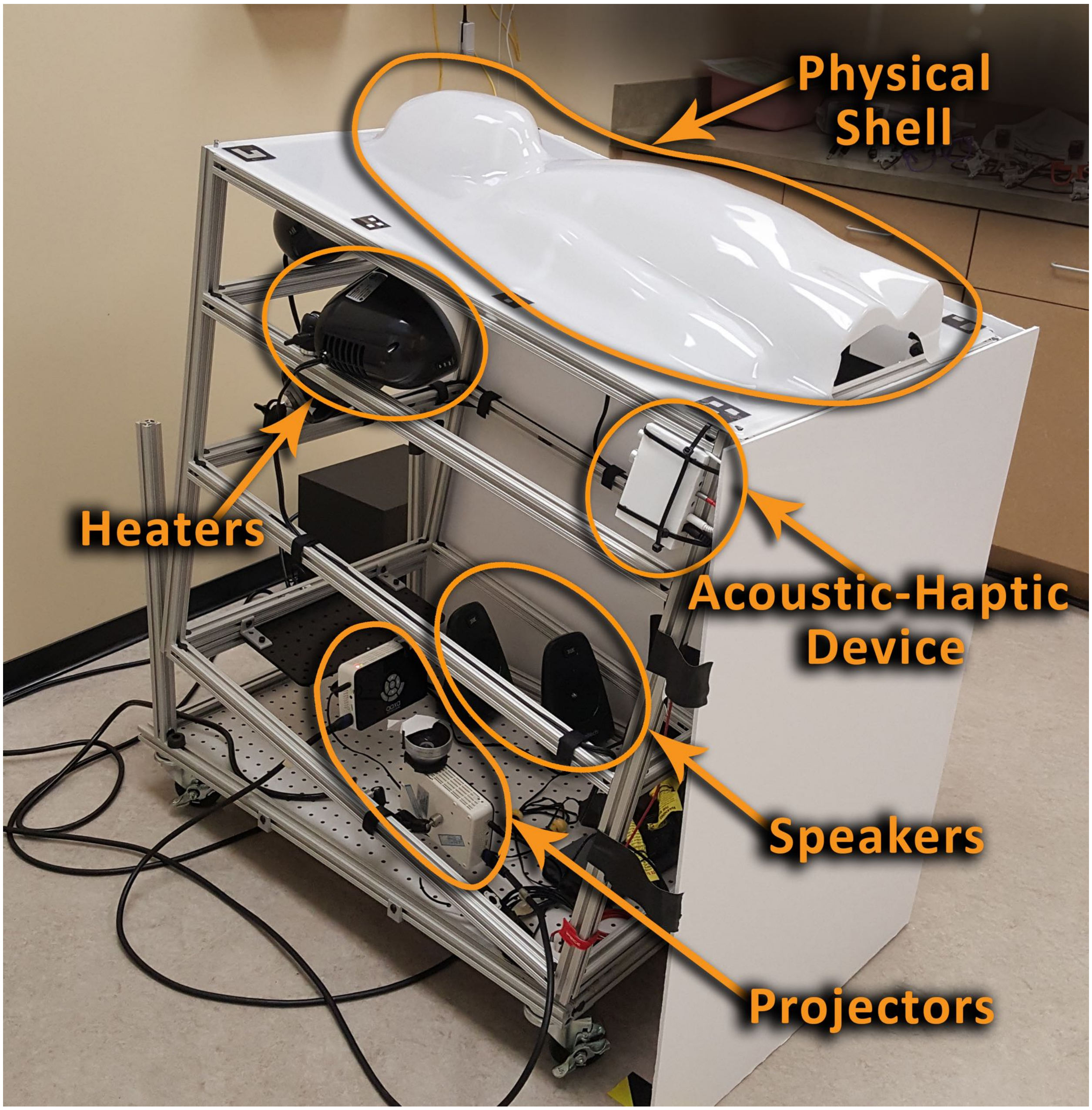


Figure 5.2: Study apparatus, showing the child-shaped PVP. Two projectors display imagery on the patient surface. Heaters and audio haptic devices provide temperature and pulse, respectively. Patient speech is played through the speakers.

In collaboration with experts and teammates, we have developed a pediatric *Physical-Virtual Patient* (PVP) simulator representing a 6-year-old child that can support integrated multi-sensory output (including visuals, audio, and touch). The PVP is a physical simulator that occupies volume, being composed of 2 AAXA P300 pico projectors (resolution 1920×1080) that rear-project onto a child-shaped shell (see Figure 5.2). Under the shell, we installed five Honeywell HEC 100B heating units to provide localized temperature to different parts of the body and two acoustic haptic devices to provide pulse, which is perceivable using a stethoscope. The virtual aspects of the simulator consist of projecting animated imagery rendered using the Unity 3D graphics engine. In line with the demands of an advanced health assessment class at University of Central Florida, I developed content for two medical scenarios (sepsis and child abuse, described in subsection 5.3.3), with virtual animations (e.g. facial expressions, head and hand movement, breathing), sounds, and verbal responses controlled via human-in-the-loop operation. The controller observes the simulation from a different room using a professional video system with 2 ceiling cameras (resolution 1920×1080) that could be remotely controlled and presses buttons to trigger audio-visual responses, as is common practice in the assessment and training of nurses and physicians with a variety of low- or high-fidelity simulators [238].

For the PVP as described above, both the shape and imagery are inherently matched—the shell is the same shape as the embodied virtual agent (a child), and the dynamic visuals of the child are co-located with the child-shaped shell. In this study, we examine the importance of these matched circumstances by comparing four variations of an embodied virtual agent for patient simulation. As shown in Figure 5.3, we either used the PVP for the co-located presentation of all cues or we separated the dynamic visuals (facial expressions, movements) from the rest of the cues (static visuals, audio, haptics). We refer to these conditions as COL and SEP, respectively. As discussed before, a separation of certain cues from the main locus of the patient is quite common in nursing and physician training. In the case of separated visuals (SEP), the dynamic visuals were displayed on a 1920×1080 television screen positioned vertically in a landscape orientation behind the head of the patient, showing a rendering of the child patient. The motivation to position the screen vertically came from the healthcare practice of adding a screen near a mannequin in simulation; also, the working space in the room would have been very limited had we positioned the TV screen horizontally. In a landscape orientation, the TV width and height were large enough to display the patient at a comparable size and orientation to what is displayed on the simulator. Next, to compare the effects of the shape of the patient simulator, we further developed a similar Flat-Virtual Patient (FVP) simulator, using the same equipment as the PVP but with a flat surface. Both the PVP and FVP use plastic rear-projection surfaces.

This study apparatus supports all 4 experimental conditions, allowing us to vary the 2 independent variables: simulator shape (human-shaped vs. flat-shaped) and the location of the dynamic visuals (co-located with the rest of the cues vs. separated on a nearby TV screen). These variables resulted in the following experimental conditions (see Figure 5.3):

- **PVP-COL**: Child-shaped shell with co-located dynamic visuals
- **FVP-COL**: Flat shell with co-located dynamic visuals
- **PVP-SEP**: Child-shaped shell with separated dynamic visuals on a TV
- **FVP-SEP**: Flat shell with separated dynamic visuals on a TV

This experimental apparatus further included equipment to monitor and assess our participants' performance. We prepared two largely identical rooms for the experiment. Two participants in each of the two rooms wore head trackers and eye trackers during the experiment. The head trackers consisted of an adjustable headgear with an HTC VIVE sensor mounted on top whose position and orientation were tracked using two VIVE lighthouse units. The eye trackers were built based on the *Do It Yourself* guide by Pupil Labs [239], for which we used 120 Hz, 1080p wide angle and fisheye lens ELP USB cameras that were mounted on glass-less frames using 3D printed camera mounts capturing each participant's left eye. Each participant's eye tracker was connected to a dedicated graphics workstation in the simulation room, using 9-meter cables to allow for unrestrained movement. For safety reasons and to prevent participants from blocking the main camera feeds through which the operator saw the room, we restricted participants from standing along the patient's right side (see Figure 5.1). In each simulation room, one of the two eye tracking computers was also used to collect head tracking data for both participants.

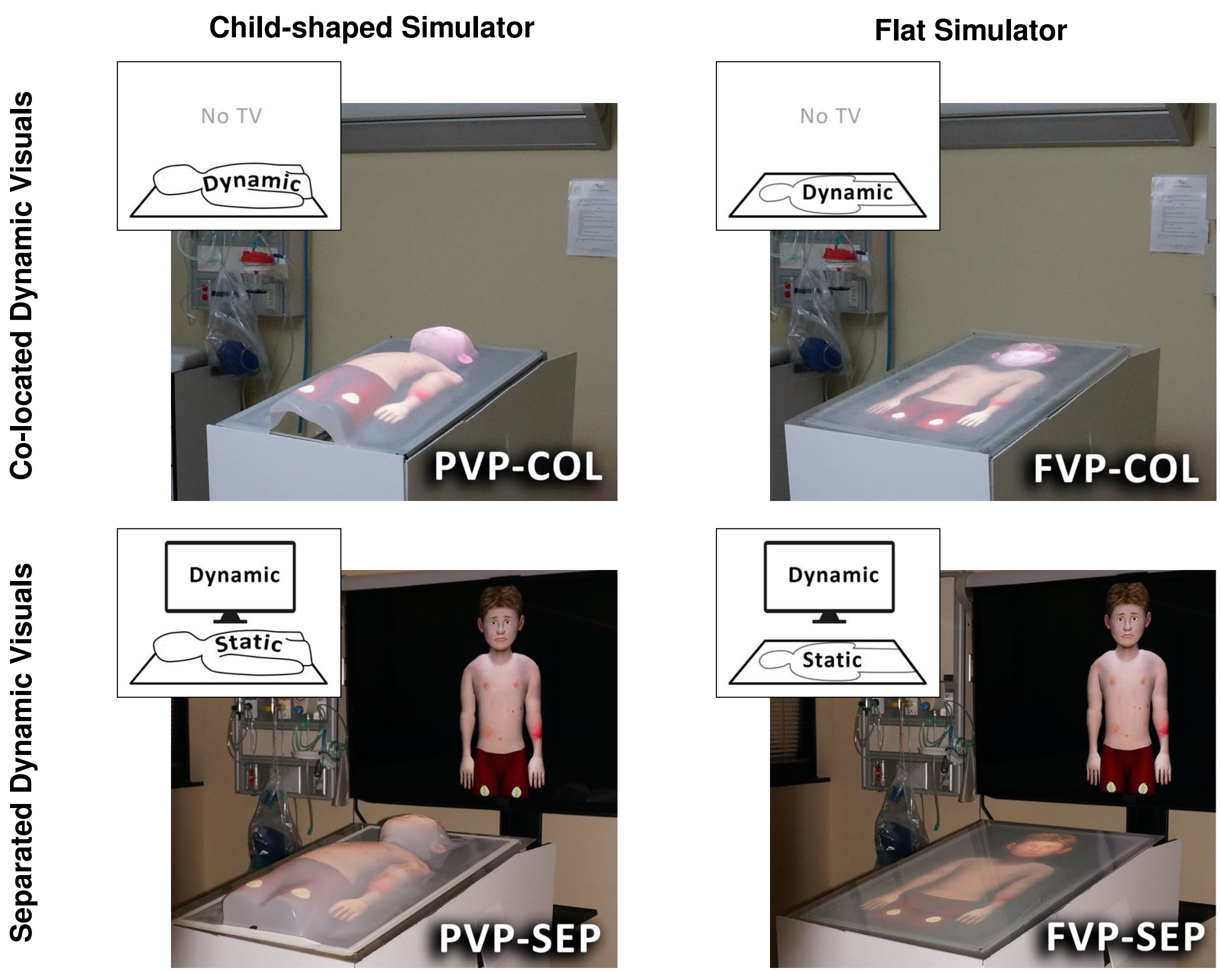


Figure 5.3: The four study conditions, which considered the presentation of dynamic visuals (co-located with or separated from the simulator) and the simulator shape (child-shaped or flat).

5.3.2 Participants

Over the course of the study, 44 graduate nurse practitioner students participated (3 male, 40 female, 1 omitted). There were 3 participants below 24 years old, 28 between 25 and 34 years old, 8 between 35 and 44, and 4 above 45. During the study, 10 participants wore glasses, and 33 did not. Only one participant reported a vision limitation that was not corrected, which in this case was blurry vision in one eye, which we did not consider to be a sufficient reason to exclude this participant. The simulation experiment was conducted as part of an advanced health assessment class at the local university's College of Nursing. In terms of prior healthcare simulation experience, 35 participants were previously exposed to physical mannequins, 42 participants were previously exposed to virtual patients, and 41 participants were previously exposed to standardized patients. Participants assessed each simulated patient in two scenarios (sepsis, child abuse) as a pair. The experiment was conducted over three days. At the end of each simulation day, the entire group of participants for that day received a collective debriefing.

5.3.3 Scenario

I modeled two simulated patients: one showing signs of child abuse and one showing signs of pediatric sepsis. Both of these scenarios involve the combination of multi-sensory cues provided by our study apparatus, including visuals, audio, temperature, and pulse. For each scenario, we developed corresponding patient histories and cues, including verbal and non-verbal responses (such as facial expressions). Both patients demonstrated awareness of their environment and the people around them as appropriate for their character and medical condition. The verbal and non-verbal responses in both scenarios were controlled using a uniform Unity-based graphical user interface for all conditions, and each scenario was controlled by a specific operator throughout the course of the study to maintain interaction consistency. Each simulator had a physical shirt. When participants removed it to examine the patient's torso, the simulation operator removed the virtual shirt on the TV. Participants were able to examine the heart sounds of the simulated patients using a stethoscope.

Both virtual male child models were modeled, textured, rigged, and animated in Maya, then imported into Unity using its legacy animation system. For the child abuse scenario, the graphical content included burn marks, bite marks, bruises, and cigarette burns. The child avoided looking at the nursing students when they asked probing questions about the burns and bruises. His voice showed slight signs of pain and sadness. The simulated pediatric patient with signs of sepsis was pale and exhibited mottled skin on his torso, blue lips, and droopy eyes. His verbal responses were quiet and slow. Aside from the characteristics of the two patients' voices, the verbal responses were largely the same, consisting of short phrases such as "yes" and "no," information about the patients' ages, and longer phrases relating to the patients' conditions (such as "it hurts to breathe"). We compared medical scenario effects for all conditions and found no statistical difference in our measures, so the scenarios are combined in our results. See Appendix C for more details about the sepsis scenario development, and Appendix D for more details about the abuse scenario development.

5.3.4 Study Design and Procedure

The study, approved by the Institutional Review Board, featured a mixed design, with both within- and between-subject components. All pairs of participants performed two simulated patient assessments: one with co-located dynamic visuals and a second with separated dynamic visuals on a TV screen. The physical form of the PVP—either child-shaped or flat—was consistent across these two assessments; patient shape was varied as a between-subject component. The two assessments represented our two simulated patient scenarios: child abuse and sepsis. Though participants assessed each patient together as a pair, they were asked to complete all questionnaires independently and to not discuss their answers with one another.

First, each participant independently completed a pre-questionnaire concerning basic demographic information, vision issues, and prior exposure to simulated patients. Before each session, participants were outfitted with the head-tracking and eye-tracking equipment and completed a short calibration process using Pupil Labs software. Next, we asked them to look at specific marked locations in the room, both to refine the eye tracker calibration and to align the eye- and head-tracking coordinate spaces. Once they were instrumented, each pair watched a short video (approximately 3 minutes) featuring an abstract sample assessment of a healthy patient on the simulator corresponding to the next study condition, demonstrating its various capabilities. After the video, the pair was provided with the patient's history and given a final chance to ask questions, if desired. The participants were instructed to assess the patient for the current condition once the simulator operator gave a verbal "begin simulation" command. Participants had a total of approximately 10 minutes to complete their assessment, which ended when the operator gave a verbal "end simulation" command.

Each medical condition had a specific operator assigned throughout the entire study. I was the operator for the Sepsis scenario while Jason Hochreiter was the operator for the Child Abuse scenario running at the same time. In addition to starting and stopping the simulation, each simulation operator was responsible for remotely observing the participants and controlling the behavior of the virtual patient, both verbal and non-verbal. The same software and controller interface was used in all four study conditions. In certain situations, additional verbal responses provided by a healthcare trainer were necessary; for example, if a participant wished to physically move the patient or assess his legs, the healthcare trainer informed participants that these actions were not possible and not needed for the assessment.

Following each assessment, participants individually completed post-questionnaires relating to their experience. These included questions regarding their evaluation of the simulator and cognitive load.

5.3.5 Measures

We asked participants to evaluate the realism of various components of the simulators, each as a Likert scale (Table 5.1). The visual realism questions asked participants to score the visual appearance of the simulator from "extremely ugly" to "extremely good looking" and the animations from "extremely unnatural" to "extremely natural" (both 1–7). In terms of face realism, participants ranked the simulator's facial expressions from "extremely exaggerated" to "extremely realistic" and its ability to show subtle facial changes from "extremely poor" to "extremely good" (both 1–7). Next, we asked participants to evaluate the haptics realism in terms of temperature and pulse and the audio realism in terms of speech and heart sounds from "very unrealistic" to "very realistic" (1–6); a seventh option was reserved for a "not applicable" rating that participants could choose if they did not notice a particular cue. Each participant also completed the NASA Task Load Index (TLX) [1], which reflected the mental, physical, and temporal demands they experienced during their assessments in a series of Likert scale questions from 1–10 (Table 5.2).

Additionally, we asked a few qualitative questions regarding the pair of simulated patients each participant assessed, which differed in the location of dynamic visuals but had a constant shape. First, we asked participants to indicate which of the two paradigms—co-located or separated dynamic visuals—they preferred and why. To answer our research question Q3 concerning the perceived identity of the patient in the case when dynamic visuals are displayed on a TV separated from the physical simulator, we asked participants which of the two entities they thought represented the actual simulated patient.

As described above, during the experiment we tracked the participants' head poses using HTC Vive sensors and eye gaze using eye trackers. Using the tracked head positions of the participants and 3D reference points around the apparatus, we computed the percentage of time participants spent at 4 defined areas around the patient: side of patient, feet of patient, corner, and inside (see Figure 5.1). The "inside" location represents times when the participants leaned over the patient. We excluded this area as it covered only a very small amount of the experiment time (2% or less) and used the remaining data as a measure for each participant's standing location throughout the simulation. To analyze the effects of simulator shape on the standing location of participants, we also excluded the "corner" area, which was ambiguous between the side and feet areas. For eye-tracking, we were interested in gaze points with fixation duration longer than 300 ms in the patient's head and torso areas. We also collected eye fixations occurring outside of the patient.

Table 5.1: Realism questions relating to visual (scale from 1 to 7), face (1 to 7), haptic (1 to 6), and sound (1 to 6) components.

	Realism
V1	Rate the simulator's visual appearance.
V2	Rate the simulator's animations (facial expressions, speech, visual changes).
F1	Rate the simulator's ability to show subtle facial changes.
F2	How exaggerated/realistic were the facial expressions of the simulator?
H1	How realistic was the patient in exhibiting pulse capabilities?
H2	How realistic was the patient in exhibiting body temperature capabilities?
S1	How realistic was the patient in exhibiting the speech capabilities?
S2	How realistic was the patient in exhibiting the heart sounds capabilities?

Table 5.2: NASA-TLX [1] (scale from 1 to 10).

	NASA-TLX
1	Mental demand: How much mental and perceptual activity was required (e.g. thinking, deciding, calculating, remembering, looking, searching, etc.)?
2	Physical demand: How much physical activity was required (e.g. pushing, pulling, turning, controlling, activating, etc.)?
3	Temporal demand: How much time pressure did you feel due to the rate or pace at which the tasks or task elements occurred?
4	Performance: How successful do you think you were in accomplishing the goals of the task set by the experimenter (or yourself)? How satisfied were you with your performance in accomplishing these goals?
5	Effort: How hard did you have to work (mentally and physically) to accomplish your level of performance?
6	Frustration: How insecure, discouraged, irritated, stressed, and annoyed versus secure, gratified, content, relaxed, and complacent did you feel during the task?

5.4 Results

We made the decision to use non-parametric statistical tests (Mann-Whitney U for independent samples and Wilcoxon signed-rank for paired samples) to analyze the Likert-scale ordinal data from the questionnaires. Some scientific disciplines have the common practice to treat Likert-type scales as interval-level measurements [224]. By using non-parametric tests, we avoid the discussion related to whether parametric statistics can be a valid method for the analysis of such data [225, 226]. A summary of the findings and hypotheses related to our research questions are shown in 5.3.

Table 5.3: Summary of the overall study findings. The numbered hypotheses correspond to our research questions.

		Hypothesis	Evidence
Co-location (COL) vs. Separation (SEP) of Dynamic Visuals	**H1**	Mean **realism evaluations** will be higher for the COL simulator than for the SEP simulator	Partially accepted: - Visual appearance rated significantly higher (COL > SEP)
	H2	**Cognitive load scores** will be higher for the SEP simulator than for the COL simulator	No evidence
	H3	When dynamic visuals are separated onto a TV, participants will consider the TV as **representing the patient** instead of the physical simulator (SIM)	Partially accepted: - Participants showed a trend to consider the separate TV as the patient representation instead of the physical simulator (TV > SIM; trend)
Child-Shaped (PVP) vs. Flat (FVP) Simulator	**H4**	Mean **realism evaluations** will be higher for the PVP simulator than for the FVP simulator	Partially rejected: - Animations rated significantly lower (PVP < FVP) - Subtle facial changes rated lower (PVP < FVP; trend) - Speech rated lower (PVP < FVP; trend) - Heart sounds rated significantly lower (PVP < FVP)
	H5	**Cognitive load scores** will be higher for the PVP simulator than for the FVP simulator	Accepted: - Cognitive load significantly higher (PVP > FVP)
	H6	Participants will spend more time **standing by the side** of the PVP than by the feet	Accepted: - Time spent by side significantly higher than time spent by feet

5.4.1 Co-location vs. Separation of Dynamic Visuals

First, I present results comparing the location of the dynamic visuals: either co-located with the physical simulator (COL) or presented on a separate TV screen (SEP). I start with aggregate measures that include both possible shapes of the simulator—child-shaped (PVP) or flat (FVP). Where appropriate, I also present results filtered by the simulator shape.

5.4.1.1 Realism

Several of the realism questions were significantly higher for the COL conditions compared to the SEP ones, supporting our hypothesis H1 and **[TS 1B]**.

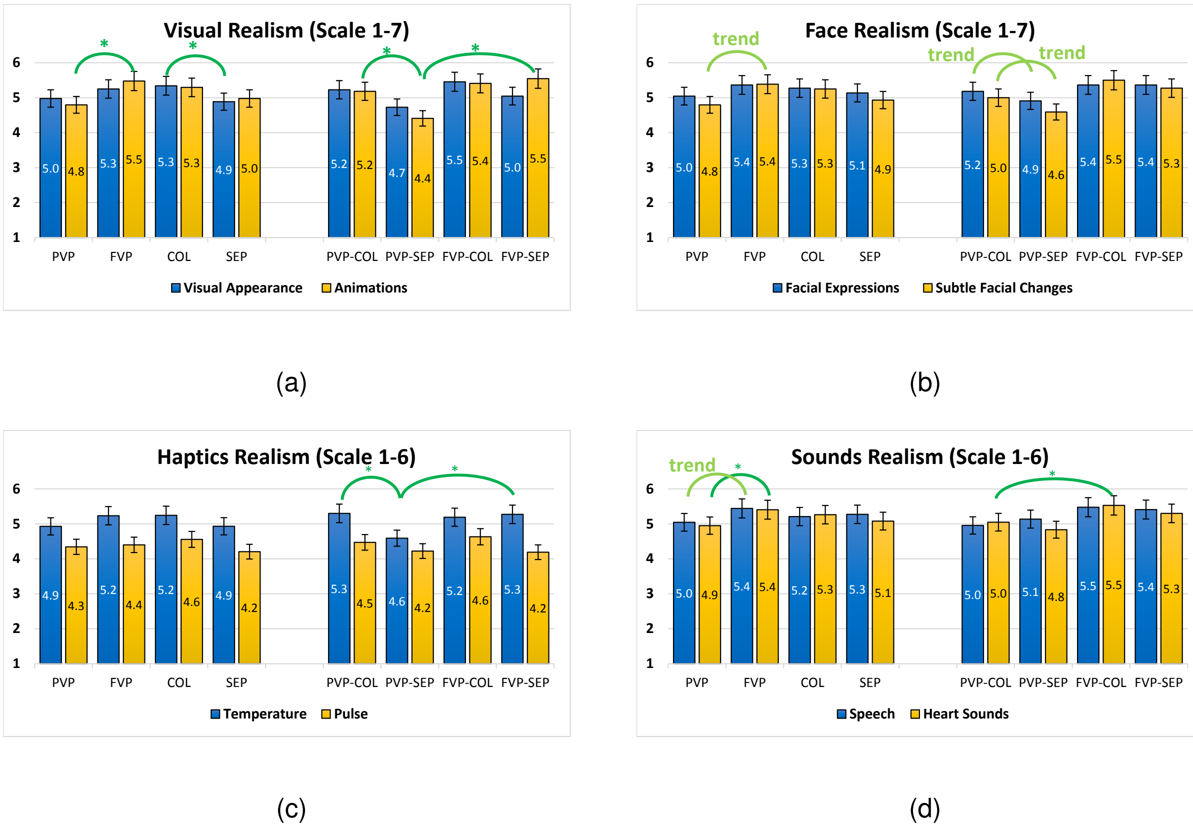


Figure 5.4: Realism results. (a) Visual realism. (b) Face realism. (c) Haptics realism. (d) Sounds realism. Stars (*) denote significant results ($p \leq 0.05$), and "trend" indicates results for which $0.05 < p \leq 0.1$.

In terms of visual realism , participants rated the visual appearance significantly higher ($W = 224$, $p = 0.031$) for the COL conditions than for the SEP conditions. There was no significant difference in visual appearance ratings between the PVP-COL and PVP-SEP or between the FVP-COL and FVP-SEP conditions. We observed no significant difference in how participants evaluated the realism of simulator animations between the COL and SEP conditions, including between FVP-COL and FVP-SEP. However, participants considered the PVP-COL to have more realistic animations than the PVP-SEP; this difference was significant ($W = 113$, $p = 0.018$).

For face realism, there were no significant differences between ratings of the realism of the facial expressions or ability to show subtle facial changes between the COL and SEP conditions, including between FVP-COL and FVP-SEP. However, we observed a trend ($W = 36$, $p = 0.095$) suggesting participants found the facial expressions more realistic for the PVP-COL than the PVP-SEP. Additionally, results show a trend ($W = 51$, $p = 0.097$) suggesting that participants considered the PVP-COL better able to show subtle facial changes than the PVP-SEP.

Haptics realism considered the temperature and pulse capabilities of each simulator. There was no significant difference in participant temperature evaluations of the COL and SEP conditions, including between FVP-COL and FVP-SEP. However, we did observe a significant difference ($W = 48$, $p = 0.037$) for the child-shaped simulators, with participants finding the temperature more realistic for the PVP-COL than for the PVP-SEP. There was no significant differences or trends in participant ratings of pulse realism between any of the COL and SEP conditions.

Finally, the sound realism was comparable regardless of whether dynamic visuals were co-located with the simulator, with no significant differences between participant evaluations of speech and heart sounds realism between any of the COL and SEP conditions.

5.4.1.2 Cognitive Load

In terms of overall self-reported cognitive load (Figure 5.5), we observed no significant differences between the COL and SEP conditions, which does not support our hypothesis H2. However, we did observe a significant difference for the fifth NASA-TLX question ($W = 314$, $p = 0.032$), indicating that participants expended more mental and physical effort when assessing the COL conditions compared to the SEP conditions. Additionally, there was a trend ($W = 92.5$, $p = 0.060$) suggesting the same increase in effort was required for the FVP-COL compared to the FVP-SEP. Furthermore, we observed a trend regarding the second NASA-TLX question ($W = 53.5$, $p =$

0.067) suggesting that participants experienced increased physical demand when assessing the FVP-COL compared to the FVP-SEP. There were no similar effects for these two NASA-TLX questions for the PVP-COL compared to the PVP-SEP.

5.4.1.3 Qualitative

Participants were asked "Having experienced two different representations of a patient, which one would you prefer to use and why?" Each pair of participants assessed two patients: one with co-located dynamic visuals (COL) and one with dynamic visuals separated on a TV (SEP), with the shape of the simulator held constant. One participant did not answer the question; as for the remaining 43 participants, the majority ($n = 34$) preferred the COL patient over the SEP patient ($n = 9$). A chi-squared test indicates the difference is significant ($z = 14.5$, $p < 0.001$). The same effect was present regardless of the shape of the simulator: there was a significant preference ($z = 5.8$, $p = 0.016$) for the PVP-COL patient ($n = 16$) over the PVP-SEP patient ($n = 5$) and a significant preference ($z = 8.9$, $p = 0.0028$) for the FVP-COL patient ($n = 18$) over the FVP-SEP patient ($n = 4$).

After each SEP condition, participants were asked "In the case where you had both a simulator and a separate TV screen, which one seemed more like the patient you were treating?" Two participants did not answer the question; from the remaining 42, the majority ($n = 27$) indicated that they perceived the dynamic visuals as being the true patient instead of the physical simulator presenting the rest of the cues ($n = 15$). The chi-squared test shows a trend close to being significant ($z = 3.4$, $p = 0.064$). We looked closer for each of the SEP cases when filtered by the physical form of the simulator (Figure 5.6). For the PVP-SEP condition, the answers were almost evenly split, with 10 participants feeling that the physical simulator with static visuals, temperature, and pulse on a child-shaped shell was more like the actual patient, whereas 11 participants considered the TV screen with dynamic imagery to be the patient. There is no statistical difference, which does not match our hypothesis H3 in this case. In contrast, for participants in the FVP-SEP condition, the majority ($n = 16$) perceived the TV screen with dynamic visuals as being the patient, while 5 participants felt that the physical simulator with static visuals, temperature, and pulse on a flat surface was the actual patient. The difference is significant ($z = 5.8$, $p = 0.016$), supporting our hypothesis H3 in this case.

5.4.1.4 Standing Location

Using the percentage of time participants stood in certain locations around the patient, we classified each participant depending on whether they spent a majority of their assessment time standing by the patient's side or by the patient's feet; we analyzed these classifications to determine whether the shape of the simulator impacted participant standing location. There was no significant difference for participant standing locations (patient's side vs. feet) for COL vs. SEP.

5.4.1.5 Participant Gaze

We analyzed the gaze points corresponding to fixations longer than 300 ms, annotating detected points as belonging to the patient's head or patient's torso or appearing outside of the apparatus. Due to low calibration accuracy, we excluded 34 of 88 sessions.

5.4.1.5.1 COL vs. SEP

We used the Wilcoxon signed-rank test to compare eye fixations between COL and SEP. We considered fixation points at the simulated patient's full body (head and torso together) and at the head and torso individually. For the SEP conditions, we consider these locations on the patient's body across both the physical simulator (SIM) and the separated television screen (TV). The percentage of fixations at either the patient's full body, head, or torso was not significantly different between participants who assessed the COL patient and participants who assessed the SEP patient.

5.4.1.5.2 COL

For the PVP-COL and FVP-COL conditions, we used the Mann-Whitney U test to compare fixation points for the patient's full body (head and torso together) and for the head and torso individually. For all of these locations, there was no significant difference in the percentage of fixations between participant assessments of the PVP-COL and the FVP-COL conditions. However, when using the Wilcoxon signed-rank test, we observed a trend ($W = 321$, $p = 0.07$) suggesting that participants fixated more on the COL patient's head ($M = 49.1$, $SD = 12.6$) than the torso ($M = 40.5$, $SD = 13.7$).

5.4.1.5.3 SEP

Within the SEP group, we used the Wilcoxon signed-rank test to compare fixation points between the static physical simulator (SIM) and the television screen with dynamic visuals (TV). For the patient's full body, head, and torso, we observed no significant differences in fixation points between the SEP-SIM and the SEP-TV. Next, we compared fixation points between the patient's head and torso across the two patient manifestations (SIM and TV), again using the Wilcoxon signed-rank test. For the SEP-SIM and the SEP-TV taken collectively, participants had significantly more fixations ($W = 50$, $p = 0.011$) on the head ($M = 37.3$, $SD = 11.9$) than on the torso ($M = 50.3$, $SD = 12.8$). Similarly, for just the SEP-SIM, participants had significantly more fixations ($W = 52$, $p = 0.002$) on the head ($M = 16.6$, $SD = 11.0$) than on the torso

($M = 24.5$, $SD = 12.9$). However, for the SEP-TV only, we observed no significant difference between participant fixations for the head compared to the torso.

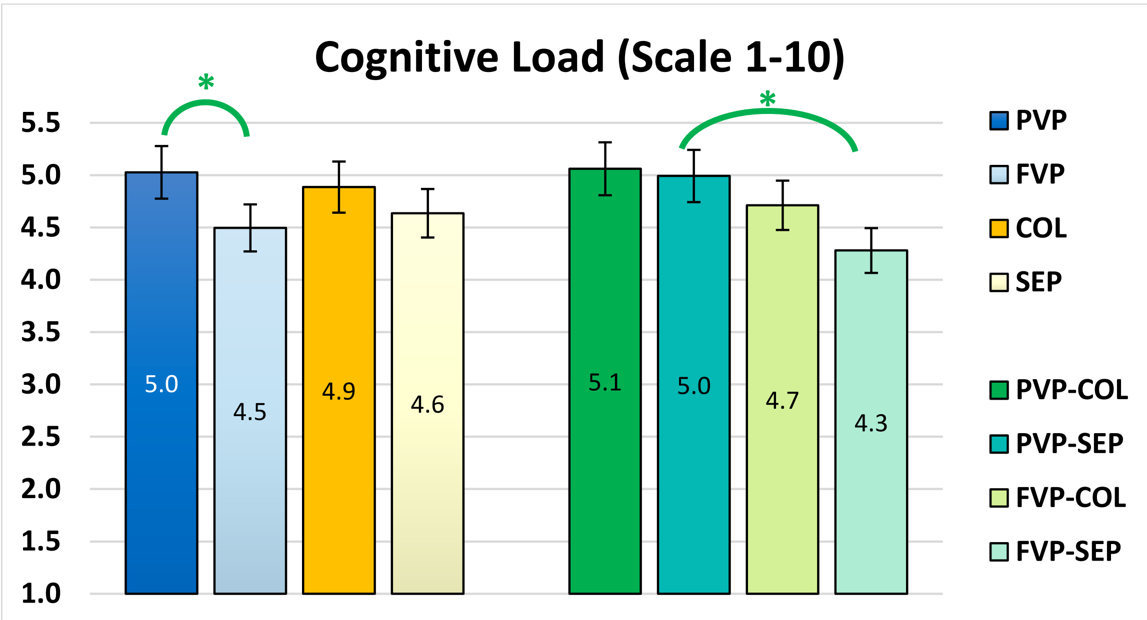


Figure 5.5: Cognitive load results (NASA-TLX).

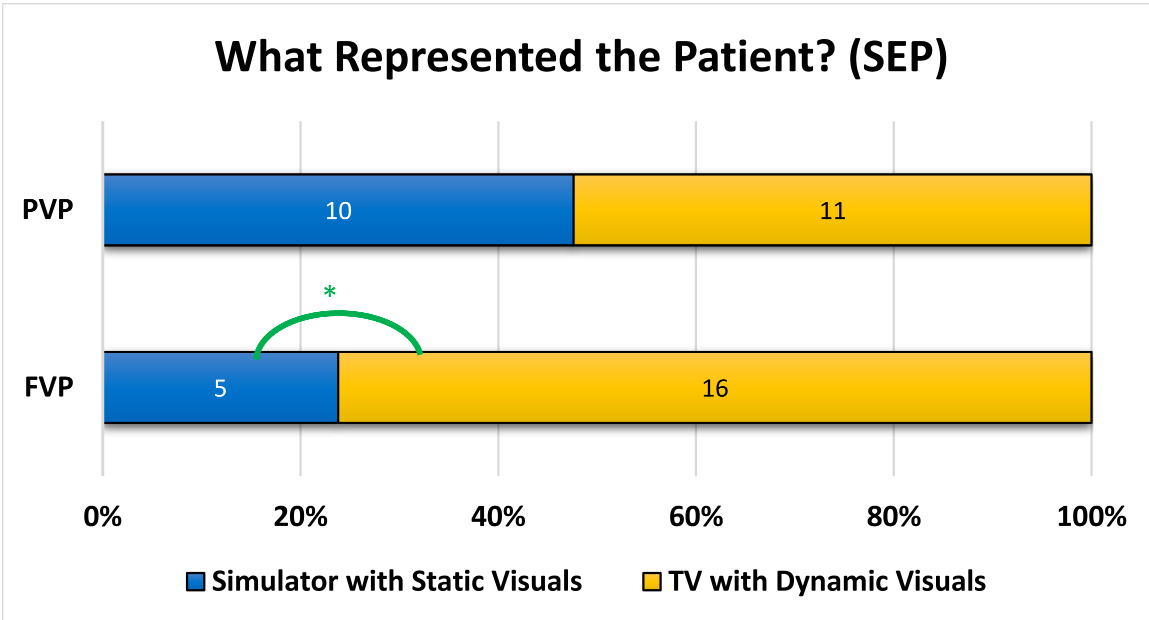


Figure 5.6: Qualitative responses for whether the physical simulator or the TV was perceived as being the patient.

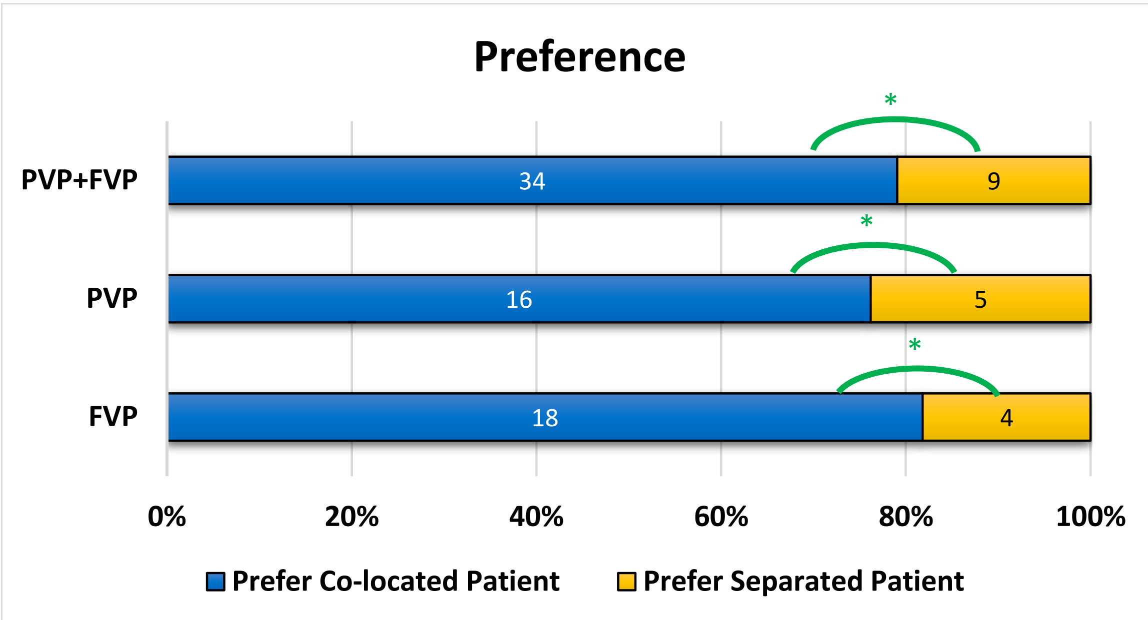


Figure 5.7: Subjective preference for assessment of COL or SEP conditions.

5.4.2 *Child-Shaped vs. Flat Simulator*

Next, we present results that consider the effects of the shape of the simulator, either child-shaped (PVP) or flat (FVP), aggregated across the location of dynamic visuals (COL and SEP). Results comparing simulator shape with respect to a single presentation of dynamic visuals are also provided where appropriate.

5.4.2.1 *Realism*

We did not observe many significant differences between realism questions for the PVP and FVP conditions, and the ones we did observe indicate that participants found the PVP to be less realistic for certain measures than the FVP, which does not match our hypothesis H4.

For visual realism , there were no significant differences in participant evaluations of the visual appearance between any of the PVP and FVP conditions. Participants rated the realism of animations significantly lower ($W = 1219$, $p = 0.031$) for the PVP than for the FVP. We also observed a significant difference ($W = 355.5$, $p = 0.006$) indicating lower perceived realism of animations specifically for the PVP-SEP compared to the FVP-SEP. However, there was no significant difference between the PVP-COL and FVP-COL conditions for animation realism evaluations.

In terms of face realism , we observed no significant differences in realism ratings of facial expressions between any of the PVP and FVP conditions. We did observe a trend ($W = 1182.5$, $p = 0.067$) suggesting that participants found the PVP less capable of displaying subtle facial changes than the FVP. There were no trends between the PVP-COL and FVP-COL or between the PVP-SEP and FVP-SEP conditions.

We observed few differences regarding the realism of simulator haptics There were no significant differences in temperature realism between the PVP and FVP conditions, including between PVP-COL and FVP-COL. However, participants rated the temperature of the PVP-SEP less realistic than the FVP-SEP. The difference was significant ($W = 333$, $p = 0.023$). We observed no significant differences in the realism evaluations of pulse between any of the PVP and FVP conditions.

Finally, regarding sound realism ratings , there was a trend ($W = 1147.5$, $p = 0.057$) suggesting participants found the speech of the PVP less realistic than the FVP. This trend was not observed between PVP-COL and FVP-COL or between PVP-SEP and FVP-SEP. Participants rated the heart sounds of the PVP less realistic than the FVP. The difference was significant ($W = 929$, $p = 0.020$). There was a similar significant difference ($W = 240.5$, $p = 0.048$) reflecting lower perceived realism for the heart sounds of the PVP-COL compared to the FVP-COL. We observed no significant difference for the realism rating of heart sounds for the PVP-SEP and the FVP-SEP.

5.4.2.2 Cognitive Load

Participants were more likely to have a higher cognitive load (Figure 5.5) in the PVP conditions than in the FVP conditions. The difference is significant ($W = 675, p = 0.014$), supporting our hypothesis H5 and **[TS 1B]**. We observed the same significant increase in cognitive load ($W = 139.5$, $p = 0.016$) for assessment of the PVP-SEP compared to the FVP-SEP. However, we observed no significant difference in overall cognitive load between the PVP-COL and FVP-COL. Moreover, there was a significant difference in the third NASA-TLX question ($W = 666.5$, $p = 0.011$) indicating participants felt greater time pressure when assessing the PVP compared to the FVP. Trends suggested this increase in time pressure was present regardless of whether dynamic visuals were co-located with ($W = 168$, $p = 0.082$) or separated from ($W = 162$, $p = 0.057$) the simulator. These trends suggest greater perceived temporal demands for assessment of the PVP-COL compared to the FVP-COL and for assessment of the PVP-SEP compared to the FVP-SEP.

5.4.2.3 Standing Location

Again, we classified participant standing locations using head-tracking results (Figure 5.8). According to a chi-squared test, participants who assessed the PVP were significantly more likely ($z = 6.4$, $p = 0.012$) to stand by the patient's side instead of his feet, supporting our hypothesis H6. We observed trends suggesting the same effect within the PVP-COL ($z = 3.6, p = 0.058$) and PVP-SEP ($z = 2.8$, $p = 0.096$) conditions. However, for the FVP, FVP-COL, and FVP-SEP conditions, there was no significant difference in standing locations.

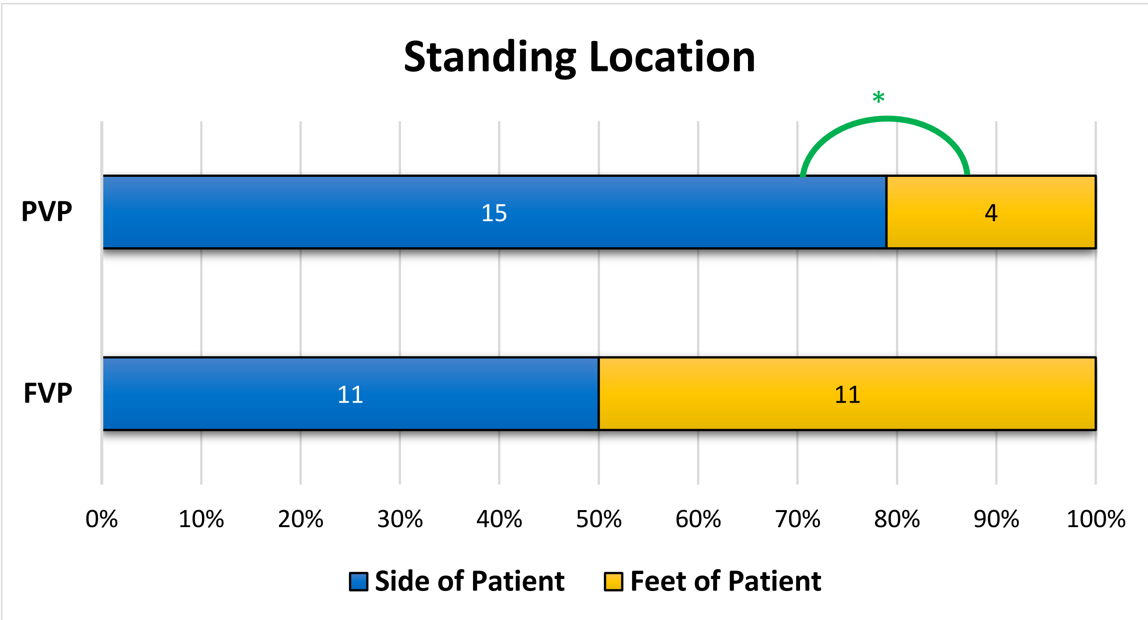


Figure 5.8: Standing location results.

5.5 Discussion

In this section we discuss the results, first focusing on the differences between co-located and separated dynamic visuals, and then focusing on the shape of the simulator. Finally, we present overall summary findings.

5.5.1 Co-location vs. Separation of Dynamic Visuals

[TS 1B] Participants rated the visual appearance of the simulator with co-located dynamic visuals (COL) higher than the simulator with dynamic visuals on a separate TV screen (SEP). In the COL case, participants were able to interact with the visual content more directly than in the

SEP case. We found no difference in participant evaluation of animation realism between COL and SEP in general. The fact that there were no significant differences in the perceived realism of speech between the COL and SEP cases is not surprising, as it was provided equivalently in both cases. While pulse and heart sounds were also provided equivalently, they were separated in space; despite this, we also found no significant difference of realism ratings of pulse and heart sounds between COL and SEP. However, animation realism was rated higher specifically between the PVP-COL and PVP-SEP conditions. There were also trends suggesting increased perceived realism for facial expressions and subtle facial changes for the PVP-COL compared to the PVP-SEP. This is interesting, as participants generally rated the animations of the PVP less realistic than for the FVP, and the animations of the PVP-SEP appeared on a flat surface (similar to the FVP). Likewise, participants found the temperature significantly more realistic specifically for the PVP-COL compared to the PVP-SEP. We think the increase in perceived realism of animations and temperature of the PVP-COL compared to the PVP-SEP was due to the physical context of these cues, which appeared in a position and orientation representative of a child in a hospital bed. Furthermore, participants tended to look at the patient's head more often in the COL cases than in the SEP cases.

We were surprised that there was no significant difference in cognitive load between the COL and SEP cases, since in the latter the patient is shown in two separate locations that compete for the participant's attention. However, there was a significant difference in one of the NASA-TLX questions, suggesting higher mental and physical effort was required for assessment of the COL conditions compared to the SEP conditions. We suspect that participants primarily focused on the television screen in the SEP cases, which is an easy and intuitive interface, and ignored the physical simulator, especially after any haptic interactions (pulse and temperature) were completed. Indeed, qualitative results indicate that participants in the SEP conditions generally considered the TV screen to be the patient as opposed to the physical simulator. This suggests that the dynamic visuals of the simulator, including lip-synced speech, had a strong influence on the perceived identity of the patient. This is likely because the dynamic visuals command more attention than the other cues: patient speech was prominent throughout the assessment, unlike cues such as pulse and temperature, which would have been assessed infrequently. Because of this, speech may have been a more obvious and direct indicator that the patient was alive and alert than other cues. However, we note that in the case when the physical simulator had a human shape (PVP-SEP), participants were equally split as to whether they considered the simulator or the television screen to be the locus of the patient, indicating the importance of physical shape on user perception.

Overall, our results show that separated visuals elicit similar cognitive load and in general similar perception as a co-located physical-virtual patient simulator, which validates their usefulness for medical training scenarios. However, we stress that training nurses on a simulator with the potential for a shift in attention away from the physical locus of a (simulated) patient towards another set of cues has the potential to hurt their performance in the long run. During healthcare assessments, nurses receive a variety of data which they must prioritize and interpret, including data from the patient and data from machines such as vital signs monitors; it is considered bad practice for nurses to pay greater attention to vitals displays than to the (real) patient they are treating. In our study, nursing students who attended primarily to the dynamic visuals display might have missed other useful cues, such as pulse and temperature. As such, our guideline for practitioners is to prefer the use of patient simulators with co-located dynamic visuals to prevent unnecessary attentional shifts and the reinforcement of improper behavior.

5.5.2 Child-Shaped vs. Flat Simulator

In general, participants rated the realism of the PVP conditions lower than the FVP conditions. Animations were considered less realistic on the PVP, and we observed a trend suggesting participants felt the PVP was less capable of exhibiting subtle facial changes than the FVP. While the addition of a physical shape to a patient simulator may increase perceptions of realism in certain aspects, it may have impacts on the perceived realism of others. For instance, when the physical shape does not exactly match the virtual content, animations might be subjected to view-dependent distortions, and the shape may cause occlusions from certain viewpoints. Additionally, we specifically used a smoothed human shape to allow for coarse animations, like head rotations; while these might have appeared less realistic compared to such animations shown on a flat surface, they would look significantly distorted and unnatural on a static human shape. Subtle facial changes may have been more difficult for participants to notice on a physical shape compared to a flat one. The animations include lip-syncing, which might explain why patient speech was considered less realistic in the PVP conditions compared to the FVP ones, even though the audio was provided equivalently in both cases. The heart sounds of the PVP were considered less realistic than those of the FVP, which may have been because it was easier to place the resonator of the stethoscope onto the flat surface than onto the child-shaped one.

[TS 2] As we expected, the overall self-reported cognitive load was higher for the PVP compared to the FVP, including along the specific aspect of temporal demand. The physical shape of the PVP required participants to assess the patient from multiple viewpoints, while all of the visuals on the FVP were visible from a single position. However, we observed no significant differences in overall cognitive load between the PVP-COL and the FVP-COL, suggesting that the addition of a human shape did not lead to more demanding interactions when cues were co-located.

Participants who assessed the PVP were more likely to stand by the simulated patient's side than by his feet, which is typical of assessments of patients lying in hospital beds. However, those who assessed the FVP were equally likely to stand by the patient's feet or his head, perhaps because interactions with flat representations of virtual humans are generally conducted from a frontal view. In training scenarios, maintaining natural behavior is important to prevent trainees from developing and reinforcing incorrect practices, which can be dangerous in healthcare settings. As the standing location of participants more closely reflects actual practice when the simulator has a human shape, we think this points to increased realism of participant behavior, which would correlate with increased training efficacy for the PVP compared to the FVP [240].

Overall, although the PVP has limitations in its visual realism, the fact that it provided a 3D physical human-like surface proved beneficial in eliciting more natural behavior related to standing/moving next to the patient bed and more natural head movements that required the participants to assess the patient realistically from different sides, which we did not observe in the flat condition. Considering the aforementioned limitations of the realism of the dynamic visuals of the PVP, we believe that future extensions of this approach could include HMDs such as a Microsoft HoloLens worn additionally by trainees to gain the benefits of both rear-projected and head-mounted dynamic visuals.

5.5.3 Summary Findings

In summary, our results show benefits when embodied virtual agents used for patient simulators present co-located cues (matched visuals) and have physical human forms (matched shape) supporting **[TS 1B]** While it is possible to augment static simulators with dynamic virtual imagery on a TV screen, this separate display might be perceived as being the patient instead of the simulator. This competition for the nurse's attention could potentially contribute to them ignoring other cues presented by the simulator, such as pulse and temperature. However, the shape of the simulator is also an important influence on such perceptions, and human-shaped simulators are less likely than flat ones to be ignored due to separated dynamic visuals. Additionally, simulators with human shape can prompt more natural behaviors from participants, such as standing by the side of the simulated patient rather than their feet. Though human-shaped simulators provide these perceptual and behavioral benefits, they require extra care regarding animated visual content and cognitive load.

5.6 Conclusion

In this paper, we presented a human-subject study concerning the matching of visual and shape characteristics of simulated patients. This study considered two independent variables: the location of dynamic visuals (co-located with or separated from the physical simulator) and the physical shape of the simulator (human-shaped or flat). Pairs of graduate nursing students assessed two patients in the scope of an advanced health assessment class: one showing signs of child abuse and one with signs of sepsis.

The results show that participants preferred the simulator with co-located dynamic visuals (matched visuals) for the two assessment scenarios compared to the simulator with dynamic visuals displayed on a separate TV screen. We observed results suggesting that participants perceived the TV screen with dynamic visuals more as the locus of the patient they were assessing than the simulated patient with temperature and pulse lying in front of them. This effect was significant when the static simulator had a flat shape, while participants with the child-shaped simulator (matched shape) were equally split between whether they considered the physical simulator or the TV to be the patient. For the child-shaped simulator, we observed greater cognitive load than for the flat simulator, which suggests a more realistic level of workload, characterized by the participants' behavior, e.g., related to moving their head around the three-dimensional patient to examine it from all sides supporting **[TS 2]** Additionally, we observed that participants who assessed the child-shaped simulator were more deliberate in their own physical location compared to those who assessed the flat simulator, tending to naturally stand by the side of the patient rather than somewhere near his feet. However, those who assessed the flat simulator showed no strong preferences for standing location.

In the future, we are planning to explore the effects of the co-location and separation of other patient cues, such as pulse and temperature, which are not always possible to easily integrate with physical simulators. Furthermore, we intend to expand the apparatus to support a full body patient simulator. We are interested in exploring other assessment scenarios on diverse simulated patients that require the ability to present co-located multi-sensory cues. In particular, we plan to investigate scenarios that would directly benefit from a physical three-dimensional matched shape of the simulator, such as those that involve size and volume estimations (for example, the estimation of burn surface areas) and those that require touch interactions (such as assessing patient reflexes).

CHAPTER 6: COGNITIVE AND TOUCH PERFORMANCE EFFECTS OF MISMATCHED 3D PHYSICAL AND VISUAL PERCEPTIONS (TOUCH STUDY)

The work in this chapter has been accepted in IEEE VR 2018 [44].

Title: Cognitive and touch performance effects of mismatched 3D physical and visual perceptions

Authors: Jason Hochreiter, Salam Daher, Gerd Bruder, Gregory Welch.

Published at: IEEE Virtual Reality, 2018.

Jason lead the research on this study focusing on the touch sensing aspects and performance. Jason and I collaborated and my contribution was in the initial idea, study design, graphics, running participants, interviewing participants, and analyzing questionnaires and interviews. My interest is in the cognitive load, usability, and preferences of participants as the physical shape and the visual mechanism can affect training. This chapter contains only the parts that reflect my contribution relevant to this dissertation while making sure it stays coherent. This chapter does NOT contain everything that was published in the paper, especially not the parts where Jason focused on the touch-sensing part or touch performance.

This chapter is in support of: **TS 2** Fidelity of the Physical Shape Can Affect the Cognitive Load

6.1 Abstract

We are interested in researching how does the mismatch between physical and visual perception affect the cognitive load, usability and preference during a task that requires touching in Aug-

mented Reality (AR). We varied the *physical fidelity* (matching vs. non-matching physical shape) and *visual mechanism* (projector-based vs. HMD-based AR) of the representation. Participants were asked to touched visual targets on a human head during four experimental conditions where we varied representation of the physical head and the visuals. We evaluated the cognitive load task by asking participants to estimate target size while performing a secondary counting task. After each experimental condition, participants were asked to complete cognitive load and usability questionnaires, and were interviewed for qualitative answers. Results indicated lower cognitive load, and increased usability in the cases where participants were exposed to rear-projected visuals and they touched a physical head-shaped surface that matched the visuals.

6.2 Introduction

Humans naturally use physical touch for common activities to experience the world around them. Many training and simulation domains require the ability to touch objects in 3D (e.g., manufacturing, healthcare). Healthcare training often requires touching a patient in precise locations, and many times the provider is under cognitive stress of performing multiple tasks at the same time. Virtual reality (VR) and augmented reality (AR) can improve the fidelity of training. On one side, there are two types to display computer generated imagery in AR: projector-based spatial augmented reality (SAR) [241] or head-mounted displays (HMDs). In parallel, there are different approaches to provide touch feedback of virtual imagery such as haptic gloves, and 3D physical surfaces [242–244].

We are interested in comparing two state-of-the-art technologies in terms of preference, usability and cognitive load.

This chapter presents human-subject study where participants are tasked with touching different dots on a on the simulator (i.e., human head). Accurately touching the patient is common in healthcare training fields.

We conducted an experiment with two independent variables resulting in four experimental conditions. The first independent variable is *physical fidelity* (i.e., physical human form, physical flat shape, no physical shape). The second independent variable was the *visual mechanism* (i.e., rear projected visuals vs. HMD-based visuals).

The following are the four resulting conditions:
rear-projected (SAR) imagery on a physical head-shaped surface that matches the visuals.
rear-projected (SAR) imagery on a physical flat-shaped surface.
Head Mounted Display (HMD) imagery on a physical head-shaped surface that matches the visuals.
Head Mounted Display (HMD) imagery without any physical surface to touch.
We are interested in comparing cognitive load, usability, and subjective ratings.

The task involved estimating the size of the dots on the surface while performing a secondary task of counting. The size estimation is similar to common tasks in healthcare training such as making decisions about how dangerous a mole is based on its size, or estimating the surface of burns on a patient's body to determine the seriousness of the condition. The counting task uses cognitive resources needed to estimate the size on the surface.

Our study shows that for AR-based touch tasks, the cognitive load increases when there a mismatch between the visual and tactile perception. Participants preferred and experienced lower cognitive load in the SAR condition with geometry that matches the virtual content. Participants subjectively found SAR easier to use than HMDs.

6.3 Experimental Setup

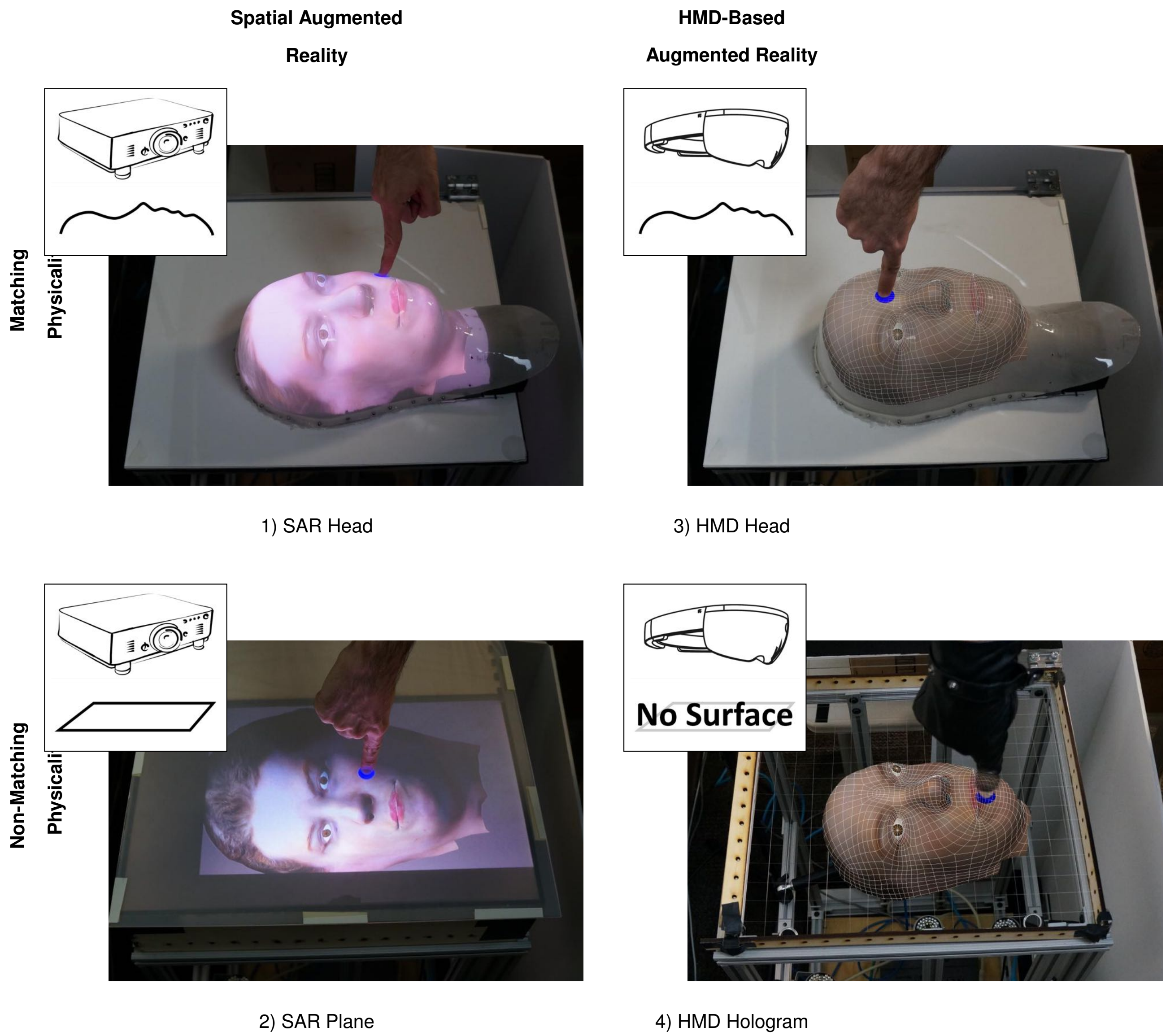


Figure 6.1: Participants experienced four study conditions, each differing in physical and virtual representation of a 3D human head model. The physical object with which they interacted either matched or did not match the virtual object; the virtual object was displayed either via a projector or through an HMD. For the HMD conditions, the imagery shown above is simulated.

This section describes the apparatus for each of the 4 conditions, all using a human head model.

We use the terms *SAR Head*, *SAR Plane*, *HMD Head*, and *HMD Hologram* (Figure 6.1).

In the SAR condition, a projector provided the imagery. In the HMD conditions a Microsoft HoloLens provided the imagery. Participants touched the physical head-shaped surface in the SAR Head and HMD Head conditions, they touched a physical flat surface in the SAR Plane, and touched in midair using a clicker in the HMD Hologram case. The clicker comes with the HoloLens and it is common for it to be used as an input device in practice. We considered other alternative to recording touches such as leaving the finger longer but that could affect other measures such as performance and could compromise the priority for the dual task. We did not include audio in increasing the cognitive load as the study was designed so the interaction is only limited to the visual and physical perception.

The study platform for the Physical-Virtual Head is described in chapter 3. The platform consisted of an aluminum rig that holds Point Grey monochrome cameras , IR filters, projector, and IR lights and was used in all four conditions (Figure 6.2).

We attached mounts to the rig to be able to switch between the head-shaped surface, flat surface, and no surface without affecting the calibration. The 3d head was attached using hinges on the frame, and when the 3D head was not in use (flipped to the side) , I designed and built a wooden support that sits on top of the rig for two purposes: (1) to support the flat surface in a raised height to match the head in the SAR Plane condition. (2) the wooden support had a wire grid to guard the equipment from being accidentally touched by participants in the case they go too far in the HMD Hologram condition.

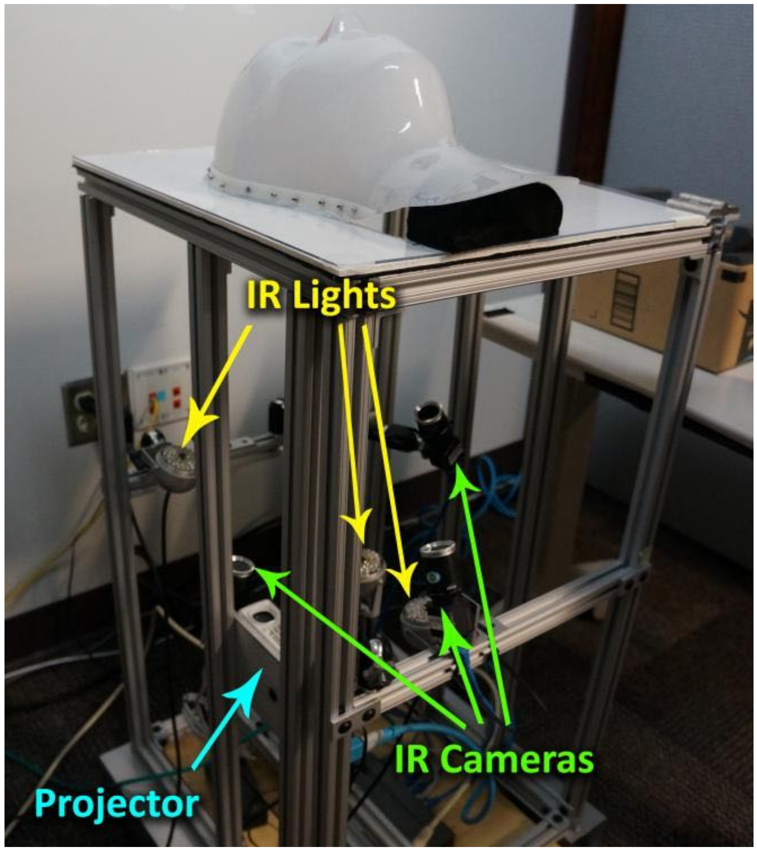


Figure 6.2: Study platform: the physical head surface, IR cameras and rear-diffused light used for touch sensing, and projector for the SAR conditions. The head surface is attached to the aluminum frame with hinges and can be replaced with a flat surface or removed for the relevant study conditions.

This chapter does not go into the technical details of calibration, touch detection, and alignment of the visuals as those fall under Jason's main contribution, not mine. You can find the details in the paper [44].

6.4 Experiment

This section describes the experiment to investigate the differences between four physical-virtual representations of the 3D human head model capable of sensing touch. I am interested in the usability effects, cognitive load, and subjective preference.

6.4.1 Participants

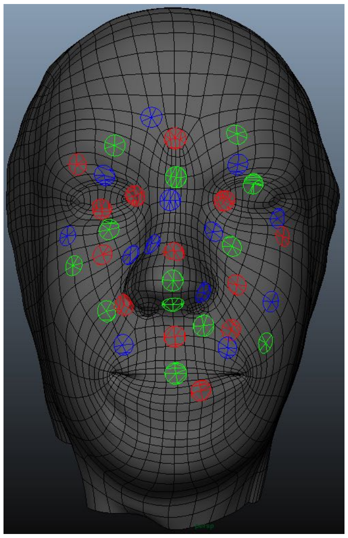
There were 24 participants (14 males and 10 females) from the University of Central Florida community with ages ranging from 18 to above 50 who participated in out study. All participants had good vision (correct or corrected, 8 wore glasses) without known visual or motor disorders. We asked for handedness, there were 22 right-handed, 2 left handed, and one left-handed participants choose to use their right hand to accommodate for a medical condition. There were 14 participants who had experience with 3D content and 3 have previously used the HoloLens. For each participant, we measured the inter-pupillary distance ($M = 6.12\,\mathrm{cm}$, $SD = 0.3\,\mathrm{cm}$) and applied it to the rending content on the HoloLens. The study had a within-subject design as we expected interpersonal differences in touch behavior and performance. We asked each participant to complete all 4 study conditions (Figure 6.1). We used Latin square design to counterbalance the order of which we presented the conditions. Each condition had two phases: *Touch Accuracy* and *Cognitive Load*.

6.4.2 Phases and Tasks

Visual targets of different sizes were displayed on the head and participants were asked to touch the medium sized target. We chose a set of 39 vertices across the 3D model as the location for the targets. (**??**(a)).

I created the circular targets in Maya by intersecting spheres (small ($5\,\mathrm{mm}$), medium ($7.5\,\mathrm{mm}$), and large ($10\,\mathrm{mm}$)) with the face mesh. The intersection left us with roughly circular targets. The targets were consistent across all conditions (i.e., if we used spheres, the HoloLens condition would have looked different than the SAR condition where the spheres would look flattened).

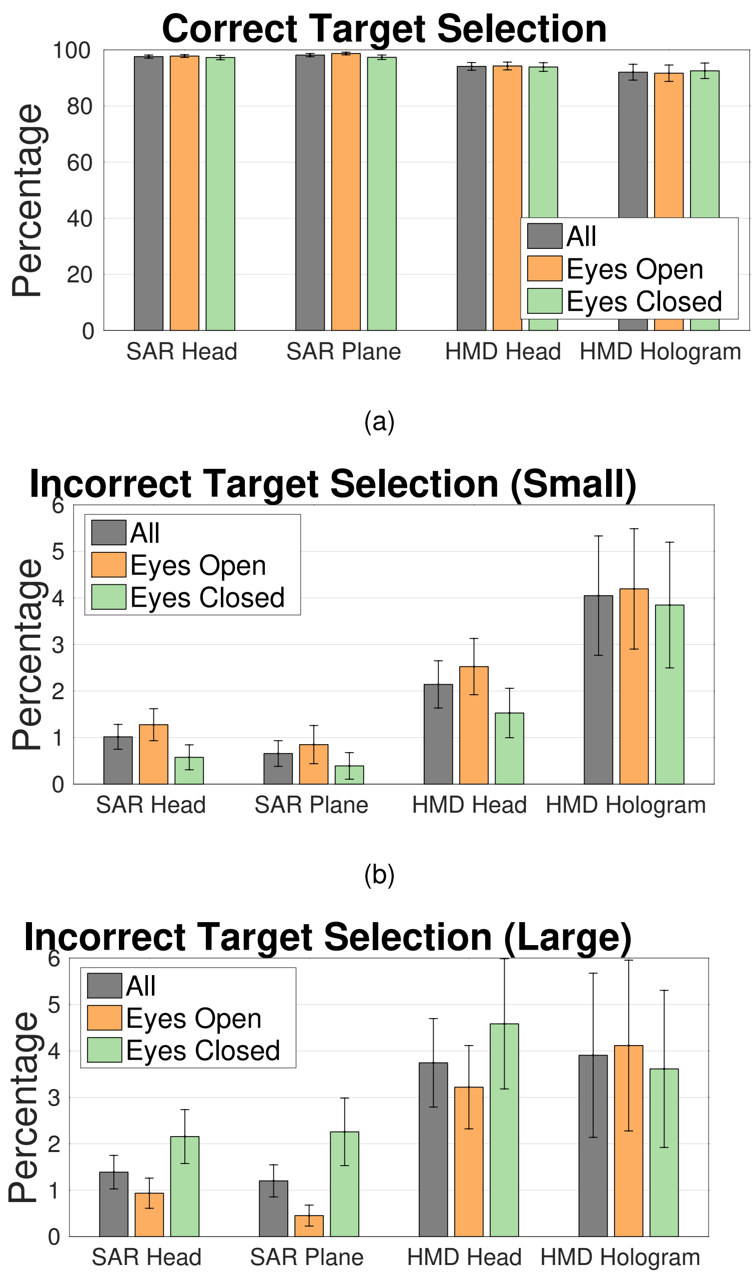
As soon as the participant touched the surface or clicked the clicker, the touch was recorded.



(a)

(b)

Figure 6.3: Visual targets. ((a)) The 39 touch targets across the 3D head. These targets were further divided into small (5 mm, shown in red), medium (7.5 mm, green) and large (10 mm, blue) groups for the *Touch Accuracy Phase*. All targets are shown above as small-sized targets for visualization purposes. ((b)) Example cognitive load trials. In all trials, one small, one medium, and one large target are shown to participants. They are tasked with touching the medium-sized target. When the virtual human's eyes are closed (right), participants must also provide a verbal response related to a secondary counting task (subtraction by 7 from a given starting number around 500).



(a)

(b)

(c)

Figure 6.4: Results for the target selection task in the *Cognitive Load Phase*. ((a)) The percentage of trials for which participants correctly touched the medium-sized targets. ((b)) and ((c)) The percentage of trials for which participants incorrectly touched either the small- or large-sized targets, respectively. We found significant main effects of display condition on medium (correct) and small (incorrect) selections. Questionnaire responses indicated significantly higher perceived task load and significantly lower usability for the HMD Hologram compared to the SAR Head (Figure 6.5).

6.4.2.1 Touch Accuracy Phase

During this phase participants were asked to touch the center of each dot displayed sequentially and to press spacebar when done to gave the same starting point for all targets. Each participant touched 78 dots from a pool of 39 targets divided into 13 small, 13 medium , and 13 large. The order of targets was the same for any given condition.

6.4.2.2 Cognitive Load Phase

In this phase we asked participants to perform a dual-task to compete with their finite cognitive resources. The main task was to pick the medium target from a set of 3 displayed on the head (one small, one medium, and one large target). We presented from 39 different variations (**??**(b)) where each of the medium ones gets shown twice. The other two dots (small and large) were there to distract the participant. In total they were shown 78 cognitive load trials. The order of targets was randomized, predetermined, and consistent across all participants. The secondary task consisted of asking the participants to count backward by 7 Participants started with a number around 500. The are asked to verbally indicate their results only if the eyes of the virtual human were closed. They are not supposed to provide a verbal feedback if the eyes are open. The eyes of the virtual human can only change with each new set of targets. The participants still had to update the count even if the eyes remained closed for two consecutive trials. The targets automatically advanced from one trial to another every 4 seconds without the need for the participant to press the spacebar. The priority was to touch the medium target before updating the count for the secondary task.

6.4.3 Study Procedure

After reading the informed consent form and completing a demographic questionnaire, participants watched a video sample (2-minute) as examples of the cognitive load tasks advancing at the same speed as the actual task. The purpose of the video is to practice counting backward by 7 when the virtual human closes her eyes, therefore the dots in the video were all the same size so the participants can focus only on practicing the counting task. The next part involved a short training with 6 examples of touch accuracy and 25 examples of cognitive load tasks. They were given a last change to ask questions before the phases start. Participants were given a 60 second break between the *Touch Accuracy Phase* and the *Cognitive Load Phase* so they can rest their arms, neck, and eyes. Participants were given a 10 seconds notice before the beginning of the *Cognitive Load Phase* Participants were given subjective questionnaires to measure cognitive load and usability at the end of each each condition. After they finished all the four conditions I interviewed the participants for a qualitative feedback and asked them to rank the conditions from easiest to hardest.

6.4.4 Measures

In the *Cognitive Load Phase* participants were asked to pick the medium target and to counts backwards when the virtual human closes her eyes. We logged the verbal responses, verified them in real time and compared to videos. We computed the percentage of correct responses (*verbal counting task response—secondary task*). After each condition, participants completed the NASA Task Load Index (TLX) questionnaire [1] to measure cognitive load, and the Simple Usability Scale (SUS) questionnaire [245] to measure usability. At the end of the experiment we interviewed the participants for qualitative answers and ranking of conditions from easiest to hardest.

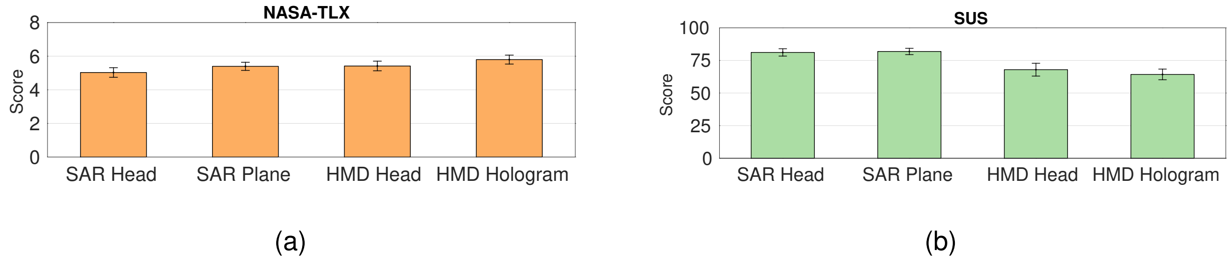


Figure 6.5: Results of the subjective questionnaires for the four experimental conditions: ((a)) NASA Task Load Index (TLX) and ((b)) Simple Usability Scale (SUS). Note that lower is better for NASA-TLX, whereas higher is better for SUS. We found significant main effects of display condition on both scores. These results align with objective user performance in terms of touch-target accuracy and response time.

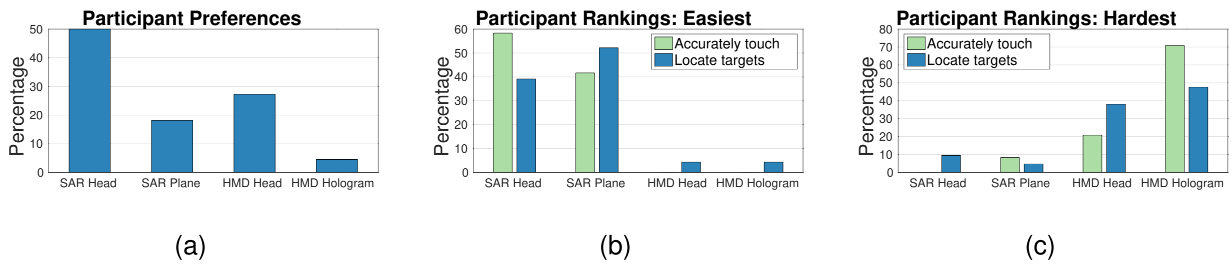


Figure 6.6: Subjective participant rankings for the four experimental conditions. ((a)) Overall preferred condition. ((b)) and ((c)) Subjectively ranked easiest and hardest conditions, respectively, for accurately touching and locating the visual targets. Participants preferred interacting with the rear-projection head, since it affords physical feedback on touch, and they found the HoloLens uncomfortable and disliked the field of view limitations. In general, participants did objectively perform the best on the conditions they subjectively found easiest **??**).

6.5 results

We used descriptive and inferential statistics for the *Cognitive Load* phase, and subjective questionnaire responses. We used repeated-measures ANOVAs and Tukey multiple comparisons with Bonferroni correction at the 5% significance level to analyze the results. We used Shapiro-Wilk tests at the 5% level and QQ plots to confirm normality. Degrees of freedom were corrected using Greenhouse-Geisser estimates of sphericity when Mauchly's test indicated that the assumption of sphericity had been violated. We analyzed the questionnaire responses using parametric statistical tests in line with the ongoing discussion in the field of psychology indicating that parametric statistics can be a valid and often more expressive method for the analysis of ordinal data measured by the experimental questionnaires [225, 226].

6.5.1 Cognitive Load

In the *Cognitive Load Phase*, participants were under cognitive pressure yet they had to choose the correct target We found significant main effects of display condition on medium (correct) selections and on small (incorrect) selections.. The full results related to touch accuracy and performance are detailed in the paper [44]

6.5.2 Subjective Responses

the Cognitive load, usability, and subjective participant rankings of easiness to accurately touch and visually locate targets were significantly different between display condition.

6.5.2.1 Task Load

Below are the results of the NASA-TLX questionnaire Figure 6.5(a).

We found a significant main effect of the display conditions on the NASA-TLX task load scores, $F(3, 69) = 3.40, p = .023, \eta_p^2 = .129$. Pairwise comparisons indicate the task load for HMD Hologram was a significantly ($p = .033$) higher compared to the SAR Head. There were no other significant pairwise effects.

6.5.2.2 Usability

Below are the results of the SUS questionnaire Figure 6.5(b).

We found a significant main effect of the display conditions on the SUS usability scores, $F(1.75, 40.22) = 9.95, p = .001, \eta_p^2 = .302$. Pairwise comparisons indicate that usability for the SAR Head a significantly higher ($p = .003$) compared to the HMD Hologram. Similarly the usability for the SAR Plane compared was significantly higher ($p = .002$) than the HMD Hologram . Results showed a trend the SAR Head had higher usability compared to the HMD Head ($p = .074$), and a trend indicating the SAR Plane had higher usability compared to the HMD Head ($p = .067$)

6.5.2.3 Preference

This section reports on the post-experiment interview where participants were asked to indicate their preference between the conditions.

There were 11 participants who preferred the SAR Head over the other conditions, the results are significant ($z = 2.357$, $p = 0.018$) (Figure 6.6(a)). Many participants indicated that the SAR Head was easier than the other conditions. Participants liked the physicality of the head surface, and did not like the complexity of wearing an HMD. Some participants were excited to try new technology (i.e., HoloLens), others had comfort related complaints regarding the HMD's weight being too heavy that it bothered their neck and nose, and was uncomfortable. The narrow field of view of the HoloLens was obvious to all participants to the extent that several limited the movement of their head movement when they were wearing the HMD; for example many avoided turning their head when pressing the spacebar key during the *Touch Accuracy Phase*. Some qualitative comments support these preferences, such as, *The SAR Head was easier to use "because [it was] physically there," The HoloLens was "uncomfortable" and "very heavy,"* and, *The SAR Plane was easier to* *use because "[I am] used to seeing [imagery] on [a] screen."*

6.5.2.4 Rankings

Participants were asked to give a ranking to the conditions from easiest to hardest in terms of visually locating the targets and accurately touching them (Figure 6.6(b)) (Figure 6.6(c))

We used to exact Clopper-Pearson confidence interval [246, 247] to calculate the ranking. There were 14 participants who ranked that the SAR Head as the easiest ($z = 3.771$, $p < 0.001$), and 17 participants who ranked the HMD Hologram as the hardest ($z = 5.185$, $p < 0.001$).

In terms of feeling that they touched the targets accurately, 14 participants felt that touch accuracy was easiest with the SAR Head ($z = 3.771$, $p < 0.001$), and 17 ranked the HMD Hologram as the hardest ($z = 5.185$, $p < 0.001$). As for visually identifying the targets, 21 expressed that it was easier with the SAR conditions ($z = 3.674$, $p < 0.001$), and 18 felt it was harder with HMD conditions than the SAR conditions ($z = 2.449$, $p = 0.014$).

6.6 Discussion

For trainers interested in AR with touch tasks, we suggest the following guidelines:

For training tasks that require accurate touching and where it is important to have a decreased cognitive load and increased usability, projector-based displays should be used as opposed to HMDs. If the projected imagery is not an option, an interaction using and HMD that augments a physical object is preferable to an HMD in free space. This discussion covers both the objective and subjective measures supporting these findings.

In general the SAR conditions were preferred, easier and more intuitive. Specifically There was a strong preference for the SAR Head condition. It is possible that this effect was because the interface was more simple and more friendly compared to HMD conditions. The SAR Plane was ranked as easiest for locating the targets, probably due to the participants familiarity with similar devices their everyday lives (e.g. smart phones and touchscreens). Also, the SAR plane offers the advantage of showing all targets at once compared to the SAR Head where participants had to look around form multiple view points to locate targets. That being said, the SAR Head was ranked as being easier to accurately touch than the SAR Plane, probably because the physical head surface matches the imagery more closely.

During the *Cognitive Load Phase*, participants made more mistakes in the two HMD conditions, indicating an increase in cognitive load over the SAR conditions. This was supported in the subjective results where participants felt the SAR Head was the easiest and the HMD Hologram the hardest during the task (i.e., remember to observe the virtual human's eyes and to update the counts); however, participants on average performed similarly in the counting task in all conditions.

Participants preferred the conditions where the imagery matched the object for interaction (SAR Head and HMD Head). It was evident in the objective and subjective data that the HoloLens adds to the physical and mental demands on the user. Also, the HoloLens overpowers non-HoloLens imagery (e.g., hands, environment) since the imagery directly displayed in front of a user's eyes. The use of the clicker in the HMD Hologram condition to submit a touch was less enjoyable and less user-friendly. The physical touch feedback from the physical surface improved the experience with the HMD compared to the HMD alone. It is possible that an HMD with a larger field of view and lighter weight could have produced a more enjoyable interaction. However, current HMDs have issues (e.g. occlusion of imagery, wearing a device) that would likely have similar effects on comparable AR touch tasks.

6.7 Conclusion and Future Work

In this chapter we compared four physical-visual representations of a touch-sensitive 3D human head model how how these representations impacted the cognitive load, and subjective preferences.

Overall, participants preferred interacting with a physical surface with geometry that matches the 3D model supporting **[TS2]** They also found that rear-projected imagery was easier as opposed to provided by an HMD. I am interested in how theses results impact the training in the medical field. We used used a state-of-the-art technology to represent patients physically and virtually which could replace or complement traditional mannequins in healthcare training.

CHAPTER 7: SIMULATOR-SCENARIO COMPATIBILITY

The work in this chapter has not been published. It is in support of **[TS 4]**

7.1 Motivation

Educators use scenarios to help achieve one or more learning objective(s). Each scenario needs multiple cues where each cue varies in importance. The importance of each cue to the scenario is determined by the educator designing or using that scenario to meet the learning objective(s).

A simulator provides several cues, each of those cues varies in fidelity. For each simulator, anyone who uses that simulator would evaluate the fidelity of each cue to indicate how well that cue is represented using this simulator.

Ideally the cues needed by a scenario are provided by the simulator. A scenario is satisfied when the needed cues are provided by the simulator. A scenario is NOT satisfied when the needed cues are NOT provided by the simulator. Educators care about the scenarios being satisfied or in other words they care about increasing the Scenario Satisfaction (S).

On the other side, the cues provided by the simulator can match the scenario's needs or they can be provided as extras. When the Simulator provides more cues than the scenario needs, this does not increase the Scenario Satisfaction, but decreases the Simulator Utilization (U) as the simulator would be under-utilized. A simulator is well utilized (high U) when the cues provided are needed without providing cues that are not needed.

On the cue level, if the cue is important for the scenario and is provided with high fidelity then there is a match. Similarly, if the cue is of low importance and it is provided with low fidelity then there is a match. Vice versa, if the cue is important and is provided with a low fidelity then this negatively affects the satisfaction of the scenario therefore this is not a match. If the cue is not important to the scenario and is provided with a high fidelity, while it does not affect the scenario satisfaction, it does negatively affect the utilization of the simulator as there is a waste of resources therefore this is NOT a match.

The aim of this method is to extend this idea from one cue to an overall simulation evaluation in a way that combines all cues.

Whether you are an educator or an administrator managing a simulation center, it is important to know how well does a specific simulator match the learning objective for a specific scenario, and how do the existing simulators compare between each other for that same scenario.

So far, there is no quantitative method to evaluate the goodness of a simulator in terms of meeting learning objectives and in terms of proper utilization of the simulation. The choice to use a specific simulator largely depends on first having access to the simulators (i.e., the simulation center bought that simulator), that simulator is suitable to teach certain learning objectives using specific scenarios, and that simulator being available at the time when it is needed.

From the perspective of someone willing to invest in a new simulator (e.g. administrator), the new simulator needs to have capabilities compatible with the learning objectives as set by the educators, and the price needs to fit into the budget. In general, a simulator that offers more capabilities would be more expensive compared to another simulator that offers less capabilities. It would make sense to want to maximize the match between what is needed for the scenario and what is provided by simulator, and at the same time try to avoid paying for simulator capabilities that are not needed.

If the simulator does not offer enough capabilities to satisfy what is needed by the scenario, there is a low overall match. If the simulator offers enough capabilities to satisfy what is needed by the scenario, there is a high match. If the simulator offers more capabilities than what is needed, there is a low utilization and a high waste in resources, so the simulator is under-utilized (only a part of what it can provide is being used). If the simulator does NOT offer any more capabilities than what is needed (it offers just what is needed), there is low waste in resources, so the simulator has a high utilization.

Every Simulation has a Scenario and a Simulator. Every Simulator provides a set of Cues with varying fidelity depending on its capabilities. Every Scenario depends on certain Cues. Some cues are more important than others (See Figure 7.1). In a simplified world, each cue on the scenario side can be of low or high importance. That same cue can either be represented by the simulator in a low fidelity manner or a high fidelity manner. Figure 7.2 shows a simple graph of Importance-Fidelity pair for each cue. For a cue to be in a good importance-fidelity match it has to have either a (low importance AND low fidelity) or (high importance AND high fidelity). For a pair to be in a bad match it has to have either a (low importance AND high fidelity) or (high importance AND low fidelity).

In the real world, the values are not discrete, therefore if you are on the 45 degree line (line of slope = 1) it means the cues are matched. If you are away from that line it means you are either not satisfying the scenario (for that cue) or under-utilizing the simulator (for that cue).

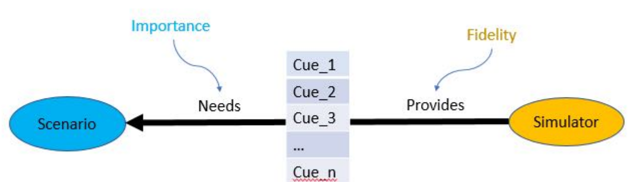


Figure 7.1: The simulator provides cues with a certain fidelity for each cue. The Scenario needs cues where each cue has a specific importance.The cues are the same for the simulator and the scenario

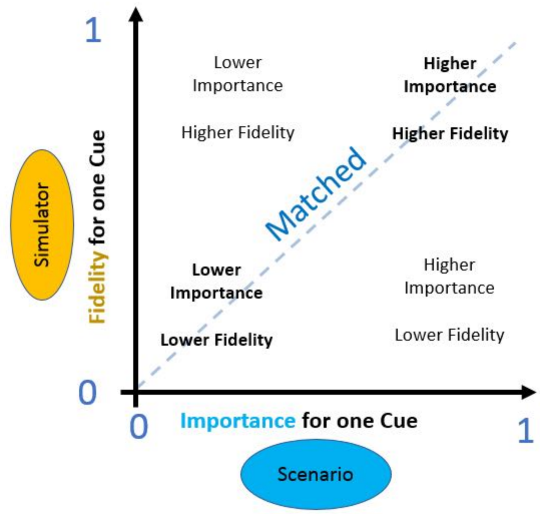


Figure 7.2: Importance-Fidelity goodness of match for one cue only. It is a good match. When the Importance is low (value closer to 0) and the Fidelity is low (value closer to 0), or when the importance is high (value closer to 1) and Fidelity is high (value closer to 1). It is a bad match (shown in red) when the Importance is high (value closer to 1) and the Fidelity is low (value closer to 0), or the Importance is low (value closer to 0) and the Fidelity is high (value closer to 1).

In the sections below I characterize the Scenario and the Simulator in terms of Cues for both importance and fidelity, then I describe a method to compute the overall Scenario Satisfaction (S), Simulator Utilization (U) for a specific scenario. One could derive Waste (W) for a specific scenario from the (U). The S, U and W are used to plot the results and visually compare simulators.

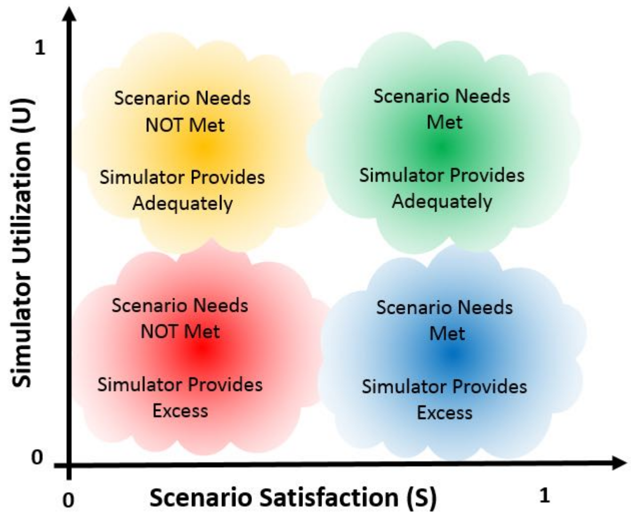


Figure 7.3: Overall Simulator Compatibility. The regions are color coded to indicate the Satisfaction and Utilization (and waste) of the simulator for a specific scenario. The red area near the origin represents the worst case where the simulator is a bad match with low scenario satisfaction and high waste from capabilities that are not needed and not being used. The Green area furthest away form the origin represents the best world where the simulator is a good match with high scenario satisfaction and it is being utilized appropriately reducing waste. The Yellow area means that the simulator is not suitable for this scenario so in that respect it is a bad match where the scenario is not satisfied, but this simulator is utilizing a significant number of its resources without wasting much, it is still not enough. The Blue area means it is a good match in term of scenario satisfaction, but in addition to that there are features of the simulators that are not being used indicating possible high waste of resources.

7.2 Characterizing the Scenario

Let n be the total Number of Cues.

Let i_{cue} be the importance of one scenario cue, with "cue" ranging from 1 to n to determine which cue is being evaluated. Experts are asked to provide a real value for i_{cue} between 0 or 1. A value of 0 indicates the lowest possible importance of this cue for this scenario (i.e., not important at all). A value of 1 represents the highest importance of this cue for this scenario (i.e., extremely important). If there are more than one expert, the average is recorded as the raw-value of i_{cue}. Raw values of i_{cue} are obtained from the experts for Cue importance of scenario with a range of real numbers from [0-1].

Cue Importance Vector $\vec{I}$ is made out of n elements of scenario cue importance i_{cue}. Each vector $\vec{I}$ represents the rating of the importance of each cue (real values between 0 and 1) in the whole scenario. For example for n cues in a scenario, the Importance Vector is

$$\vec{I} = \{i_1, i_2, \ldots, i_n\}$$

The unit for i_{cue} is in *imp*.

7.3 Characterizing the Simulator

Let n be the total Number of Cues.

Let f_{cue} be the fidelity of one simulator cue, with "cue" ranging from 1 to n to determine which cue is being evaluated. Experts are asked to provide a real value for f_{cue} between 0 and 1. A value of 0 indicates the lowest possible fidelity of this cue for this simulator (i.e., non existent). A value of 1 represents the highest possible fidelity of this cue for this simulator (i.e, very realistic, no different than a real patient). If there are more than one expert, the average is recorded as the raw-value of

f_{cue}. Raw values of f_{cue} are obtained from the experts for Cue fidelity of the simulator with real values ranging from [0-1].

Cue Fidelity Vector $\vec{F}$ is made out of n elements of simulator cue fidelity f_{cue}. Each vector $\vec{F}$ represents the rating of the fidelity of each cue (real value between 0 and 1) in the whole simulator. For example for n cues in a simulator, the Simulator Cue Fidelity Vector is

$$\vec{F} = \{f_1, f_2, \ldots, f_n\}$$

Units for f_{cue} are in *fid*

7.4 Method to assess compatibility

Total Number of Cues = n

Scenario Cue Importance Vector $\vec{I} = \{i_1, i_2, \ldots, i_n\}$

Simulator Cue Fidelity Vector $\vec{F} = \{f_1, f_2, \ldots, f_n\}$

Scenario-Satisfaction for a specific Simulator

$$S = \frac{(i_1 * f_1) + (i_2 * f_2) + \ldots + (i_{cue} * f_{cue})}{i_1 + i_2 + \ldots + i_{cue}} \tag{7.1}$$

$$S = \frac{\vec{I}.\vec{F}}{\sum\limits_{cue=1}^{n} i_{cue}} \tag{7.2}$$

Simulator Utilization for a specific Scenario

$$U = \frac{(i_1 * f_1) + (i_2 * f_2) + \ldots + (i_{cue} * f_{cue})}{f_1 + f_2 + \ldots + f_{cue}} \tag{7.3}$$

$$U = \frac{\vec{I}.\vec{F}}{\sum_{cue=1}^{n} f_{cue}} \tag{7.4}$$

Simulator Utilization + Simulator Waste = 1

Simulator Waste for a Scenario

$$W = 1 - U \tag{7.5}$$

The values of S, and U are used to compute two indicators in a polar coordinate system. The primary indicator is an indicator of overall Goodness:

$$Goodness = \frac{\sqrt{S^2 + U^2}}{\sqrt{2}} \tag{7.6}$$

The secondary indicator is an indicator of Balance between what is needed and what is provided

$$Balance = \frac{U}{S} \tag{7.7}$$

7.5 Method Interpretation

It this section I define thresholds as real numbers between 0 and 1 used to help interpret the results. The Goodness Threshold T_G helps classify the simulator-scenario pairs by an overall goodness of match. The larger the G the better, if it is above the T_G then it is considered good. If G is below the T_G the simulator-scenario pair are considered not good.

The two Balance Thresholds: Balance Threshold for Lack of Resources T_{BL} and Balance Threshold for Waste of Resources T_{BW} help classify the simulator-scenario pairs as optimal in terms of both satisfaction and utilization, or not optimal and differentiate which is not optimal (the satisfaction or the utilization).

For illustration purposes these numbers are chosen in the middle of the visual space as default values (T_G = 0.5), (T_{BL} = 0.5), (T_{BW} = 2). In real life these thresholds would be moved to lower or higher delimiting the colored regions (Green, Red, Yellow, and Blue) in Figure 7.4 to better fit specific needs, but the idea is still the same.

IF G < T_G

    Low Satisfaction, Low Utilization (high waste)

IF G >= T_G

    IF B > = T_{BL}

        Low Satisfaction, High Utilization (simulator well utilized but still not good enough)

    IF T_{BL} > B >= T_{BW}

        High Satisfaction, High Utilization (low waste)

        (Simulator is a Perfect Match)

    IF B < T_{BW}

        High Satisfaction, Low Utilization (high waste)

        (Expensive Simulator, Waste of Resources)

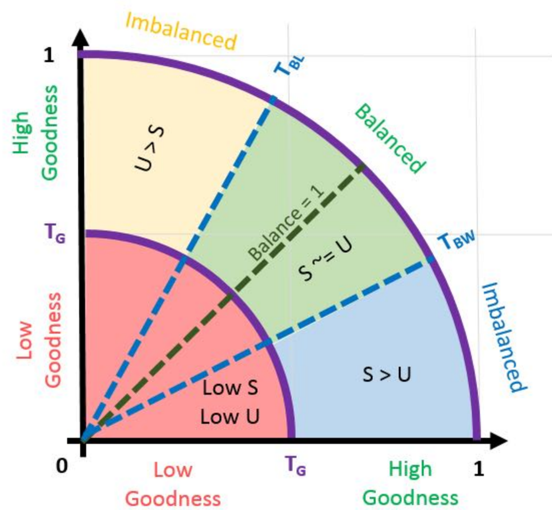


Figure 7.4: Graph showing colored regions for each category. The red area near the origin represents the worst case where the simulator is a bad match and has a high waste from capabilities that are not needed and not being used. The Green area is the opposite as it represents the best world where the simulator is a good match and it is being fully utilized appropriately reducing waste. The Yellow area means that the simulator is not good enough so in that respect it is a bad match. The simulator in yellow is utilizing a significant number of its resources without wasting much, it is still not enough. The Blue area means it is a good match and it satisfies the needs of the scenario, in addition to that there are features of the simulators that are not being used indicating possible high waste of resources.

7.6 Examples

This section contains examples, first some general examples to illustrate how a simulator can end up in one category or region in the graph. I follow that with simplified examples that relate back to the scenarios and simulators mentioned in Chapters 4 and 5.Note that the numbers used in the examples are 0 or 1 just for illustration purposes. The same calculations can be made with real numbered between 0 and 1 inclusive.

7.6.1 General Example A: Good Satisfaction, High Utilization, Low Waste

Scenario $\vec{I}$[1,1,1,0,0,0]

Simulator $\vec{F}$[1,1,1,0,0,0]

$S = \frac{(1*1)+(1*1)+(1*1)+(0*0)+(0*0)+(0*0)}{1+1+1+0+0+0}$

$S = \frac{3}{3} = 1$

$U = \frac{(1*1)+(1*1)+(1*1)+(0*0)+(0*0)+(0*0)}{1+1+1+0+0+0}$

$U = \frac{3}{3} = 1$

$W = 1 - U$

$W = 0$

$G = \frac{\sqrt{S^2+U^2}}{\sqrt{2}}$

$G = \frac{\sqrt{1^2+1^2}}{\sqrt{2}}$

$G = 1$

$B = \frac{U}{S}$

$B = \frac{1}{1}$

$B = 1$

Result = Perfect match, No Waste

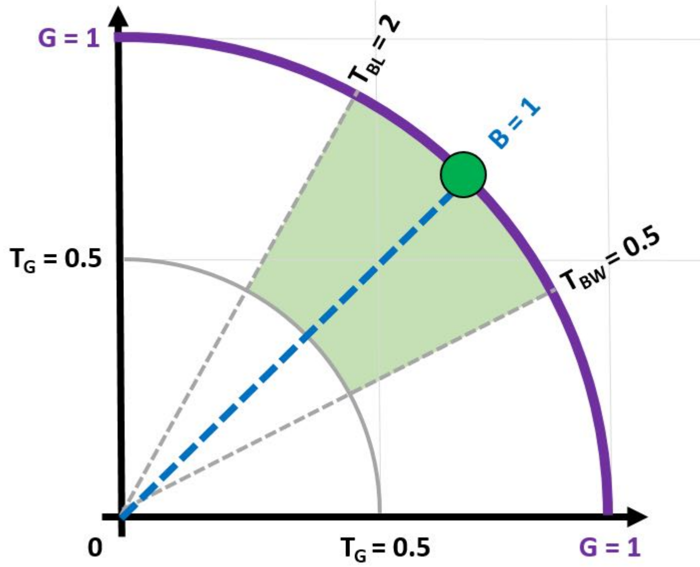


Figure 7.5: The green dot is in the green region indicating a good match where the scenario is satisfied, the simulator has high utilization, and low Waste.

7.6.2 General Example B: Low Satisfaction, High Utilization, Low Waste

Scenario $\vec{I}$[1,1,1,0,0,0]

Simulator $\vec{F}$[1,0,0,0,0,0]

$S = \frac{(1*1)+(1*0)+(1*0)+(0*0)+(0*0)+(0*0)}{1+1+1+0+0+0}$

$S = \frac{1}{3} = 0.33$

$U = \frac{(1*1)+(1*0)+(1*0)+(0*0)+(0*0)+(0*0)}{1+0+0+0+0+0}$

$U = \frac{1}{1} = 1$

$W = 1 - U$

$W = 0$

$G = \frac{\sqrt{S^2+U^2}}{\sqrt{2}}$

$G = 0.75$

$B = \frac{U}{S}$

$B = \frac{3}{1}$

$B = 3$

Even though there is no waste and it is fully utilized, this simulator is NOT Good Enough for this scenario.

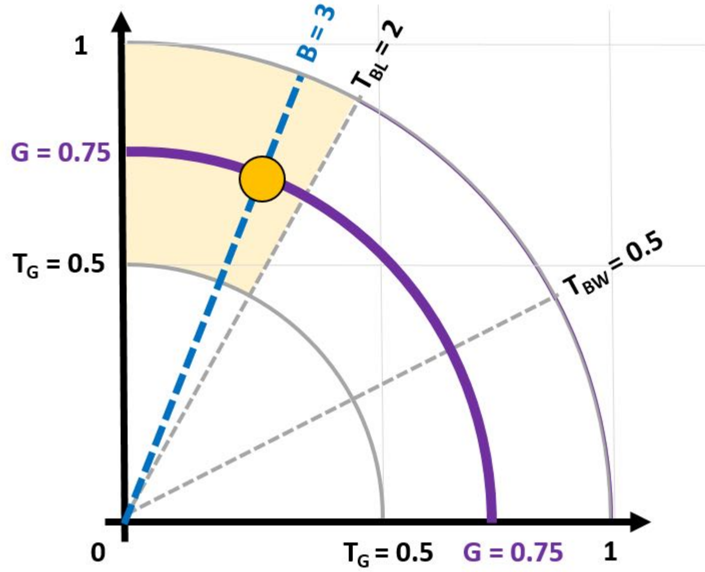


Figure 7.6: The yellow dot is in the yellow region indicating a bad match, high utilization, and low Waste.

7.6.3 General Example C: High Satisfaction, Low Utilization, High Waste

Scenario $\vec{I}$[1,1,0,0,0,0]

Simulator $\vec{F}$[1,1,1,1,1,1]

$S = \frac{(1*1)+(1*1)+(0*1)+(0*1)+(0*1)+(0*1)}{1+1+0+0+0+0}$

$S = \frac{2}{2} = 1$

$U = \frac{(1*1)+(1*1)+(0*1)+(0*1)+(0*1)+(0*1)}{1+1+1+1+1+1}$

$U = \frac{2}{6} = 0.33$

$W = 1 - U$

$W = 0.67$

$G = \frac{\sqrt{S^2+U^2}}{\sqrt{2}}$

$G = 0.75$

$B = \frac{U}{S}$

$B = 0.33$

Results = Simulator is a good match for the scenario, but it provides capabilities that are not needed, therefore wasting resources for this scenario.

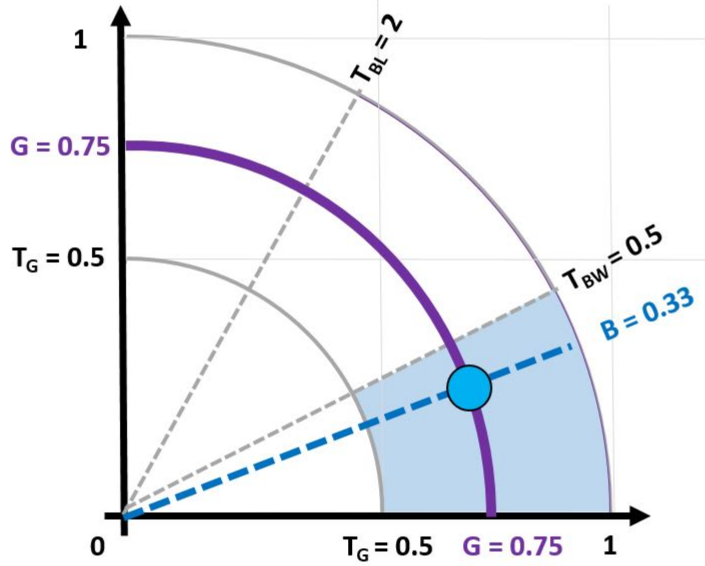


Figure 7.7: The blue dot is in the blue region indicating a good match, low utilization, and high waste.

7.6.4 General Example D: Low Satisfaction, Low Utilization, High Waste

Scenario $\vec{I}$[1,1,1,1,0,0]

Simulator $\vec{F}$[0,0,0,1,1,1]

$S = \frac{(1*0)+(1*0)+(1*0)+(1*1)+(0*1)+(0*1)}{1+1+1+1+0+0}$

$S = \frac{1}{4} = 0.25$

$U = \frac{(1*0)+(1*0)+(1*0)+(1*1)+(0*1)+(0*1)}{0+0+0+1+1+1}$

$U = \frac{1}{3} = 0.33$

$W = 1 - U$

$W = 0.67$

$G = \frac{\sqrt{S^2+U^2}}{\sqrt{2}}$

$G = 0.29$

$B = \frac{U}{S}$

$B = 1.33$

Result = Simulator is a bad match and wastes resources for this scenario.

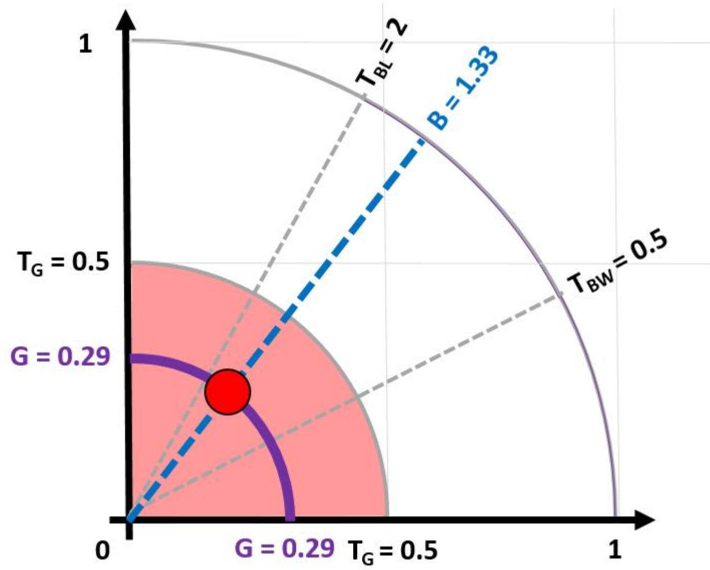


Figure 7.8: The red dot is in the red region indicating a bad match, low utilization, and high waste.

Here are a few simplified examples for stroke to match the simulators and scenario in chapter 4. These examples are for illustration purposes only. The rating of the cues importance was not validated by subject matter experts.

7.6.5 *Stroke Scenario Example for PVHead vs Mannequin*

Cues	Facial	Tongue Movement	Pupil Reactions	Speech Quality	Realtime Touch Reactions	Lift Arms	Pulse	Heart Sounds
Importance	1	1	1	1	1	1	0	0
PVHead Fidelity	1	1	1	1	1	0	0	0
Mannequin Fidelity	0	0	1	1	0	0	1	1

Figure 7.9: Table showing simplified example of a stroke scenario with important cues and simulator fidelity for each of these cues. These values are for illustration only and they are not externally validated.

7.6.5.1 *Score for Stroke-PVPHead pair*

Scenario $\vec{I}$ [1,1,1,1,1,1,0,0]

Simulator $\vec{F}$ [1,1,1,1,1,0,0,0]

$S = 5/6 = 0.83$

$U = 5/5 = 1$

$W = 0$

$G = 0.92$

$B = 1.2$

Result = Good match, No Waste

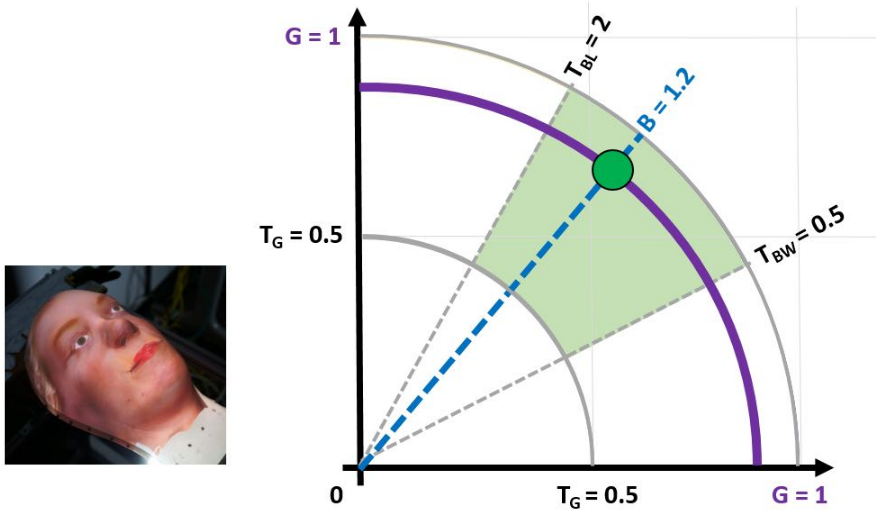


Figure 7.10: Graph showing a good match, no waste.

7.6.5.2 Score for Stroke-Mannequin pair

Scenario $\vec{I}$ [1,1,1,1,1,1,0,0]

Simulator $\vec{F}$ [0,0,1,1,0,0,1,1]

$S = 2/6 = 0.33$

$U = 2/4 = 0.50$

$W = 0.50$

$G = 0.42$

$B = 1.51$

Result = Bad match + Waste

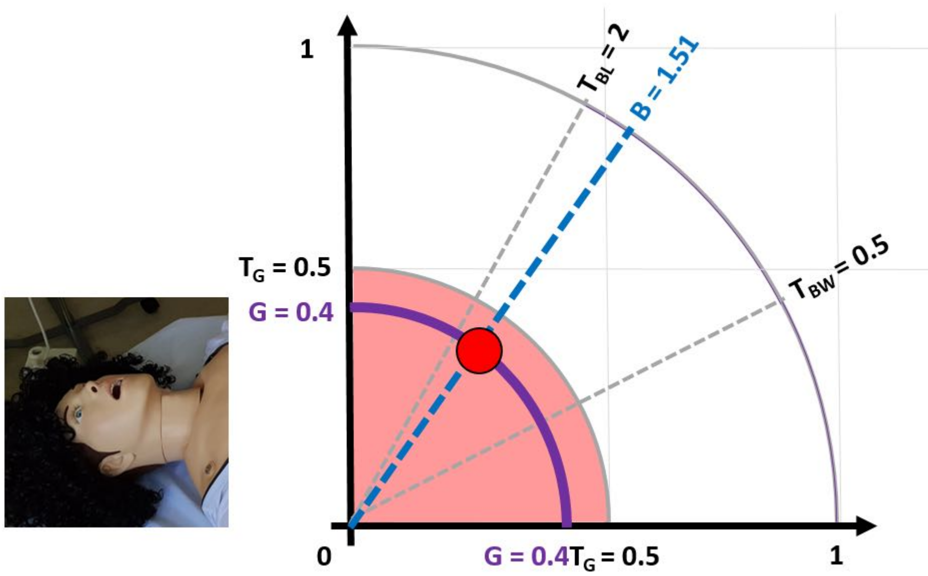


Figure 7.11: Graph showing simplified example of a stroke-Mannequin pair falling in the red area indicating a bad match, low utilization, and high waste.

7.6.6 Sepsis Scenario Example for PVPbed vs Mannequin vs Computer

Cues	Pallor	Cyanosis	Eyes	Mottling	Temperature	Pulse	Heart Sounds	Breathing	Capillary Refill	Voice
Importance	1	1	1	1	1	1	1	1	1	1
PVP Bed Fidelity	1	1	1	1	1	1	1	1	1	1
Mannequin Fidelity	0	0	0	0	0	1	1	1	0	1
Computer-Based	1	1	1	1	0	0	1	1	1	1

Red = Bad Match
Green = Good Match

Figure 7.12: Table showing simplified example of a sepsis scenario with important cues and simulator fidelity for each of these cues. These values are for illustration only and they are not externally validated.

7.6.6.1 Score for Sepsis-Mannequin pair

Scenario $\vec{I}$[1,1,1,1,1,1,1,1,1,1]

Simulator $\vec{F}$[0,0,0,0,0,1,1,1,0,1]

$S = 4/10 = 0.4$

$U = 1/1 = 1$

$W = 0$

$G = 0.76$

$B = 2.5$

Bad match, Fully Utilized, No Waste

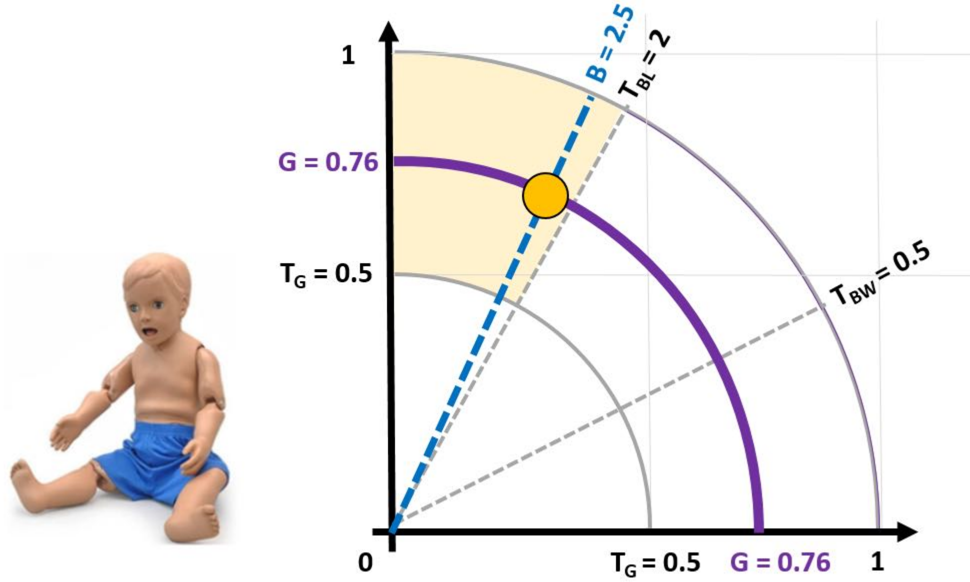


Figure 7.13: Example plot for Sepsis-Mannequin pair

7.6.6.2 Score for Sepsis-Computer Based pair

Scenario $\vec{I}$[1,1,1,1,1,1,1,1,1,1]

Simulator $\vec{F}$[1,1,1,1,0,0,1,1,1,1]

$S = 8/10 = 0.8$

$U = 1/1 = 1$

$W = 0$

$G = 0.9$

$B = 1.25$

Good match, No Waste

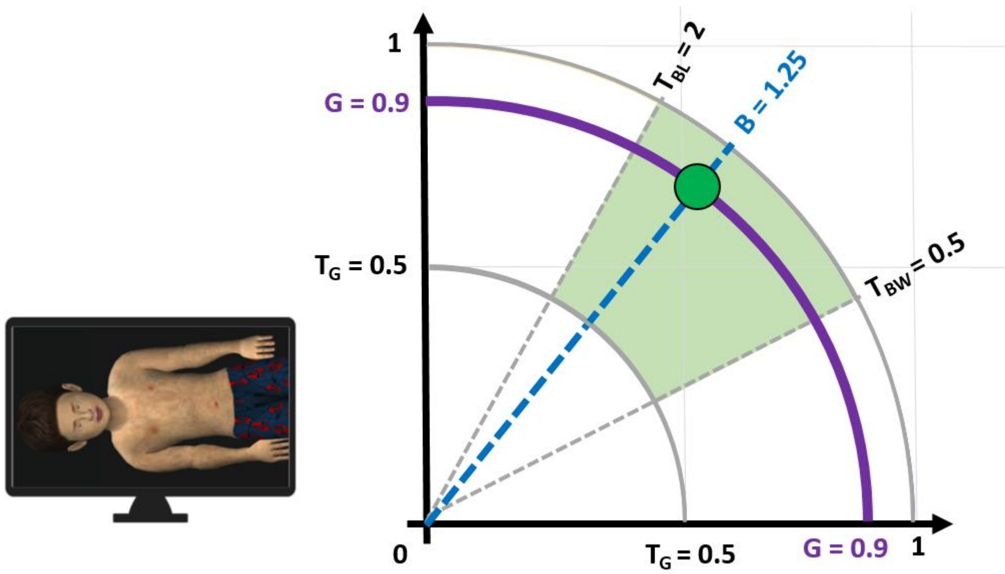


Figure 7.14: Example plot for Sepsis-Computer-based pair

7.6.6.3 *Score for Sepsis-PVPBed pair*

Scenario $\vec{I}$ [1,1,1,1,1,1,1,1,1,1]

Simulator $\vec{F}$ [1,1,1,1,1,1,1,1,1,1]

$S = 10/10 = 1$

$U = 1/1 = 1$

$W = 0$

$G = 1$

$B = 1$

Good match, No Waste

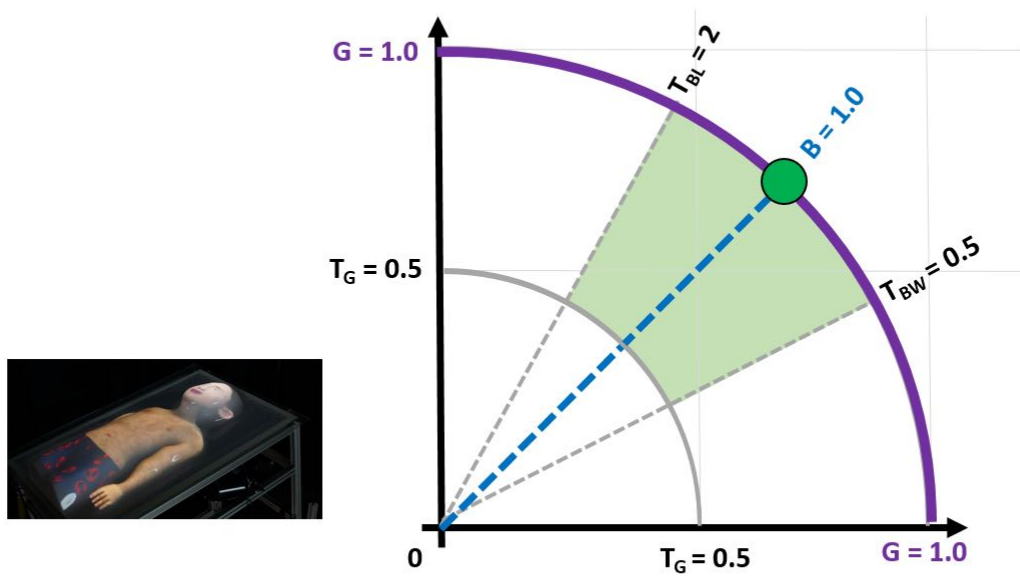


Figure 7.15: Example plot for Sepsis-PVP pair

7.6.7 *Child Abuse Scenario Example for PVPbed vs Mannequin vs Computer*

Cues	Story	Attitude	Facial Expressions	Skin Marks	Capillary Refill	Arm Fracture
Importance	1	1	1	1	1	1
PVP Bed Fidelity	1	1	1	1	1	1
Mannequin Fidelity	1	1	0	1	0	0
Computer-Based	1	1	1	1	1	1

Red = Bad Match
Green = Good Match

Figure 7.16: Table showing simplified example of a stroke scenario with important cues and simulator fidelity for each of these cues. These values are for illustration only and they are not externally validated.

7.6.7.1 *Score for Abuse-Mannequin pair*

Scenario $\vec{I}$[1,1,1,1,1,1]

Simulator $\vec{F}$[1,1,0,1,0,0]

$S = 3/6 = 0.5$

$U = 1/1 = 1$

$W = 0$

$G = 0.79$

$B = 2$

Bad match, Fully Utilized, No Waste

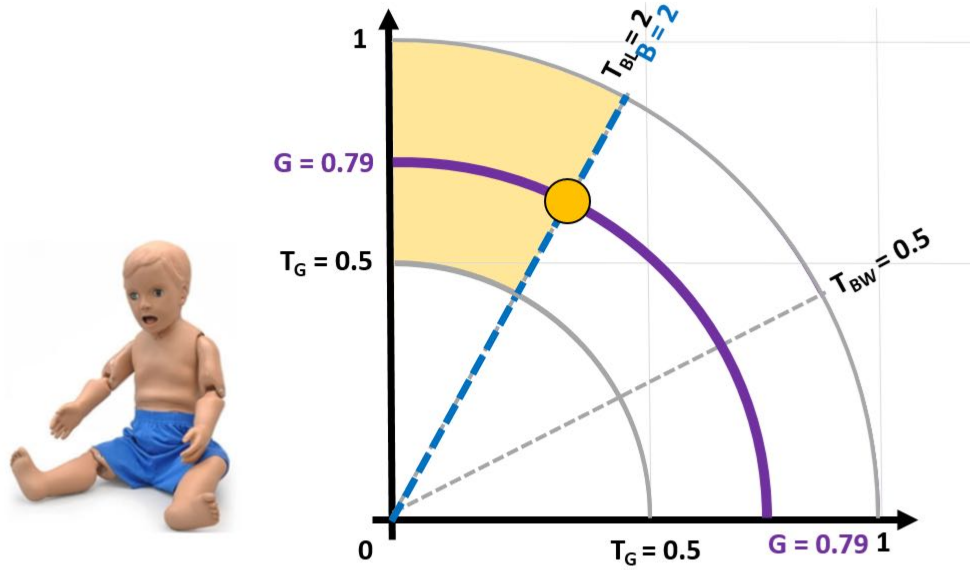


Figure 7.17: Example plot for Abuse-Mannequin pair

7.6.7.2 Score for Abuse-Computer Based pair

Scenario $\vec{I}$[1,1,1,1,1,1]

Simulator $\vec{F}$[1,1,1,1,1,1]

$S = 6/6 = 1$

$U = 6/6 = 1$

$W = 0$

$G = 1$

$B = 1$

Good match, No Waste

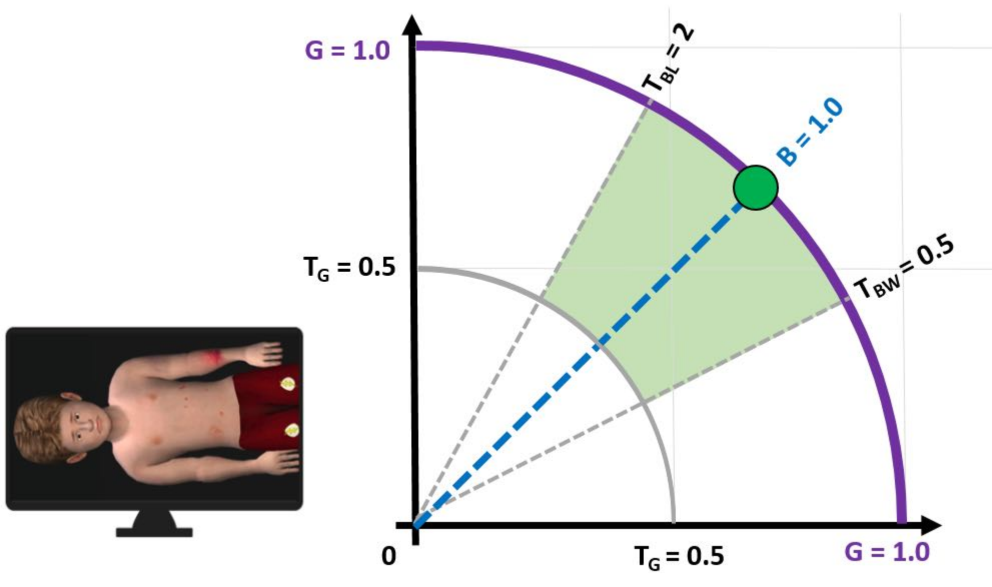


Figure 7.18: Example plot for Abuse-Computer-Based pair

7.6.7.3 Score for Abuse-PVPBed pair

Scenario $\vec{I}[1,1,1,1,1,1]$

Simulator $\vec{F}[1,1,1,1,1,1]$

$S = 6/6 = 1$

$U = 6/6 = 1$

$W = 0$

$G = 1$

$B = 1$

Good match, No Waste

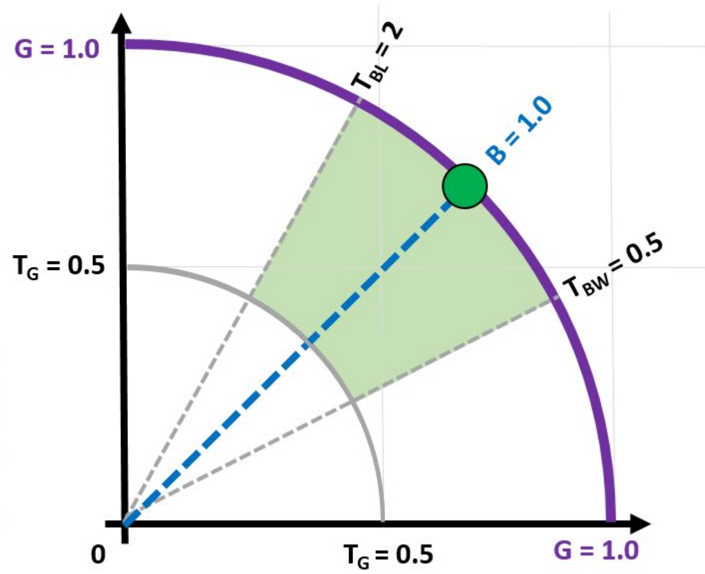


Figure 7.19: Example plot for Abuse-PVP pair

7.7 Discussion

The work in this chapter focuses on the evaluation in terms of Scenario Satisfaction and Simulator Utilization without looking closely at cost, and assuming that a higher fidelity would be associated with higher cost. To evaluate cost, one has to take into consideration many factors including but not limited to the cost of development, cost of equipment, cost of space, and cost of labor. Among the different types of costs there are initial costs, incremental costs, recurring costs, and maintenance costs. All of these should be taken into consideration when considering the cost of a new simulator. This would be explored in future work.

In addition, the learner's capacity to handle higher fidelity simulation is assumed to be determined by the educator while determining the importance of certain cues for a scenario. The same scenario can have different variations depending on the learning objective for a specific audience (e.g., novice vs advanced). If the learners are novice, a high fidelity representation of a certain cue could be overwhelming therefore a variation of the same scenario where that cue has lower importance could be more appropriate. Similarly, if the learner is more advanced then a variation of that same scenario where that same cue is rated as more important therefore a higher fidelity for that cue would be a better match and could be more beneficial for that learner.

CHAPTER 8: EFFECTS OF SOCIAL PRIMING ON SOCIAL PRESENCE WITH INTELLIGENT VIRTUAL AGENTS (SOCIAL PRIMING STUDY)

The work in this chapter has been presented at Intelligent Virtual Agent 2017 and published [19].

Title: Effects of Social Priming on Social Presence with Intelligent Virtual Agents

Authors: Salam Daher, Kangsoo Kim, Myungho Lee, Ryan Schubert, Gerd Bruder, Jeremy Bailenson, Gregory Welch.

Published at: Intelligent Virtual Agents, 2017

This work supports **[TS 4]**:

Social Presence Priming Can be Used to Improve a Participant's Perception of a Simulated Human.

I lead the design, development, execution, and analysis of the work described below. I acknowledge the effort of co-authors and collaborators.

8.1 Abstract

This chapter explores whether witnessing an Intelligent Virtual Agent (IVA) in what appears to be a socially engaging discussion with a Confederate Virtual Agent (CVA) prior to a direct interaction, can prime a person to feel and behave more socially engaged with the IVA in a subsequent interaction. To explore this social priming phenomenon, we conducted an experiment in which participants in a control group had no priming while those in an experimental group were briefly exposed to an engaging social interaction between an IVA and a nearby CVA (i.e. a virtual actor). The participants primed by exposure to the brief CVA-IVA interaction reported being significantly more excited and alert, perceiving the IVA as more responsive, and showed significantly higher measures of Co-Presence, Attentional Allocation, and Message Understanding dimensions of so-

cial presence for the IVA, compared to those who were not primed.

8.2 Introduction

An intelligent virtual agent (IVA) can provide a flexible and versatile means to communicate verbal and spatial information with real humans. IVAs can be especially valuable when the presence of an actual human is not safe or feasible, such as in medical emergencies or military training. An IVA can be embedded not only in immersive virtual environments but also in the real world via augmented reality technologies to share the physical space with real humans [167]. For IVAs, it is desirable to facilitate a high sense of presence, co-presence, and social presence in order to elicit behavior in real humans that matches what can be observed between humans in the real world [248]. Lombard and Ditton define *presence* as the sense of non-mediation, which means that one can perceive presence via a technological medium if one can be oblivious to the existence of the medium [249]. There are many interpretations of the terms *social presence* and *co-presence*, e.g., see [171]. Goffman et al. indicate that *co-presence* exists when people sensed that they were able to perceive others and that others were able to actively perceive them [250]. Blascovich et al. define *social presence* both as a "psychological state in which the individual perceives himself or herself as existing within an *interpersonal* environment" (emphasis added) and "the degree to which one believes that he or she is in the presence of, and interacting with, other veritable human beings." [251, 252]. Harms and Biocca illustrated co-presence as one of several dimensions that make up social presence, and they evaluated the validity of their social presence measures with questionnaires [54]. While there is no universal agreement on the definitions of these terms, for the purpose of this chapter we consider *social presence* to be one's sense of being socially connected with the other, and *co-presence* to be one's sense of the other person's presence.

Most previous research on interaction with IVAs focused on the perceived behavioral realism *while* directly interacting with the IVA. However, we believe that the observed behaviors *prior to* such direct interaction will have an important influence on the initial and perhaps lasting impression of the IVA. For example, there is evidence from psychology that perceptions of intelligence and disposition can be influenced by observations of a person's behavior prior to an interaction and an individual's apparent mood can be "contagious"—transferred to another person via implicit nonverbal behaviors [138, 140, 141].

In this chapter, we explore the question of whether social presence can also be contagious. We use the word "Confederate" to indicate that the person is intentionally part of the experiment even though the participants may *not* think of that person as part of the experiment. Specifically, we used a *confederate virtual agent* or CVA. We present an experiment in which we test whether perceiving a socially engaging interpersonal discussion between an IVA and a CVA—i.e., exhibiting apparent social presence—can subsequently lead to the participant feeling increased excitement, alertness, and social presence with respect to the IVA.

This paper is structured as follows: Section 8.3 provides background information on IVAs, behavioral models, priming, and presence. Section 8.4 describes our experiment in which we analyze effects of an initial interaction between an IVA and a CVA on the subsequent perception of social presence with the IVA. Section 8.5 presents the results, which are discussed in Section 8.6. Section 8.7 concisely summarizes our experiment, the results, and presents our conclusions.

8.3 Background

While IVAs can be used as a replacement for real humans in certain situations, people usually do not treat an IVA exactly as they would treat a real human. For instance, in studies where medical students interacted with either an IVA or a real human pretending to have the same symptoms, participants appeared less engaged, sincere, and interested, and had a poorer attitude towards the IVA [71]. In an experiment with a computer graphics representation of an IVA, its advice was more rarely sought out compared to a physically present robot [72]. Often people treat IVAs as mere pixels instead of replacements for humans, even when compared to robots that occupy a physical space. One explanation for this phenomenon is the low sense of presence, social presence, and co-presence induced by the IVA. In this chapter we aim to increase the sense of social presence by exposing participants to a "social priming" *before* the interaction with the IVA.

Bailenson et al. studied participants' sense of co-presence in a multi-user shared immersive virtual environment while manipulating the non-verbal behavior of their virtual self-representations. The participants reported a higher sense of co-presence in a condition with head movement compared to the other conditions [168]. Garau et al. evaluated participants' responses, including presence, co-presence, and physiological signals, with respect to an IVA's degree of responsiveness. Their results did not show a significant relationship between perceived co-presence and the IVA's degree of responsiveness. However, they did suggest a link between higher levels of co-presence and participants who reported using computers less [170]. We took these findings into consideration while designing the study and analyzing the data.

While there are multiple possibilities for how the sense of social presence or co-presence of an IVA can be improved through modifications to its behavior *during* an interaction [170, 173, 174], the motivation for our work comes from exploring what could be done with the IVA *prior to* such a direct interaction.

Mood and even racial biases can be "contagious", i.e., transferred to other humans via implicit nonverbal behaviors [138, 139]. We wondered if exposure to a social presence priming could also be contagious. In general, *priming* can be seen as the incidental activation of a person's knowledge structure, which can lead the person to specific behaviors and attitudes [141]. It can affect social judgment [142], as well as goal-driven tasks, as Bargh et al. demonstrated by showing that primed participants performed comparatively better in an intellectual task [141]. Dijksterhuis and Bargh indicate that perception itself can prime or activate a behavioral tendency. Apart from perceiving observables of what is literally present, people make trait inferences and activate social stereotypes as forms of social perception that elicit the tendency to imitate in the social perceiver [143]. Qu identified three main elements in a conversation between a real human and an IVA: the surrounding environment, the virtual conversation partner, and the virtual bystanders [144]. Qu used images and videos to prime participants. Primed participants mentioned more keywords related to the priming content. Qu showed that priming with surrounding media content had a guidance effect in both the real world and the virtual world [145]. Similarly, we explore exposing our participants to a social priming and compare the effects on their social presence. Various studies have examined the concept of priming, some related to virtual reality [146, 147], but most of them explore the theory underlying the priming phenomenon. Researchers explored racial biases, gender, and IVA personality in virtual environments [148–150]. To our knowledge, there are no studies that use priming in the context of supporting social presence of an IVA.

8.4 Experiment

In this section we present an experiment in which we analyzed the effects of introducing social priming behavior between an IVA and a CVA.

8.4.1 Material

We built a room-sized experimental setup (approx. 3m $\times$ 3.6m) where a virtual character was presented as sitting behind a shared physical-virtual desk between two bookshelves (see Fig. **??**).

We modeled and animated a 3D virtual character, named "Katie," in Autodesk Maya. Katie was designed with animations for facial expressions, speaking, and body gestures. She had a mostly neutral, serious, and polite demeanor during the interaction (i.e., designed to not be too warm or cold towards the participant). We then imported the model into Unity3D where we added a graphical user interface allowing an operator to trigger specific body gestures or pre-recorded phrases with corresponding speaking animations, in order to play a game of twenty questions and to carry out other limited responses as needed before or after the game. With this human-in-the-loop experimental setup, the operator pressed buttons behind the scenes to trigger Katie's responses. Katie's image was rear projected onto the screen behind the physical desk using an Optoma TW610ST projector. The participants were recorded using 5 Logitech c920 webcams (2 close ups and 3 wide angles) observing the space from different positions.The CVA, which we call "Michael," was presented on a Panasonic TC-P65VT30 screen.

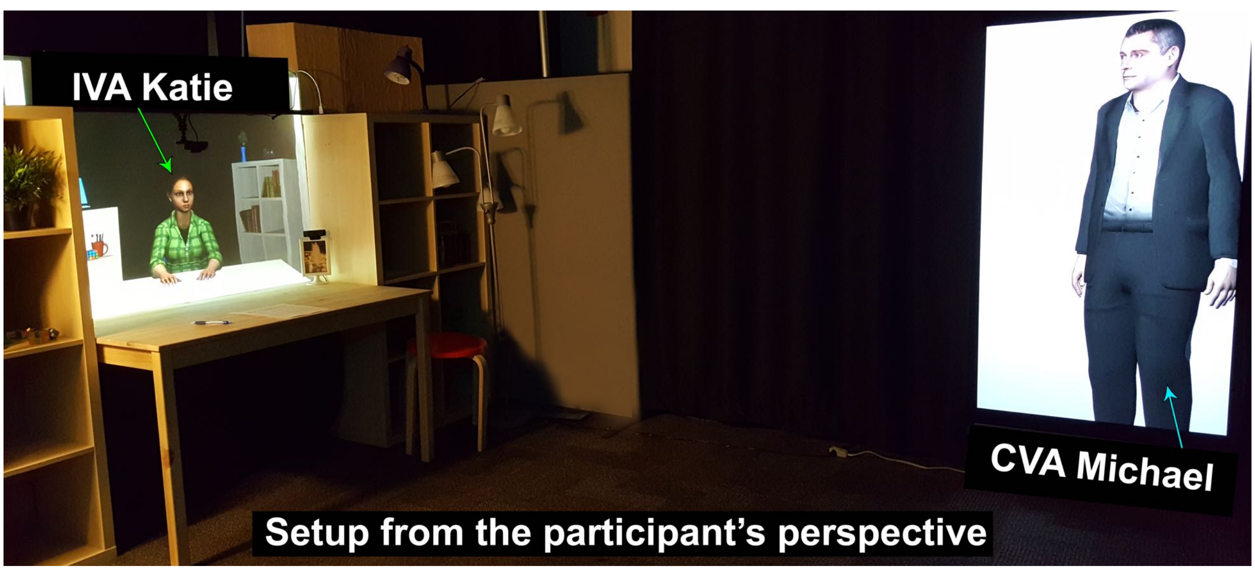


Figure 8.1: Experimental setup: Participants were exposed to a brief conversation between the IVA (Katie) on the left and the CVA (Michael) on the right. The virtual elements in the scene were rendered from the participant's viewpoint.

8.4.2 Methods

We used a between-participants design for this experiment. To investigate the effect of social priming on social presence with an IVA we defined two groups: (i) *control group* and (ii) *social priming group*. Participants in both groups were asked to play a game of *twenty questions* with the IVA (Katie). Participants in the social priming group perceived a short interaction between Katie and the CVA Michael before they played the game. Participants in the control group were not primed with this social interaction before playing the game. Figure 8.2 illustrates the overall procedure which is comprised of three steps: (a) preparation, (b) priming (only for the social priming group), and (c) interaction with a twenty questions game.

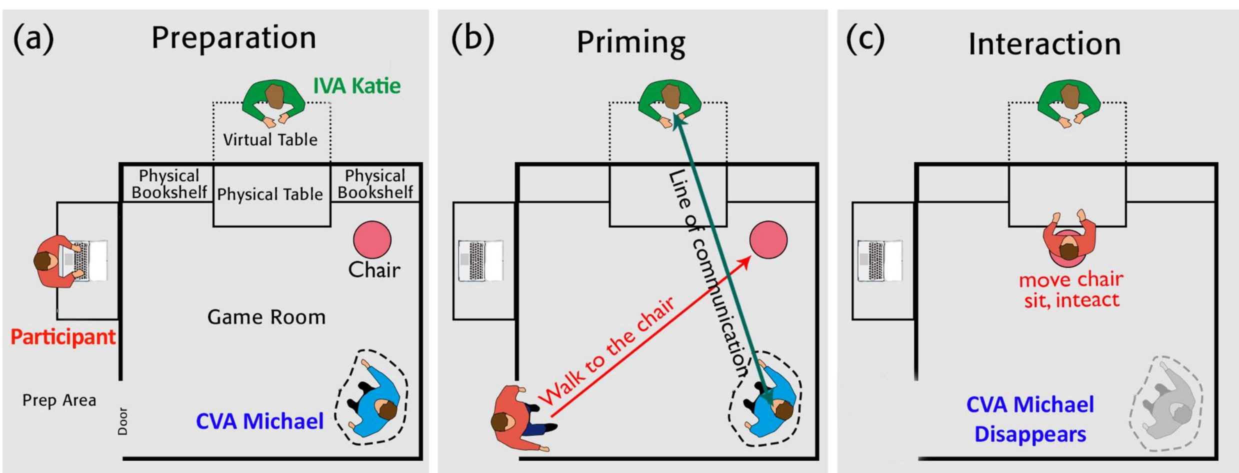


Figure 8.2: Setup and procedure in the two experimental conditions. The dotted line around the CVA (Michael) indicates that he was there only in the social priming condition but not in the control condition. The other elements were the same in both conditions.

In (a), before the participants entered the room, all participants read the informed consent and completed a demographics questionnaire. Then, the experimenter briefed them on the rules of the twenty questions game:

> *"You are going to play two games of twenty questions. During the first game, Katie has an object in mind that you have to guess; you can ask her questions that have YES or NO answers. During the game you cannot ask her open ended questions. You have a maximum of 20 questions to guess the object. In the second game, the same rules apply, except that Katie will ask you the YES or NO questions and you have to answer."*

Participants were asked to pick their object for the second game from a deck of cards before starting the interaction. They were asked to write down the answers to questions and to record the winner of each game during the interaction with the IVA. While individual interaction times might vary between participants due to the nature of the game (i.e., some people took more time to think before asking questions), the exposure times to the IVA were comparable across both groups.

In (b), the participants in the *social priming condition* saw Katie seated at the table and the CVA (Michael) standing in the corner of the room. They were then exposed to an interaction between Katie and Michael. Michael acted as if he had just finished a game with Katie. As soon as the participant entered, Michael looked at the participant, then at Katie, and said *"Oh, you've got visitors. I'll leave you two to play,"* then Katie and Michael exchanged phrases such as *"It was nice playing with you! Thanks for your time. See you later!"* This short exchange constituted the social presence priming. Participants in the *control condition* were not exposed to this social interaction.

In (c), the participants then entered the room and interacted directly with Katie. Katie was seated at the table, and she initiated the conversation with phrases like *"Hello, how are you? Nice to meet you!"* and then moved on to playing the game. She ended the interaction with *"It was nice playing with you! Thanks for your time. See you later, bye!"*

During the experiment, participants were video recorded from multiple angles and observed for verbal and non-verbal behavior. Specifically, we observed whether the participants waited for the IVA and CVA to finish in (b), or if instead they walked straight to the chair, breaking the line of communication between Katie and Michael.

After completing the experiment, participants were asked to fill out a set of post-questionnaires including a social presence questionnaire [54], a presence questionnaire [253], an affective attraction questionnaire [187], and their subjective emotions state using Wilhelm's Mood Rating Questionnaire which consists of six questions (How did your interaction with the other player (Katie) make you feel: anxious, excited, tense, alert, in control, desire to leave the situation) [188]. Since we used a mixed reality setup rather than an immersive virtual environment (cf. [253]), we used a subset of the original questions, removing those inappropriate for our environment.

The exact questions for social presence dimensions are shown in Table **??** with *Co-Presence* being the degree the participant thinks he/she is not alone, *Attentional Allocation* is the amount of attention the participant allocates to and receives from the IVA, *Perceived Message Understanding* is the ability of the participant to receive a message from the IVA and for the IVA to understand their message, *Perceived affective understanding* is the ability of the participant to understand the IVA's emotional and attitudinal states and for the IVA to understand the participant's emotional and attitudinal states, *Perceived Affective Interdependence* is the extent to which the participant's emotional and attitudinal state affects and is affected by the emotional and attitudinal states of the IVA , *Perceived Behavioral Interdependence* is the extent to which a participant's behavior affects and is affected by the IVA's behavior [54]. The means for each dimension are computed by adding the scores for these questions and dividing by the total number of questions ($N = 6$). The social presence questions are on a 1 to 7 Likert scale. Questions marked with a star were inverted by negating the answer and adding 8. The individual questions from the *presence* questionnaires that showed significantly different answers are shown in Table 8.1.

Table 8.1: Measurement for select Presence questions regarding Responsiveness, and Involvement that showed significant differences.

Responsive	How responsive were the other player (Katie) and her environment to actions that you initiated (or performed)?
Involved	How much did the visual aspects of the other player (Katie) and her environment involve you?

8.4.3 Participants

58 participants (35 males and 23 females from multiple ethnicities) were randomly assigned to the control ($n = 29$) or social priming ($n = 29$) experimental group. Participants were recruited from our university community (students, employees, and alumni from various colleges within the university) via web postings and email lists. Participants' experience with IVAs varied. Thirteen participants had never interacted with an IVA before, while the others reported varying familiarity with the concept of Virtual Humans (VH), from having encountered some sort of VH at some point, e.g. while playing video games, to four indicating being involved in some type of VH development at some point in their lives. None of the participants had prior experience with the IVAs used in this experiment.

8.5 Results

8.5.1 Qualitative Results

In the social priming condition, most participants commented that they did not pay much attention to the CVA Michael. Fourteen participants acknowledged that they ignored Michael, with comments such as "I didn't pay him much attention," or "I completely ignored him." Five participants did not expect to see the CVA Michael and expressed being surprised or confused. Three participants felt that they "interrupted Katie and Michael's conversation." Six participants expressed positive reactions regarding Michael such as he was "friendly and heartwarming" and that he set

Table 8.2: Questionnaires for Social Presence: CoPresence (CoP), Attentional Allocation (Atn), Perceived Message Understanding (MsgU), Perceived Affective Understanding (Aff), Perceived Emotional Interdependence (Emo), Perceived Behavioral Interdependence (Behv).

CoP-Q1	I noticed the other player (Katie).
CoP-Q2	The other player (Katie) noticed me.
CoP-Q3	The other player (Katie)'s presence was obvious to me.
CoP-Q4	My presence was obvious to the other player (Katie).
CoP-Q5	The other player (Katie) caught my attention.
CoP-Q6	I caught the other player (Katie)'s attention.
Atn-Q1*	I was easily distracted from the other player (Katie) when other things were going on.
Atn-Q2*	The other player (Katie) was easily distracted from me when other things were going on.
Atn-Q3	I remained focused on the other player (Katie) throughout our interaction.
Atn-Q4	The other player (Katie) remained focused on me throughout our interaction.
Atn-Q5*	The other player (Katie) did *not* receive my full attention.
Atn-Q6*	I did *not* receive the other player (Katie)'s full attention.
MsgU-Q1	My thoughts were clear to the other player (Katie).
MsgU-Q2	The other player (Katie)'s thoughts were clear to me.
MsgU-Q3	It was easy to understand the other player (Katie).
MsgU-Q4	The other player (Katie) found it easy to understand me.
MsgU-Q5*	Understanding the other player (Katie) was difficult.
MsgU-Q6*	The other player (Katie) had difficulty understanding me.
Aff-Q1	I could tell how the other player (Katie) felt.
Aff-Q2	The other player (Katie) could tell how I felt.
Aff-Q3	The other player (Katie)'s emotions were not clear to me.
Aff-Q4	My emotions were not clear to the other player (Katie).
Aff-Q5	I could describe the other player (Katie)'s feelings accurately.
Aff-Q6	The other player (Katie) could describe my feelings accurately.
Emo-Q1	I was sometimes influenced by the other player (Katie)'s moods.
Emo-Q2	The other player (Katie) was sometimes influenced by my moods.
Emo-Q3	The other player (Katie)'s feelings influenced the mood of our interaction.
Emo-Q4	My feelings influenced the mood of our interaction.
Emo-Q5	The other player (Katie)'s attitudes influenced how I felt.
Emo-Q6	My attitudes influenced how the other player (Katie) felt.
Behv-Q1	My behavior was often in direct response to the other player (Katie)'s behavior.
Behv-Q2	The behavior of the other player (Katie) was often in direct response to my behavior.
Behv-Q3	I reciprocated the other player (Katie)'s actions.
Behv-Q4	The other player (Katie) reciprocated my actions.
Behv-Q5	The other player (Katie)'s behavior was closely tied to my behavior.
Bhv-Q6	My behavior was closely tied to the other player (Katie)'s behavior.

the tone as “more realistic” and made participants “more excited,” or put them “in a good mood.”

Ten participants gave positive comments regarding the IVA Katie’s friendliness and expressiveness such as saying she “was expressive” and that she gave off a “friendly vibe,” and six participants gave comments suggesting improving the IVA Katie’s emotions and expressiveness.

In the control condition, there were more mixed comments regarding the IVA Katie. Nine participants gave positive comments related to Katie’s realism and character such as saying she was “very realistic and friendly” while 11 participants gave comments suggesting improvements for the IVA Katie’s character, emotions and expressions such as “[she] could have been nicer and more friendly” and that she does not show much “emotion” or “her face doesn’t show feeling” or they felt “a little distant” from her.

8.5.2 Quantitative Results

We decided to use non-parametric statistical tests to analyze the Likert scale ordinal data from the questionnaires [254] comparing the priming condition with the control condition. While it is common practice in some scientific disciplines to treat Likert-type scales as interval-level measurements [224], we avoid the discussion on whether parametric statistics can be a valid method for the analysis of non-parametric data [225, 226] by using non-parametric (Mann–Whitney U) tests.

Social Presence: We aggregated scores for the questions per category to cover all dimensions of social presence. We found a significantly higher social presence in the social priming condition for co-presence ($U = 256.0$, $p = 0.009$, $r = -0.664$), attentional allocation ($U = 288.0$, $p = 0.039$, $r = -0.662$), and perceived message understanding ($U = 276.0$, $p = 0.024$, $r = -0.527$). There was no significant difference for perceived affective interdependence ($U = 386.0$, $p = 0.596$, $r = -0.235$), perceived emotion interdependence ($U = 431.0$, $p = 0.876$, $r = -0.016$), and perceived behavioral interdependence ($U = 365$, $p = 0.391$, $r = -0.246$). Figure 8.3 shows the results for the dependent variables in the experiment.

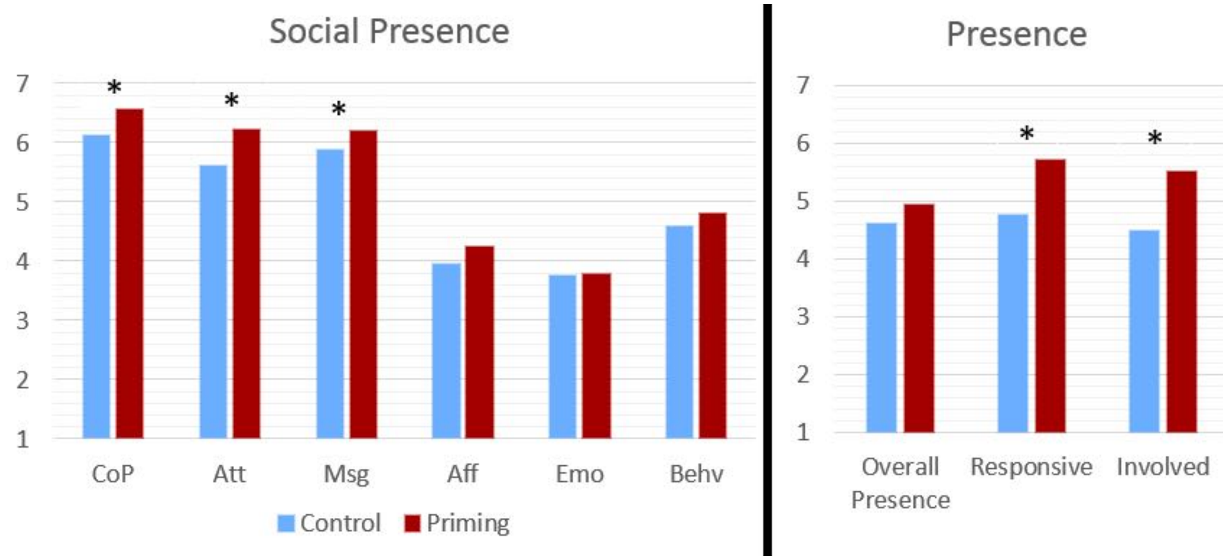


Figure 8.3: Results for social presence dimensions (left) and for Presence (right). Star "*" indicates significant results

Mood Rating: Participants were asked to rate their interaction with the IVA Katie for the following Mood Ratings: anxious, excited, tense, alert, in control, and desire to leave the situation. The results show that participants in the social priming condition were more excited ($U = 294.5$, $p = 0.047$, $r = -0.572$) and more alert ($U = 267.5$, $p = 0.017$, $r = -0.651$) compared to those in the control condition. There was no significant difference for feelings of being anxious ($U = 385.5$, $p = 0.586$, $r = -0.182$), tense ($U = 371.5$, $p = 0.442$, $r = -0.134$), in control ($U = 415.0$, $p = 0.944$, $r = 0.000$), or desire to leave the situation ($U = 490.5$, $p = 0.208$, $r = 0.257$).

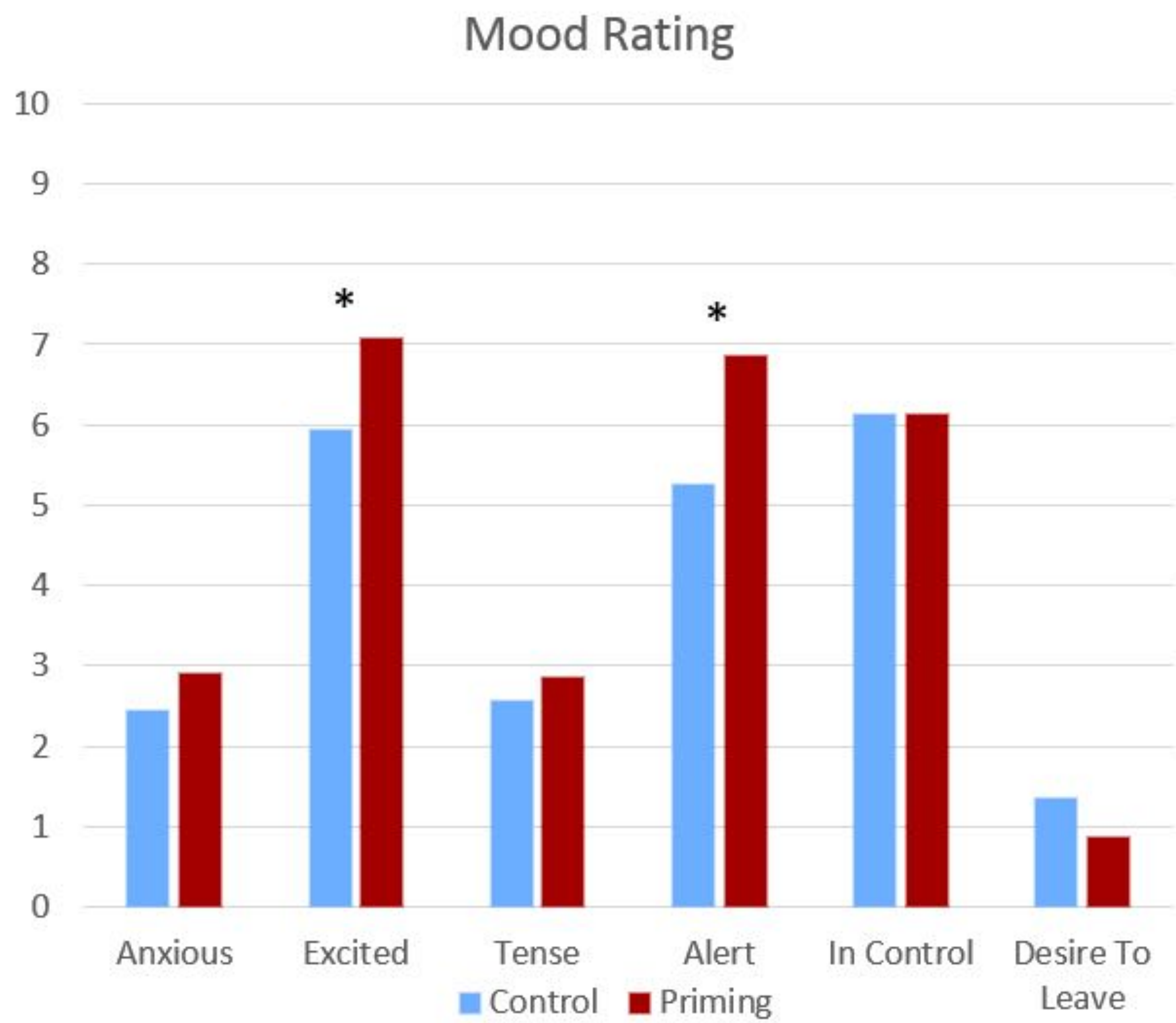


Figure 8.4: Results for Mood Ratings. Star "*" indicates significant results.

Presence: When comparing individual Presence questions, the results show that participants in the social priming condition perceived the IVA Katie as more responsive to their actions ($U = 278.5$, $p = 0.024$, $r = -0.641$) and that they were more involved in the visual aspects of the IVA ($U = 254.5$, $p = 0.008$, $r = 0.772$). There was no significant difference for the presence (all $p > 0.05$), affect attraction (all $p > 0.05$), and virtual environment (all $p > 0.05$) questions.

8.6 Discussion

Overall, our results indicate that subtle social priming, by means of perceiving a short interaction between an IVA and a CVA, can result in a significant increase in social presence (specifically the co-presence, attentional allocation, and perceived message understanding dimensions of social presence), an increase in the sense of feeling more excited and more alert, an increase in the sense of feeling more involved in the visual aspects, and a perception of the IVA as more responsive compared to those who were not primed supporting **[TS 4]**.

The results are particularly interesting as the social priming in our experiment was a rather short pre-scripted/pre-programmed conversation between our IVA and CVA. Indeed, because the IVA-CVA interaction occurs prior to, and independent of, the participant's interaction with the IVA, the IVA-CVA interaction can be pre-programmed and yet still look spontaneous. One could create a library of such short IVA-CVA vignettes to be invoked periodically when it appears that the circumstances are currently unlikely to yield direct human-IVA interaction. Such vignettes could then also be transferred to other characters and setups, without adding strict requirements in terms of space or interactivity.

Interpretation of the results within the social presence dimensions: We expected *Co-Presence* to be significantly different since Co-Presence is the degree the participant thinks he/she is not alone, and the fact of adding a CVA reinforces this idea. We also expected that the *Attentional Allocation* would be higher as the CVA is expected to draw attention to the unexpected priming conversation. Given that the priming was designed to be the end of the previous player's game, it is possible that the increase in the *Perceived Message Understanding* came from the appearance that the IVA knows what she is doing since she played that game before and it was a pleasant experience for the CVA. We did *not* expect to see a significant change in the dimensions that are related to affect (*Perceived affective understanding*, and *Perceived Affective Interdependence*) as both the IVA and the CVA were designed to have a neutral, polite, and professional character (not too warm, not too cold). The means of the Control and Priming groups were almost identical. We did not design the priming in a way to intentionally affect the behavior (i.e., the CVA did not give a hint about the game, nor did he ask the participant to perform any specific action). We were not surprised that the *Perceived Behavioral Interdependence* was not significantly different. We did notice that *during* the priming, a few participants chose to wait while others chose to keep on walking but that change in behavior did *not* affect the results.

Interpretation of the Mood Ratings results: An argument can be made in either direction whether it is good or bad to be more Alert and more Tense. According to The Positive and Negative Affect Schedule (PANAS), "Excited" and "Alert" are classified as positive affects [184]. According to the Cognitive-Affective Theory of Learning with Media, the multimedia learning process is mediated by the learner's mood, and positive mood has a facilitating effect on multimedia learning [182]. We conclude that using social priming can have a positive effect on the mood and possibly on learning using multimedia such as an IVA.

Are there other ways to prime social presence? Other stimuli might also be powerful social presence priming tools. For example, the IVA could exhibit "human-like" traits or characteristics, such as engaging in humor, referencing a recent real world event, or reacting to stimuli in the environment, showing awareness of the person and their surroundings. Likewise, it may be possible to strengthen (or weaken) the priming approach used in this study (e.g., making the perceived conversation appear more exciting). The CVA could engage in an unpleasant conversation such as scolding the IVA which could result in a positive priming (i.e., sympathize with the IVA) or negative priming (take sides with the CVA). Similarly, it may be possible to negatively prime participants if the confederate intentionally ignores the IVA. Future research directions may include experimenting with variations of other aspects of the IVA and CVA, such as attire, gender, or ethnicity as well as replacing the CVA with a real human confederate.

What are the long-term effects? We do not know how long the effects of social priming last or the effectiveness when the same priming situation is re-used multiple times. It may be the case that the effects stay until something else changes them or it could be the case that the effects fade over time and a reminder or "booster" may be periodically required, for which one could pseudo-randomly select one of multiple canned priming sequences. Social priming in this way could be particularly useful to many applications employing IVAs because of its relatively low cost and independence with the actual direct interaction.

8.7 Conclusion

In this chapter we presented a novel method to increase social presence with IVAs. It is generally believed that a higher sense of presence with an IVA has the potential to make applications such as training more effective, which can translate into increased performance in teams in a real environment [248]. Observing the behaviors of another human *prior to* interaction with that human can influence perceptions of that person's intelligence and disposition. To see if this effect could be used to increase the perceived realism of an IVA, we explored what we call *social presence priming*, where we exposed participants to an IVA participating in a seemingly spontaneous but actually pre-programmed socially engaging interaction with a confederate virtual agent *before* the participants interacted directly with the IVA. In the condition where the participants were socially primed, the co-presence, attentional allocation, and perceived message understanding dimensions of social presence were found to be significantly higher compared to the control condition. Participants also felt more excited and alert and perceived the IVA as more responsive. The results of this study are encouraging for the use of relatively low-cost social priming in existing and future IVA applications, since the proposed confederate virtual agents only need a limited functionality to complete a short canned interaction with the IVA.

CHAPTER 9: CONCLUSION

This chapter summarizes the work presented in this dissertation. I presented a new type of physical-virtual simulator that allows healthcare providers to interact with a patient that has physical form, can dynamically change appearance (e.g., verbal and non-verbal communication, capillary refill), and can present audio and tactile physiology (e.g., breathing, pulse, localized skin temperature.)

The Physical-Virtual Patients (PVP Head and PVP Bed) can represent a combination of symptoms corresponding to a variety of conditions. I used the PVPs with three different scenarios (stroke, sepsis, and child abuse) in different experiments, displaying subtle combinations of signs/symptoms.

The natural interface allows a direct interaction between the participants and the patient, minimizing the need for the controller of the simulation to intervene or provide additional information, which can interfere with any evolving engagement with the simulated patient and potentially participant learning or assessment.

I presented experiments where the PVPs were used to explore variations in the representation of dynamic visuals, and physical shape, and the match between these representations during a patient simulation.

In summary, adding realistic dynamic visuals on a human physical shape *near* a mannequin improves the perception of realism, social presence, communication, engagement, sense of urgency, and learning supporting **[TS 1A]**. When the visuals are separated in a manner where the dynamic visuals are displayed on a TV screen near the rest of the visually static simulator containing haptic cues (temperature and pulse), what participants perceived as the patient depends on the physical form of the simulator. When the simulator was flat, the locus of the patient follows the location of the dynamic visuals (i.e., on the TV screen). When the simulator had a physical human shape, the physicality competes over the dynamic visuals showing equal importance for what represents the patient. It is preferred to combine both the dynamic visuals and the physical human form are co-located in one space. Participants perceived the co-located patient as more visually realistic than the separated patient **[TS 1B]**. Ideally the dynamic visuals match human physical form, especially that people are sensitive when a simulator becomes close to looking, feeling, and behaving like a human which increases the expectations. While adding dynamic visuals to a physical human form is important, it comes with a cautionary warning that the content development becomes more challenging to keep up with the humans' expectations.

Real humans are physical beings that have an inherent physical shape. Healthcare simulation recognizes the importance of that physicality by using mannequins in the majority of the simulations. Since mannequins lack the flexibility and richness that virtual simulators provide, often flat computer-based monitors are used instead. The flat representation of humans simplifies the complexity of the human, making the training easier than it would have been on a real patient. In a controlled experiment that kept the same visual content constant and varied the physical shape, we found that the cognitive load is higher and the behavior was more natural when the patient had a human physical shape compared to when the patient was flat supporting **[TS 2]**. This raises the question whether the current training using flat patients is good enough.

In tasks that require touching a human, a physical surface with geometry that matches the 3D model is preferred over a mismatching one **[TS 2]**. Also for touching tasks, people prefer projection-based spatial augmented reality (SAR) over optical see through head mounted displays (HMDs) augmented reality.

In this dissertation, I presented a general objective method that applies to any and every patient simulator to evaluate the compatibility of any scenario-simulator pair in a quantitative way. For every scenario I used a cue importance vector comprising a normalized importance to score each cue. For every simulator I used a cue fidelity vector comprising a normalized fidelity score for each cue. For every scenario-simulator combination I developed a simulation compatibility score formed by dot product of cue importance vector and cue fidelity vector to evaluate the match, utilization, and waste. This score can be used to compare simulator-scenario pairs in an objective way that can have educational, economical, and logistical impact **[TS 3]**.

Finally, regardless of what type of simulator you currently have, one can improve the experience of the participants by priming them before the simulation. In this context priming consists of exposing participants to a short engaging conversation with the simulated agent prior to the simulation. This brief exposure improves social presence and the perception of the simulated human, which could improve performance. This method could be applied to any simulation including healthcare simulation. **[TS 4]**

APPENDIX A: APPENDIX A: PATIENT SIMULATORS THROUGHOUT HISTORY

This chapter presents a brief historical overview of simulators and specifically medical simulators as well as a brief technology review. First the methodology for finding this information is described then a detailed chronological history followed by a brief summary of the same information viewed from a different view and linking back to the chronological section for more details. Hash tags with short names are given to each simulator in history between curly brackets for a quick link back, for example {#SimulatorShortName}.

Methodology

The methodology used to gather the data about simulators came from two different approaches depending on the type of the search. It started when I was searching to check if anyone is using a patient simulator with both physical and virtual features and then the search progressed to exploring simulators throughout history. I did an exhaustive search of going through all the titles of published articles in “Simulation in Healthcare” [255] from 2006 to 2014, and JDMS [256]from 2004 to 2014 journals to find any related works. Also in parallel I used keyword searched in the UCF library, Google scholar search, Google search for word combinations that contain the following: physical, virtual, physical-virtual, simulator, human patient, medical, nursing, simulation, head, stroke, culture, neurology, socio-cultural, socio-psychological, history, 3D character, 3d avatar, . . . etc. I did not find many resources describing patients similar the physical-virtual patient described in chapter 3. After that I searched for the other types of virtual patient simulators and for physical simulators through history starting with the keyword search, I found the article “Early use of simulation in medical education” of Prof Owen [28], and “history of medical simulation” by Rosen [257] and “the comprehensive textbook in healthcare simulation” by Levine [55] . I followed the original sources of the information in those articles and their branches storing relevant information in an XLS spreadsheet so I can reorganize and filter the data by date, country, and type of simulation, technology used. I also noted to a lesser detail history of flight simulation and how it overlaps with medical simulation (for this document, the focus is on the history of healthcare simulation). In addition, I knew about the more recent simulators after visiting IITSEC and IMSH and talking with exhibitors and interviewing subject matter experts at UCF from College of Nursing and from IST. After finding the information in text format, I looked for images of those simulators through history. Sometimes those images were directly in the resources, sometimes they were found through a Google search or search inside a mentioned museum. The images and text were put into a Google spreadsheet and put into a timeline using timeline.js javascript library [258]. The link with a summary of the information is [259] http://salamdaher.net/UCF/dissertation/healthcareSimulators/timeline.html

Chronologic History of Simulators

Throughout history, people have been trying to model and simulate other humans:

{#VenusWillendorf} Between 24000 BC and 22000 BC, the earliest reference for a human model is Venus of Willendorf, a carve figure model of a human showing exaggerated breasts and hips. It is believed that it could have been used as an artistic representation. Some were found in Europe and others in Siberia [260–262]. (See Fig A.1.)

Figure A.1: Venus of Willendorf
Image Source: User:MatthiasKabel [GFDL (http://www.gnu.org/copyleft/fdl.html), CC-BY-SA-3.0 (http://creativecommons.org/licenses/by-sa/3.0/) or CC BY 2.5 (https://creativecommons.org/licenses/by/2.5)], from Wikimedia Commons]

Wooden Anatomical Models

{#WoodenAnatomical} In 3400BC Egyptians created wooden anatomical models but those were not used for educational purposes, in fact they were used for display in private parties to remind people to enjoy themselves and drink before they die [263]. (See Figure A.2)

Figure A.2: Wooden Anatomical Models
Image Source: as they appear in Regis Olry, Wax, Wooden, Ivory, Cardboard, Bronze, Fabric, Plaster, Rubber and Plastic Anatomical Models: Praiseworthy Precursors of Plastinated Specimens , J Int Soc Plastination Vol 15, No 1: 30-35, 2000 [263]]

MayanHead

{#MayanHead} Between 300-600AD Mayans in Central America carved heads where half the head looks healthy and the other half is sick or possibly dead, they used clay as a "memento mori" to inspire people that death cannot be avoided and that life is not long [264]. (See Fig A.3.)

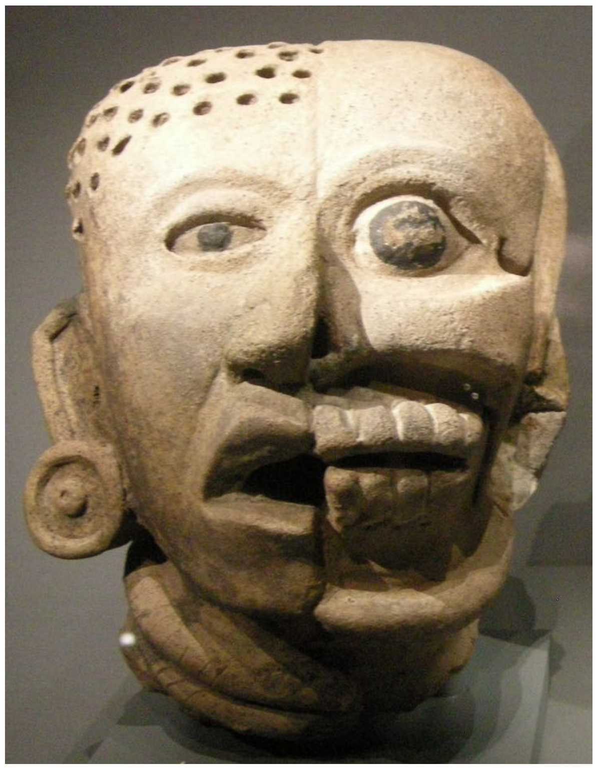


Figure A.3: Mayan Head
Image Source: I, Sailko [GFDL (http://www.gnu.org/copyleft/fdl.html), CC-BY-SA-3.0 (http://creativecommons.org/licenses/by-sa/3.0/) or CC BY 2.5 (https://creativecommons.org/licenses/by/2.5)]

{#BronzeStatues} Between 987-1067 in China, a life size basic patient simulator statue made of bronze was used by physician Wang Wei-Yi to teach acupuncture in a standardized way. The model has 354 open holes, is filled with liquid and covered with wax. Depending on which hole you poke you get a liquid feedback to indicate correct or incorrect practice [265]. (See Fig A.4.)

Figure A.4: Bronze Statues.
Image Source: Bronze acupuncture figure, in silk covered box bearing text Wellcome L0057614.jpg [CC BY 4.0 (https://creativecommons.org/licenses/by/4.0)], via Wikimedia Commons

{#EyeTrauma} In 1559 in France king Henry II was injured during a tournament when lance pierced his helmet. Experiments of thrusting sticks into eyes of corpses were made in an attempt to determine the exact injury to help find a cure. The experiments failed [28].(See Fig A.5.)

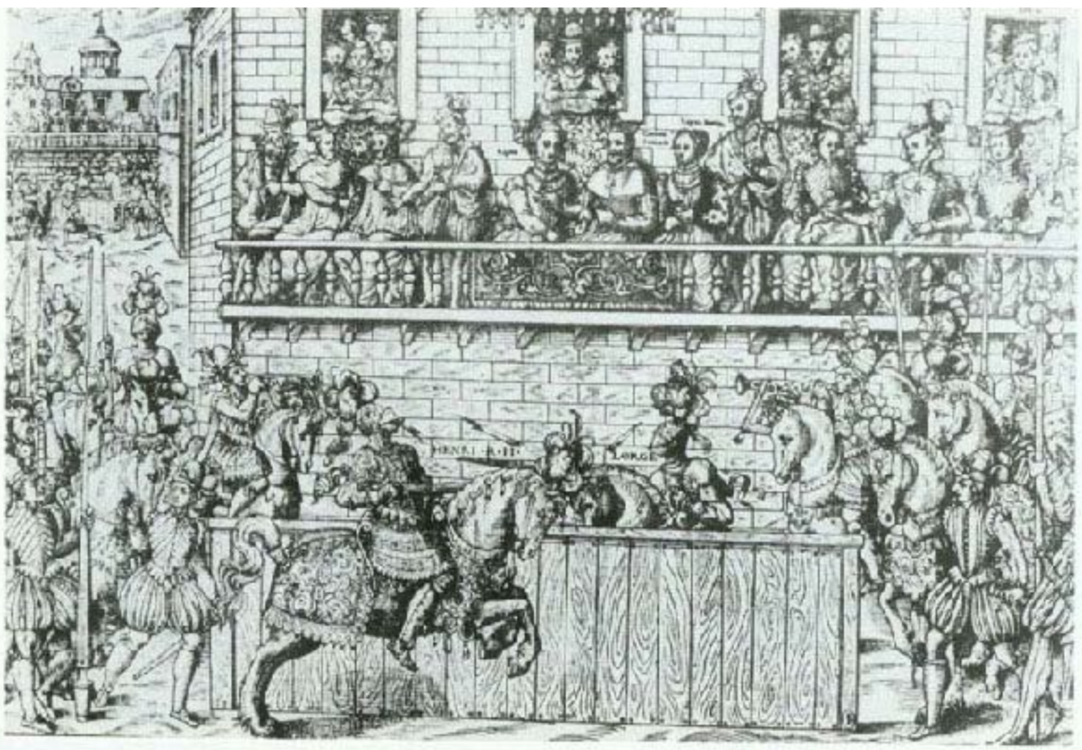


FIG. 2. Engraving (circa 1560) of the famous joust in which King Henry II was mortally wounded. A portion of the splintered wooden lance is seen projecting from his right orbit. (Reproduced by courtesy of the Fotomas Index, London.)

Figure A.5: Eye Trauma
Image Source: The Death of Henry II of France, Journal of Neurosurgery, Dec 1992 [266]

Iron Skeletal Simulators

{#IronSkeletal} Between 1570 and 1700 in Italy, there were musculoskeletal iron simulators similar to Fabricius' Opera Chirurgica's illustration (1582). They feature major joints to teach articulation and dislocation of the limbs and corresponding treatment. One of these models is currently in the Science Museum in London [267].

Cigolis Anatomy

{#CigolisAnatomy} In 1598 in Florence, Cigoli's Anatomy or "Anatomia del Cigoli" is a bronze copy of the first wax anatomic model with the purpose of studying anatomy [268]. Sometime between the 17th and 18th century in Europe the Flayed Man (écroché) was used to teach medical students. The anatomic simulator shows the blood vessels, muscles and the skeleton sometimes [269].

In 1609, the first working eye simulator was made out of strips of leather to simulate the muscles. The pieces could be taken apart to show the physiology and anatomy [28]

Doctor's Lady

{#DoctorsLady} Between 1644 and 1912 in China, the "doctor's lady" is a miniature of a lying down nude female is used as a basic patient simulator. Under the moral code of the Ching Dynasty, women could not be directly examined by a male physician, and instead female patients would use the carved model to point to the location of symptoms [270–272].(See Fig A.6.)

Figure A.6: Doctor's Lady
Image Source: Mountshang at the English language Wikipedia [GFDL (http://www.gnu.org/copyleft/fdl.html) or CC-BY-SA-3.0 (http://creativecommons.org/licenses/by-sa/3.0/)]

{#Zumbo} In the 18th century in Geneva, the Italian sculptor Gaetano Giulio Zumbo created a realistic life-size patient simulator made out of polychromatic wax for the purpose of making an educational anatomic model [273, 274]. (See Fig A.7.)

Figure A.7: Zumbo
Image Source: I, Sailko [GFDL (http://www.gnu.org/copyleft/fdl.html) or CC BY-SA 3.0 (https://creativecommons.org/licenses/by-sa/3.0)]

Lelli's Wax Models

{#Lelli} Between 1730 and 1739 in Bologna, the painter Ecrole Lelli excelled in the domain of human body and its anatomy. He prepared artistic wax display. Giovanni Manzolini, a wax modeler and anatomist and Lellis' assistant, and his wife Anna Morandi became well known makers of wax anatomical models for medical education. Anna Morandi was known for her obstetric anatomy models as well [275–278]. (See Fig A.8.)

Figure A.8: Lelli
Image Source: Warburg [CC BY-SA 3.0 (https://creativecommons.org/licenses/by-sa/3.0) or GFDL (http://www.gnu.org/copyleft/fdl.html)], from Wikimedia Commons

Abraham Chovets' Advertised Simulator

{#AbrahamChovets} In 1733 in London, Abraham Chovet advertised a new figure for circulation simulator with a woman is opened alive showing the circulation of the blood through glass veins and arteries. The circulation from the mother to the child and back to the mother is shown with systolic and diastolic motion to the heart, and action of the lungs. The simulator did NOT exist and models were destroyed in 1888 [279].

Gregoir's Obstetric Simulators

{#Gregoir} In 1739 in Paris, dead babies were used by father and son surgeon accoucheurs Gregoir for obstetric simulators. Pelvis was used to show a variety of positions for delivery [280].

Vaucanson Plan

{#Vaucanson} In 1741 in France, the inventor Jacques de Vaucanson submitted a plan to simulate all animal operations such as circulation of the blood, respiration, digestion, the movement of muscles, tendon and nerves. His plan to use elastic gum to simulate veins as rubber tubes was approved by the king in 1761 but the project lapsed. Famous images of a defecating duck show the duck's ability to eat a grain from one end and get it out from the other [281].

Smellie

{#Smellie} In 1742 in the UK, Scot William Smellie's Improved a Phantom. Smellie's machine is described to have real human bones covered with fine smooth leather and stuffed with nice soft substance. The fetus is made of wood and rubber with articulating limbs and a placenta. All the parts seem very natural both to look and touch; the contents of the abdomen, intestine, kidney and large vessels look very natural. The uterus externum and internum are made to contract and dilate to vary the difficulty of the delivery. It is not known what Smellie's machine looks like [282, 283].

Giovanni Gall's Birthing Simulator

{#GiovanniGall} In the 1750s in Bologna, Giovanni Antonio Gall created a birthing simulator made of a glass uterus in a pelvis. It has a flexible fetus. The simulator is used to train surgeons and midwives for childbirth [260].

Du Coudray's Illustrated Book

{#DuCoudray} Between1759 and 1785 in France Angélique Marguerite Le Boursier du Coudray was a midwife and decided to transfer her knowledge by making an illustrated book "Abrege de l'art des accouchements" and used a life size mannequin with removable uterus. It simulated cervical dilation, newborn, a 7 month fetus and twins. Later models simulated blood and amniotic fluids [284]. (See Fig A.9.)

Figure A.9: Du Coudray
Image Source: Ji-Elle [CC BY-SA 4.0 (https://creativecommons.org/licenses/by-sa/4.0)], from Wikimedia Commons

Biheron's Wax Anatomy Models

{#Biheron} Between 1770 and 1771 in Paris, Marie-Catherine Biheron was an anatomist and the only woman in Europe to be an independent wax anatomy modeler. She modeled whole bodies and she was known to model a pregnant woman to demonstrate coping with dangerous deliveries. She reproduced the stages of birthing. The simulator supported modeling parts with a cervix that dilates and closes on demand and removable infants [285, 286].

Medical Venerina

{#MedicalVenerina} In 1771 in Florence, Medical Venerina (Venus) were realistic life-size patient simulator showing female anatomy made to look more sensual than cadaveric [260].(See Fig A.10.)

Figure A.10: Medical Venerina
Image Source: La Rocaille [CC BY-SA 2.0 (https://creativecommons.org/licenses/by-sa/2.0)], via Wikimedia Commons

In the 19th century, in Florence La Specula made wax models cast directly from patients. The female models were beautiful while the male anatomy was grotesque. When the wax models show manifestation of a disease they were called medical moulage [287].

Misemono

{#Misemono} During the same period, in Japan, unusually realistic Japanese models of pregnant women and fetuses were made from carved wood. A set of 7 detailed uterus carvings depict the different stages of fetal growth. They were known as “Misemono” [288]. Also during the same period in England, head models were made out of porcelain to teach phrenology which later on was proven to be useless [289].

Ziegler’s Wax Models

{#Ziegler} Starting in 1820 and during the 19th century, in Germany, specialized wax models became important part of academic research with Adolf Ziegler. He helped students learn how the fetus develops through his embryology models. The models of tiny complex and fast changing objects were very magnified which provided an indispensable help for students to grasp difficult subjects [290]. (See Fig A.11.)

Figure A.11: Ziegler
Image Source: Hill, M.A. (2018, November 13) Embryology Ziegler model 05.jpg. Retrieved from $https$: $//embryology.med.unsw.edu.au/embryology/index.php/File$: $Ziegler_{m}odel_{0}5.jpg$

Sachs's Ophtalmo-Phantom

{#Sachs} In 1820 in Berlin, Dr. Albert Sachs created an Ophtalmo-Phantom. The ophtalmo-phantom was used to teach ophthalmic surgery. The orbs are held in place by spring-loaded concave disks that push them against a ring of prongs. Both of these removable sockets are present. The set screw at the center of the throat allows the head to be tilted back to various positions [291].

Zeiller's Anatomic Models

{#Zeiller} Between 1820 and 1893, in Germany, Paul Zeiller worked on making anatomic models from wax even though he did not have a medical degree. He was opposed by scientists as models have no role in science. Zeiller had enough medical support to realize his vision in a private anatomy and anthropology museum. Medical students visited the museum as a visual aid but not as replacement for dissection [290].

Azoux's Anatomy Body Model

{#Azoux} In 1822 in France, Auzoux, a young French medical student developed a full size anatomy body model using papier maché which is cheap. The motivation is that when human bodies are used, they decay quickly making it impossible to use during warmer climates when refrigeration is not available [263, 292, 293]. (See Fig A.12.)

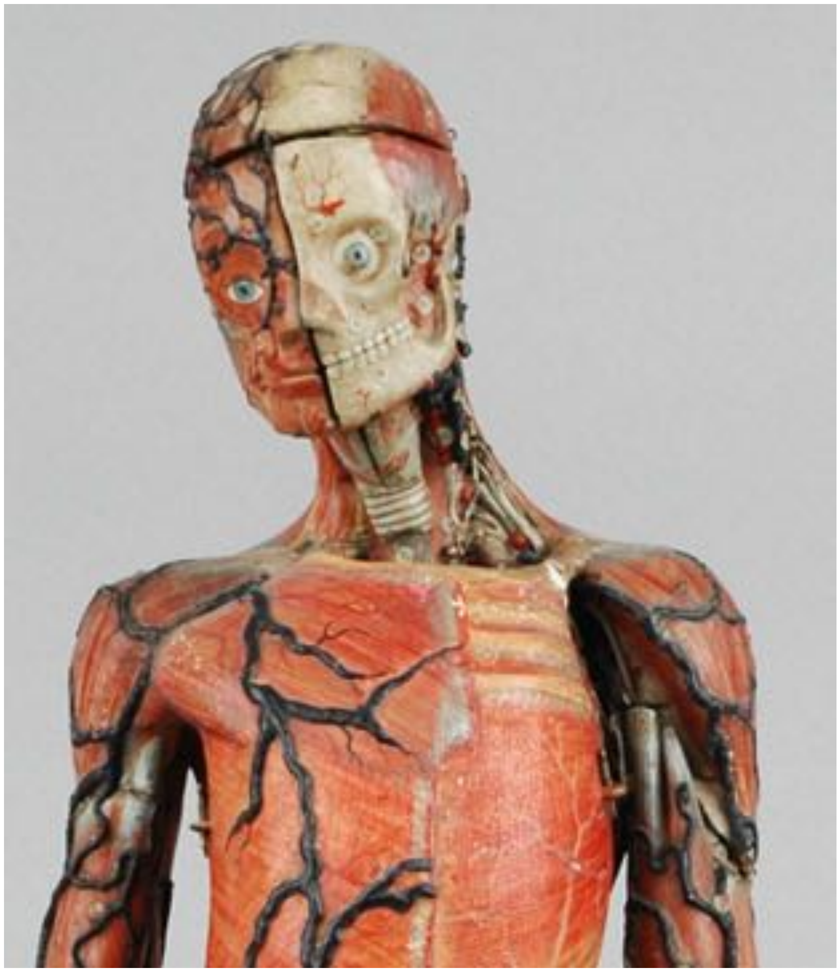


Figure A.12: Azoux's papier-mache models
Image Source: copyright of the Whipple Museum of the History of Science (Wh.5893) [293]. http://www.sites.hps.cam.ac.uk/whipple/explore/models/drauzouxsmodels/

Steger Anatomy Simulator

{#Steger} Between 1845 and 1938 in Germany, Steger's anatomy simulators were made from plaster which is cheaper than wax or papier maché and more accurate since frozen specimens were used to make casts [294, 295]. (See Fig A.13.)

Figure A.13: Steger
Image Source: Jon Cornwall and Chris Smith, Anatomical models by F.J. Steger (1845-1938): the University of Otago Collection, Department of Anatomy, University of Otago, Dunedin, New Zealand,European Journal of Anatomy 2014

{#Labus} In 1879 a surgical simulator by Labus' Laryngoscope, as described by Bacchi, had an electric bell to give feedback to the user [28].

{#Sushruta} Sometime after 1890, and originating from India, Sushruta Samhita's book was found. The book is an important text on medicine with emphasis on the importance of practice. In the book, surrogate material such as vegetables and leather bags are used to practice incision and excision [28]. (See Fig A.14.)

Figure A.14: Sushruta
Image Source: Alokprasad at en.wikipedia [CC BY-SA 3.0 (https://creativecommons.org/licenses/by-sa/3.0) or GFDL (http://www.gnu.org/copyleft/fdl.html)]

Schultze Phantom

{#Schultze} In 1890 in Germany, Schultze Phantom consists of a model of a female pelvis made of leather and metal stand. It demonstrates the mechanism of childbirth and applying forceps [296].

Resusci Anne

{#ResusciAnne} In 1960 in Norway, Resusci Anne was the first CPR mannequin created by Peter Safar, an Austrian physician, and Asmund Laerdal, a Norwegian toy maker. In the 19th century a young girl was pulled from the River Seine in Paris, generations later Asmund Laerdal started to develop a realistic and effective training aid to teach mouth to mouth resuscitation. The face of the mannequin was adopted from the death mask of the girl [29, 297]. (See Fig A.15.)

Figure A.15: Resusci Anne, the "most kissed" face of all time.
Image Source: aorta [CC BY 2.0 (https://creativecommons.org/licenses/by/2.0)], via Wikimedia Commons

In parallel, Gaumard had a rescue breathing and cardiac massage simulator including IV arm, catheterization and colonic irrigation [61].

Sim One

{#SimOne} In 1967 in California, SimOne, the 1st realistic anesthesia simulator was developed by Dr. Stephen Abrahamson and Dr. Judson Denson at the University of Southern California in collaboration with Sierra Engineering Company to train anesthesiologists, especially for endotracheal intubation and induction of anesthesia. The patient mannequin has a head, neck, thorax, upper abdomen, and arms. It was permanently mounted on a table and contained electromechanical and pneumatic actuators. SimOne did not simulate any electronic monitors. It had palpable pulses, heart sounds and movement. A variety of electronic sensors were used to detect the clinician's actions. A mask can be sensed on the mannequin's face by tiny relays embedded in the plastic flesh of the face. The position of a special magnetic endotracheal tube could be determined by magnetic sensors. SimOne could automatically recognize the identity and amount injected of four drugs. A mainframe computer program provided the control logic. There was no true modeling of pharmacokinetics or pharmacodynamics. The IOS provided a variety of preprogrammed events, including cardiac arrest, bucking, increased and decreased blood pressure or heart rate, changes in respiratory rate, and occlusion of a main-stem bronchus [298].

Harvey

{#Harvey} In 1968 in Florida, Harvey a cardiac training mannequin at University of Miami that Dr. Michael Gordon named after his mentor Dr. Proctor Harvey. Harvey can breathe, generate audible heartbeats, generate palpable pulses at multiple locations (right ventricle, left ventricle, aorta, pulmonary artery, carotids, and jugular vein), and blood Pressure. The heart sounds are in synch with the pulses. It can show 50 disease states [299]. (See Fig A.16.)

Figure A.16: Harvey
Image Source: Gene Hobbs [CC BY-SA 3.0 (https://creativecommons.org/licenses/by-sa/3.0)], from Wikimedia Commons

Gaumard's Female Nursing Simulator

{#FemaleNursing} In 1970 Gaumard introduced a female nursing simulator with dilating pupils if CPR is correctly administered. (See Fig A.17.)

Figure A.17: Female Nursing Simulator by Gaumard Scientific

In 1975 Gaumard introduced a family of GYN simulators [61].

{#Fontana} In 1976 in Florence, Fontana developed Musculoskeletal simulators made out of 3000 pieces of wood. Each piece can be moved and the model can be made to male or female [260, 263, 273].

Gaumard's Endoscopic Visualization

{#EndoscopicVis} In 1980, Gaumard introduced simulators for endoscopic visualization [61].(See Fig A.18.)

Figure A.18: Endoscopic Visualization by Gaumard Scientific

CASE

{#CASE} In 1986, the Comprehensive Anesthesia Simulation Environment mannequin system (CASE 0.5) prototype was created by Dr. Gaba and others in collaboration with CAE-Link. It had a minimal mannequin and a minimal audiovisual system. In 1988 the CASE 1.2 added physiologic monitoring. The physiologic simulators of ECG, invasive blood pressure, temperature, and oximetry displayed patient data on a monitor. The simulator was able to produce measurable CO2, had breath sounds, and produced clinically relevant pressures when ventilated with an anesthesia machine. Noninvasive blood pressure was simulated as well as Urine output and fluid infusion. The CASE 1.2 was used for staging of exercises in a real operating room to mimic critical incidents. This 1.2 prototype used a stock mannequin torso "Eddie Endo" from Armstrong Industries [300]. In 1995 the CASE 2.0 was there.

GAS: Gainesville Anesthesia Simulator

{#GAS} In 1988, Drs Michael Good and JS Gravenstein developed the Gainesville Anesthesia Simulator (GAS), which later became the prototype for the Medical Education Technologies Inc. (METI) simulator. GAS comprises a patient mannequin, anesthesia gas machine, and a full set of normally operating monitoring instruments. The patient can spontaneously breathe, has audible heart and breathe sounds, and palpable pulses. The lung model consumes and eliminates gas according to physiological principles.

Gaumard's Gynecologic Simulator ZOE

{#ZOE} In 1990, Gaumard presents ZOE, a Gynecologic Simulator that combines the ability to demonstrate multiple gynecologic procedures as well as practice laparoscopic examination and mini laparotomy [61]. (See Fig A.19.)

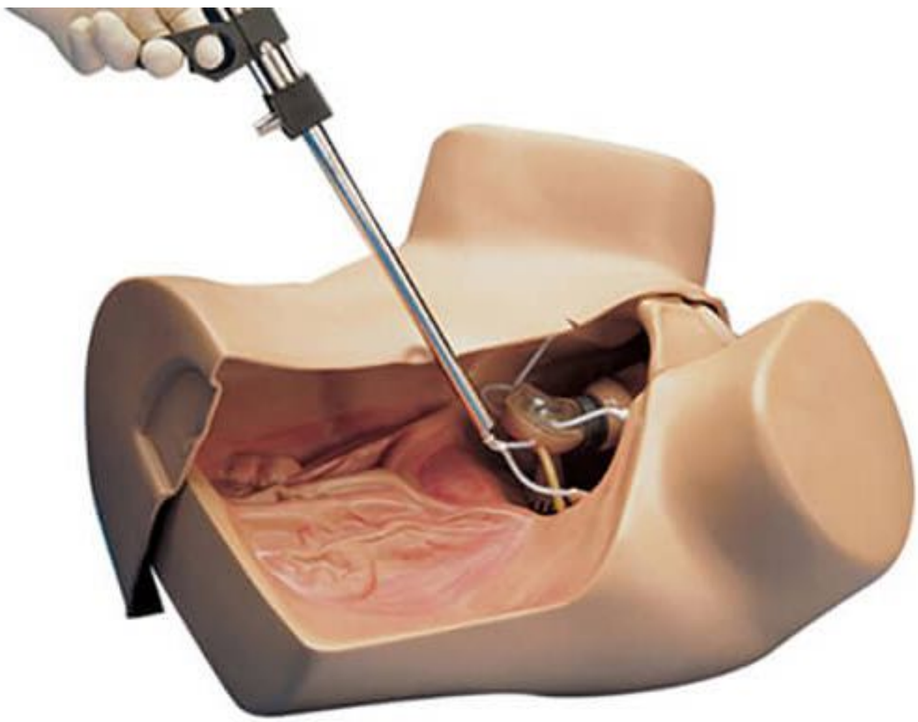


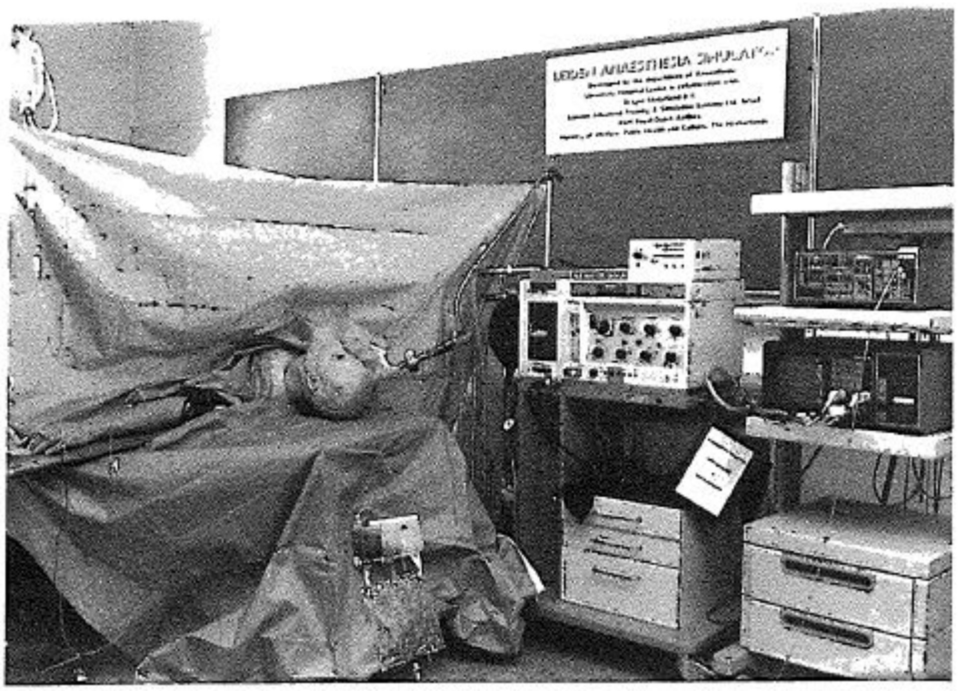
Figure A.19: ZOE **Image Source:** by Gaumard Scientific

PatSim for Anesthesia Training

{#PatSim} In 1990 at Stavanger College in Norway, PatSim was developed for training anesthesia and intensive care personnel. It consists of a mannequin controlled by a standard personal computer; it can be ventilated by an anesthesia machine or a ventilator, and can be connected to intravenous pumps. ECG, invasive and non-invasive blood pressure, airway pressure and CO2 are supported and data can be displayed on any monitor and existing medical equipment during simulation scenario. The mannequin is capable of spontaneous breathing. Typical clinical symptoms can be created during the simulation such as laryngospasm, change of lung compliance or airway resistance, pneumothorax, leakage of the intubation tube cuff, blocking of the breathing sounds from one lung, secretion, gastric regurgitation and diuresis [300, 301].

Leiden Anesthesia Simulator

{#Leiden} In 1992, the Leiden anesthesia simulator can be used with most commercial monitors and anesthesia equipment as they would when connected to a patient. It consists of commercially available head and thorax mannequins with artificial arm. The mannequin has anatomically correct airway, Laryngospasm, breath sounds, fluids can be in tubes through simulated veins, and urine output [302]. (See Fig A.20.)



The Leiden anaesthesia simulator.

Figure A.20: Leiden Anesthesia Simulator
Image Source: from The Leiden anaesthesia simulator,British Journal of Anaesthesia, 1994 [302]

Sophus

{#Sophus} In 1993 in Denmark, the Sophus group started in 1992 researching human error in anesthesiology. They developed a simulation environment with a PC user interface to the SOPHUS anesthesia simulator. The scripting language made making scenarios possible. The hardware consists of an interface between the PC and the anesthesia monitor, non-invasive blood pressure, a printer and loud speakers in the chest of the mannequin [303].

ACCESS

{#ACCESS} In 1994 at University of Wales, The Anesthesia Computer Controlled Emergency Situation Simulator (ACESS) was designed to simulate anesthetic emergencies to train junior doctors. The simulations are based on standard anesthetic equipment, with a microcomputer providing an image of commonly used instruments. Teachers present problems as scenarios to test the skills of the pupil [304, 305].(See Fig A.21.)

The ACCESS system, mannequin and anaesthetic equipment.

Figure A.21: ACCESS
Image Source: as it appears in M.W. P. Goodwin, G. W. G. French, Simulation as a training and assessment tool in the management of failed intubation in obstetrics, International Journal of Obstetric Anesthesia (2001) [305]

{#Ultarsim} In 1995, Ultrasim was an ultrasound simulation mannequin based on real US patient data sets. It started with replicated abdominal pathology relevant to obstetrics and gynecology and later expanded to the representation of diverse intra-abdominal problems. It was one of the first mannequin products by CAE-Link to include educational material, instruction manuals, and clinical case presentations [2].

Gaumard's Code Blue III

{#CodeBlueIII} Also in 1995 Gaumard presents Code Blue III, a family of computer interactive simulators with virtual instruments [61]. (See Fig A.22.)

Figure A.22: Code Blue III by Gaumard Scientific

Laerdal's Little Anne

{#LittleAnne} During the same period Laerdal introduces Little Anne as a supplemental trainer to meet the need for a lower student-mannequin ratio [29].

METI ‘s Human Patient Simulator HPS

{#HPS} In 1996, METI introduces HPS (Human Patient Simulator). HPS has true oxygen and CO2 gas exchange. It is designed specifically with anesthesia, respiratory and critical care programs in mind. The HPS’s cardiovascular, respiratory, neurological and pharmacological modeling and lung mechanics represent complex surgical, critical care and drug interaction scenarios. The HPS can interface with real patient monitors [306, 307].

{#MedSimEagle} In 1997, in Stanford CA, the MedSim-Eagle Patient Simulator is a realistic, simulator of the anesthetized or critically ill patient where the clinical environment and the patient are represented as real physical objects. It has 3 components: A patient mannequin based on a commercially available medical training patient mannequin. An interface cart controls the electrical and pneumatic components. It is connected to the simulation computers where mathematical models are running and to the instructor control station that has a graphics user interface. The MedSim-Eagle Patient Simulator includes features such as regional anesthesia, drug recognition system, cardiopulmonary bypass, computer-controlled eyes, arm movement, arm/leg swelling, drug customization, concentrations, pharmacokinetics, and pharmacodynamics. It has a more complete airway management training head and neck, better breath sounds and heart sounds as wells as a built-in gas analyzer, and pre-defined "abnormal" patients [308].

METI's PediaSim

{#PediaSim} In 1999, Business Wire announces PediaSim, the first pediatric Critical care simulator (by METI).

Gaumard's Noelle

{#Noelle} In 2000, The NOELLE maternal simulator is introduced by Gaumard. It can provide a complete birthing experience before, during, and after delivery. It features a full-size articulating female mannequin, intubatable airway with chest rise, IV arm for meds/fluids, removable stomach cover, practice Leopold maneuvers, multiple fetal heart sounds, automatic birthing system, measure head descent and cervical dilation, multiple placenta locations, replaceable dilating cervices, articulating birthing baby with placenta, and resuscitation baby with intubatable airway and umbilical catheter site [309].(See Fig A.23.)

Figure A.23: Noelle by Gaumard Scientific

Laerdal's Sim Man

{#SimMan} Also in 2000, SimMan was the 1st computerized commercial mannequin launched by Laerdal after purchasing Medical Plastics Laboratories. It had most of the features of previous models at a fraction of the cost. It did not have any computer modeling of physiology. An instructor drives the responses. It sold about 6000 units and retired after 10 years to be replaced with a wireless version [55, 310].

METI's Emergency Care Simulator ECS

{#ECS} In 2001, METI introduces the Emergency Care Simulator (ECS) which is similar in appearance to the HPS but it is more mobile and has autonomous capabilities such as de-saturating spontaneously during apneic episodes [55, 62].

Simulab's Trauma Man

{#TraumaMan} Also in 2001, TraumaMan is developed by Simulab as a surgical training mannequin for Advanced Trauma Life Support courses. The TraumaMan System allows students to practice several advanced surgical procedures including: cricothyroidotomy, percutaneous tracheostomy, needle decompression, chest tube insertion, pericardiocentesis, diagnostic peritoneal lavage, intravenous cut down. TraumaMan's bleeding tissues mimic human tissue from incisions to suturing [311, 312].

Gaumard's Pedi

{#Pedi} In 2002, Gaumard introduces PEDI is a pediatric simulator that include intubatable airway, and an injection training arm, an intra-osseous injection leg and appropriate arterial/venous simulation [61].

Gaumard's Premie

{#Premie} In 2003, Gaumard introduces Premie, a premature newborn, a 28 week articulating simulator. It has a realistic airway with tongue, vocal cords, trachea, esophagus, realistic internal organs for CPR. It supports "heel stick" simulation, Oral and nasal Intubation, suction procedures, bilateral lung expansion with realistic chest rise, peripheral and central cyanosis, changing rates of improvement and deterioration [313]. (See Fig A.24.)

Figure A.24: Premie by Gaumard Scientific

METI's Pelvic ExamSIM

{#ExamSim} Also in 2003, METI introduces Pelvic ExamSIM which is a simulator from Limbs & Things for training and assessing female pelvic examination. It has anatomical fidelity and soft tissue realism and combines sensor technology designed and validated by Dr Carla Pugh of Stanford and Northwestern Universities, USA. Pelvic ExamSIM is designed to represent variations of normal anatomy as well as pathology. Trainees can practice palpations with accuracy and learn the subtleties of appropriate touch pressures [314].

{#Hal} In 2004, Gaumard's Hal® S3000 marks the beginning of tether-less simulation where the technology, communications, compressor, and power supply is inside HAL®, eliminating external tubes, wires, and compressors. This allows for training during transport from an accident scene to the ER, to the ICU, while care providers diagnose and treat the patient's condition using real monitoring and resuscitation equipment. Hal can be controlled from up to 300 meters and between rooms and floors of conventional buildings [315]. (See Fig A.25.)

Figure A.25: Hal by Gaumard Scientific

METI's BabySim

{#BabySim} Also in 2004, METI introduces their infant simulator BabySIM. BabySIM is an infant-sized simulator with advanced physiology. The infant simulator has eyes that blink, variable pupil size, cooing, crying, tearing and secretions from the ears, eyes and mouth as well as bulging fontanel capability. BabySIM can produce heart, bowel and breath sounds, including bilateral chest excursion and seesaw breathing [316].

Gaumard's NoelleS555

{#NoelleS555} In 2006, NOELLE S555 and S565 by Gaumard are birthing simulators where each delivery can be precisely controlled while devices track student actions. The fetus may be manipulated to resolve a delivery dilemma. Students can see instant feedback of force, torque on the fetus, and the fetus head position. The data is graphed and synchronized with a fetal monitor for debriefing and evaluation. The fetus can be released on command [317]. (See Fig A.26.)

Figure A.26: NoelleS555 by Gaumard Scientific

Laerdal's SimBaby

{#SimBaby} Also in 2006, Laerdal introduces SimBaby, an advanced infant patient simulator ideal for training in all aspects of infant care. It has realistic anatomy and clinical functionality and supports airway management, breathing, circulation, defibrillation, vascular access, and anatomy [318].

Gaumard's NOELLE S575

{#NoelleS575} In 2007, Gaumard presents NOELLE S575 is a tether-less Maternal and Neonatal birthing simulator. Each delivery can be precisely controlled while devices track student actions. The fetus may be manipulated to resolve a delivery dilemma, to get instant feedback of force, torque on the fetus, and head position. This data is graphed and synchronized. The fetus is released on command [319].

METI's iStan

{#iStan} Also in 2007, METI presents iStan, an advanced wireless patient. It has internal robotics that mimics human cardiovascular, respiratory and neurological systems. When iStan bleeds, his blood pressure, heart rate and other clinical signs change automatically, and he responds to treatment with minimal input from an instructor. iStan has integrated physiological responses that closely resemble those of a human patient, multiple pulse points and fully independent right and left lungs that respond automatically to treatment. It can simulate flail chest, cyanosis on fingers and toes and difficult airway challenges [320].

Gaumard's Pediatric HAL

{#PediatricHal} In 2008, Gaumard presents Pediatric HAL®, a tether-less pediatric patient simulator, and PremieHAL, a premature patient simulator. They both remain fully functional while being moved from place to place [321, 322].

Laerdal's SimMan3G

{#SimManThreeG} Also in 2008, Laerdal introduces SimMan 3G, an advanced patient simulator that can display neurological and physiological symptoms, and automatic drug recognition [323].

Laerdal's ALS

{#ALS} In 2009, Laerdal introduces the ALS Simulator designed for emergency care training in both pre-hospital and in-hospital environments [324].

Prompt Task Trainer Birthing Simulator

{#Prompt} They also introduced the PROMPT task trainer Birthing Simulator which allows instructors to teach the complexities associated with birthing. It has a detachable abdominal and perineal skin that enables visualization of internal maneuvers and fetal positioning during training [325].

Laerdal's SimNewB

{#SimNewB} Also in 2009, SimNewB was introduced. It is a newborn simulator designed by Laerdal with the American Academy of Pediatrics to meet the training requirements of neonatal emergency medicine and resuscitation courses. SimNewB has realistic anatomy with chest rise and breathing rates up to 100 breaths per minute, lung compliance can be altered, and has intra-osseous access in both legs [326].

METI's MetiMan

{#MetiMan} During the same year, METI introduces METIman, a wireless realistic model-driven simulation built to withstand a wide variety of real-life indoor or outdoor learning environments. With automatic physiological responses, METI man is designed for teaching basic nursing and pre-hospital skills offering scenarios and features specifically tailored to each discipline [327].

Gaumard's Susie

{#Susie} In 2010, Gaumard introduces Susie® , a patient simulator that has tether-less technology where communications, compressor, and power supply are all inside her, eliminating external tubes, wires, and compressors. It can be controlled at distances up to 300 meters and between rooms and floors of conventional buildings. Susie smoothly transitions between physiologic states in response to commands from a wireless PC [328].

{#Clinispace} In 2011, CliniSpace presents Virtual Sim Center is a 3D, immersive, computer application for practicing patient care and clinical management with interactive devices affecting the health of virtual patients where multiple learners can collaborate, authors can create their own patient scenarios. It supports interaction with dynamic patient & medical objects, practice documentation and use of EMR. Learners' performance can be tracked [73].

Gaumard Tetherless Hal

{#Hal3201} In 2011, Gaumard presents Hal® 3201, a patient simulator that has tether-less technology allowing the communications, compressor, and power supply to be inside HAL®, eliminating external tubes, wires, and compressors. HAL® can be controlled at distances up to 300 meters and between rooms and floors of conventional buildings. HAL® smoothly transitions between physiologic states in response to commands from a wireless PC [329].

Laerdal's Mama Natalie

{#MamaNatalie} Also in 2011, Laerdal introduces MamaNatalie, a birthing simulator that comes with NeoNatalie. It is strapped onto the operator, who takes the role of the mother, and manually controls the training scenario featuring bleeding, positioning and delivery of the baby, delivery of placenta, fetal heart sounds, cervix landmark, urine bladder catheterization, uterine massage, and uterine compression [330].

Laerdal's Sim Junior

{#SimJunior} During the same year, SimJunior, a full-body interactive simulator for pediatric emergencies, was designed by Laerdal with the American Academy of Pediatrics to meet the education and training needs of healthcare providers. He represents a 6 year old boy that simulates a wide range of conditions from a healthy, talking child to an unresponsive, critical patient with no vital signs [331].

ShadowHealth

{#ShadowHealth} In 2012, Shadow Health offers digital clinical experiences for nursing programs, pharmacy and physician assistants. Shadow Health's patient cases are designed for students to practice communicating with and examining their virtual patients, as well as providing them the opportunity to document their findings, all while synthesizing the data and information they have discovered during the interaction. Concept labs are immersive tutorials that illustrate complex nursing subjects, they include examples of real body sounds and realistic, 3D anatomical body models to compare and contrast normal and abnormal findings. The interactive interface of these concept labs can be used to explore body systems at an in-depth level [74]. (See Fig A.27.)

Figure A.27: Shadow Health
Image Source: Permission Granted by Shadow Health to use this image via email from Brooke Rowe.

Gaumard's Airway Hal

{#AirwayHal} In 2013, Gaumard presents an adult airway trainer HAL® powered by BVM without the need for electrical power to ventilate airway and compress chest. It has an intubatable programmable airway and supports bilateral needle decompression and drainage. Either lung can be disabled, and it has interchangeable tracheal inserts [332]. (See Fig A.28.)

Figure A.28: Airway Hal by Gaumard Scientific

Laerdal's SimMom

{#SimMom} Also in 2013, Laerdal introduces SimMom, a full body birthing simulator with accurate anatomy and functionality to facilitate multi-professional obstetric training of birth management, with both manual and automatic delivery modes. SimMom supports the use of ultrasound, automatic delivery and wireless connectivity. Scenarios can be customized and real-time instructor can control and adapt the scenarios. Pre-programmed scenarios include normal and operative vaginal deliveries of infants, vaginal delivery with shoulder dystocia encounters and breech infants. Scenarios for the Automatic Delivery include obstetric emergency, normal vaginal delivery, shoulder dystocia for SimMom, and obstetrical emergencies for hospital nursing [333].

{#iHuman} In 2013, i-Human Patients are virtual patients that simulate a complete medical patient encounter with animated avatars, human physiology and pathophysiology, virtual histopathology and 3D anatomy to improving students' diagnostic reasoning skills and patient outcomes [75]. (See Fig A.29.)

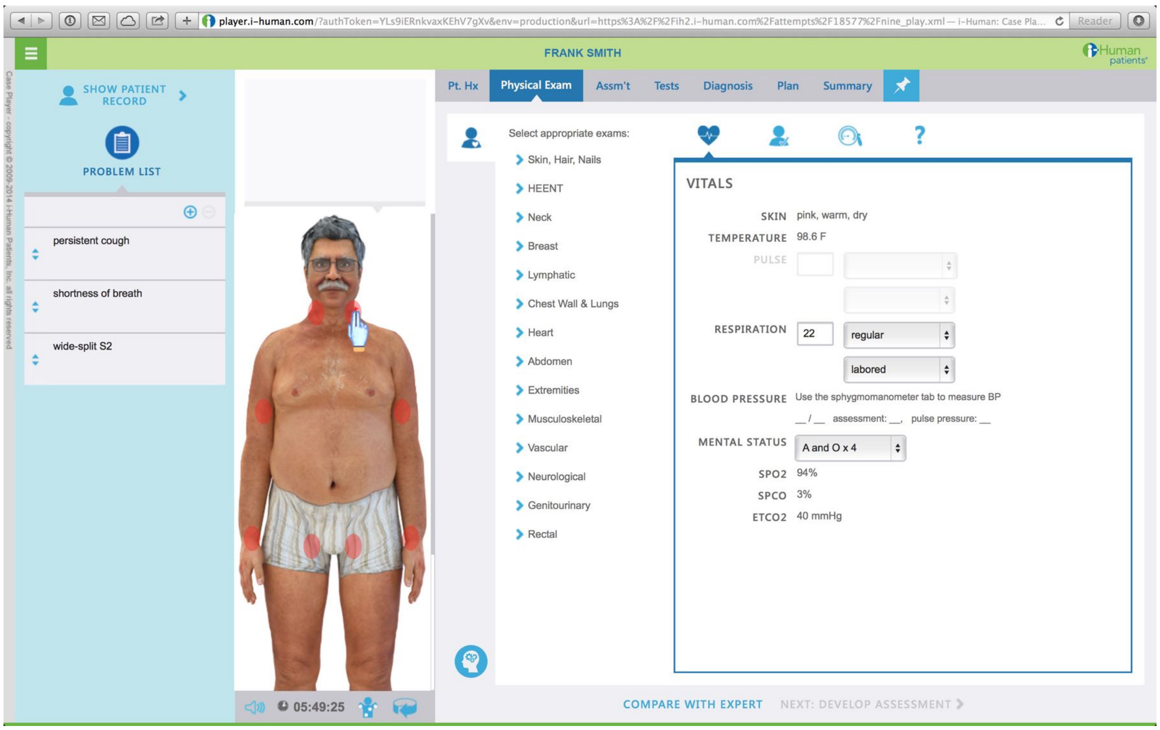


Figure A.29: iHuman **Image Source:** Image from www.i-human.com used with permission from Doug Miller.

CAE's Trauma Simulator Caesar

{#Caesar} In 2014, CAE Healthcare presents Caesar, a trauma simulator designed for disaster response and combat casualty care. Caesar is a life-sized wireless simulator with modeled physiology for point-of-injury training. Caesar is the most rugged patient simulator available today that is durable and water resistant; he remains tough-skinned and resilient through tourniquet placements, patient decontamination, and extreme temperatures and conditions. Caesar is built with modular limbs and has articulation of his neck, back, shoulder, elbows, forearms and wrists as well as directional eye movement and speech patterns that reflect his level of consciousness. He can present dramatic bleeding from up to six sites and produce automatic physiological responses to tourniquet application [64].

Laerdal's SimMan3GTrauma

{#SimManThreeGTrauma} During the same year, Laerdal presents SimMan 3G Trauma which is designed to train military and civilian emergency medical personnel in trauma situations, such as hemorrhage control both in classroom and in the battlefield [323].

Gaumard's Victoria

{#Victoria} Also in 2014, Gaumard presents Victoria, the newest member of Gaumard's NOELLE® family of birthing simulators. She provides the most advanced and realistic simulation experience that offers more immersive simulations that result in better learning experiences for the simulation participants [334].

Gaumard's Lucina

{#Lucina} In 2015, CAE Healthcare introduces Fidelis Lucina Maternal Fetal Simulator, a birthing simulator that offers training for childbirth maneuvers and emergency response. It is the only childbirth simulator with validated maternal-fetal physiology that allows learners to monitor and manage both patients without instructor intervention. Learners gain hands-on experience performing normal deliveries and pelvic exams to recognize cervical dilation, effacement, presentation, position and station, and maternal emergencies from cardiac and respiratory arrest to shoulder dystocia and breech deliveries [63]. (See Fig A.30.)

Figure A.30: Lucina by Gaumard Scientific

History of Simulators by Medical Area

Looking at the information from the chronological section, this section groups the simulators by type and refers to them using the short nametags chronologically under each medical area.

Basic Adult Patient Models and Simulators

{#VenusWillendorf}, {#WoodenAnatomical}, {#MayanHead}, {#BronzeStatues}, {#IronSkeletal}, {#CigolisAnatomy}, {#FlayedMan}, {#DoctorsLady}, {#Zumbo}, {#MedicalVenerina}, {#Zeiller}, {#Azoux}, {#Steger}, {#Fontana}, {#CodeBlueIII}, {#SimMan}, {#HALS3000}, {#SimManThreeG}, {#METIman}, {#SusieS2000}, {#HAL}

Obstetric and Birthing Simulators

{#Gregoire}, {#Smellie}, {#GiovanniGall}, {#DuCoudray}, {#Biheron}, {#Schultze}, {#Misemono}, {#EndoscopicVis}, {#Kny}, {#GYNsimulators}, {#ZOE}, {#Noelle}, {#ExamSim}, {#NoelleS555}, {#NoelleS575}, {#PROMPT}, {#MamaNatalie}, {#SimMom}, {#Victoria}, {#Lucina}

Anesthesia Simulators

{#SimOne}, {#CASE}, {#GAS}, {#PatSim}, {#Leiden}, {#Sophus}, {#ACCESS}, {#HPS}, {#MedSimEagle}.

Pediatric and Newborn Simulators

{#PediaSim}, {#PEDI}, {#Premie}, {#BabySIM}, {#SimBaby}, {#PediatricHAL}, {#SimNewB}, {#SimJuniror}

Surgical Simulators

{#EyeTrauma}, {#Fabricius}, {#Sachs}, {#Sushruta}, {#Cystoscopic}

CPR Heart Sounds Respiration Part Task Trainers

{#Labus}, {#ResusciAnne}, {#Harvey}, {#LittleAnne}, {#AirwayTrainer}

Trauma Simulators

{#TraumaMan}, {#Caesar}, {#SimManThreeGTrauma}

Emergency Care Simulators

{#ECS}, {#ALSSimulator}

Virtual Simulators

{#clinispace}, {#ShadowHealth}, {#iHuman}

Medical Simulation by Geographic Area or Major Company

This section classifies the early contribution by geographic areas. The more recent simulators are classified by company in the next section instead of geographic area since many current companies have multiple locations around the world.

Region: Africa

In Egypt 3400BC, they had wooden anatomical models of humans used in display during private parties to remind people to enjoy themselves and drink before they die {#WoodenAnatomical}

Region: Asia

The very first carved human model Venus of Willendorf is believed to have originated from Siberia but it was found in Europe as well {#VenusWillendorf}. Sushruta' s book where the first part-task trainers for practicing sutures on vegetable and leather bags were first described originated from India{#Sushruta}. Life-size bronze human body was used in to practice acupuncture China {#BronzeStatues}, also models of female body were used for female patients to point to the physician is male {#DoctorsLady}. In Japan, they had realistic models of pregnant women and fetuses {#Misemono}.

Region: Europe

Denmark

The Anesthesia simulator Sophus was developed Denmark {#Sophus}.

France

Azoux developed life size anatomic models using cheap papier maché {#Azoux}. The 1st eye trauma simulation was motivated by the injury of the king Henry II but the experiments failed {#EyeTrauma}. In obstetrics, Marie-Catherine Biheron modeled full bodies that include pregnant women and reproduced the stages of birthing {#Biheron}. Madame Du Coudray created the first illustrated book and used a life size mannequin to transfer her knowledge and experience as a mid-wife {#DuCoudray}. Surgeon accoucheurs Gregoire and sons used pelvis simulator to demonstrate different delivery positions {#Gregoire}.

Germany

Anatomic models from wax {#Zeiller}, and specialized models showing fetus development {#Ziegler}, and cheaper plaster models {#Steger} were developed. Schultze developed obstetric simulators to demonstrate the mechanism of childbirth and applying forceps {#Schultze}. Dr. Albert Sachs developed the Ophtalmo-Phantom, a surgical simulator {#Sachs}

Italy

Italy was known for different creating anatomical models in different regions of the country such as wax models Genova {#Zumbo}, and in Bologna {#Lelli}. Also in Bologna Geovanni Antonia Gall created a glass uterus in a pelvis with flexible fetus to help surgeons and midwives for childbirth {#GiovanniGall}. Life-size simulators showing female anatomy {#MedicalVenerina} and realistic models cast from patients showing manifestation of the diseases (medical moulage) such as La Specola's wax models, and Fontana's wood figures in Florence. Also in Italy they found iron skeletal models to teach joint articulations and dislocation of limbs {#IronSkeletal}.

Norway

The very first CPR part task trainer Resusci Anne was created by a Norwegian toy maker Asmund Laerdal. Laerdal is now one of the major companies in healthcare simulation {#ResusciAnne}. (See the rest of Laerdal simulators under the company section)

UK

Scot William Smellie improved on obstetric simulators {#Smellie}. The Anesthesia Computer Controlled Emergency Situation ACCESS was based on standard equipment and came from the University of Wales in the UK. {#ACCESS}

Region: America

Central Americal

Memento Mori carving model of a head where one half is healthy and the other half is sick were found in Central America to remind people that we are all mortals {#MayanHead}.

USA

In the USA Obstetric simulators were advertised by Kny-sheerer Company {#Kny} and New Adams Company. In Chicago, a cyctoscopic phantom was listed in a catalog of medical instruments in 1893. In 1968 Harvey was the 1st simulator that simulates heart sounds; it was invented by Dr. Michael Gordon at University of Miami. Starting in 1987, anesthesia simulators were invented on the west coast, specifically in California {#CASE}, {#SimOne}, {#MedSimEagle} and on the east coast in Florida {#GAS}. Many of the simulators started as research in a university and then became a joint effort with simulation companies such as CAE-Link and METI which later became CAE healthcare. (See the rest of the simulators by METI and CAE healthcare under the company section).

Company: CAE Healthcare / METI / CAE Link

CAE-Link presented a prototype for Comprehensive Anesthesia Simulation Environment mannequin system prototype that was created by Dr. Gaba and others {#CASE}. Also UltraSim is an ultrasound simulator based on real US patients' datasets {#UltraSim}. METI has simulators in different areas: In Anesthesia simulation, GAS which later became a prototype for METI simulators{#GAS} and the Human Patient Simulator can interface with real monitors {#HPS}. In Pediatric Simulation PediaSim was the first pediatric critical care simulator {#PediaSim}. In Emergency Care Simulation, ECS is similar in appearance to the EPS but I more mobile and autonomous {#ECS}. In obstetric Simulation, Pelvic ExamSim for training and assessing female pelvic examination {#ExamSim}. In Infant simulation, BabySIM is an infant simulator with advanced physiology {#BabySIM}. In basic patient simulators, iSTAN and METIman are wireless simulators {#iSTAN}, {#METIman}. METI was acquired by CAE healthcare which now offers a trauma simulator Caesar {#Caesar} and a birthing simulator Fidelis Lucina {#Lucina}.

Company: Laerdal

Laerdal was a toy company and created the 1st CPR simulator Resusci Anne {#ResusciAnne}, later on introduces a supplement Little Anne {#LittleAnne}. Laerdal is knows for their link of patient simulators: SimMan, SimManThreeG, SimMan Essential and a trauma Simulator SimManThreeG Trauma {#SimMan}, {#SimManThreeG}, {#SimManEssential}, {#SimManThreeG-

Trauma}. They offer ALS, an emergency care simulator {#ALS} and birthing simulators such as PROMPT, Mama Natalie, and SimMom {#PROMPT}, {#MamaNatalie}, {#SimMom}. They also have Sim Junior, a pediatric simulator {#SimJunior} and Infant and newborn simulators such as Sim Baby and SimNewB {#SimBaby}, {#SimNewB}.

Company: Gaumard

Gaumard had adult simulators such as Code Blue III and HALS3000, HALS3201 and other HAL variations as well as SusieS2000 {#CodeBlueIII}, {#HALS3000}, {#HAL}. The HAL series was the beginning of tether less simulation. They have pediatric simulators such as PEDI and Pediatric HAL {#PEDI}, {#PediatricHAL}, and new born simulators such as Premie, and PremieHAL {#Premie}. Gaumard is known for their birthing simulators such as Noelle, Noelle S555 and S575. The latest is NoelleS2200-Victoria. {#Noelle}, {#NoelleS555}, {#NoelleS575}, {#Victoria}

Technology Progression through History

People used different material to model and represent the human body, and later on simulate its functionality; they started with carved models as a pure representation. Sometimes that was used for artistic purposes and sometimes for a functional purpose such as pointing to specific regions. Often anatomical models were created for teaching purposes, especially for medical students. The medium of creation varied throughout history, starting with carved figures {#VenusWillendorf} {#MayanHead}, wooden anatomical models {#WoodenAnatomical} and wooden obstetric simulators {#Misemono}, bronze statues filled with liquid for acupuncture simulation {#BronzeStatues} or for studying anatomy {#CigolisAnatomy}, iron models that have moving parts {#IronSkeletal}, wax anatomical models {#Zumbo} {#Lelli} {#Biheron} {#Ziegler} {#Zeiller}. Plans for mechanical simulators were presented as well as plans for using elastic gum and rubber tubes to simulate circulation {#Vaucanson}. Obstetric simulators used more soft materials such as leather {#Smellie} and glass {#GiovanniGall}. Sometimes porcelain was used to model heads, especially for phrenology which was proven to be useless. Springs and disks were used to create Ophtalmo-Phantoms {#Sachs}. Cheaper materials were a motivation to create models from paper using papier maché {#Azoux} and later on plaster was an even cheaper material {#Steger}. Electricity was a new technology with the "Laryngo-fantome" made by Labus where they used an electric bell to give feedback {#Labus}. Current physical simulators use mechanical and electronic parts to simulate certain body functions such as breathing with chest rise and fall, pupil dilation, palpable pulse. Anesthesia simulators extend that to actually using chemical reactions so the simulator "can actually breathe". Computer advancements allow vitals to be simulated using mathematical models. Tetherless simulators such as the HAL series {#HAL} provide a more mobile patient simulator. Virtual simulators are more common now with the advancements in computer graphics and game engines {#clinispace} [75] {#shadowHealth}, this extends to multiple applications in virtual reality and possibly soon to a new generation of physical-virtual patients. In today's technology we can go beyond the flat computer screen to display visuals. Some interesting visual displays include head mounted displays [335] and projector based graphics using one or more projectors on parametric [336] and non-parametric surfaces [197]. Also there is research going on using nano-scale material that can change its appearance. Interaction interfaces can provide haptic and tactile feedback as an output of a simulation, at the same time touch can be sensed on a surface and sent to a simulation system as an input to the simulation.

Observations after Looking at the Data

After researching historical and current simulators around the world and in multiple domains the following information is worth mentioning. Developing good simulators can be expensive, people have tried to create cheaper alternatives {#Azoux} {#Steger} . Sometimes the simulator was way ahead of its time that the cost was too high and came in the way of being used more {#SimOne}. In few instances plans for simulation were approved and when people realized the cost they changed their mind so those simulator plans never got implemented. It is worth mentioning that looking at the timeline you can see that the density of simulators increased exponentially. This can be partially because many of the simulators were not documented, another reason may be that technology is advancing fast and the interested in simulation is increasing as well. Throughout this search, I found very little literature (if any) regarding neurological patient simulators, psycho-social patient simulators and cultural patient simulators. Physical-Virtual simulators may be able to help filling this gap.

APPENDIX B: APPENDIX B: SCENARIO 1 (STROKE)

This scenario was used in the stroke study in chapter 4. The clinical section below was provided by Prof. Laura Gonzalez, it is provided here to support more details about the Stroke scenario we used in chapter 4. My contribution is NOT in the clinical part. I contributed in the software development section of this chapter which includes coming up with script for the patient, recording audio, animating the patient to the audio, and software development for the scenario control.

Stroke Scenario: Clinical

this section contains the clinical aspects

Objective:

1) Perform focused assessment to include: affect, cognition, speech, vision, touch, sensation, movement 2) Call provider with findings using SBAR format

Report off: Situation:

The patient is a XX year-old woman brought in to the ED by ambulance with right upper and lower extremity hemiplegia. She also has a right visual field loss in both eyes. The patient's sister, with whom she lives, states she noticed a change in the patient approximately one hour prior to calling EMS. At this time, the patient was last at her baseline or symptom-free state one hour and 20 minutes ago. Admission orders have been written.

Background:

The patient has a history of transient ischemic attacks, the last one occurring six months ago. At that time, she experienced right-sided weakness, which completely resolved. She has atrial fibrillation, coronary artery disease, hypertension and hyperlipidemia. She is awake and responds appropriately.

Assessment

: Vital Signs: HR 94 and irregular, BP 180/120, RR 18, SpO2 94% on room air, Temperature 37.3C

General Appearance: appears stated age

Cardiovascular: Atrial fibrillation

Respiratory: Breath sounds clear

GI: Normoactive, abdomen soft and flat

GU: Has not voided

Extremities: Motor function in legs: Able to lift both legs slightly off bed. No limb ataxia

Skin: Pink, warm and dry

Neurological: Alert and oriented to person, place and time. Pupils are unequal, round, and reactive to light. Right visual field loss in both eyes. Mild sensory loss in right arm and right leg. Answers questions appropriately. Speech slurred but understandable. Smile asymmetry to right side. Decreased sensation to right side

IVs: 20-gauge saline lock in the right forearm, patent and non-reddened

Labs: Lab values are pending

Fall Risk: High-risk

Pain: Denies pain

Overview:

The patient is a X-year-old woman who was brought in to the Emergency Department (ED) by ambulance with right upper and lower extremity hemiplegia. She also has a right visual field loss in both eyes. The patient's sister, with whom she lives, states she noticed a change in the patient approximately one hour prior to calling emergency medical services (EMS). Until that time, the patient was symptom-free. The patient has a history of transient ischemic attacks (TIAs), and the last one occurred six months ago. At that time, she experienced right-sided weakness, which completely resolved. After stabilization in the ED, the patient is transferred to the Neuro Intensive Care Unit (NICU).

State 1 Initial Assessment, the patient exhibits an irregular HR in the 80s to 110s, BP in the 160s to 190s/120s to 130s, RR in the mid-teens, SpO2 in the low 90s and a temperature of 37.3C. Her breath sounds are clear. The cardiac rhythm is atrial fibrillation. Her skin is pink, warm and dry. Her bowel sounds are normoactive and her abdomen is soft and flat. She is alert and awake, and oriented to person, place and time. She answers questions appropriately with slurred speech. She responds to commands appropriately but exhibits mild dysarthria. She repeatedly complains of thirst. Her pupils are unequal, round and SLOW to react to light. Her gaze is normal, but there is right visual field loss in both eyes with complete hemianopia and visual inattention to the right side. She has right sided ptosis and smile asymmetry. Her grip is strong on left side, but there is no grasp on her right side. She is able to lift both legs slightly off the bed. There is no limb ataxia. She exhibits mild sensory loss in her right arm and right leg.

State 2 The patient's HR is in the 80s to 100s, BP is in the 150s to 180s/90s to 120s, RR is in the mid 20s, SpO2 is in the mid 90s on oxygen at 2 LPM via nasal cannula and temperature is 37.3C. Breath sounds are clear. Her cardiac rhythm is still atrial fibrillation. Her bowel sounds are normoactive. She is confused to place and time, but she is agitated and complaining of being thirsty. There is no urine output. The neurological exam: Her pupils are slow to react to light. She has decreased sensation of right side of face. Ptosis of right eye. Speech remains slurred.

Student should prepare to call provider with detailed SBAR

Orders:

Diagnosis: Rule out brain attack

Full code

NPO

Swallow evaluation

Bedrest

VS with neurological checks every 15 minutes

Notify healthcare provider with acute changes

Continuous cardiac and SpO2 monitoring

Titrate O2 to maintain SpO2 greater than 92%

Nitroglycerin ointment 1 inch for BP greater than 185/110

IV NS at 30 mL/hour

Computed tomography (CT) scan of head without contrast STAT

12-lead ECG

Capillary blood glucose STAT

Labs: CBC, INR, PT, PTT, serum glucose, Na, K, Cl, CO2, BUN, creatinine, troponin STAT

Anti-embolic stockings

Insert urinary catheter

Intake and output every 8 hours

Stroke Scenario: Development

This section contains elements related to the development, and additional images from the developed software.

Script

The following is a list of audio recorded to serve as patient's answers.

Numbers: Zero to 20, 24, 30, 40, 42, 43, 48, 49, 50, 60, 70, 80, 90, 100

Years: 1973, 1968, 2017

2 hours

(deep breath)

(humor sound)

A little

Aaah

About two hours ago

Back

Black

Blonde

Brown

By ambulance

Bye

Can you say that again

Cheek

Chin

Down

Dull

Eye

Forehead

Front

Good

Green

Hello

Here

here

Hold on

Hurts

I am not feeling well

I am not sure

I can’t do that

I can’t feel it

I can’t hear you

I can’t answer the question

I can’t see you

I don’t know how to answer that

I don’t remember

I don’t know

I don’t understand

I feel funny

I feel it

I forgot

I have a headache

I have a headache and tingling in my face

I hear it

I see it

I took the bus

I’m confused

I’m scared

I’m Thirsty

In the clinic

In the hospital

It was sudden

it’s on the chart

It’s not clear

Jaw

Leave me alone (agitated)

Left

Maybe

Mouth

My memory is vague

None

None that I noticed

No

No problem

Nose

Not so good

Ok

Please come closer

Please

Rarely

Right

See you later

Sharp

Side

Soft

Sometimes

Sorry

Teen

Thank you

Trump

Tuesday

uh

Umh about a 6

ummm

Up

Vera Real

wait

what’s going on?

whole head

why?

with my sister

Would you please let me know before you touch me

Yes

Images from the software

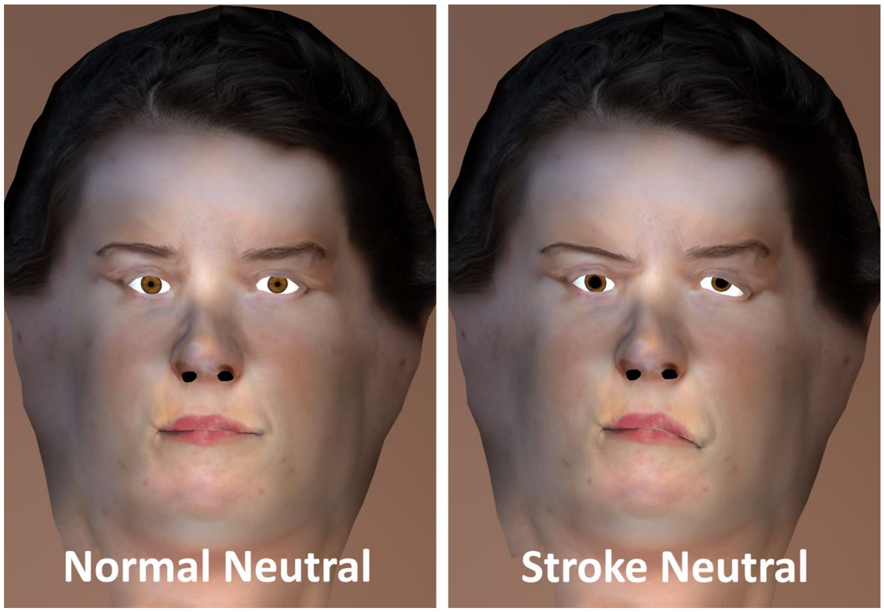


Figure B.1: Neutral

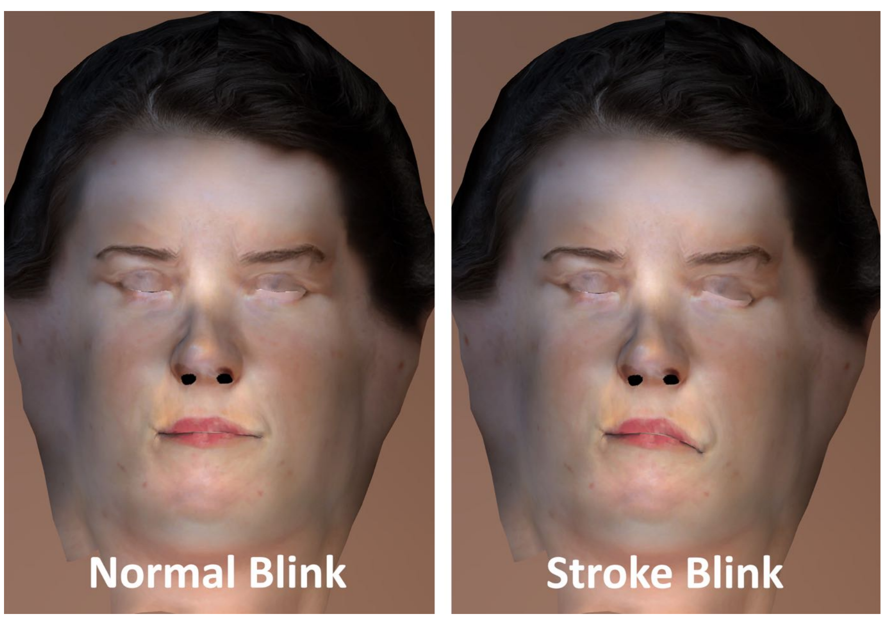


Figure B.2: Blink

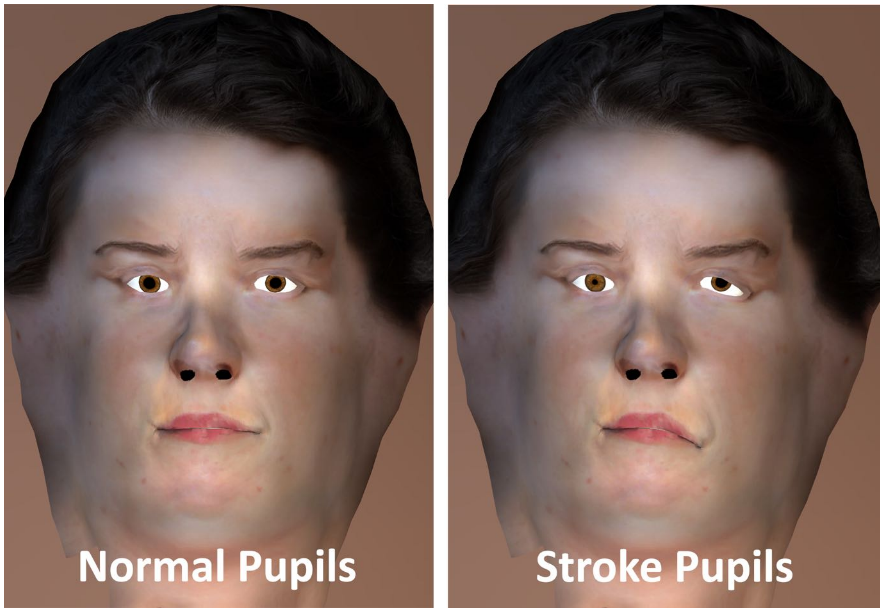


Figure B.3: Pupils

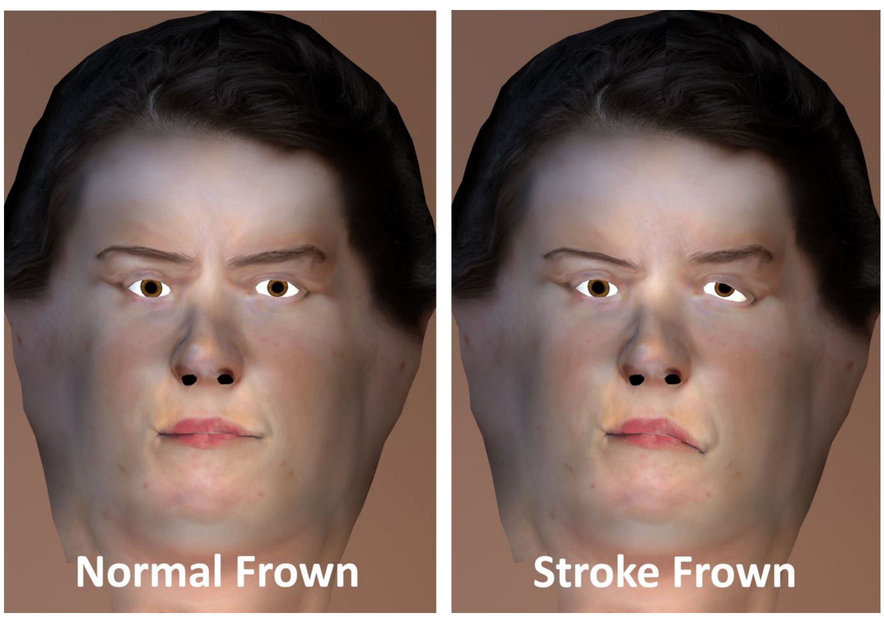


Figure B.4: Frown

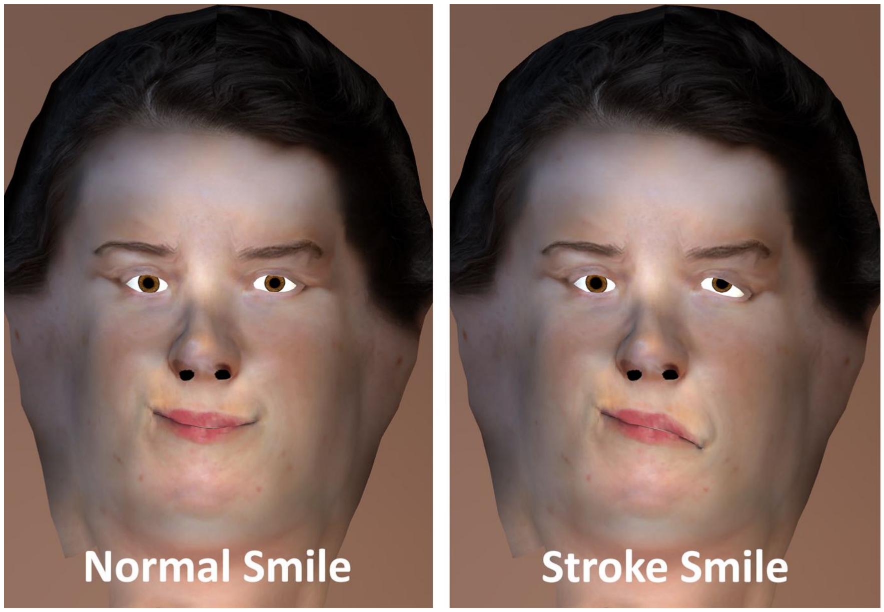


Figure B.5: Smile

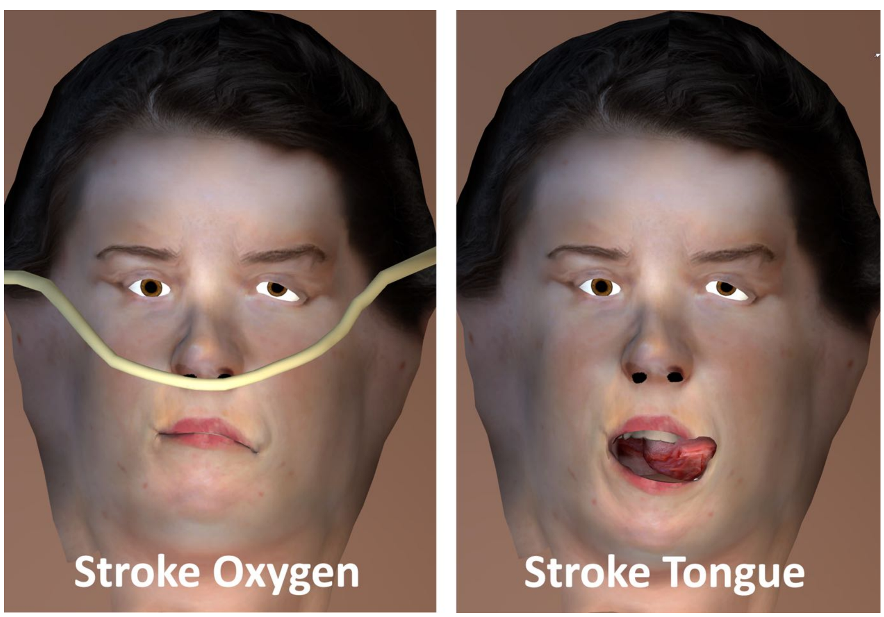


Figure B.6: Stroke tongue and O2

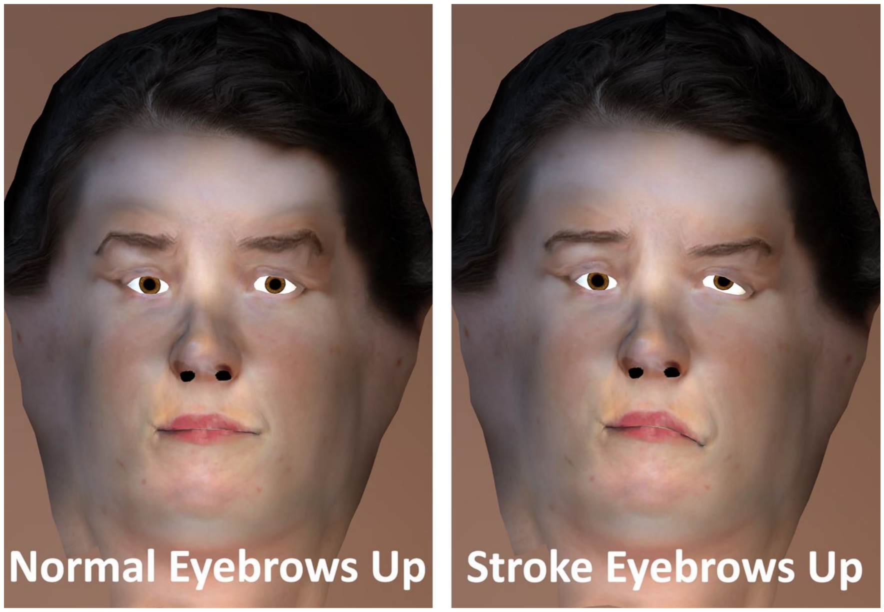


Figure B.7: Eyebrow up

APPENDIX C: APPENDIX C: SCENARIO 2 (SEPSIS)

This scenario was used in chapter 5. The clinical section below was provided as a collaboration with Professors from College of Nursing (Prof. Laura Gonzalez, Prof. Mindi Anderson, and Prof. Desiree Diaz) and College of Medicine (Prof. Juan Cendan). I contributed in the software development section of this chapter which includes coming up with script for the patient, recording audio, animating the patient to the audio, and software development for the scenario control.

Sepsis Scenario: Clinical

Report:

Patient's Name: Pauly Gee Date of Birth: 1 Nov 2012

5-year-old comes to the pediatric office with fever and lethargy, mom states he had a cold approx. 2 days ago. With runny nose, cough and congestion. She states he "felt" warm. Today he is slow to respond to verbal commands. Normally the parent would be in the room with the child, but the mom had to step away, please go ahead and assess the patient

Vitals

RR = 30 per min HR – Pulse = 110 bpm BP = 88 / 53 Temp = 101.4 F O2 Sat = 93

Controller Notes

Notes to Controller (not shown to participant) He is lethargic, eyes half cast. Mild discoloration to circumoral region. He is hot to touch. Skin exam reveals mottling over torso and extremities.

Voice = sleepy Patient coughs Ear pain = no

What hurts: hurts to breathe Pain Level = 7 Trigger 1: When is my mom coming back Trigger 2: I'm thirsty. Trigger 3: I'm Tired. Trigger 4: Cough + clear throat + Do you have a tissue. Trigger 5: is there a tv?

Sepsis Scenario: Development

Script

We developed the same combined list of responses and recorded the super-set in 3 tones: normal, lethargic, and in pain. each tone what used in a different scenario. The lethargic one was used for sepsis. While all the audio responses were technically available, we grayed out the ones that make sense for child abuse.

Rank Numbers: 1st,2nd,3rd,4th,5th,6th,7th,8th,9th,10th

Numbers: 1,2,3,4,5,6,7,8,9,10,17

Date: 22-May

Years: 2009,2010,2011,2012,2013,2017

A lot

A little

Aaah

Ask dad

Ask mom

Awww [pain]

Bye

Can you scratch my arm?

Dad

Do you have a tissue?

Do you have toys?

Do you have video games?

Doctor’s.

Down

Dull

far

Fell off my bike

Friday

Gerry

Go away

Good

Help me

Here [point to chest]

Hi

Hospital.

Hurts to breathe.

I can’t hear you

I can’t tell you

I don’t know

I don’t think so

I don’t want to

I fall a lot

I fell

I like to play

I think so

I wanna go home

I want daddy

I want mommy

I was bad

I'm hungry

I'm scared

I'm thursty

I'm tired

Is my nose bleeding?

Is there a TV

It hurts

It hurts here

It just hurts

It really really hurts.

It's a secret

Jo

Just happened

Kindergarten

Leave me alone.

Left

Mary

Maybe

Mom

My arm

My arm is itchy

My nose is bleeding

My nose is itchy

My tummy hurts

Near

Not sure what happened

November 1st

No

November

Office

OK

Orlando

Ouch

Pauly Gee

Pedro Johnson

Playing outside

Right

Riverdale Elementary

Say again?

See you later

Sharp

Sneeze

Sneeze2

Sometimes

Sorry

Spiderman

Stop

Sudden

Superheroes

Thank you

That one

the flash

The other one

This one

Thursday

Today

Uh

Ummm

Up

Uuhh [whiney sound effect]

When is my mom coming back?

Where is dad?

Where is mom?

Yes

Yesterday

[clear throat]

[crying2]

[crying]

[deep breath2]

[deep breath]

[laugh2]

[laugh]

[multiple cough]

[SFX after sneeze2]

[SFX after sneeze]

[single cough]

[sleepy tone] Bad.

[sleepy tone] I don’t feel good.

[sleepy tone] Yucky.

[whining2]

[whining]

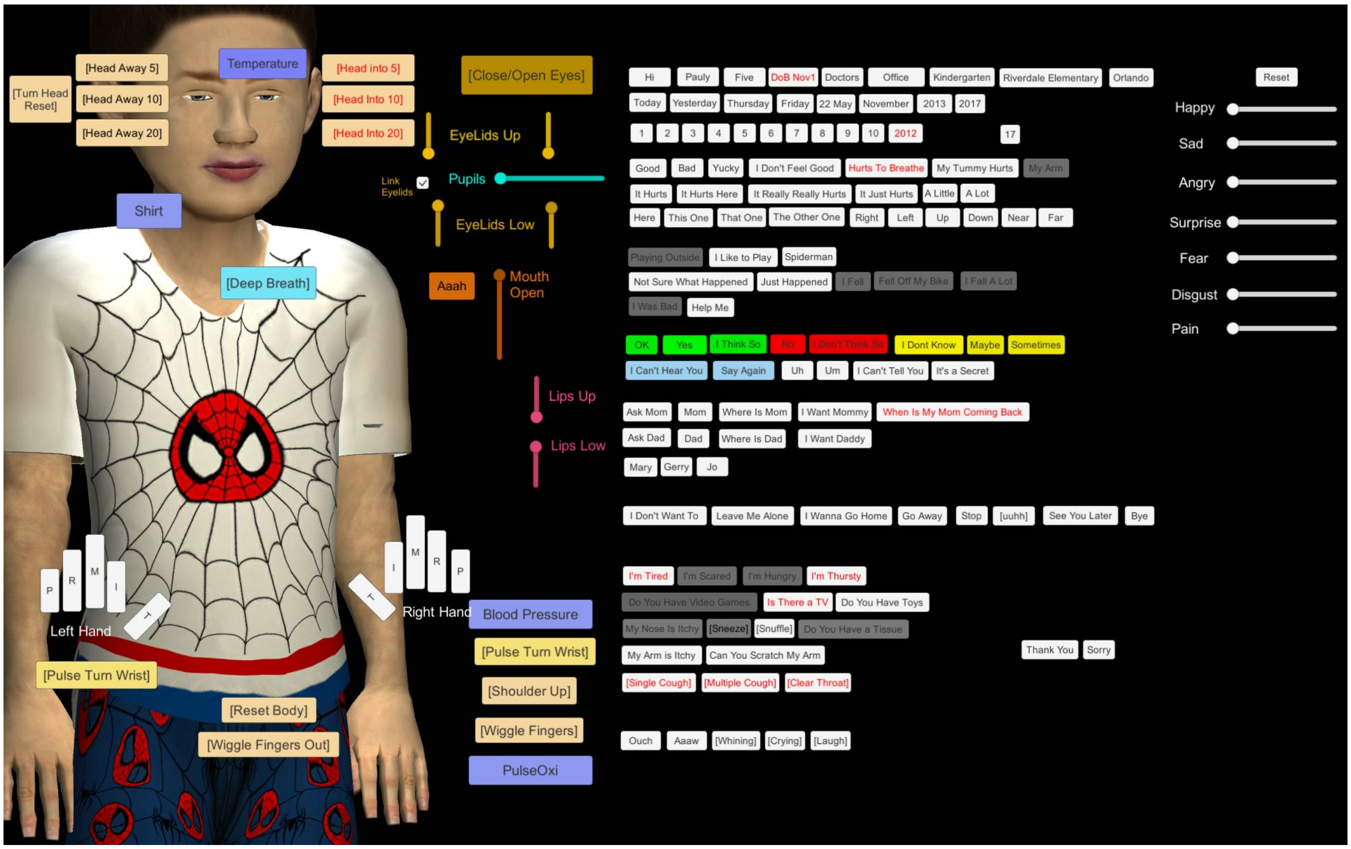


Figure C.1: GUI sepsis

Figure C.2: Sepsis Head

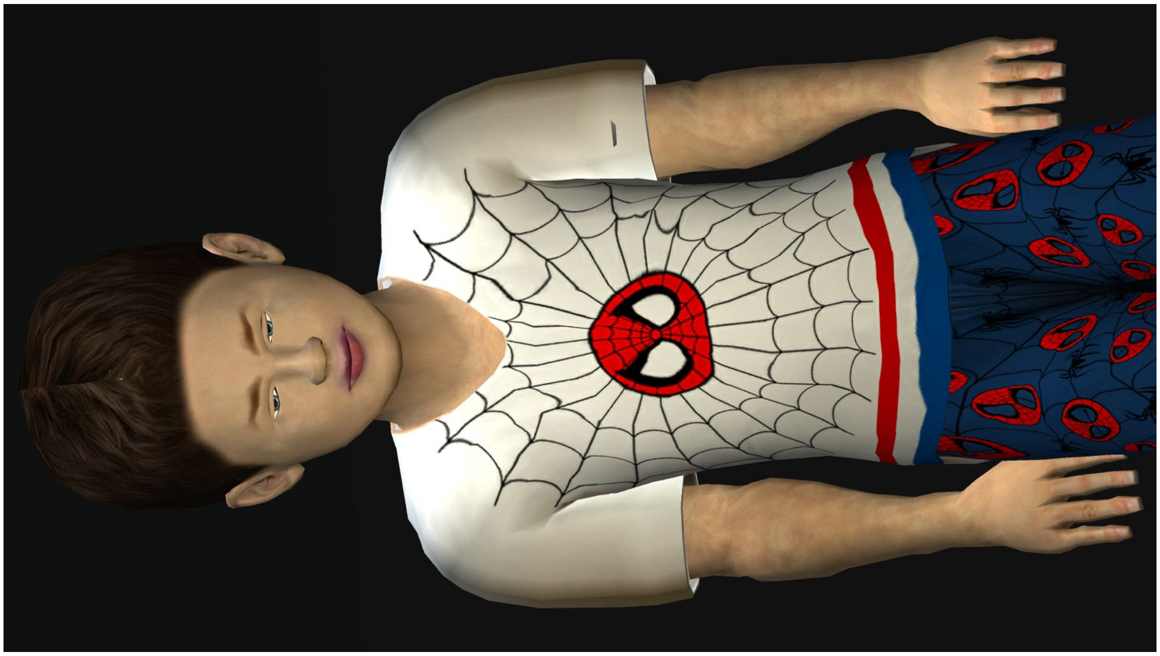


Figure C.3: Sepsis body with shirt

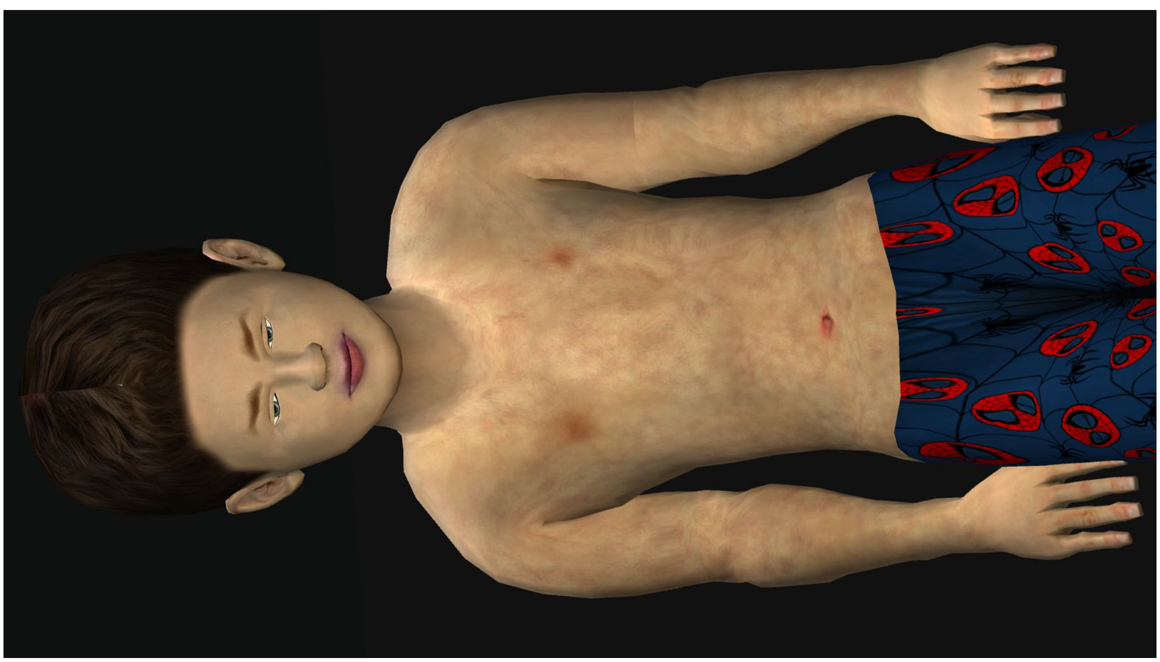


Figure C.4: Sepsis body no shirt

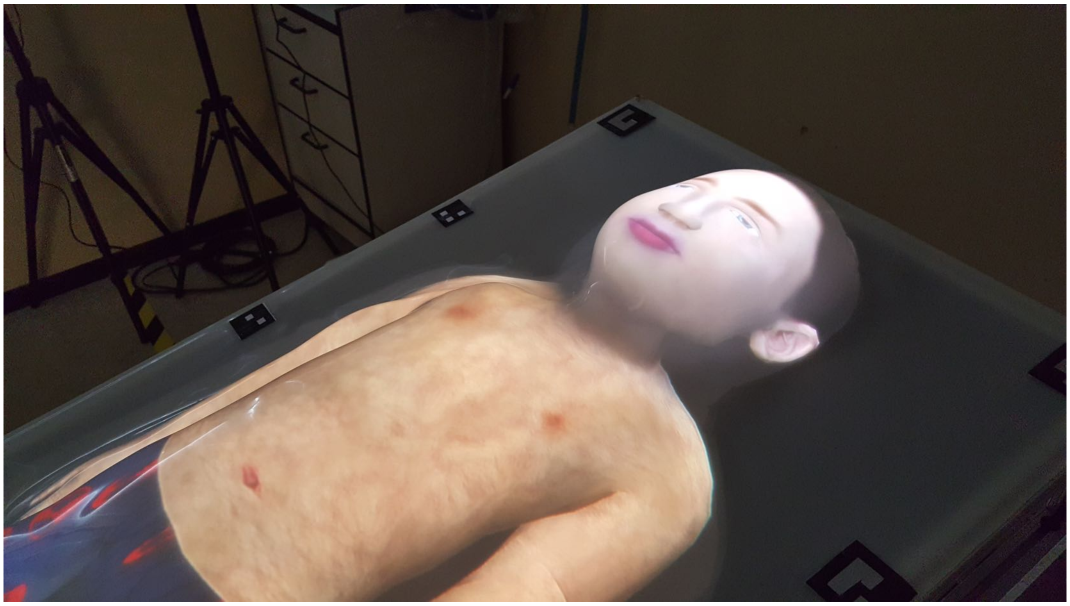


Figure C.5: Sepsis closeup

APPENDIX D: APPENDIX D: SCENARIO 3 (ABUSE)

This scenario was used in chapter 5. The clinical section below was provided as a collaboration with Professors from College of Nursing (Prof. Laura Gonzalez, Prof. Mindi Anderson, and Prof. Desiree Diaz) and College of Medicine (Prof. Juan Cendan). I contributed in the software development section of this chapter which includes coming up with script for the patient, recording audio, animating the patient to the audio, and software development for the scenario control.

Abuse Scenario: Clinical

Report:

Patient's Name: Joe Johnson Date of Birth = 22 May 2011

6-year-old male comes to the pediatric office with guardian. Guardian states child fell off sofa while jumping on it and injured his arm. Child states he fell off his bike. Normally the guardian is in the room, but the guardian had to step away, please go ahead and assess the patient.

Vitals

RR = 18 per min HR – Pulse = 80 bpm BP = 103/68 Temp = 97.8 F O2 Sat = 99

Controller Notes

On exam left radius and ulna deviated. Able to move distal fingers with good capillary refill <3 secs. Nurse request child remove shirt for a more detailed exam and notices cigarette burns, teeth marks and bruising in different states of healing.

What hurts: My arm. My tummy hurts. Pain Level = 7 Trigger 1: When is my mom coming back Trigger 2: I'm Hungry. Trigger 3: I'm Scared. Trigger 4: Sneeze + Snuffle + Do you have a tissue. Trigger 5: Do you have video games?

Abuse Scenario: Development

Script

We developed the same combined list of responses and recorded the super-set in 3 tones: normal, lethargic, and in pain. each tone what used in a different scenario. The pain one was used for abuse. While all the audio responses were technically available, we grayed out the ones that make sense for child abuse.

Rank Numbers: 1st,2nd,3rd,4th,5th,6th,7th,8th,9th,10th

Numbers: 1,2,3,4,5,6,7,8,9,10,17

Date: 22-May

Years: 2009,2010,2011,2012,2013,2017

A lot

A little

Aaah

Ask dad

Ask mom

Awww [pain]

Bye

Can you scratch my arm?

Dad

Do you have a tissue?

Do you have toys?

Do you have video games?

Doctor’s.

Down

Dull

far

Fell off my bike

Friday

Gerry

Go away

Good

Help me

Here [point to chest]

Hi

Hospital.

Hurts to breathe.

I can’t hear you

I can’t tell you

I don’t know

I don’t think so

I don’t want to

I fall a lot

I fell

I like to play

I think so

I wanna go home

I want daddy

I want mommy

I was bad

I'm hungry

I'm scared

I'm thursty

I'm tired

Is my nose bleeding?

Is there a TV

It hurts

It hurts here

It just hurts

It really really hurts.

It's a secret

Jo

Just happened

Kindergarten

Leave me alone.

Left

Mary

Maybe

Mom

My arm

My arm is itchy

My nose is bleeding

My nose is itchy

My tummy hurts

Near

Not sure what happened

November 1st

No

November

Office

OK

Orlando

Ouch

Pauly Gee

Pedro Johnson

Playing outside

Right

Riverdale Elementary

Say again?

See you later

Sharp

Sneeze

Sneeze2

Sometimes

Sorry

Spiderman

Stop

Sudden

Superheroes

Thank you

That one

the flash

The other one

This one

Thursday

Today

Uh

Ummm

Up

Uuhh [whiney sound effect]

When is my mom coming back?

Where is dad?

Where is mom?

Yes

Yesterday

[clear throat]

[crying2]

[crying]

[deep breath2]

[deep breath]

[laugh2]

[laugh]

[multiple cough]

[SFX after sneeze2]

[SFX after sneeze]

[single cough]

[sleepy tone] Bad.

[sleepy tone] I don't feel good.

[sleepy tone] Yucky.

[whining2]

[whining]

Supporting Images

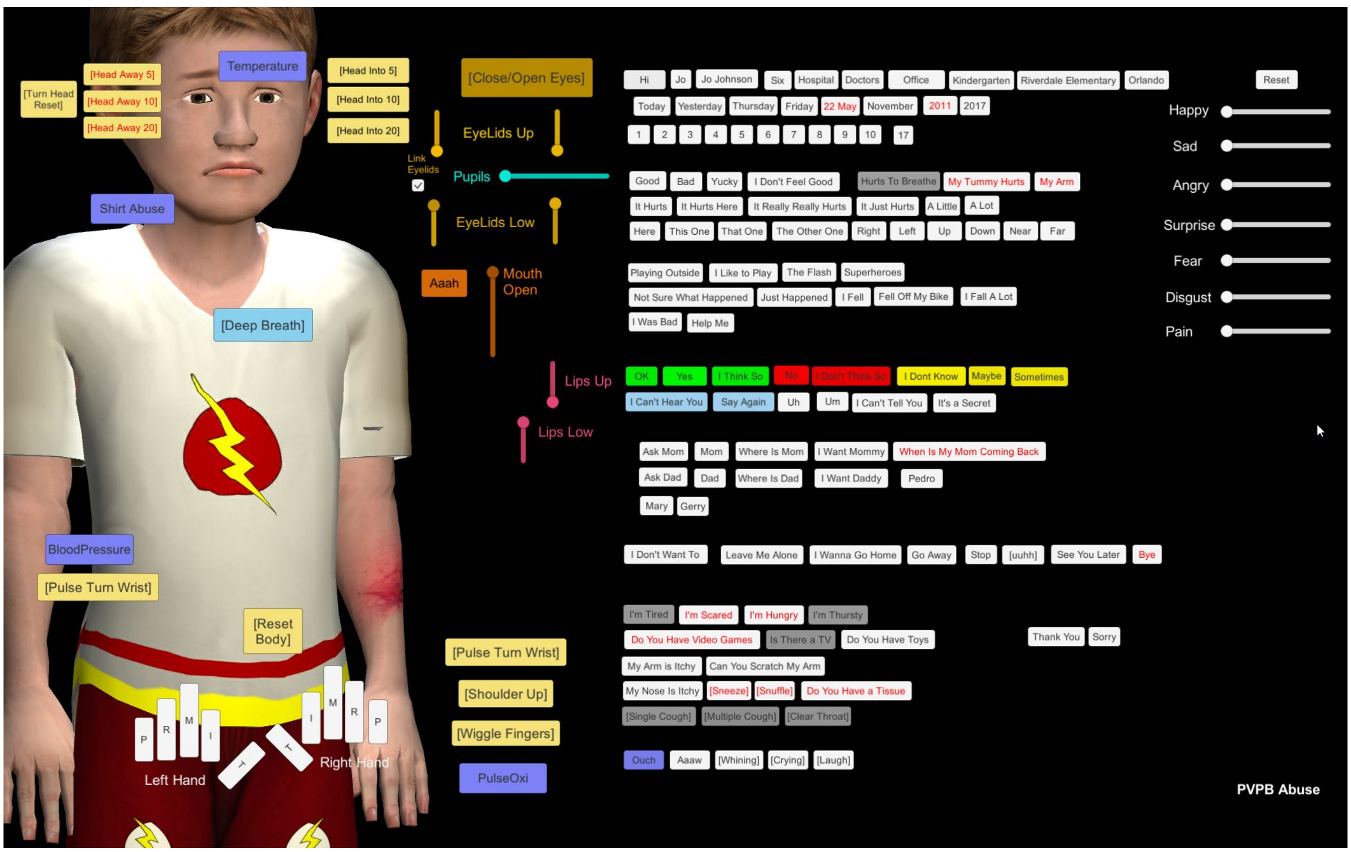


Figure D.1: GUI for child abuse scenario

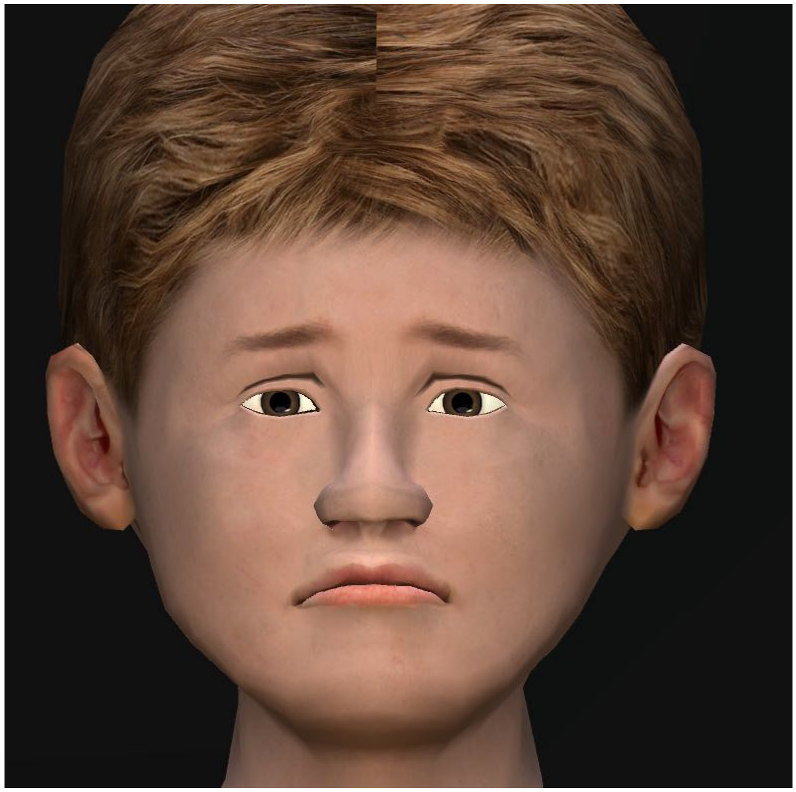


Figure D.2: Abuse Head

Figure D.3: Abuse body with shirt

Figure D.4: Abuse body no shirt

Figure D.5: Abuse closeup

APPENDIX E: APPENDIX E: SCENARIO 4 (BURNS)

This Appendix shows images for burns developed for the Physical-Virtual child patient. This was was developed to explore different variations of appearance. This was not used in a study.

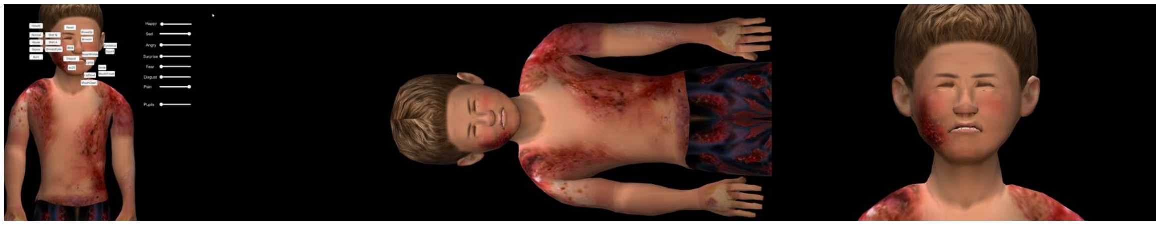


Figure E.1: GUI interface for burnt patient.

Figure E.2: Burn Patient's Face

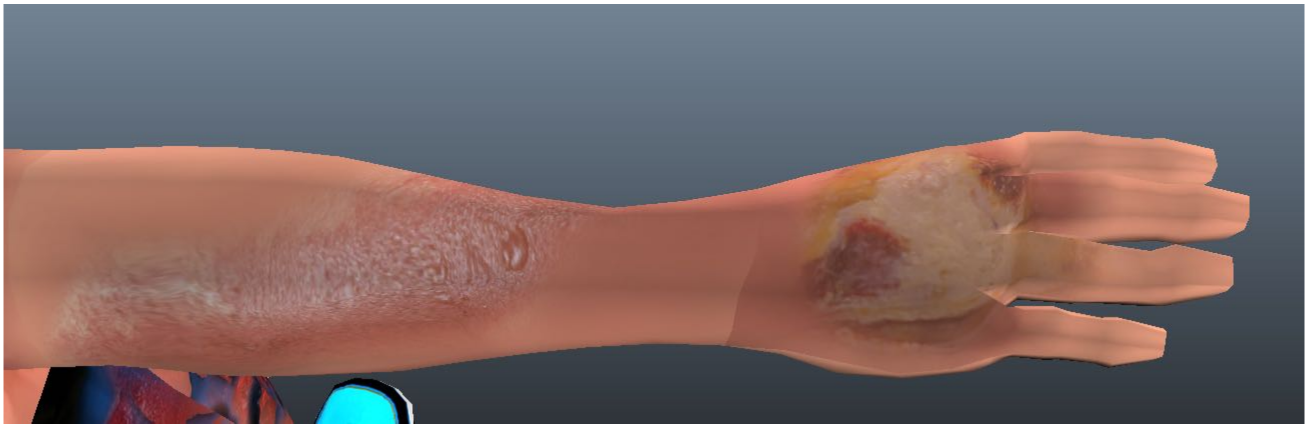


Figure E.3: Burn Patient's Arm

Figure E.4: Burn Patient's Back

Figure E.5: Front view of the Burn Patient

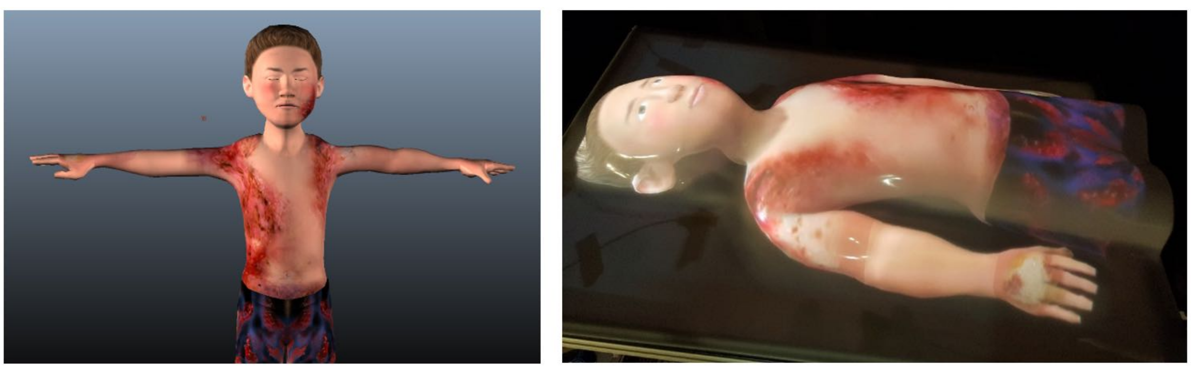


Figure E.6: Burn Patient projected on the PVPbed

APPENDIX F: APPENDIX F: IRB LETTERS

University of Central Florida Institutional Review Board
Office of Research & Commercialization
12201 Research Parkway, Suite 501
Orlando, Florida 32826-3246
Telephone: 407-823-2901 or 407-882-2276
www.research.ucf.edu/compliance/irb.html

Approval of Human Research

From: **UCF Institutional Review Board #1**
FWA00000351, IRB00001138

To: **Salam Daher** and Co-PIs: **Greg Welch, Laura Gonzalez**

Date: **May 16, 2016**

Dear Researcher:

On 05/16/2016, the IRB approved the following human participant research until 05/15/2017 inclusive:

Type of Review:	IRB Continuing Review Application Form Expedited Review
Project Title:	Assessment of Neurologic Symptomatology Using an Interactive Physical-Virtual Head with Touch.
Investigator:	Salam Daher
IRB Number:	SBE-15-11364
Funding Agency:	
Grant Title:	
Research ID:	N/A

The scientific merit of the research was considered during the IRB review. The Continuing Review Application must be submitted 30days prior to the expiration date for studies that were previously expedited, and 60 days prior to the expiration date for research that was previously reviewed at a convened meeting. Do not make changes to the study (i.e., protocol, methodology, consent form, personnel, site, etc.) before obtaining IRB approval. A Modification Form **cannot** be used to extend the approval period of a study. All forms may be completed and submitted online at https://iris.research.ucf.edu .

If continuing review approval is not granted before the expiration date of 05/15/2017, approval of this research expires on that date. When you have completed your research, please submit a Study Closure request in iRIS so that IRB records will be accurate.

Use of the approved, stamped consent document(s) is required. The new form supersedes all previous versions, which are now invalid for further use. Only approved investigators (or other approved key study personnel) may solicit consent for research participation. Participants or their representatives must receive a copy of the consent form(s).

All data, including signed consent forms if applicable, must be retained and secured per protocol for a minimum of five years (six if HIPAA applies) past the completion of this research. Any links to the identification of participants should be maintained and secured per protocol. Additional requirements may be imposed by your funding agency, your department, or other entities. Access to data is limited to authorized individuals listed as key study personnel.

In the conduct of this research, you are responsible to follow the requirements of the Investigator Manual.

On behalf of Sophia Dziegielewski, Ph.D., L.C.S.W., UCF IRB Chair, this letter is signed by:

University of Central Florida Institutional Review Board
Office of Research & Commercialization
12201 Research Parkway, Suite 501
Orlando, Florida 32826-3246
Telephone: 407-823-2901 or 407-882-2276
www.research.ucf.edu/compliance/irb.html

Approval of Human Research

From: **UCF Institutional Review Board #1**
FWA00000351, IRB00001138

To: **Salam Daher** and Co-PIs **Greg Welch and Laura Gonzalez**

Date: **March 15, 2017**

Dear Researcher:

On 03/15/2017 the IRB approved the following modifications to human participant research until 05/15/2017 inclusive:

Type of Review:	Submission Response for IRB Addendum and Modification Request Form
Modification Type:	Updated Timelines, Removed Compensation, Updated Barbara Lee's Role, Updated Sample Size to n-140, and Revised Protocol and Consent.
Project Title:	Assessment of Neurologic Symptomatology Using an Interactive Physical-Virtual Head with Touch.
Investigator:	Salam Daher
IRB Number:	SBE-15-11364
Funding Agency:	
Grant Title:	
Research ID:	N/A

The scientific merit of the research was considered during the IRB review. The Continuing Review Application must be submitted 30days prior to the expiration date for studies that were previously expedited, and 60 days prior to the expiration date for research that was previously reviewed at a convened meeting. Do not make changes to the study (i.e., protocol, methodology, consent form, personnel, site, etc.) before obtaining IRB approval. A Modification Form **cannot** be used to extend the approval period of a study. All forms may be completed and submitted online at https://iris.research.ucf.edu .

If continuing review approval is not granted before the expiration date of 05/15/2017, approval of this research expires on that date. When you have completed your research, please submit a Study Closure request in iRIS so that IRB records will be accurate.

Use of the approved, stamped consent document(s) is required. The new form supersedes all previous versions, which are now invalid for further use. Only approved investigators (or other approved key study personnel) may solicit consent for research participation. Participants or their representatives must receive a copy of the consent form(s).

All data, including signed consent forms if applicable, must be retained and secured per protocol for a minimum of five years (six if HIPAA applies) past the completion of this research. Any links to the identification of participants should be maintained and secured per protocol. Additional requirements may be imposed by your funding agency, your department, or other entities. Access to data is limited to authorized individuals listed as key study personnel.

In the conduct of this research, you are responsible to follow the requirements of the Investigator Manual.

University of Central Florida Institutional Review Board
Office of Research & Commercialization
12201 Research Parkway, Suite 501
Orlando, Florida 32826-3246
Telephone: 407-823-2901 or 407-882-2276
www.research.ucf.edu/compliance/irb.html

Approval of Human Research

From: **UCF Institutional Review Board #1**
FWA00000351, IRB00001138

To: **Jason Hochreiter** and Co-PIs: **Gerd Bruder, Greg Welch, Salam Daher**

Date: **August 25, 2017**

Dear Researcher:

On 08/25/2017 the IRB approved the following human participant research until 08/24/2018 inclusive:

Type of Review:	UCF Initial Review Submission Form Expedited Review
Project Title:	Exploring Effects of Physical and Virtual Representation of 3D Objects on Touch Interactions
Investigator:	Jason Hochreiter
IRB Number:	SBE-17-13377
Funding Agency:	National Science Foundation
Grant Title:	CHS: Medium: Physical-Virtual Patient Bed for Healthcare Training and Assessment
Research ID:	1059356

The scientific merit of the research was considered during the IRB review. The Continuing Review Application must be submitted 30days prior to the expiration date for studies that were previously expedited, and 60 days prior to the expiration date for research that was previously reviewed at a convened meeting. Do not make changes to the study (i.e., protocol, methodology, consent form, personnel, site, etc.) before obtaining IRB approval. A Modification Form **cannot** be used to extend the approval period of a study. All forms may be completed and submitted online at https://iris.research.ucf.edu .

If continuing review approval is not granted before the expiration date of 08/24/2018, approval of this research expires on that date. When you have completed your research, please submit a Study Closure request in iRIS so that IRB records will be accurate.

Use of the approved, stamped consent document(s) is required. The new form supersedes all previous versions, which are now invalid for further use. Only approved investigators (or other approved key study personnel) may solicit consent for research participation. Participants or their representatives must receive a copy of the consent form(s).

All data, including signed consent forms if applicable, must be retained and secured per protocol for a minimum of five years (six if HIPAA applies) past the completion of this research. Any links to the identification of participants should be maintained and secured per protocol. Additional requirements may be imposed by your funding agency, your department, or other entities. Access to data is limited to authorized individuals listed as key study personnel.

In the conduct of this research, you are responsible to follow the requirements of the Investigator Manual.

On behalf of Sophia Dziegielewski, Ph.D., L.C.S.W., UCF IRB Chair, this letter is signed by:

University of Central Florida Institutional Review Board
Office of Research & Commercialization
12201 Research Parkway, Suite 501
Orlando, Florida 32826-3246
Telephone: 407-823-2901 or 407-882-2276
www.research.ucf.edu/compliance/irb.html

Determination of Exempt Human Research

From: **UCF Institutional Review Board #1**
FWA00000351, IRB00001138

To: **Salam Daher** and Co-PIs: **Desiree A. Diaz, Gerd Bruder, Gregory Welch, Laura Gonzalez, Mindi A. Anderson**

Date: **November 07, 2017**

Dear Researcher:

On 11/07/2017, the IRB reviewed the following activity as human participant research that is exempt from regulation:

Type of Review:	Exempt Determination
Project Title:	Effects of Manipulating Cues Related to Assessment of Simulated Patients
Investigator:	Salam Daher
IRB Number:	SBE-17-13399
Funding Agency:	National Science Foundation
Grant Title:	Physical-Virtual Patient Bed for Healthcare Training and Assessment
Research ID:	1059356

This determination applies only to the activities described in the IRB submission and does not apply should any changes be made. If changes are made and there are questions about whether these changes affect the exempt status of the human research, please contact the IRB. When you have completed your research, please submit a Study Closure request in iRIS so that IRB records will be accurate.

In the conduct of this research, you are responsible to follow the requirements of the Investigator Manual.

This letter is signed by:

Signature applied by Gillian Morien on 11/07/2017 03:51:31 PM EST

Designated Reviewer

University of Central Florida Institutional Review Board
Office of Research & Commercialization
12201 Research Parkway, Suite 501
Orlando, Florida 32826-3246
Telephone: 407-823-2901 or 407-882-2276
www.research.ucf.edu/compliance/irb.html

Approval of Human Research

From: **UCF Institutional Review Board #1**
FWA00000351, IRB00001138

To: **Greg Welch** and Co-PIs: **Andrew Brian Raij, Charles E. Hughes**

Date: **July 13, 2015**

Dear Researcher:

On 07/13/2015, the IRB approved the following human participant research until 07/12/2016 inclusive:

Type of Review:	UCF Initial Review Submission Form Expedited Review Category #4, 6, and 7 This approval includes a Waiver of Written Documentation of Consent
Project Title:	The Effects of Realism Cues on Interactions with Human Surrogates
Investigator:	Greg Welch
IRB Number:	SBE-15-11405
Funding Agency:	Office of Naval Research
Grant Title:	Human-Surrogate Interaction
Research ID:	1056687

The scientific merit of the research was considered during the IRB review. The Continuing Review Application must be submitted 30days prior to the expiration date for studies that were previously expedited, and 60 days prior to the expiration date for research that was previously reviewed at a convened meeting. Do not make changes to the study (i.e., protocol, methodology, consent form, personnel, site, etc.) before obtaining IRB approval. A Modification Form **cannot** be used to extend the approval period of a study. All forms may be completed and submitted online at https://iris.research.ucf.edu .

If continuing review approval is not granted before the expiration date of 07/12/2016, approval of this research expires on that date. When you have completed your research, please submit a Study Closure request in iRIS so that IRB records will be accurate.

Use of the approved, stamped consent document(s) is required. The new form supersedes all previous versions, which are now invalid for further use. Only approved investigators (or other approved key study personnel) may solicit consent for research participation. Participants or their representatives must receive a copy of the consent form(s).

All data, including signed consent forms if applicable, must be retained and secured per protocol for a minimum of five years (six if HIPAA applies) past the completion of this research. Any links to the identification of participants should be maintained and secured per protocol. Additional requirements may be imposed by your funding agency, your department, or other entities. Access to data is limited to authorized individuals listed as key study personnel.

In the conduct of this research, you are responsible to follow the requirements of the Investigator Manual.

On behalf of Sophia Dziegielewski, Ph.D., L.C.S.W., UCF IRB Chair, this letter is signed by:

APPENDIX G: APPENDIX G: PERMISSION LETTERS

Consent to Publish

Lecture Notes in Computer Science

Springer

Title of the Book or Conference Name: . . . .

Volume Editor(s): . . . .

Title of the Contribution: . . . .

Author(s) Name(s): . . . .

Corresponding Author's Name, Address, Affiliation and Email: . . . .

. . . .

. . . .

When Author is more than one person the expression "Author" as used in this agreement will apply collectively unless otherwise indicated.

§ 1 Rights Granted

Author hereby grants and assigns to Springer International Publishing AG, Cham (hereinafter called Springer) the exclusive, sole, permanent, world-wide, transferable, sub-licensable and unlimited right to reproduce, publish, distribute, transmit, make available or otherwise communicate to the public, translate, publicly perform, archive, store, lease or lend and sell the Contribution or parts thereof individually or together with other works in any language, in all revisions and versions (including soft cover, book club and collected editions, anthologies, advance printing, reprints or print to order, microfilm editions, audiograms and videograms), in all forms and media of expression including in electronic form (including offline and online use, push or pull technologies, use in databases and networks for display, print and storing on any and all stationary or portable end-user devices, e.g. text readers, audio, video or interactive devices, and for use in multimedia or interactive versions as well as for the display or transmission of the Contribution or parts thereof in data networks or seach engines), in whole, in part or in abridged form, in each case as now known or developed in the future, including the right to grant further time-limited or permanent rights. For the purposes of use in electronic forms, Springer may adjust the Contribution to the respective form of use and include links or otherwise combine it with other works. For the avoidance of doubt, Springer has the right to permit others to use individual illustrations and may use the Contribution for advertising purposes.

The copyright of the Contribution will be held in the name of Springer. Springer may take, either in its own name or in that of copyright holder, any necessary steps to protect these rights against infringement by third parties. It will have the copyright notice inserted into all editions of the Contribution according to the provisions of the Universal Copyright Convention (UCC) and dutifully take care of all formalities in this connection in the name of the copyright holder.

§ 2 Regulations for Authors under Special Copyright Law

The parties acknowledge that there may be no basis for claim of copyright in the United States to a Contribution prepared by an officer or employee of the United States government as part of that person's official duties. If the Contribution was performed under a United States government contract, but Author is not a United States government employee, Springer grants the United States government royalty-free permission to reproduce all or part of the Contribution and to authorize others to do so for United States government purposes.

If the Contribution was prepared or published by or under the direction or control of Her Majesty (i.e., the constitutional monarch of the Commonwealth realm) or any Crown government department, the copyright in the Contribution shall, subject to any agreement with Author, belong to Her Majesty.

If the Contribution was created by an employee of the European Union or the European Atomic Energy Community (EU/Euratom) in the performance of their duties, the regulation 31/EEC, 11/EAEC (Staff Regulations) applies, and copyright in the Contribution shall, subject to the Publication Framework Agreement (EC Plug), belong to the European Union or the European Atomic Energy Community.

If Author is an officer or employee of the United States government, of the Crown, or of EU/Euratom, reference will be made to this status on the signature page.

§ 3 Rights Retained by Author

Author retains, in addition to uses permitted by law, the right to communicate the content of the Contribution to other scientists, to share the Contribution with them in manuscript form, to perform or present the Contribution or to use the content for non-commercial internal and educational purposes, provided the Springer publication is mentioned as the

16.02.2016
10:50

ACM Copyright and Audio/Video Release

Title of the Work: Physical-Virtual Agents for Healthcare Simulation
Submission ID:27
Author/Presenter(s): Salam Daher;Jason Hochreiter;Nahal Norouzi;Laura Gonzalez;Gerd Bruder;Gregory Welch
Type of material: Full Paper

Publication and/or Conference Name: IVA '18: International Conference on Intelligent Virtual Agents Proceedings

I. Copyright Transfer, Reserved Rights and Permitted Uses

* Your Copyright Transfer is conditional upon you agreeing to the terms set out below.

Copyright to the Work and to any supplemental files integral to the Work which are submitted with it for review and publication such as an extended proof, a PowerPoint outline, or appendices that may exceed a printed page limit, (including without limitation, the right to publish the Work in whole or in part in any and all forms of media, now or hereafter known) is hereby transferred to the ACM (for Government work, to the extent transferable) effective as of the date of this agreement, on the understanding that the Work has been accepted for publication by ACM.

Reserved Rights and Permitted Uses

(a) All rights and permissions the author has not granted to ACM are reserved to the Owner, including all other proprietary rights such as patent or trademark rights.

(b) Furthermore, notwithstanding the exclusive rights the Owner has granted to ACM, Owner shall have the right to do the following:

(i) Reuse any portion of the Work, without fee, in any future works written or edited by the Author, including books, lectures and presentations in any and all media.

(ii) Create a "Major Revision" which is wholly owned by the author

(iii) Post the Accepted Version of the Work on (1) the Author's home page, (2) the Owner's institutional repository, (3) any repository legally mandated by an agency funding the research on which the Work is based, and (4) any non-commercial repository or aggregation that does not duplicate ACM tables of contents, i.e., whose patterns of links do not substantially duplicate an ACM-copyrighted volume or issue. Non-commercial repositories are here understood as repositories owned by non-profit organizations that do not charge a fee for accessing deposited articles and that do not sell advertising or otherwise profit from serving articles.

(iv) Post an "Author-Izer" link enabling free downloads of the Version of Record in the ACM Digital Library on (1) the Author's home page or (2) the Owner's institutional repository;

(v) Prior to commencement of the ACM peer review process, post the version of the Work as submitted to ACM ("Submitted Version" or any earlier versions) to non-peer reviewed servers;

(vi) Make free distributions of the final published Version of Record internally to the Owner's employees, if applicable;

(vii) Make free distributions of the published Version of Record for Classroom and Personal Use;

(viii) Bundle the Work in any of Owner's software distributions; and

Salam Daher

From: Scott Stachowiak <Scott.Stachowiak@russopartnersllc.com>
Sent: Tuesday, October 9, 2:45 2018 PM
To: Salam Daher
Cc: ChristopherJ@Gaumard.com; FathiaL@gaumard.com
Subject: RE: Media Asset Request Form

Hi Salam,
Thanks for the snapshot. Use the images with the photo credits.
Good luck,
Scott

From: Salam Daher <salam@Knights.ucf.edu>
Sent: Tuesday, October 9, 2018 2:38 PM
To: Scott Stachowiak <Scott.Stachowiak@russopartnersllc.com>
Cc: ChristopherJ@Gaumard.com; FathiaL@gaumard.com
Subject: RE: Media Asset Request Form

Dear Scott,
Thank you for your quick reply.
I attached a snapshot of the images I plan to use in my dissertation, they will all be credited Gaumard Scientific.
Please let me know if these are OK for me to use in my dissertation.
Thank you
Salam Daher

From: Scott Stachowiak <Scott.Stachowiak@russopartnersllc.com>
Sent: Tuesday, October 9, 2018 1:54 PM
To: Salam Daher <salam@Knights.ucf.edu>
Cc: ChristopherJ@Gaumard.com; FathiaL@gaumard.com
Subject: RE: Media Asset Request Form

Hi Salam,
Congratulations on your dissertation idea. If you could make a list of images that you're planning to use, that'd be helpful for us.
They should be credited to Gaumard Scientific.

Thanks for reaching out.

Scott Stachowiak
Vice President
RussoPartners/LLC
Office (646) 942-5630 Mobile (646) 300-3590
scott.stachowiak@russopartnersllc.com

Salam Daher

From: Brooke Rowe <Brooke@shadowhealth.com>
Sent: Tuesday, November 13, 10:24 2018 AM
To: Salam Daher
Cc: Shadow Health Communications
Subject: Re: permission to use image

Good morning, Salam!

You absolutely have our permission to use this image in your research. We'd be happy to provide other images for you to use, as well.

We are honored that you are including us, and would love to see the end result. Please feel free to send any materials our way!

Sincerely,

Brooke Rowe
Product Owner
800-860-3241
ShadowHealth.com

On Mon, Nov 12, 2018 at 5:57 PM Salam Daher <salam@knights.ucf.edu> wrote:

> My name is Salam Daher, I am a PhD student in Modeling and Simulation at University of Central Florida. As part of my dissertation I am working on a review about humans models and simulators throughout history. Can I have your permission to use an image from Shadow Health in my dissertation?
>
> I attached the image I would like to have permission to use.
>
> Thank you
>
> Salam

Salam Daher

From: Salam Daher
Sent: Tuesday, November 13, 5:29 2018 PM
To: 'Doug Miller'
Subject: RE: permission to use image

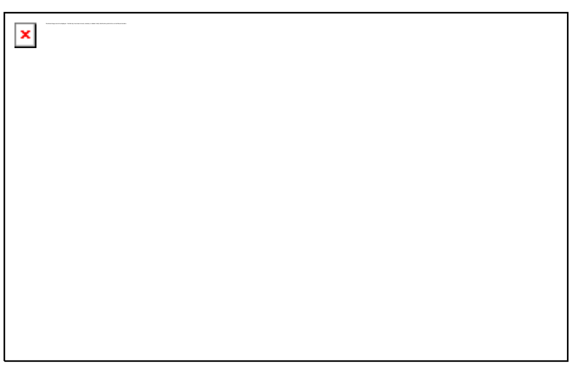
Thank you so much.
-Salam

From: Doug Miller <doug.miller@kaplan.com>
Sent: Tuesday, November 13, 2018 5:26 PM
To: Salam Daher <salam@Knights.ucf.edu>
Subject: Re: permission to use image

Hello Salam,
Thank you for your email and request to use an image of the i-Human program. Yes, you have our approval to use the screenshot you have sent in your email. This approval is for the image only, please do NOT include any links to the assignment or the case website. Please feel free to use our public website, www.i-human.com, for any references.
Best,
Doug Miller
Customer & Sales Support Manager | **Need support?** Click the following link to submit a support ticket or call 503-451-6292.
i-Human Patients, a part of Kaplan Test Prep
doug.miller@kaplan.com
www.i-human.com

On Mon, Nov 12, 2018 at 5:52 PM Salam Daher <salam@knights.ucf.edu> wrote:

> Hello,
>
> My name is Salam Daher, I am a PhD student in Modeling and Simulation at University of Central Florida. As part of my dissertation I am working on a review about humans models and simulators throughout history. Can I have your permission to use an image from iHuman software in my dissertation.
>
> Something like the image below:



LIST OF REFERENCES

[1] Sandra G Hart and Lowell E Staveland. Development of NASA-TLX (task load index): Results of empirical and theoretical research. In *Advances in psychology*, volume 52, pages 139–183. Elsevier, 1988.

[2] Kathleen R Rosen. The history of medical simulation. *Journal of critical care*, 23(2):157–166, 2008.

[3] Akira Nishisaki, Ron Keren, and Vinay Nadkarni. Does simulation improve patient safety? self-efficacy, competence, operational performance, and patient safety. *Anesthesiology Clinics*, 25(2):225 – 236, 2007.

[4] Jeffrey B Cooper. The role of simulation in patient safety. *Simulators in critical care and beyond*, pages 20–24, 2004.

[5] Malcolm Gladwell. *Outliers: The story of success*. Hachette UK, 2008.

[6] Society of simulation in healthcare. `http://www.ssih.org/About-SSH/About-Simulation`.

[7] George E Miller. The assessment of clinical skills/competence/performance. *Academic medicine*, 65(9):S63–7, 1990.

[8] K. Anders Ericsson, Ralf Thomas Krampe, and Clemens Tesch-Römer. The role of deliberate practice in the acquisition of expert performance. *Psychological review*, 100(3):363–406, 1993.

[9] Jimmy Kyaw Tun, Guillaume Alinier, Jessica Tang, and Roger L Kneebone. Redefining simulation fidelity for healthcare education. *Simulation & Gaming*, 46(2):159–174, 2015.

[10] National Council of State Boards of Nursing, National Council of State Boards of Nursing, et al. Report of findings from the effect of high-fidelity simulation on nursing students' knowledge and performance: a pilot study. *NCSBN Research Brief*, 40, 2009.

[11] Jin Lee and Pok Ja Oh. Effects of the use of high-fidelity human simulation in nursing education: A meta-analysis. *Journal of Nursing Education*, 54(9):501–507, 2015.

[12] Kathie Lasater. High-fidelity simulation and the development of clinical judgment: Students' experiences. *Journal of Nursing Education*, 46(6), 2007.

[13] S Barry Issenberg, William C Mcgaghie, Emil R Petrusa, David Lee Gordon, and Ross J Scalese. Features and uses of high-fidelity medical simulations that lead to effective learning: a beme systematic review. *Medical teacher*, 27(1):10–28, 2005.

[14] Marina Krcmar, Kirstie Farrar, and Rory McGloin. The effects of video game realism on attention, retention and aggressive outcomes. *Computers in Human Behavior*, 27(1):432–439, 2011.

[15] Peter Dieckmann, David Gaba, and Marcus Rall. Deepening the theoretical foundations of patient simulation as social practice. *Simulation in Healthcare*, 2(3):183–193, 2007.

[16] Masahiro Mori. The uncanny valley. *Energy*, 7(4):33–35, 1970.

[17] David Swapp, Vijay Pawar, and Céline Loscos. Interaction with co-located haptic feedback in virtual reality. *Virtual Reality*, 10(1):24–30, 2006.

[18] R Viciana-Abad and A Reyes-Lecuona. Effects of co-location and crossmodal interaction between haptic, auditory and visual cues in presence. haptics: Perception, devices and scenarios, lncs 5024: 832–837. *Berlin: Springer. doi*, 10(1007):978–3, 2008.

[19] Salam Daher, Kangsoo Kim, Myungho Lee, Ryan Schubert, Gerd Bruder, Jeremy Bailenson, and Greg Welch. Effects of social priming on social presence with intelligent virtual agents. In *International Conference on Intelligent Virtual Agents*, pages 87–100. Springer, 2017.

[20] Salam Daher, Kangsoo Kim, Myungho Lee, Gerd Bruder, Ryan Schubert, Jeremy Bailenson, and Gregory F Welch. Can social presence be contagious? effects of social presence priming on interaction with virtual humans. In *3D User Interfaces (3DUI), 2017 IEEE Symposium on*, pages 201–202. IEEE, 2017.

[21] Kangsoo Kim, Arjun Nagendran, Jeremy Bailenson, and Greg Welch. Expectancy violations related to a virtual human's joint gaze behavior in real-virtual human interactions. In *28th Annual Conference on Computer Animation and Social Agents*, 2015.

[22] Adam I Levine, Samuel DeMaria Jr, Andrew D Schwartz, and Alan J Sim. *The comprehensive textbook of healthcare simulation*. Springer Science and Business Media, 2013.

[23] Howard S Barrows. An overview of the uses of standardized patients for teaching and evaluating clinical skills. *Academic Medicine*, 68(6):443–51, 1993.

[24] Paul G Ramsey, MPH J Randall Curtis MD, Douglas S Paauw, and Marjorie D Wenrich. History-taking and preventive medicine skills among primary care physicians: an assessment using standardized patients. *The American journal of medicine*, 104(2):152–158, 1998.

[25] Graceanne Adamo. Simulated and standardized patients in osces: achievements and challenges 1992-2003. *Medical teacher*, 25(3):262–270, 2003.

[26] Tsuen-Chiuan Tsai. Using children as standardised patients for assessing clinical competence in paediatrics. *Archives of disease in childhood*, 89(12):1117–1120, 2004.

[27] Leonard H Glantz. Conducting research with children: Legal and ethical issues. *Journal of the American Academy of Child & Adolescent Psychiatry*, 35(10):1283–1291, 1996.

[28] Harry Owen. Early use of simulation in medical education. *Simulation in Healthcare*, 7(2):102–116, 2012.

[29] Laerdal. `http://www.laerdal.com/us/doc/367/History`, 2015.

[30] David M Gaba and Abe DeAnda. A comprehensive anesthesia simulation environment: re-creating the operating room for research and training. *Anesthesiology*, 69(3):387–394, 1988.

[31] JB Cooper and VR Taqueti. A brief history of the development of mannequin simulators for clinical education and training. *Quality and Safety in Health Care*, 13(suppl 1):i11–i18, 2004.

[32] Salam Daher, Jason Hochreiter, Nahal Norouzi, Laura Gonzalez, Gerd Bruder, and Greg Welch. Physical-virtual agents for healthcare simulation. In *Proceedings of IVA 2018*, 2018.

[33] Colin F Mackenzie, Ben D Harper, and Yan Xiao. Simulator limitations and their effects on decision-making, 1996.

[34] `https://shadowhealth.com/`, May 2018. Accessed: 05/10/2018.

[35] Virtual sim center. http://virtualsimcenter.clinispace.com/. (Accessed on 10/17/2018).

[36] ihuman. http://www.i-human.com.

[37] Cae vimedix ultrasound simulator | cae healthcare. https://caehealthcare.com/ultrasound-simulation/vimedix/. (Accessed on 10/05/2018).

[38] Jannick Rolland, Larry Davis, Felix Hamza-Lup, Jason Daly, Yonggang Ha, Glenn Martin, Jack Norfleet, Richard Thumann, and Celina Z Imielinska. Development of a training tool for endotracheal intubation: Distributed augmented reality, 2003.

[39] Larry Davis, Felix G Hamza-Lup, Jason Daly, Yonggang Ha, Seth Frolich, Catherine Meyer, Glenn Martin, Jack Norfleet, Kuo-Chi Lin, Celina Imielinska, et al. Application of augmented reality to visualizing anatomical airways. In *Helmet-and Head-Mounted Displays VII*, volume 4711, pages 400–406. International Society for Optics and Photonics, 2002.

[40] Joseph T Samosky, Douglas A Nelson, Bo Wang, Russell Bregman, Andrew Hosmer, Brandon Mikulis, and Robert Weaver. Bodyexplorerar: enhancing a mannequin medical simulator with sensing and projective augmented reality for exploring dynamic anatomy and physiology. In *Proceedings of the Sixth International Conference on Tangible, Embedded and Embodied Interaction*, pages 263–270. ACM, 2012.

[41] LOK Benjamin. Shader lamps virtual patients: The physical manifestation of virtual patients. *Medicine Meets Virtual Reality 19: NextMed*, 173:372, 2012.

[42] Microsoft hololens | the leader in mixed reality technology. https://www.microsoft.com/en-us/hololens. (Accessed on 10/05/2018).

[43] Meta augmented reality. https://www.metavision.com/. (Accessed on 10/05/2018).

[44] J. Hochreiter, S. Daher, G. Bruder, and G. Welch. Cognitive and touch performance effects of mismatched 3D physical and visual perceptions. In *2018 IEEE Conference on Virtual Reality and 3D User Interfaces (VR)*, pages 379–386, March 2018.

[45] Dieter Schmalstieg, Anton Fuhrmann, Gerd Hesina, Zsolt Szalavári, L Miguel Encarnaçao, Michael Gervautz, and Werner Purgathofer. The studierstube augmented reality project. *Presence: Teleoperators and Virtual Environments*, 11(1):33–54, 2002.

[46] DWF Van Krevelen and Ronald Poelman. A survey of augmented reality technologies, applications and limitations. *International journal of virtual reality*, 9(2):1, 2010.

[47] Dieter Schmalstieg and Tobias Hollerer. *Augmented reality: principles and practice*. Addison-Wesley Professional, 2016.

[48] Anton Fuhrmann, Gerd Hesina, François Faure, and Michael Gervautz. Occlusion in collaborative augmented environments. *Computers and Graphics*, 23(6):809–819, 1999.

[49] HR Champion and AG Gallagher. Surgical simulation—a 'good idea whose time has come'. *British journal of surgery*, 90(7):767–768, 2003.

[50] Jane B Paige and Karen H Morin. Simulation fidelity and cueing: A systematic review of the literature. *Clinical Simulation in Nursing*, 9(11):e481–e489, 2013.

[51] JO Lopreiato, D Downing, W Gammon, L Lioce, B Sittner, V Slot, AE Spain, Terminology & Concepts Working Group, et al. Healthcare simulation dictionary, 2016.

[52] Colleen Meakim, Teri Boese, Sharon Decker, Ashley E Franklin, Donna Gloe, Lori Lioce, Carol R Sando, and Jimmie C Borum. Standards of best practice: Simulation standard i: Terminology. *Clinical Simulation in Nursing*, 6(9):S3–S11, 2013.

[53] Carol F Durham. The international nursing association for clinical simulation and learning (inacsl), a community of practice for simulation. *Clinical Simulation in Nursing*, 9(8):e275–e276, 2013.

[54] Chad Harms and Frank Biocca. Internal consistency and reliability of the networked minds measure of social presence. In *Annual International Presence Workshop*, pages 246–251, 2004.

[55] Adam I Levine, Samuel DeMaria Jr, Andrew D Schwartz, and Alan J Sim. *The comprehensive textbook of healthcare simulation*. Springer Science & Business Media, 2013.

[56] Howard S Barrows. An overview of the uses of standardized patients for teaching and evaluating clinical skills. aamc. *Academic Medicine*, 68(6):443–51, 1993.

[57] Adam I Levine, Samuel DeMaria Jr, Andrew D Schwartz, and Alan J Sim. *The comprehensive textbook of healthcare simulation*. Springer Science and Business Media, 2013.

[58] Adam I Levine, Samuel DeMaria Jr, Andrew D Schwartz, and Alan J Sim. *The comprehensive textbook of healthcare simulation*. Springer Science and Business Media, 2013.

[59] JJ Stunt, PH Wulms, GM Kerkhoffs, J Dankelman, CN Van Dijk, and GJM Tuijthof. How valid are commercially available medical simulators? *Advances in medical education and practice*, 5:385, 2014.

[60] Janice C Palaganas, Juli C Maxworthy, Chad A Epps, and Mary E Mancini. *Defining excellence in simulation programs*. Lippincott Williams & Wilkins, 2014.

[61] Gaumard simulators. http://www.gaumard.com/aboutsims/, 2015.

[62] http://www.mcc.edu/ims/lab_features.shtml, 2015.

[63] Cae healthcare birthing simulator | childbirth simulator | maternal fetal simulator. http://www.caehealthcare.com/eng/patient-simulators/maternal-fetal-childbirth-simulator, 2015.

[64] Cae healthcare - product and services: Caesar. http://www.caehealthcare.com/eng/patient-simulators/caesar-simulator, 2015.

[65] National Council of State Boards of Nursing. Report of findings from the effect of high-fidelity simulation on nursing students' knowledge and performance: a pilot study. *NCSBN Research Brief*, 40, 2009.

[66] Jacqueline Vaughn, Michael Lister, and Ryan J Shaw. Piloting augmented reality technology to enhance realism in clinical simulation. *CIN: Computers, Informatics, Nursing*, 34(9):402–405, 2016.

[67] Cynthia L. Foronda, Diana-Lyn Baptiste, and Jessica Ockimey. Research brief: As simple as black and white: The presence of racial diversity in simulation product advertisements. *Clinical Simulation in Nursing*, 13:24 – 27, 2017.

[68] Eric Baldrighi, Nicholas Thayer, Michael Stevens, Sonya Ranson Echols, and Shashank Priya. Design and implementation of the bio-inspired facial expressions for medical mannequin. *International Journal of Social Robotics*, 6(4):555–574, 2014.

[69] Elodie Bonnetain, Jean-Michel Boucheix, Maël Hamet, and Marc Freysz. Benefits of computer screen-based simulation in learning cardiac arrest procedures. *Medical education*, 44(7):716–722, 2010.

[70] http://secondlife.com/, May 2018. Accessed: 05/10/2018.

[71] Andrew B Raij, Kyle Johnsen, Robert F Dickerson, Benjamin C Lok, Marc S Cohen, Margaret Duerson, Rebecca Rainer Pauly, Amy O Stevens, Peggy Wagner, and D Scott Lind. Comparing interpersonal interactions with a virtual human to those with a real human. *IEEE transactions on visualization and computer graphics*, 13(3):443–457, 2007.

[72] Ye Pan and Anthony Steed. A comparison of avatar, video, and robot-mediated interaction on users' trust in expertise. *Frontiers in Robotics and AI*, 3:12, 2016.

[73] Virtual sim center. http://virtualsimcenter.clinispace.com/home.html, 2015.

[74] Shadow health. http://www.shadowhealth.com/, 2015.

[75] ihuman. http://www.i-human.com.

[76] Anatomage. http://www.anatomage.com, 2015.

[77] Atsusi Hirumi, Andrea Kleinsmith, Kyle Johnsen, Stacey Kubovec, Michael Eakins, Kenneth Bogert, Diego J. Rivera-Gutierrez, Ramsamooj Javier Reyes, Benjamin Lok, and Juan Cendan. Advancing virtual patient simulations through design research and interplay: part i: design and development. *Educational Technology Research and Development*, 64(4):763–785, 2016.

[78] Neil Gesundheit, Pauline Brutlag, Patricia Youngblood, William T. Gunning, Nabil Zary, and Uno Fors. The use of virtual patients to assess the clinical skills and reasoning of medical students: initial insights on student acceptance. *Medical Teacher*, 31(8):739 – 742, 2009.

[79] Maryam Moosaei, Michael J. Gonzales, and Laurel D. Riek. Naturalistic pain synthesis for virtual patients. In *International Conference on Intelligent Virtual Agents*, pages 295–309. Springer, 2014.

[80] Maryam Moosaei, Sumit K. Das, Dan O. Popa, and Laurel D. Riek. Using facially expressive robots to calibrate clinical pain perception. *ACM/IEEE International Conference on Human-Robot Interaction*, pages 32 – 41, 2017.

[81] Margaret Bearman. Is virtual the same as real? medical students' experiences of a virtual patient. *Academic Medicine: Journal Of The Association Of American Medical Colleges*, 78(5):538 – 545, 2003.

[82] John Wiecha, Robin Heyden, Elliot Sternthal, and Mario Merialdi. Learning in a virtual world: experience with using second life for medical education. *Journal of medical Internet research*, 12(1):e1, 2010.

[83] Dohyun Ahn, Youngnam Seo, Minkyung Kim, Joung Huem Kwon, Younbo Jung, Jungsun Ahn, and Doohwang Lee. The effects of actual human size display and stereoscopic presentation on users' sense of being together with and of psychological immersion in a virtual character. *Cyberpsychology, Behavior, and Social Networking*, 17(7):483 – 487, 2014.

[84] Feng Zhou, Henry Been-Lirn Duh, and Mark Billinghurst. Trends in augmented reality tracking, interaction and display: A review of ten years of ismar. In *Proceedings of the 7th IEEE/ACM International Symposium on Mixed and Augmented Reality*, pages 193–202. IEEE Computer Society, 2008.

[85] Paul Milgram and Fumio Kishino. A taxonomy of mixed reality visual displays. *IEICE TRANSACTIONS on Information and Systems*, 77(12):1321–1329, 1994.

[86] Ramesh Raskar, Greg Welch, Kok-Lim Low, and Deepak Bandyopadhyay. Shader lamps: Animating real objects with image-based illumination. In *Rendering Techniques 2001*, pages 89–102. Springer, 2001.

[87] Ramesh Raskar, Greg Welch, and Henry Fuchs. Spatially augmented reality. In *First IEEE Workshop on Augmented Reality (IWAR'98)*, pages 11–20. Citeseer, 1998.

[88] Diego J Rivera-Gutierrez, Greg Welch, Peter Lincoln, Mary C Whitton, Juan Cendan, David A Chesnutt, Henry Fuchs, and Benjamin Lok. Shader lamps virtual patients: the physical manifestation of virtual patients. In *MMVR*, pages 372–378, 2012.

[89] A. Kotranza and Ieee. *Real-Time In-Situ Visual Feedback Of Task Performance In Mixed Environments For Learning Joint Psychomotor-Cognitive Tasks*. IEEE, 2015.

[90] Melissa. Locker. "this robotic butt could save your life. *TimeCom November 1 Business Source Premier EBSCOhost accessed November 11*, 14 SRC - GoogleScholar, 2015.

[91] Sebastian Ullrich and Torsten Kuhlen. Haptic palpation for medical simulation in virtual environments. *IEEE Transactions on Visualization and Computer Graphics, Visualization and Computer Graphics, IEEE Transactions on, IEEE Trans. Visual. Comput. Graphics*, 18(4):617, 2012.

[92] Mark J Garside, Matthew P Rudd, and Christopher I Price. Stroke and tia assessment training: A new simulation-based approach to teaching acute stroke assessment. *Simulation in Healthcare*, 7(2):117–122, 2012.

[93] Diego Rivera-Gutierrez. "shader lamps virtual patients: The physical manifestation of virtual patients. *Studies In Health Technology And Informatics 173 MEDLINE Web 11 Nov*, pages 372–378 SRC – GoogleScholar, 2015.

[94] Christoph Bichlmeier, Ben Ockert, Oliver Kutter, Mohammad Rustaee, Sandro Michael Heining, and Nassir Navab. The visible Korean human phantom: Realistic test & development environments for medical augmented reality. In *International Workshop on Augmented environments for Medical Imaging including Augmented Reality in Computer-aided Surgery (AMI-ARCS 2008), New York, USA*, 2008.

[95] Aaron Kotranza and Benjamin Lok. Virtual human+ tangible interface= mixed reality human an initial exploration with a virtual breast exam patient. In *Virtual Reality Conference, 2008. VR'08. IEEE*, pages 99–106. IEEE, 2008.

[96] Borko Furht. *Handbook of augmented reality*. Springer Science & Business Media, 2011.

[97] Chen Lin, Candelario Laserna, D. Scott Lind, Chandni Kalaria, Tia Aryal, Benjamin Lok, Kyle Johnsen, Aaron Kotranza, Juan Cendan, and J. Ned Pruitt. User response to the simulation of a virtual patient with cranial nerve injury. *Bio-Algorithms and Medical Systems*, 8(1):1, 2012.

[98] Teresa R. Johnson, Rebecca Lyons, Regis Kopper, Kyle J. Johnsen, Benjamin C. Lok, and Juan C. Cendan. Virtual patient simulations and optimal social learning context: a replication of an aptitude-treatment interaction effect. *Medical Teacher*, 36(6):486 – 494, 2014.

[99] Jason Hochreiter, Salam Daher, Arjun Nagendran, Laura Gonzalez, and Greg Welch. Touch sensing on non-parametric rear-projection surfaces: A physical-virtual head for hands-on healthcare training. In *Virtual Reality (VR), 2015 IEEE*, pages 69–74. IEEE, 2015.

[100] Jason Hochreiter, Salam Daher, Arjun Nagendran, Laura Gonzalez, and Greg Welch. Optical touch sensing on nonparametric rear-projection surfaces for interactive physical-virtual experiences. *PRESENCE: Teleoperators and Virtual Environments*, 25(1):33–46, 2016.

[101] Norman E Lane and Earl A Alluisi. Fidelity and validity in distributed interactive simulation: Questions and answers. Technical report, DTIC Document, 1992.

[102] NJ Maran and RJ2003 Glavin. Low-to high-fidelity simulation–a continuum of medical education? *Medical education*, 37(s1):22–28, 2003.

[103] Robert B Miller. Psychological considerations in the design of training equipment. Technical report, DTIC Document, 1954.

[104] Albert J Rehmann, Robert D Mitman, and Michael C Reynolds. A handbook of flight simulation fidelity requirements for human factors research. Technical report, DTIC Document, 1995.

[105] S Barry Issenberg and Ross J Scalese. Simulation in health care education. *Perspectives in biology and medicine*, 51(1):31–46, 2008.

[106] Michael T Curtis, Deborah DiazGranados, and Moshe Feldman. Judicious use of simulation technology in continuing medical education. *Journal of Continuing Education in the Health Professions*, 32(4):255–260, 2012.

[107] Rebecca J Sherwood and Gary Francis. The effect of mannequin fidelity on the achievement of learning outcomes for nursing, midwifery and allied healthcare practitioners: Systematic review and meta-analysis. *Nurse education today*, 2018.

[108] Lori Lioce, Colleen H Meakim, Mary K Fey, Joyce Victor Chmil, Bette Mariani, and Guillaume Alinier. Standards of best practice: Simulation standard ix: Simulation design. *Clinical Simulation in Nursing*, 11(6):309–315, 2015.

[109] Hans Rystedt and Björn Sjöblom. Realism, authenticity, and learning in healthcare simulations: rules of relevance and irrelevance as interactive achievements. *Instructional science*,

40(5):785–798, 2012.

[110] Huong Q Dinh, Neff Walker, Larry F Hodges, Chang Song, and Akira Kobayashi. Evaluating the importance of multi-sensory input on memory and the sense of presence in virtual environments. In *Virtual Reality, 1999. Proceedings., IEEE*, pages 222–228. IEEE, 1999.

[111] Paul Rousseau. The physical exam and touch., 2016.

[112] Joon Hao Chuah, Andrew Robb, Casey White, Adam Wendling, Samsun Lampotang, Regis Kopper, and Benjamin Lok. Exploring agent physicality and social presence for medical team training. *Presence: Teleoperators and Virtual Environments*, 22(2):141–170, 2013.

[113] Joon Hao Chuah, Andrew Robb, Casey White, Adam Wendling, Samsun Lampotang, Regis Kopper, and Benjamin Lok. Increasing agent physicality to raise social presence and elicit realistic behavior. In *Virtual Reality Short Papers and Posters (VRW), 2012 IEEE*, pages 19–22. IEEE, 2012.

[114] Aaron Kotranza, Benjamin Lok, Carla M Pugh, and D Scott Lind. Virtual humans that touch back: enhancing nonverbal communication with virtual humans through bidirectional touch. In *Virtual Reality Conference, 2009. VR 2009. IEEE*, pages 175–178. IEEE, 2009.

[115] John G. Bruhn. The doctor's touch: Tactile communication in the doctor-patient relationship, 01 1979.

[116] Mark T Wallace, GE Roberson, W David Hairston, Barry E Stein, J William Vaughan, and Jim A Schirillo. Unifying multisensory signals across time and space. *Experimental Brain Research*, 158(2):252–258, 2004.

[117] Marc O Ernst, Martin S Banks, and Heinrich H Bülthoff. Touch can change visual slant perception. *Nature neuroscience*, 3(1):69–73, 2000.

[118] Marco Congedo, Anatole Lécuyer, and Edouard Gentaz. The influence of spatial delocation on perceptual integration of vision and touch. *Presence: Teleoperators and Virtual Environments*, 15(3):353–357, 2006.

[119] Michael D Proctor and Lillian Campbell-Wynn. Effectiveness, usability, and acceptability of haptic-enabled virtual reality and mannequin modality simulators for surgical cricothyroidotomy. *Military medicine*, 179(3):260–264, 2014.

[120] S2225 Pediatric HAL | Gaumard Scientific. `https://www.gaumard.com/s2225`. Accessed: 2018-09-10.

[121] David L Rodgers. The effect of high-fidelity manikin-based human patient simulation on educational outcomes in advanced cardiovascular life support courses, 2007.

[122] Robert Riener and Matthias Harders. Virtual reality for rehabilitation. In *Virtual Reality in Medicine*, pages 161–180. Springer, 2012.

[123] Colin Ware, Kevin Arthur, and Kellogg S Booth. Fish tank virtual reality. In *Proceedings of the INTERACT'93 and CHI'93 conference on Human factors in computing systems*, pages 37–42. ACM, 1993.

[124] Oliver Bimber and Ramesh Raskar. *Spatial augmented reality: merging real and virtual worlds*. CRC press, 2005.

[125] Carolina Cruz-Neira, Daniel J Sandin, Thomas A DeFanti, Robert V Kenyon, and John C Hart. The cave: audio visual experience automatic virtual environment. *Communications of the ACM*, 35(6):64–73, 1992.

[126] Ronald T Azuma. A survey of augmented reality. *Presence: Teleoperators and virtual environments*, 6(4):355–385, 1997.

[127] Jannick P Rolland and Henry Fuchs. Optical versus video see-through head-mounted displays in medical visualization. *Presence: Teleoperators and Virtual Environments*, 9(3):287–309, 2000.

[128] Robert W Proctor and Trisha Van Zandt. *Human factors in simple and complex systems*. CRC press, 2008.

[129] David E Kieras and David E Meyer. An overview of the epic architecture for cognition and performance with application to human-computer interaction. *Human-computer interaction*, 12(4):391–438, 1997.

[130] John Sweller, Jeroen JG Van Merrienboer, and Fred GWC Paas. Cognitive architecture and instructional design. *Educational psychology review*, 10(3):251–296, 1998.

[131] Jeroen JG Van Merrienboer and John Sweller. Cognitive load theory and complex learning: Recent developments and future directions. *Educational psychology review*, 17(2):147–177, 2005.

[132] Jeroen JG Van Merriënboer and John Sweller. Cognitive load theory in health professional education: design principles and strategies. *Medical education*, 44(1):85–93, 2010.

[133] Jeremy M Wolfe, Keith R Kluender, Dennis M Levi, Linda M Bartoshuk, Rachel S Herz, Roberta L Klatzky, Susan J Lederman, and Daniel M Merfeld. *Sensation & perception*. Sinauer Sunderland, MA, 2006.

[134] James K Bowmaker and HJk Dartnall. Visual pigments of rods and cones in a human retina. *The Journal of physiology*, 298(1):501–511, 1980.

[135] Vasilios G Chouvardas, Amalia N Miliou, Miltiadis K Hatalis, et al. Tactile displays: a short overview and recent developments. In *Proceedings of the 5th International Conference on*

Technology and Automation (ICTA'05), pages 246–251, 2005.

[136] Herve Delingette. Toward realistic soft-tissue modeling in medical simulation. *Proceedings of the IEEE*, 86(3):512–523, 1998.

[137] Miyuki Morioka, Darren J Whitehouse, and Michael J Griffin. Vibrotactile thresholds at the fingertip, volar forearm, large toe, and heel. *Somatosensory & motor research*, 25(2):101–112, 2008.

[138] Thomas Sy, Stéphane Côté, and Richard Saavedra. The contagious leader: Impact of the leader's mood on the mood of group members, group affective tone, and group processes. *Journal of Applied Psychology*, 90(2):295 – 305, 2005.

[139] Greg Willard, Kyonne-Joy Isaac, and Dana R. Carney. Some evidence for the nonverbal contagion of racial bias. *Organizational Behavior and Human Decision Processes*, 128:96 – 107, 2015.

[140] Tanya L. Cahrtrand and John A. Bargh. The chameleon effect: The perception–behavior link and social interaction. *Journal of Personality and Social Psychology*, 76(6):893–910, 1999.

[141] John A. Bargh, Mark Chen, and Lara Burrows. Automaticity of social behavior: Direct effects of trait construct and stereotype activation on action. *Journal of Personality and Social Psychology*, 71(2):230–244, 1996.

[142] Thomas K Srull and Robert S Wyer. The role of category accessibility in the interpretation of information about persons some determinants and implications. *Journal of Personality and Social psychology*, 37(10):1660, 1979.

[143] Ap Dijksterhuis and John A Bargh. The perception-behavior expressway: Automatic effects of social perception on social behavior. *Advances in experimental social psychology*, 33:1–40, 2001.

[144] Chao Qu. *Talking with a Virtual Human: Controlling the Human Experience and Behavior in a Virtual Conversation*. TU Delft, Delft University of Technology, 2014.

[145] Chao Qu, Willem-Paul Brinkman, Pascal Wiggers, and Ingrid Heynderickx. The effect of priming pictures and videos on a question–answer dialog scenario in a virtual environment. *Presence*, 22(2):91–109, 2013.

[146] David Nunez and Edwin Blake. Conceptual priming as a determinant of presence in virtual environments. In *Proceedings of the 2nd international conference on Computer graphics, virtual Reality, visualisation and interaction in Africa*, pages 101–108. ACM, 2003.

[147] Jorge Peña, Jeffrey T Hancock, and Nicholas A Merola. The priming effects of avatars in virtual settings. *Communication Research*, 36(6):838–856, 2009.

[148] Xueni Pan, Marco Gillies, and Mel Slater. Male bodily responses during an interaction with a virtual woman. In *International Workshop on Intelligent Virtual Agents*, pages 89–96. Springer, 2008.

[149] Xueni Pan, Marco Gillies, and Mel Slater. Virtual character personality influences participant attitudes and behavior–an interview with a virtual human character about her social anxiety. *Frontiers in Robotics and AI*, 2:1, 2015.

[150] Tabitha C Peck, Sofia Seinfeld, Salvatore M Aglioti, and Mel Slater. Putting yourself in the skin of a black avatar reduces implicit racial bias. *Consciousness and cognition*, 22(3):779–787, 2013.

[151] Douglas J Herrmann. *Applied cognitive psychology: A textbook*. Psychology Press, 2006.

[152] George A Miller. The magical number seven, plus or minus two: some limits on our capacity for processing information. *Psychological review*, 63(2):81, 1956.

[153] Nelson Cowan. Metatheory of storage capacity limits. *Behavioral and brain sciences*, 24(01):154–176, 2001.

[154] Pierre Barrouillet, Sophie Bernardin, Sophie Portrat, Evie Vergauwe, and Valérie Camos. Time and cognitive load in working memory. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 33(3):570, 2007.

[155] John Sweller. Element interactivity and intrinsic, extraneous, and germane cognitive load. *Educational psychology review*, 22(2):123–138, 2010.

[156] John Sweller, Slava Kalyuga, and Paul Ayres. *Cognitive load theory. [electronic resource]*. Explorations in the learning sciences, instructional systems and performance technologies. New York : Springer, c2011., 2011.

[157] John Sweller. Cognitive load theory, learning difficulty, and instructional design. *Learning and instruction*, 4(4):295–312, 1994.

[158] John Sweller and Paul Chandler. Why some material is difficult to learn. *Cognition and instruction*, 12(3):185–233, 1994.

[159] John Sweller, Paul Chandler, Paul Tierney, and Martin Cooper. Cognitive load as a factor in the structuring of technical material. *Journal of Experimental Psychology: General*, 119(2):176, 1990.

[160] Matthew Lombard and Theresa Ditton. At the heart of it all: The concept of presence. *Journal of Computer-Mediated Communication*, 3(2):0–0, 1997.

[161] Erving Goffman. *Behavior in public places; notes on the social organization of gatherings.* New York, Free Press of Glencoe 1969, 1969.

[162] Jim Blascovich. Social influence within immersive virtual environments. In *The social life of avatars*, pages 127–145. Springer, 2002.

[163] Jim Blascovich, Jack Loomis, Andrew C. Beall, Kimberly R. Swinth, Crystal L. Hoyt, and Jeremy N. Bailenson. Immersive virtual environment technology as a methodological tool for social psychology. *Psychological Inquiry*, 13(2):103–124, 2002.

[164] Chad Harms and Frank Biocca. Internal consistency and reliability of the networked minds measure of social presence, 2004.

[165] Kwan Min Lee, Younbo Jung, Jaywoo Kim, and Sang Ryong Kim. Are physically embodied social agents better than disembodied social agents?: The effects of physical embodiment, tactile interaction, and people's loneliness in human–robot interaction. *International Journal of Human-Computer Studies*, 64(10):962–973, 2006.

[166] Kangsoo Kim, Arjun Nagendran, Jeremy N. Bailenson, Andrew Raij, Gerd Bruder, Myungho Lee, Ryan Schubert, Xin Yan, and Gregory F. Welch. A large-scale study of surrogate physicality and gesturing on human-surrogate interactions in a public space. *Frontiers in Robotics and AI*, 4:1–20, 2017.

[167] Benjamin Lok, Joon Hao Chuah, Andrew Robb, Andrew Cordar, Samsun Lampotang, Adam Wendling, and Casey White. Mixed-reality humans for team training. *IEEE Computer Graphics and Applications*, 34(3):72–75, 2014.

[168] Jeremy N. Bailenson, Andrew C. Beall, and Jim Blascovich. Gaze and task performance in shared virtual environments. *The Journal of Visualization and Computer Animation*, 13(5):313–320, December 2002.

[169] Jeremy N Bailenson, Nick Yee, Dan Merget, and Ralph Schroeder. The Effect of Behavioral Realism and Form Realism of Real-Time Avatar Faces on Verbal Disclosure, Nonverbal Disclosure, Emotion Recognition, and Copresence in Dyadic Interaction. *Presence: Teleoperators and Virtual Environments*, 15(4):359–372, 2006.

[170] Maia Garau, Mel Slater, David-Paul Pertaub, and Sharif Razzaque. The Responses of People to Virtual Humans in an Immersive Virtual Environment. *Presence Teleoperators and Virtual Environments*, 14(1):104–116, 2005.

[171] Saniye Tugba Bulu. Place presence, social presence, co-presence, and satisfaction in virtual worlds. *Computers & Education*, 58(1):154–161, jan 2012.

[172] A. Robb, A. Cordar, S. Lampotang, C. White, A. Wendling, and B. Lok. Teaming up with virtual humans: How other people change our perceptions of and behavior with virtual teammates. *Visualization and Computer Graphics, IEEE Transactions on*, 21(4):511–519, April 2015.

[173] Jeremy N. Bailenson, Andrew C. Beall, Jack Loomis, Jim Blascovich, and Matthew Turk. Transformed Social Interaction, Augmented Gaze, and Social Influence in Immersive Virtual Environments. *Human Communication Research*, 31(4):511–537, 2005.

[174] Lixing Huang, Louis-Philippe Morency, and Jonathan Gratch. Virtual Rapport 2.0. In HannesHögni Vilhjálmsson, Stefan Kopp, Stacy Marsella, and KristinnR. Thórisson, editors, *Intelligent Virtual Agents (Lecture Notes in Artificial Intelligence)*, volume 6895 of *Lecture Notes in Computer Science*, pages 68–79. Springer Berlin Heidelberg, 2011.

[175] Fred Paas, Alexander Renkl, and John Sweller. Cognitive load theory and instructional design: Recent developments. *Educational psychologist*, 38(1):1–4, 2003.

[176] Kristin L Fraser, Paul Ayres, and John Sweller. Cognitive load theory for the design of medical simulations. *Simulation in Healthcare*, 10(5):295–307, 2015.

[177] Gabriel B Reedy. Using cognitive load theory to inform simulation design and practice. *Clinical Simulation in Nursing*, 11(8):355–360, 2015.

[178] Marie-Laurence Tremblay, Alexandre Lafleur, Jimmie Leppink, and Diana HJM Dolmans. The simulated clinical environment: Cognitive and emotional impact among undergraduates. *Medical teacher*, 39(2):181–187, 2017.

[179] Choon Looi Bong, Kristin Fraser, and Denis Oriot. Cognitive load and stress in simulation. In *Comprehensive healthcare simulation: pediatrics*, pages 3–17. Springer, 2016.

[180] Kristin Fraser, Irene Ma, Elise Teteris, Heather Baxter, Bruce Wright, and Kevin McLaughlin. Emotion, cognitive load and learning outcomes during simulation training. *Medical Education*, 46(11):1055 – 1062, 2012.

[181] Tuulikki Keskitalo and Heli Ruokamo. Students' emotions in simulation-based medical education. *Journal of Interactive Learning Research*, 28(2):149–159, 2017.

[182] Tze Wei Liew and Tan Su-Mae. The effects of positive and negative mood on cognition and motivation in multimedia learning environment. *Journal of Educational Technology & Society*, 19(2):104, 2016.

[183] Sook Jung Kang and Hae Young Min. Psychological safety in nursing simulation. *Nurse educator*, 2018.

[184] David Watson, Lee Anna Clark, Auke Tellegen, D. Watson, L. A. Clark, and A. Tellegen. Positive and negative affect schedule. *Journal of Personality and Social Psychology*, 54(6):1063 – 1070, 1988.

[185] Tamara Power, Claudia Virdun, Haidee White, Carolyn Hayes, Nicola Parker, Michelle Kelly, Rebecca Disler, and Amanda Cottle. Plastic with personality: Increasing student engagement with manikins. *Nurse education today*, 38:126–131, 2016.

[186] Brennen W Mills, Owen B-J Carter, Cobie J Rudd, Louise A Claxton, Nathan P Ross, and Natalie A Strobel. Effects of low-versus high-fidelity simulations on the cognitive burden and performance of entry-level paramedicine students: a mixed-methods comparison trial using eye-tracking, continuous heart rate, difficulty rating scales, video observation and interviews. *Simulation in Healthcare*, 11(1):10–18, 2016.

[187] Kenneth C Herbst, Lowell Gaertner, and Chester a Insko. My head says yes but my heart says no: cognitive and affective attraction as a function of similarity to the ideal self. *Journal of personality and social psychology*, 84(6):1206–1219, 2003.

[188] Frank H Wilhelm and Walton T Roth. Ambulatory assessment of clinical anxiety, 1996.

[189] Jonathan Lazar, Jinjuan Heidi Feng, and Harry Hochheiser. *Research methods in human-computer interaction*. Morgan Kaufmann, 2017.

[190] Michael Voit and Rainer Stiefelhagen. Deducing the visual focus of attention from head pose estimation in dynamic multi-view meeting scenarios. In *Proceedings of the 10th international conference on Multimodal interfaces*, pages 173–180. ACM, 2008.

[191] Sileye O Ba and Jean-Marc Odobez. Head pose tracking and focus of attention recognition algorithms in meeting rooms. In *International Evaluation Workshop on Classification of*

Events, Activities and Relationships, pages 345–357. Springer, 2006.

[192] Stylianos Asteriadis, Paraskevi Tzouveli, Kostas Karpouzis, and Stefanos Kollias. Estimation of behavioral user state based on eye gaze and head pose—application in an e-learning environment. *Multimedia Tools and Applications*, 41(3):469–493, 2009.

[193] Cathal J Breen, Raymond Bond, and Dewar Finlay. An evaluation of eye tracking technology in the assessment of 12 lead electrocardiography interpretation. *Journal of electrocardiology*, 47(6):922–929, 2014.

[194] Peter O'Meara, Graham Munro, Brett Williams, Simon Cooper, Fiona Bogossian, Linda Ross, Louise Sparkes, Mark Browning, and Mariah McClounan. Developing situation awareness amongst nursing and paramedicine students utilizing eye tracking technology and video debriefing techniques: A proof of concept paper. *International emergency nursing*, 23(2):94–99, 2015.

[195] Norie Suetsugu, Masafumi Ohki, and Tsunehisa Kaku. Quantitative analysis of nursing observation employing a portable eye-tracker. *Open Journal of Nursing*, 6(01):53, 2016.

[196] Jonathan Currie, Raymond R Bond, Paul McCullagh, Pauline Black, Dewar D Finlay, and Aaron Peace. Eye tracking the visual attention of nurses interpreting simulated vital signs scenarios: mining metrics to discriminate between performance level. *IEEE Transactions on Human-Machine Systems*, 48(2):113–124, 2018.

[197] Jason Hochreiter, Salam Daher, Arjun Nagendran, Laura Gonzalez, and Greg Welch. Touch sensing on non-parametric rear-projection surfaces: A physical-virtual head for hands-on healthcare training. In *Virtual Reality (VR), 2015 IEEE*, pages 69–74. IEEE, 2015.

[198] Aaxa p300 pico projector - dlp hand-held mini projector - led pocket projector. `http:`

//www.aaxatech.com/products/p300_pico_projector.htm. (Accessed on 10/05/2018).

[199] Kelly L Murdock and Eric Allen. *Edgeloop Character Modeling for 3D Professionals Only*. John Wiley & Sons, Inc., 2006.

[200] Autodesk maya,. http://www.autodesk.com/products/maya/overview, 2018 10. (Accessed on 10/05/2018).

[201] Unity. unity3d. https://unity3d.com, 10 2018. (Accessed on 10/05/2018).

[202] Salam Daher, Jason Hochreiter, Ryan Schubert, Laura Gonzalez, Juan Cendan, Mindi Anderson, Desiree Giaz, and Gregory Welch. The physical-virtual patient simulator: A physical human form with virtual appearance and behavior. *Simulation in Healthcare*, tbd(tbd):tbd, 2019.

[203] Cheryl D Fryar, Qiuping Gu, and Cynthia L Ogden. Anthropometric reference data for children and adults: United states, 2007-2010., 2012.

[204] Jane Bear-Lehman, Michelle Kafko, Liana Mah, Liliana Mosquera, and Barbara Bairbre Reilly. An exploratory look at hand strength and hand size among preschoolers. *Journal of Hand Therapy*, 15(4):340–346, 2002.

[205] B Hohendorff, C Weidermann, KJ Burkhart, PM Rommens, KJ Prommersberger, and MA Konerding. Lengths, girths, and diameters of children's fingers from 3 to 10 years of age. *Annals of Anatomy-Anatomischer Anzeiger*, 192(3):156–161, 2010.

[206] Family practice notebook height measurement in children. https://fpnotebook.com/Endo/Exam/HghtMsrmntInChldrn.htm. (Accessed on 10/06/2018).

[207] Acrylic sheet product, polymer & coating solutions | plaskolite. https://plaskolite.com/. (Accessed on 10/06/2018).

[208] Honeywell hce100b heat bud mini home heater | honeywell store. https://www.honeywellstore.com/store/products/heat-bud-ceramic-portable-mini-heater-hce100-series.htm. (Accessed on 10/06/2018).

[209] techtile. http://www.techtile.org/en/techtiletoolkit/. (Accessed on 10/06/2018).

[210] Adobe photoshop cc best photo, image, and design editing software. https://www.adobe.com/products/photoshop.html. (Accessed on 10/06/2018).

[211] Voice changing software - online gaming and utilities. https://screamingbee.com/Product/MorphVOX.aspx. (Accessed on 10/06/2018).

[212] Rogo digital - lipsync. https://lipsync.rogodigital.com/. (Accessed on 10/06/2018).

[213] Nils Dahlbdck, Arne Jdnsson, and Lars Ahrenberg. Wizard of oz studies—why and how. In *Proceedings of the 1993 International Workshop on Intelligent User Interfaces: January 4-7, 1993, Orlando, Florida*, page 193. Association for Computing Machinery, 1992.

[214] Toni Bloodworth Pence, Lauren Cairco Dukes, Larry F. Hodges, Nancy K. Meehan, and Arlene Johnson. The effects of interaction and visual fidelity on learning outcomes for a virtual pediatric patient system. In *Healthcare Informatics (ICHI), 2013 IEEE International Conference on*, pages 209–218. IEEE, 2013.

[215] Aaron Kotranza, Benjamin Lok, Adeline Deladisma, Carla M. Pugh, and D. Scott Lind. Mixed reality humans: Evaluating behavior, usability, and acceptability. *IEEE Transactions on Visualization and Computer Graphics*, 15(3):369–382, 2009.

[216] Myungho Lee, Kangsoo Kim, Salam Daher, Andrew Raij, Ryan Schubert, Jeremy Bailenson, and Greg Welch. The wobbly table: Increased social presence via subtle incidental movement of a real-virtual table. In *Virtual Reality (VR), 2016 IEEE*, pages 11–17. IEEE, 2016.

[217] Myungho Lee, Gerd Bruder, and Gregory F. Welch. Exploring the effect of vibrotactile feedback through the floor on social presence in an immersive virtual environment. In *Virtual Reality (VR), 2017 IEEE*, pages 105–111. IEEE, 2017.

[218] Gianluca De Leo, Leigh A. Diggs, Elena Radici, and Thomas W. Mastaglio. Measuring sense of presence and user characteristics to predict effective training in an online simulated virtual environment. *Simulation in Healthcare*, 9(1):1–6, 2014.

[219] Fabrizia Mantovani and Gianluca Castelnuovo. The sense of presence in virtual training: enhancing skills acquisition and transfer of knowledge through learning experience in virtual environments, 2003.

[220] Tze Wei Liew, Su-Mae Tan, and Hishamuddin Ismail. Exploring the effects of a non-interactive talking avatar on social presence, credibility, trust, and patronage intention in an e-commerce website. *Human-centric Computing and Information Sciences*, 7(1):42, 2017.

[221] Simmom | overview. `http://www.laerdal.com/us/SimMom`, 2015.

[222] Kenneth C. Herbst, Lowell Gaertner, and Chester A. Insko. My head says yes but my heart says no: Cognitive and affective attraction as a function of similarity to the ideal self.

Journal of personality and social psychology, 84(6):1206–1219, 2003.

[223] Sören Huwendiek, Bas A De Leng, Andrzej A Kononowicz, Romy Kunzmann, Arno MM Muijtjens, Cees PM Van Der Vleuten, Georg Friedrich Hoffmann, Burkhard Tönshoff, and Diana HJM Dolmans. Exploring the validity and reliability of a questionnaire for evaluating virtual patient design with a special emphasis on fostering clinical reasoning. *Medical teacher*, 37(8):775–782, 2015.

[224] Norman Blaikie. *Analyzing quantitative data: From description to explanation*. Sage, 2003.

[225] Thomas R. Knapp. Treating ordinal scales as interval scales: an attempt to resolve the controversy. *Nursing research*, 39(2):121–123, 1990.

[226] William M. Kuzon Jr, Melanie G. Urbanchek, and Steven McCabe. The seven deadly sins of statistical analysis. *Annals of plastic surgery*, 37(3):265–272, 1996.

[227] Nathan J. Emery. The eyes have it: the neuroethology, function and evolution of social gaze. *Neuroscience & Biobehavioral Reviews*, 24(6):581–604, 2000.

[228] Stephen R. H. Langton, Roger J. Watt, and Vicki Bruce. Do the eyes have it? cues to the direction of social attention. *Trends in cognitive sciences*, 4(2):50–59, 2000.

[229] David I. Perrett and Nathan J. Emery. Understanding the intentions of others from visual signals: Neurophysiological evidence, 1994.

[230] Rainer Stiefelhagen. Tracking focus of attention in meetings. In *Multimodal Interfaces, 2002. Proceedings. Fourth IEEE International Conference on*, pages 273–280. IEEE, 2002.

[231] Ben Benfold and Ian D. Reid. Guiding visual surveillance by tracking human attention. In *BMVC*, volume 2, page 7, 2009.

[232] Samira Sheikhi and Jean-Marc Odobez. Recognizing the visual focus of attention for human robot interaction. In *International Workshop on Human Behavior Understanding*, pages 99–112. Springer, 2012.

[233] Benoît Massé, Silèye Ba, and Radu Horaud. Tracking gaze and visual focus of attention of people involved in social interaction. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2017.

[234] J Cassell et al. Embodiment in conversational interfaces: Rea. chi'99, 1999.

[235] Barry Issenberg, Michael S Gordon, David Lee Gordon, Robert E Safford, and Ian R Hart. Simulation and new learning technologies. *Medical teacher*, 23(1):16–23, 2001.

[236] MJ Pywell, E Evgeniou, E Pitt, and CM Estela. High fidelity, low cost moulage as a valid simulation tool to improve burns education. *Burns*, 42(4):844–852, 2016.

[237] RN Jobeth Pilcher EdD, MSN Heather Goodall, RN Cynthia Jensen, MSHCA Cordelia Jewell RNC-NIC, RNC-NIC Regina Reynolds BSN, et al. Simulation-based learning: It's not just for NRP. *Neonatal Network*, 31(5):281, 2012.

[238] Beth M Nagle, Jeanne M McHale, Gail A Alexander, and Brian M French. Incorporating scenario-based simulation into a hospital nursing education program. *The Journal of Continuing Education in Nursing*, 40(1):18–25, 2009.

[239] Pupil Labs DIY. `https://docs.pupil-labs.com/#diy`, 2018. Accessed: 2018-09-30.

[240] INACSL Standards Committee. INACSL standards of best practice: Simulation(sm): Simulation design. *Clinical Simulation in Nursing*, 12:S5–S12, 2016.

[241] Oliver Bimber and Ramesh Raskar. *Spatial Augmented Reality: Merging Real and Virtual Worlds*. A. K. Peters, Ltd., Natick, MA, USA, 2005.

[242] O Ariza, P Lubos, F Steinicke, and G Bruder. Ring-shaped haptic device with vibrotactile feedback patterns to support natural spatial interaction. In *Proceedings of the 25th International Conference on Artificial Reality and Telexistence and 20th Eurographics Symposium on Virtual Environments, ICAT-EGVE*, volume 15, pages 175–181, 2015.

[243] Simon Schätzle, Tobias Ende, Tilo Wüsthoff, and Carsten Preusche. Vibrotac: An ergonomic and versatile usable vibrotactile feedback device. In *RO-MAN, 2010 IEEE*, pages 670–675. IEEE, 2010.

[244] Anh Nguyen and Amy Banic. *3DTouch: A wearable 3D input device for 3D applications*. IEEE, 2015.

[245] John Brooke. SUS–A quick and dirty usability scale. *Usability Evaluation in Industry*, 189(194):4–7, 1996.

[246] Charles J Clopper and Egon S Pearson. The use of confidence or fiducial limits illustrated in the case of the binomial. *Biometrika*, pages 404–413, 1934.

[247] Joseph L Fleiss, Bruce Levin, and Myunghee Cho Paik. *Statistical methods for rates and proportions*. John Wiley & Sons, 2013.

[248] G De Leo, LA Diggs, E Radici, and TW Mastaglio. Measuring sense of presence and user characteristics to predict effective training in an online simulated virtual environment. *Simulation in healthcare: journal of the Society for Simulation in Healthcare*, 9(1):1 – 6, 2014.

[249] Matthew Lombard and Theresa Ditton. At the Heart of It All: The Concept of Presence. *Journal of Computer-Mediated Communication*, 3(2), June 1997.

[250] Erving Goffman. *Behavior in Public Places: Notes on the Social Organization of Gatherings*. The Free Press (a Division of Simon and Schuster, Inc.), New York, NY USA, 1963.

[251] Jim Blascovich. Social Influence within Immersive Virtual Environments. In Ralph Schroeder, editor, *The Social Life of Avatars*, Computer Supported Cooperative Work, pages 127–145. Springer London, 2002.

[252] Jim Blascovich, Jack Loomis, Andrew C. Beall, Kimberly R. Swinth, Crystal L. Hoyt, and Jeremy N. Bailenson. Immersive virtual environment technology as a methodological tool for social psychology. *Psychological Inquiry*, 13(2):103–124, 2002.

[253] Bob G. Witmer and Michael J. Singer. Measuring Presence in Virtual Environments: A Presence Questionnaire. *Presence: Teleoperators and Virtual Environments*, 7(3):225–240, June 1998.

[254] R Likert. A technique for the measurement of attitudes, 1932.

[255] Simulation In Healthcare. http://www.ssih.org.

[256] JDMS. http://www.scs.org/jdms.

[257] Kathleen R Rosen. The history of medical simulation. *Journal of critical care*, 23(2):157–166, 2008.

[258] https://timeline.knightlab.com.

[259] Salam Daher. http://salamdaher.net/ucf/dissertation/healthcaresimulators/timeline.html.

[260] Danica Marković and Bojana Marković-Živković. Development of anatomical models—chronology. *Acta Medica Medianae*, 49(2):56–62, 2010.

[261] Gillian M Morriss-Kay. The evolution of human artistic creativity. *Journal of anatomy*, 216(2):158–176, 2010.

[262] Wikipedia: Venus figurines. `https://en.wikipedia.org/wiki/Venus_figurines`, 2015.

[263] Regis Olry. Wax, wooden, ivory, cardboard, bronze, fabric, plaster, rubber and plastic anatomical models: Praiseworthy precursors of plastinated specimens. *J Int Soc Plastination*, 15(1):30–35, 2000.

[264] Harry Owen. Early use of simulation in medical education. *Simulation in Healthcare*, 7(2):102–116, 2012.

[265] Claus C Schnorrenberger. Anatomical roots of chinese medicine and acupuncture. *SCHWEIZERISCHE ZEITSCHRIFT FUR GANZHEITS MEDIZIN*, 20(3):163, 2008.

[266] Miguel A. Faria Jr. The death of henry ii of france | hacienda publishing. `https://haciendapublishing.com/articles/death-henry-ii-france`, 12 1992. (Accessed on 11/12/2018).

[267] Science museum brought to life exploring the history of medicine. `http://www.sciencemuseum.org.uk/broughttolife/objects/display.aspx?id=5026`, 2015.

[268] Le cere anatomiche di clemente susini dell universita di cagliari. `http://pacs.unica.it/oldmedicina/cere/mono02_it.htm`, 2015.

[269] Kenneth Fitzpatrick Russell. Ivory anatomical manikins. *Medical history*, 16(02):131–142, 1972.

[270] Kenneth Fitzpatrick Russell. Ivory anatomical manikins. *Medical history*, 16(02):131–142, 1972.

[271] Medical antiques archives. `http://antiquescientifica.com/archive17.htm`, 2015.

[272] Medical antiques archives. `http://antiquescientifica.com/archive54.htm`, 2015.

[273] Roberta Ballestriero. Anatomical models and wax venuses: art masterpieces or scientific craft works? *Journal of anatomy*, 216(2):223–234, 2010.

[274] Giovanni E Orlandini and Ferdinando Paternostro. Anatomy and anatomists in tuscany in the 17th century. *Italian Journal of Anatomy and Embryology*, 115(3):167–174, 2010.

[275] Danica Marković and Bojana Marković-Živković. Development of anatomical models—chronology. *Acta Medica Medianae*, 49(2):56–62, 2010.

[276] Science museum brought to life exploring the history of medicine: Anna morandi manzolini (1716-74). `http://www.sciencemuseum.org.uk/broughttolife/people/annamorandimanzolini.aspx?keywords=Anna+Morandi`, 2015.

[277] Wikipedia: Ercole lelli. `https://en.wikipedia.org/wiki/Ercole_Lelli`.

[278] Wikipedia: Giovanni manzolini. `https://en.wikipedia.org/wiki/Giovanni_Manzolini`.

[279] Alan W Bates. Dr kahn's museum: obscene anatomy in victorian london. *Journal of the Royal Society of Medicine*, 99(12):618–624, 2006.

[280] On the trail of the machine: William smellie's celebrated apparatus. `http://dittrick.blogspot.com/2013/04/on-trail-of-machine-william-smellies.html`, 2015. http://dittrick.blogspot.com/2013/04/on-trail-of-machine-william-smellies.html.

[281] Jessica Riskin. Eighteenth-century wetware. *Representations*, 83(1):97–125, 2003.

[282] Two obstetric dolls and doctor smellie's mechanical labor device. `http://www.steflenk.com/OBSTETRICdolls/08_mechanicalLaborDevice.html`, 2015.

[283] Mother machine: an uncanny valley in the eighteenth century. `http://theappendix.net/issues/2013/4/mother-machine-an-uncanny-valley-in-the-eighteenth-century`, 2015.

[284] Bricogne Stéphan www.bricogne.com. Enfore : Encyclopédie des faits ordinaires extraordinaires, 2015.

[285] Regis Olry. Wax, wooden, ivory, cardboard, bronze, fabric, plaster, rubber and plastic anatomical models: Praiseworthy precursors of plastinated specimens. *J Int Soc Plastination*, 15(1):30–35, 2000.

[286] cambridge university. Skeletons in the cupboard of medical science. `http://www.cam.ac.uk/research/features/skeletons-in-the-cupboard-of-medical-science`, 2014.

[287] Michael Sticherling and Uta Euler. The collection of dermatologic wax moulages at the university of kiel, germany. *International journal of dermatology*, 40(9):586–592, 2001.

[288] Misemono: Meiji-era sideshow pregnancy dolls. http://daddytypes.com/2009/05/14/misemono_meiji-era_sideshow_pregnancy_dolls.php, 2015.

[289] Set of sixty miniature heads used in phrenology, manchester, england, 1831. http://www.sciencemuseum.org.uk/broughttolife/objects/display.aspx?id=5000&keywords=phrenological+head, 2015.

[290] Nick Hopwood. Model politics. *Lancet (London, England)*, 372(9654):1946–1947, 2008.

[291] Medical antiques archives. http://antiquescientifica.com/archive65.htm, 2015.

[292] Artificial anatony papier mache anatomical models. http://americanhistory.si.edu/anatomy/history/nma03_history_main.html, 2015.

[293] Dr. auzoux's models. http://www.hps.cam.ac.uk/whipple/explore/models/drauzouxsmodels/, 2015.

[294] L Spencer. Chance, circumstance and folly: Richard berry and the plaster anatomical collection of the harry brookes allen museum of anatomy and pathology. *University of Melbourne Collections*, 2:3–10, 2008.

[295] Jon Cornwall and Chris Smith. Anatomical models by fj steger (1845-1938): the university of otago collection. *Eur. J. Anat*, 18(3):209–211, 2014.

[296] Schultes medacta. Medacta models. http://englisch.schultes-medacta.com/aboutus.html, 2015.

[297] About laerdal. `http://www.laerdal.com/about/default.htm`, 2015.

[298] David M Gaba and VA Palo Alto HCS. Smith b, gaba d. simulators. in: Lake c, blitt c, hines r, eds. clinical monitoring: Practical application. new york, ny: Wb saunders company. in press, 2000 brian e. smith, md staff anesthesiologist va palo alto health care system.

[299] Gordon center for research in medical education. `http://www.gcrme.miami.edu/harvey_features.php`, 2015.

[300] David M Gaba and Abe DeAnda. A comprehensive anesthesia simulation environment: re-creating the operating room for research and training. *Anesthesiology*, 69(3):387–394, 1988.

[301] Rettedal Arne, Freyer Ståle, Kleppa Ragna, and Larsen Petter. Patsim—simulator for practising anaesthesia and intensive care. *International journal of clinical monitoring and computing*, 13(3):147–152, 1996.

[302] V Chopra, FHM Engbers, MJ Geerts, WR Filet, JG Bovill, and J Spierdijk. The leiden anaesthesia simulator. *British Journal of Anaesthesia*, 73(3):287–292, 1994.

[303] Ulrik J Christensen, Soren F Andersen, John Jacobsen, Per F Jensen, and Helle Ording. The sophus anaesthesia simulator v. 2.0. a windows 95 control-center of a full-scale simulator. *International journal of clinical monitoring and computing*, 14(1):11–16, 1997.

[304] AJ Byrne, PJ Hilton, and JN Lunn. Basic simulations for anaesthetists a pilot study of the access system. *Anaesthesia*, 49(5):376–381, 1994.

[305] MWP Goodwin and GWG French. Simulation as a training and assessment tool in the management of failed intubation in obstetrics. *International journal of obstetric anesthesia*, 10(4):273–277, 2001.

[306] Cae hps | human patient simulator. https://caehealthcare.com/patient-simulation/hps, 2015.

[307] Uf's powertools. http://floridaphysician.med.ufl.edu/2013/02/features/ufs-powertools/, 2015.

[308] Medsim-eagle patient simulator - simulation center - stanford university school of medicine. http://med.stanford.edu/VAsimulator/medsim.html, 2015.

[309] Wayne state university: Noelle maternal and neonatal simulator. https://nursing.wayne.edu/simlab/labs/noelle.php, 2015.

[310] Laerdal: Sim man. http://www.laerdal.com/us/doc/86/SimMan, 2015.

[311] About the traumaman system | simulab corporation. http://www.simulab.com/about-traumaman-system, 2015.

[312] Lance Baily. Simulab traumaman & surgical abdomen product review. http://healthysimulation.com/4937/simulab-traumaman-surgical-abdomen-product-review-and-breakdown, 2015.

[313] S108 premie blue simulator with smartskin technology. http://www.gaumard.com/s108, 2015.

[314] Meti pelvic examsim. http://limbsandthings.com/uk/products/meti-pelvic-examsim.

[315] Critical care - patient simulators gaumard scientific. http://www.gaumard.com/products/critical-care, 2015.

[316] Cae healthcare - product and services: Babysim. http://www.caehealthcare.com/eng/patient-simulators/babysim, 2015.

[317] Noelle s555.100 moes - mobile obstetrics emergencies simulator gaumard scientific. http://www.gaumard.com/s555-100, 2015.

[318] Simbaby overview. http://www.laerdal.com/us/SimBaby, 2015.

[319] Noelle s575.100 tetherless maternal and neonatal birthing simulator gaumard. http://www.gaumard.com/s575-100, 2015.

[320] Cae healthcare - product and services istan. http://www.caehealthcare.com/eng/patient-simulators/istan, 2015.

[321] Pedi hal s3004 tetherless 1-year pediatric simulator | gaumard scientific. http://www.gaumard.com/s3004?____SID=U, 2015.

[322] Premie hal s3009 tetherless premature newborn simulator gaumard scientific. http://www.gaumard.com/s3009, 2015.

[323] Simman 3g advanced patient simulator | overview. http://www.laerdal.com/us/SimMan3GTrauma, 2015.

[324] Als simulator overview. http://www.laerdal.com/us/ALS, 2015.

[325] Prompt birthing simulator | overview. http://www.laerdal.com/us/doc/224/PROMPT-Birthing-Simulator, 2015.

[326] Simnewb | overview. http://www.laerdal.com/us/doc/88/SimNewB, 2015.

[327] Cae healthcare - product and services: Metiman. `http://www.caehealthcare.com/eng/patient-simulators/metiman`, 2015.

[328] Susie s2000 tetherless nursing simulator | gaumard scientific. `http://www.gaumard.com/s2000`, 2015.

[329] Hal s3201 advanced tetherless patient simulator | gaumard scientific. `http://www.gaumard.com/s3201`, 2015.

[330] Mamanatalie birthing simulator | overview. `http://www.laerdal.com/us/mamaNatalie`, 2015.

[331] Simjunior | overview. `http://www.laerdal.com/us/SimJunior`, 2015.

[332] Hal s315.400 adult multipurpose airway and cpr trainer | gaumard scientific. `http://www.gaumard.com/s315-400`, 2015.

[333] Simmom | overview. `http://www.laerdal.com/us/SimMom`, 2015.

[334] Noelle s2200 - victoria tetherless maternal and neonatal birthing simulator | gaumard scientific. `http://www.gaumard.com/s2200`, 2015.

[335] D. Schmorrow, J. Cohn, D. Nicholson, Praeger Security International, and Praeger Security International Staff. *The PSI Handbook of Virtual Environments for Training and Education: Developments for the Military and Beyond*. Praeger Security International, 2009.

[336] M. Brown, A. Majumder, and Yang Ruigang. Camera-based calibration techniques for seamless multiprojector displays. *Visualization and Computer Graphics, IEEE Transactions on*, 11(2):193–206, 2005.