escola superior de tecnologia e gestão
instituto politécnico de leiria

Instituto Politécnico de Leiria
Escola Superior de Tecnologia e Gestão
Departamento de Engenharia Informática

Mestrado em Eng.ª Informática – Computação Móvel

MULTIMODAL INTERACTION SYSTEM SUPPORTED BY DIGITAL HUMANS

CAROLINA DA SILVA COSTA PEREIRA

Leiria, September 2025

Instituto Politécnico de Leiria
Escola Superior de Tecnologia e Gestão
Departamento de Engenharia Informática

Mestrado em Eng.ª Informática – Computação Móvel

MULTIMODAL INTERACTION SYSTEM SUPPORTED BY DIGITAL HUMANS

CAROLINA DA SILVA COSTA PEREIRA

Número: 2232644

Dissertação realizada sob orientação do Professor Doutor Nuno Carlos Sousa Rodrigues (nunorod@ipleiria.pt), Professor Doutor Alexandrino José Marques Gonçalves (alex@ipleiria.pt), Professora Doutora Anabela Gonçalves Rodrigues Marto (anabela.marto@ipleiria.pt), Professor Doutor Roberto Aguiar Ribeiro (roberto.ribeiro@ipleiria.pt)

Leiria, September 2025

Originality and Copyright

This project report is original, made only for this purpose, and all authors whose studies and publications were used to complete it are duly acknowledged.

Dedication

I would like to dedicate this work to my parents and my sister, whose constant support has been invaluable throughout this process. Their encouragement provided me with the stability and perseverance needed to complete this thesis.

Acknowledgments

I would like to formally express my gratitude to the Polytechnic University of Leiria for providing academic support throughout my studies. I am especially thankful for my supervisors for their guidance and insightful feedback during the development of my thesis. I also wish to extend my thanks to the Computer Science and Communication Research Centre (CIIC) for their support through the research grant, through the Portuguese Foundation for Science and Technology (FCT) under the project UIDB/04524/2020, for granting access to the centre's facilities, and for providing the necessary equipment during a significant part of the project's development.

Abstract

As digital services continue to grow in popularity, there is an increasing need for systems that replicate human interaction while prioritizing user satisfaction and engagement. Early implementations offered interactions that felt unnatural, mainly due to the lack of realistic facial expressions and movement in character animations. Over time, however, significant advancements—particularly from the video game industry—have driven major improvements in this field. Titles such as Senua's Saga: Hellblade II and Black Myth: Wukong illustrate how these developments have enabled the creation of immersive characters with highly realistic facial animation.

The Metaverse has emerged as a key area of interest, offering immersive virtual environments where users interact within a shared digital space. This evolution has increased the demand for personalized Digital Humans—high-fidelity, computer-generated avatars capable of expressing empathy in real time.

This study examines how Digital Humans can enhance human–computer interaction by reducing the emotional disconnect commonly associated with automated systems. Such avatars show potential across remote meetings, customer service, online education, and Metaverse platforms, fostering more natural and engaging interactions.

Emotionally expressive Digital Humans were created using the MetaHuman framework and Unreal Engine, incorporating multimodal MoCap based on computer vision and RGB camera input. Two user tests were conducted: one focused on facial expressions and another combining facial expressions with full-body movement, involving a total of 40 participants. Empathy levels were assessed using the Toronto Empathy Questionnaire (TEQ), administered before and after interaction.

One-Way ANOVA analyses showed no statistically significant differences in elicited empathy between default and custom animations. In the facial-animation test, participants' average TEQ scores were 47.4 for default animations and 45.0 for custom ones, both within or above the general population average (40–45). In the combined full-body test, mean scores were 47.6 for human body movement and 45.5 for MetaHuman body motion.

Although personalization did not significantly outperform default animations, the results highlight the essential role of realistic body movement in shaping emotional perception and interaction quality. The findings confirm that emotionally expressive Digital Humans can be effectively integrated into digital platforms, while showing that facial personalization alone must be complemented by contextual and narrative elements to maximize empathetic impact. This work provides a solid foundation for future research on deploying Digital Humans in real-world interactive systems.

Keywords: Digital Environment, Digital Human, Emotion, Empathy, Human-computer interaction, MetaHuman

Resumo

À medida que os serviços digitais continuam a crescer em popularidade, torna-se cada vez mais necessário desenvolver sistemas capazes de replicar interações humanas, dando prioridade à satisfação e ao envolvimento dos utilizadores. Inicialmente, estas interações eram pouco naturais, devido à falta de realismo nas expressões faciais e nos movimentos das animações. No entanto, avanços significativos — sobretudo na indústria dos videojogos — têm impulsionado melhorias notáveis nesta área. Jogos como Senua's Saga: Hellblade II e Black Myth: Wukong demonstram como estas evoluções permitiram criar personagens imersivas com expressões faciais altamente realistas.

O Metaverso surge como uma área de interesse central, ao proporcionar ambientes virtuais imersivos onde os utilizadores interagem num mundo digital partilhado. Este crescimento aumentou a procura por Humanos Digitais personalizados — avatares gerados por computador e de elevada fidelidade — capazes de expressar empatia em tempo real.

Este estudo investiga de que forma os Humanos Digitais podem melhorar a interação humano-computador, reduzindo a distância emocional frequentemente associada a sistemas automatizados. Estes avatares têm potencial em reuniões remotas, serviços de atendimento ao cliente, educação/formação online e plataformas de Metaverso, promovendo interações mais naturais e envolventes.

Foram desenvolvidos Humanos Digitais emocionalmente expressivos utilizando a framework MetaHuman e o Unreal Engine, com captura de movimento multimodal através de visão computacional e uma câmara RGB para criar animações personalizadas. Realizaram-se dois testes com utilizadores: um focado nas expressões faciais e outro que combinou expressões faciais com movimento corporal, envolvendo um total de 40 participantes. Os níveis de empatia foram avaliados com o TEQ, aplicado antes e depois da interação.

As análises ANOVA de um fator (One-Way ANOVA) não revelaram diferenças estatisticamente significativas na empatia evocada entre animações padrão e personalizadas. No primeiro teste, centrado na animação facial, os resultados médios do TEQ foram de 47,4 para as animações padrão e 45,0 para as personalizadas, ambos dentro ou acima da média populacional (40–45). No segundo teste, que integrou movimento corporal, os valores médios foram 47,6 para o movimento corporal humano e 45,5 para o movimento corporal digital.

Apesar da ausência de diferenças significativas, os resultados confirmam que movimento corporal realista desempenha um papel essencial na perceção emocional e na qualidade da interação. Este trabalho demonstra que Humanos Digitais emocionalmente expressivos podem ser integrados com sucesso em plataformas digitais, destacando que a personalização facial, por si só, deve ser complementada por fatores contextuais e narrativos para maximizar o impacto empático. Estes resultados constituem uma base sólida para investigações futuras sobre a integração de Humanos Digitais em sistemas interativos reais.

Palavras-Chave: Ambiente Digital, Empatia, Emoção, Humano Digital, Interação Humano-Computador, MetaHuman

Contents

Originality and Copyright iii

Dedication iv

Acknowledgments v

Abstract vi

Resumo viii

List of Figures xiii

List of Tables xiv

List of Abbreviations and Acronyms xvi

1. Introduction 1

2. Background 11

2.1. Digital Human Conceptualization 11

2.2. Game Engine Development and Real-Time Rendering 11

2.3. MetaHuman Framework 13

2.4. Motion Capture (MoCap) 13

2.5. Video Conferencing Platforms and Privacy Considerations 14

3. Related Work 17

3.1. Studies with user testing 17

3.2. Studies without user testing 18

3.3. Technological trends and challenges 19

3.4. Tools to evaluate empathy 19

4. Methodology 21

4.1. Phase 1 — Related Work Review 21

4.2. Phase 2 — Proposal and Requirements Analysis 23

4.3. Phase 3 — Development Process 23

4.4. Phase 4 — User Testing 24

4.5. Phase 5 — Data Analysis **25**
5. Development **27**
5.1. Equipment **27**
5.2. Digital Human Creation **28**
5.3. Facial Animation **32**
5.4. Lip Sync Animation **33**
5.5. Body Animation **34**
5.6. Development of the Case Study **35**
6. Tests and Results **37**
6.1. Participants **37**
6.1.1. User Test 1 – Facial Expressions 37
6.1.2. User Test 2 – Body Movement 38
6.2. Measures **38**
6.2.1. Toronto Empathy Questionnaire (TEQ) 38
6.2.2. Post-Interaction Survey 38
6.3. Procedure **39**
6.4. Provisional Test **43**
6.5. Results **43**
7. Analysis and Discussion **49**
7.1. User Test 1 – Facial Expressions **49**
7.1.1. Linear Regression 49
7.1.2. Shapiro-Wilk Normality 52
7.1.3. Pearson Correlation 57
7.1.4. One-Way ANOVA 59
7.2. User Test 2 – Body Movement **61**
7.2.1. Linear Regression 62

7.2.2. Shapiro-Wilk Normality 65

7.2.3. Pearson Correlation 69

7.2.4. One-Way ANOVA 71

7.3. Discussion 73

7.3.1. RQ1: How does user perception of a Digital Human compare to that of a real person in terms of realism and emotional expressiveness? 74

7.3.2. RQ2: How do users perceive the realism and emotional expressiveness of a Digital Human's body movements compared to those of a real person performing the same actions? 74

7.3.3. RQ3: Can personalized facial animation of emotion increase the feeling of empathy compared to motion capture-based facial animation? 75

7.3.4. RQ4: How do users perceive the emotional authenticity and effectiveness of default facial expressions compared to custom-created facial expressions in a Digital Human? 75

7.3.5. RQ5: What are the technical challenges and limitations in developing Digital Humans with emotional intelligence? 76

8. Conclusion and Future Work 77

Bibliography 80

List of Figures

Figure 1 – Hellblade: Senua's Sacrifice realistic character using MoCap and photogrammetry. 4

Figure 2 – Senua's Saga: Hellblade II main character from the Unreal Engine 5 MetaHuman Demo 4

Figure 3 – Sophia Lillis' character from the upcoming video game OD (Overdose) reveal trailer. 5

Figure 4 – Hunter Schafer' character from the upcoming video game OD (Overdose) reveal trailer. 5

Figure 5 – Udo Kier character from the upcoming video game OD (Overdose) reveal trailer. 6

Figure 6 – Sophia Lillis' character conveying shock and vulnerability at the end of the trailer. 7

Figure 7 – Methodology Workflow 21

Figure 8 – Related Work Review excel with "Content" category. 22

Figure 9 – Related Work Review excel with "Technology" category. 22

Figure 10 – Related Work Review excel with "Evaluation" category. 22

Figure 11 – Steps to MetaHuman creation. 23

Figure 12 – 3D model created from the *Polycam* application. 29

Figure 13 – 3D model before and after cleanup on Blender. 29

Figure 14 – MetaHuman Identity component with the 3D model of the face cleaned up. 30

Figure 15 – Visual representation of the MetaHuman mesh. 31

Figure 16 – MetaHuman Identity mesh comparison. 31

Figure 17 – MetaHuman on the "MetaHuman Creator" website. 32

Figure 18 – MetaHuman displaying FACS emotions. 33

Figure 19 – TEQ results for the first user tests and respective averages 45

Figure 20 – PIS results for the first user tests and respective averages 46

Figure 21 – TEQ results for the second user tests and respective averages 47

Figure 22 – PIS results for the second user tests and respective averages 48

Figure 23 – Shapiro-Wilk Normal Q-Q Plot of Unstandardized Residual for the Custom Expressions 55

Figure 24 – Shapiro-Wilk Normal Q-Q Plot of Unstandardized Residual for the Default Expressions 57

Figure 25 – Shapiro-Wilk Normal Q-Q Plot of Unstandardized Residual for the MetaHuman Body Movement 67

Figure 26 – Shapiro-Wilk Normal Q-Q Plot of Unstandardized Residual for the Human Body Movement... 69

List of Tables

Table 1 – Comparison between default and custom expressions 41

Table 2 – Comparison between Human and MetaHuman body movement 41

Table 3 – PIS adapted from TEQ comparison. 42

Table 4 – Summary of TEQ scores from different validation studies 44

Table 5 – Linear Regression model for the Custom Expressions 50

Table 6 – ANOVA results for the Custom Expressions 50

Table 7 – Coefficient results for the Custom Expressions 51

Table 8 – Linear Regression model for the Default Expressions 51

Table 9 – ANOVA results for the Default Expressions 52

Table 10 – Coefficient results for the Default Expressions 52

Table 11 – Shapiro-Wilk Normality results for the Custom Expressions 54

Table 12 – Shapiro-Wilk Normality results for the Default Expressions 56

Table 13 – Pearson Correlation results for the Custom Expressions 58

Table 14 – Pearson Correlation results for the Default Expressions 59

Table 15 – One-Way ANOVA descriptives for the First User Test 59

Table 16 – One-Way ANOVA Levene's Test for the First User Test 60

Table 17 – ANOVA for the First User Test 60

Table 18 – ANOVA Effect Sizes for the First User Test 61

Table 19 – Linear Regression model for the MetaHuman Body Movement 62

Table 20 – ANOVA results for the MetaHuman Body Movement 62

Table 21 – Coefficient results for the MetaHuman Body Movement 63

Table 22 – Linear Regression model for the Human Body Movement 63

Table 23 – ANOVA results for the Human Body Movement 64

Table 24 – Coefficient results for the Human Body Movement 64

Table 25 – Shapiro-Wilk Normality results for the MetaHuman Body Movement 66

Table 26 – Shapiro-Wilk Normality results for the Human Body Movement 68

Table 27 – Pearson Correlation results for the MetaHuman Body Movement 70

Table 28 – Pearson Correlation results for the Human Body Movement 71

Table 29 – One-Way ANOVA descriptives for the Second User Test 71

Table 30 – One-Way ANOVA Levene's Test for the Second User Test 72

Table 31 – ANOVA for the Second User Test 72

Table 32 – ANOVA Effect Sizes for the Second User Test 73

List of Abbreviations and Acronyms

2D/3D	Two-Dimensional/Three Dimensional
AI	Artificial Intelligence
ANOVA	Analysis of Variance
EC	Empathy Concern
FACs	Facial Action Coding System
FS	Fantasy
HCI	Human-Computer Interaction
IoT	Internet of Things
IRC	Internet Relay Chat
IRI	Interpersonal Reactivity Index
MoCap	Motion Capture
PD	Personal Distress
PIS	Post-Interaction Survey
SD	Standard Deviation
TEQ	Toronto Empathy Questionnaire
VC	Virtual Character
VR	Virtual Reality

1. Introduction

In recent years, rapid progress in digital technologies has led to an increasing dependence on interconnected systems that support communication, enhance operational efficiency, and improve daily convenience. Amid this broad technological landscape, the Internet of Things (IoT) has become a key driver of change, influencing a variety of sectors such as healthcare, transportation, industrial automation, and logistics [1]. In the sphere of digital communication, this influence is evident in the progressive refinement of real-time interaction, which appears increasingly shaped by the imperative to deliver fluid, high-quality exchanges that meet the expectations of both personal and professional contexts. These improvements are evident across different areas and sectors, such as digital health services [2], helping patients and citizens living at home through video visits or consultations, remote monitoring and transmitting recorded information, as well as in governmental digital transformation [3], which has been shown to enhance efficiency through streamlined processes and improved interdepartmental coordination.

Over time, digital communication has undergone a marked transformation, moving away from early text-based platforms, such as Internet Relay Chat (IRC) and email, towards advanced video-conferencing technologies capable of enabling real-time, face-to-face interaction across international boundaries [4]. Advancements in network infrastructure, the broad implementation of cloud computing, and the deployment of AI technologies aimed at refining audio and visual quality have collectively supported this development [5].

The widespread adoption of digital communication tools has been further accelerated by global events, most notably the COVID-19 pandemic, which necessitated the rapid transition to remote work, online education, and virtual collaboration. Studies indicate that the pandemic led to a significant increase in the use of video conferencing platforms, highlighting both the necessity and limitations of current digital communication technologies [6]. In certain situations, it was harder to stay focused on work or academic tasks due to the increased home-based distractions, such as family demands and household tasks. Furthermore, employees tended to have more difficulty staying engaged in long virtual meetings than in long in-person meetings, prompting managers to deliberately shorten meetings to prevent employee overload during the transition to remote work [7].

Despite ongoing enhancements in network infrastructure and interface technologies, video communication still encounters considerable challenges. The lack of detailed non-verbal cues, such as fine gestures and spatial context, diminishes the richness of interactions, compelling users to invest more mental effort to understand emotional and communicative nuances. Additionally, the phenomenon known as “Zoom fatigue” captures the mental exhaustion resulting from prolonged exposure to unnaturally close visual proximity, sustained eye contact, and the pressure to remain constantly on camera. These factors, combined with limitations in interaction fluidity and technical constraints, highlight a critical need for more immersive and less cognitively demanding communication tools [8].

In summary, while digital communication technologies have enabled unprecedented levels of remote interaction, their current limitations — particularly the loss of nuanced emotional cues, increased cognitive strain, and diminished sense of presence — underscore the need for innovative solutions capable of bridging the experiential gap between virtual and in-person communication. Humans convey emotions through multiple channels, including facial expressions, gestures, body posture, and vocal tone, which together communicate affective states and intentions in social interactions [11], [12]. Recognizing and replicating these cues is crucial for digital systems aiming to create realistic and empathetic virtual characters.

In parallel with these interaction and usability challenges, the rapid deployment of chatbots and digital humans in digital communication platforms introduces a secondary, but equally critical problem space: privacy, security, and ethical concerns.

Chatbots, defined as automated conversational agents designed to interact with users via text or voice interfaces, have become widespread tools in customer service, task automation, and information retrieval across multiple industries [9], [10]. Their growing use is driven by their efficiency and availability; however, they often rely on collecting sensitive personal information to provide tailored services. This dependence on user data raises significant concerns about how this information is stored, processed, and protected, particularly considering increasing cyber threats and data breaches [11].

For example, chatbots dealing with financial transactions or private facts present attractive objectives for cyberattacks that might have excessive results for customers and businesses alike. While essential countermeasures which include data encryption, stable authentication, and continuous monitoring exist, many innovative implementations lack sufficient safeguards, exposing customers to potential privacy violations and safety breaches [12].

Beyond technical vulnerabilities, the rise of automated digital systems has sparked concerns regarding the erosion of personal interaction and the emergence of algorithmic biases. Users may feel uneasy interacting with systems lacking human empathy and the capacity to fully understand complex emotional or contextual cues. The inability to perceive subtleties such as tone, body language, or emotional undertones can lead to misunderstandings, frustration and further diminish user satisfaction and trust [13], [14].

Therefore, while chatbots and digital humans promise a more efficient and personalized service experience, their potential for privacy violations, security breaches, and the loss of human touch must be carefully addressed to ensure responsible, trustworthy deployment.

One promising direction for addressing these limitations lies in the adaptation of real-time rendering and avatar technologies originating from the gaming industry. Game engines, which have evolved from rudimentary graphics systems to advanced real-time rendering frameworks, now produce highly detailed virtual environments and lifelike digital characters [15].

The Unreal Engine, a leading real-time rendering engine, demonstrates the potential to create high-fidelity digital humans through its MetaHuman framework, which enables detailed character modelling and animation with realistic facial expressions and body language [16]. By leveraging these technologies, digital communication platforms can enhance user interaction, fostering more natural and engaging remote experiences.

In August 2017, the game development company *Ninja Theory* released *Hellblade: Senua's Sacrifice* [17], a groundbreaking 3D action-adventure game that made a significant impact on the gaming industry. The game garnered attention not only for its compelling narrative and innovative gameplay but also for its stunning, hyper-realistic graphics. These visuals, exemplified in Figure 1, set a new standard in real-time facial capture and MoCap technology, enabling ultra-realistic character performances that were unprecedented at the time.

Figure 1 – Hellblade: Senua's Sacrifice realistic character using MoCap and photogrammetry.

In May 2024, the long-awaited sequel, *Senua's Saga: Hellblade II* [18], was released, widely regarded as one of the most visually ambitious titles created with Unreal Engine 5 to date. The game displays the power of Unreal Engine 5's capabilities, pushing the boundaries of realism with highly detailed environments and character models. This time the developers incorporated Epic Games' MetaHuman framework (version 5.2) for advanced facial animations. The use of MetaHuman technology extended to the main character, Senua, whose facial rendering in the game was derived from the MetaHuman framework, taking her realism to new heights, as shown in Figure 2.

Figure 2 – Senua's Saga: Hellblade II main character from the Unreal Engine 5 MetaHuman Demo

In December 2023, Kojima Productions shared a reveal trailer for their upcoming horror game, *OD (Overdose)* [19], described by Hideo Kojima as an experimental “new form of media” that combines cinema with video games. The project utilizes Unreal Engine 5, as demonstrated in the promotional material [20], presenting high-fidelity, real-time actor captures consistent with MetaHuman and Unreal Engine 5 virtual production workflows. *OD* features characters based on well-known actors such as Sophia Lillis (Figure 3), Hunter Schafer (Figure 4) and Udo Kier (Figure 5), captured and presented in the trailers by translating live performances into digital character assets.

Figure 3 – Sophia Lillis' character from the upcoming video game OD (Overdose) reveal trailer.

Figure 4 – Hunter Schafer' character from the upcoming video game OD (Overdose) reveal trailer.

Figure 5 – Udo Kier character from the upcoming video game OD (Overdose) reveal trailer.

Through the video, we can see the three characters perform the same scripted scenario, yet each delivers it in a distinct emotional state. Sophia appears calm and relaxed, her tone suggesting quite satisfaction with the story. Hunter conveys fear, her body language and vocal delivery betraying an effort to control her emotions. In contrast, Udo delivers his lines with an intensity that exudes anger, bordering on intimidation, as if attempting to unsettle the listener. Towards the end of the trailer, Sophia Lillis’s demeanour shifts dramatically as she appears visibly frightened, culminating in a highly expressive facial reaction that conveys shock and vulnerability, as shown in Figure 6. This deliberate variation in emotional states illustrates how recent advancements in performance capture and real-time rendering technologies enable metahumans to replicate emotions with a striking degree of naturalness and realism, bridging the gap between digital characters and human expressiveness.

Figure 6 – Sophia Lillis' character conveying shock and vulnerability at the end of the trailer.

The integration of digital humans — avatars that replicate human appearance and motion with high fidelity — has gained traction in various fields, including entertainment, training simulations, and virtual reality experiences. Digital humans offer the potential to bridge the gap between traditional video calls and immersive interactions, enabling users to convey emotions, gestures, and social cues in ways that traditional 2D video conferencing cannot achieve. For instance, a study evaluating the use of digital humans in virtual co-presence environments highlights their effectiveness in enhancing user engagement and communication dynamics [21]. Moreover, research on intelligent interaction systems for digital humans further emphasizes their applications in virtual reality and HCI, supporting their role in creating more immersive and expressive communication platforms [22].

A crucial component in enhancing these experiences is MoCap technology, which enables precise tracking of human movements and expressions. This allows avatars to reflect real-time facial and body gestures [23]. The application of such technologies in video conferencing can significantly improve engagement, particularly in scenarios demanding higher expressiveness, such as remote collaboration, education, and customer interactions [24].

The incorporation of real-time 3D-rendered avatars and digital humans in communication platforms directly addresses the earlier identified limitations by increasing user engagement, reducing cognitive load, and restoring a sense of presence in virtual interactions. The application of game engines and real-time rendering in video conferencing represents an innovative step towards more natural and fluid remote communication. [25].

This study aims to investigate how Digital Humans, more precisely MetaHumans, are perceived by enhancing emotional expression and empathy in virtual environments. Specifically, it seeks to evaluate whether high-fidelity digital avatars can convey emotions effectively and to assess user engagement and empathy when interacting with these avatars.

To address these aims in a structured manner, the study is guided by the following series of research questions:

RQ1 - How does user perception of a Digital Human compare to that of a real person in terms of realism and emotional expressiveness?

RQ2 - How do users perceive the realism and emotional expressiveness of a Digital Human's body movements compared to those of a real person performing the same actions?

RQ3 - Can personalized facial animation of emotion increase the feeling of empathy compared to predefined facial animation?

RQ4 - How do users perceive the emotional authenticity and effectiveness of default facial expressions compared to custom-created facial expressions in a Digital Human?

RQ5 - What are the technical challenges and limitations in developing Digital Humans with empathy?

These questions form the basis for the study's working hypotheses, which anticipate how distinctive design choices in Digital Human creation may influence user perception and empathy.

1. Customized facial animations tailored to an individual's expressions will result in higher levels of user empathy compared to default animations provided by Unreal Engine.
2. The use of customized emotional expressions will lead to a stronger emotional engagement between users and Digital Humans, as reflected in both quantitative measures (e.g., TEQ scores) and qualitative feedback.

3. Users who experience interactions with Digital Humans exhibiting custom emotional expressions will report greater satisfaction and trust in the system than those interacting with Digital Humans displaying default expressions.
4. Although users will generally perceive the emotional expressions of a real person in video as more authentic, Digital Humans performing the same movements will still be perceived as emotionally engaging and realistic.

To assess these hypotheses and directly address the research questions, this study uses a mixed-methods research approach, integrating both qualitative and quantitative analysis. The research involves creating a high-fidelity digital avatar using Unreal Engine and MetaHuman. MoCap technology will be used to capture real-time facial expressions and body movements, ensuring realistic animation and natural interaction. The avatar will be designed to reflect human-like expressiveness and responsiveness, aiming to enhance the immersive experience in digital communication.

The research evaluates participants' emotional responses when observing and listening to both a real person and a Digital Human delivering the same content. It focuses on two key test scenarios: the first compares a Digital Human animated with custom facial expressions to one using default expressions provided by Unreal Engine; the second compares a real person's video to a Digital Human replicating the same body movements. By measuring the empathy felt by participants across these scenarios, the study seeks to understand how emotional authenticity and expression in Digital Humans influence user engagement, empathy, and perceived realism.

The study was divided into three phases. First, participants completed the TEQ to measure their baseline empathy levels. Second, they visualized a video of the digital avatar in a controlled environment. Finally, they completed a survey adapted from the TEQ to evaluate their experience, and statistical analysis was performed to assess changes in empathy levels and interaction quality.

The effectiveness of Digital Humans in fostering engagement and expressiveness in virtual interactions was assessed using multiple tools. The TEQ, a validated instrument for measuring empathy, was used before and after participants interact with the digital human. A PIS, adapted from the TEQ, explored how users' empathy levels change after visualizing the MetaHuman video. These questions were scored on a 5-point Likert scale to maintain consistency with the original TEQ.

The collected data was analysed using statistical methods, including linear regression analysis to examine relationships between empathy levels and user engagement with digital humans, the Shapiro-Wilk normality test to assess the normality of the dataset before applying further statistical analyses, and Pearson correlation to measure the strength and direction of relationships between variables, particularly focusing on empathy responses and the effectiveness of digital human interaction.

This master’s project was developed during a 24-month research grant managed and funded by CIIC — the Computer Science and Communications Research Centre, called METATEW – Metaverse The Easy Way.

The structure of the thesis is detailed as follows: Section 2 presents the background, outlining key concepts such as Digital Human conceptualization, game engine development, the Metahuman framework, MoCap techniques, and considerations about video conferencing platforms and privacy considerations. Section 3 reviews the related work in the field, while Section 4 details the methodology adopted for this research. Section 5 describes the development process, including the creation of the Digital Human, facial and body animation, lip sync integration and the development of the case study. Section 6 reports the tests and results, covering participant details, measurement instruments, procedures, provisional testing and the analysis of results from the TEQ and PIS. Section 7 presents the analysis and discussion, including data analysis methods, the discussion of findings concerning the research questions and their implications. Finally, Section 8 concludes the thesis by summarizing the main contributions and outlining directions for future work. By addressing these key areas, this project aims to contribute valuable insights into the potential of MetaHumans in digital communication, paving the way for more immersive and empathetic virtual interactions.

2. Background

This section presents a detailed overview of the key concepts and technological foundations relevant to the development and application of Digital Humans in the context of digital communication. It explores the evolution of related technologies, such as game engines, digital human creation tools, MoCap systems, and existing video communication platforms, providing the necessary background for understanding their integration into immersive communication solutions.

2.1. Digital Human Conceptualization

Digital Humans, sometimes referred to as virtual humans or synthetic avatar, represent a convergence of 3D modelling, animation, and AI to simulate human-like presence in digital environments [26]. These entities are designed to closely resemble real humans not only in appearance but also in behaviour, emotional expression, and interactivity. Unlike conventional avatars or game characters, Digital Humans aim to emulate the complexity of human communication, including verbal and non-verbal cues such as facial expressions, micro-gestures, tone of voice, and body language.

The implementation of Digital Humans spans various domains, including virtual customer service agents, educational tutors, brand ambassadors, healthcare assistants, and characters in virtual or augmented reality experiences. Their realism is achieved through detailed mesh sculpting, advanced shading and lighting models, and behavioural logic programmed through AI or controlled via human input. As emotional intelligence becomes a critical factor in HCI, the use of Digital Humans offers a promising avenue for creating more natural, intuitive, and empathetic user experiences.

2.2. Game Engine Development and Real-Time Rendering

Game engines serve as the foundation for building interactive digital environments. Over the past two decades, they have evolved from simple 2D rendering frameworks to sophisticated 3D engines capable of producing cinematic-quality visuals in real time. Examples include Unreal Engine, developed by Epic Games [27], Unity, developed by Unity Technologies [28], and several proprietary engines such as REDengine [29] (*Cyberpunk*

2077) [29], RE Engine (*Resident Evil*, *Monster Hunter*) [30] [31], and Decima Engine (*Horizon Zero Dawn*, *Death Stranding*) [32] [33].

While proprietary engines can produce highly realistic digital characters, they are closed source and not publicly available for broader experimentation, making them unsuitable for academic or independent research workflows. In contrast, Unreal Engine is freely accessible for non-commercial use, has extensive documentation, and offers a flexible licensing model, making it ideal for prototyping and deploying Digital Humans in research contexts.

Unreal Engine's capabilities have also been embraced beyond games, including virtual production, simulation training, architecture, automotive design, and immersive media experiences. The latest version, Unreal Engine 5, introduces several innovative technologies that significantly enhance the realism and responsiveness of digital environments. For instance, Nanite, a virtualized geometry system, allows developers to use high-poly models without performance degradation [34], while Lumen, a real-time global illumination system, simulates complex lighting interactions dynamically [35]. These features are particularly useful for creating realistic skin, eye reflections, and environmental lighting in scenes involving Digital Humans.

Game engines also play a key role in metaverse-oriented experiences. Unreal Engine has been used by fashion brands such as Balenciaga to create interactive runway experiences [36], while Unity has explored metaverse integration across industries [37]. Furthermore, notable examples of realistic digital characters in commercial games include Aloy in *Horizon Forbidden West* [38], Adam Jensen in *Deus Ex: Mankind Divided* [39], and the characters of *Cyberpunk 2077* [29], all demonstrating industry-level character fidelity that research in Digital Humans seeks to replicate or adapt.

Moreover, Unreal Engine's real-time capabilities make it possible to render scenes and characters at interactive frame rates, enabling seamless communication scenarios where digital avatars can respond to user inputs or live performance data. These capabilities are essential for deploying Digital Humans in dynamic communication contexts, such as video calls or virtual meetings, where latency and realism significantly impact user experience.

2.3. MetaHuman Framework

The MetaHuman Creator, launched by Epic Games, is a revolutionary tool that significantly lowers the barrier to creating highly realistic digital human characters [40]. This cloud-based application allows users to design, customize, and deploy digital avatars using an intuitive graphical interface. The MetaHuman framework offers a library of pre-rigged facial meshes, skin textures, hairstyles, and clothing options, all of which are procedurally generated to ensure both realism and performance optimization.

What distinguishes MetaHuman from traditional 3D character creation pipelines is its high degree of facial detail, compatibility with MoCap systems, and integration with Unreal Engine's animation and rendering tools. Characters created using MetaHuman are rigged with over 50 facial blend shapes and muscle movements, allowing for nuanced expression of emotions and speech.

In 2024, with the release of MetaHuman Animator as part of Unreal Engine 5.2, the framework became even more accessible. Users can now record facial performance using an iPhone's front-facing camera and automatically apply the data to MetaHuman characters. This innovation enables small teams and researchers to animate digital humans with impressive realism, supporting use cases such as virtual assistants, interview bots, and educational avatars without the need for expensive facial capture equipment. In Unreal Engine 5.6 (released in June 2025), the MetaHuman Creator became fully embedded within the engine [41].

2.4. Motion Capture (MoCap)

MoCap refers to the process of digitally recording human movements and translating them into animations that can be applied to 3D characters [42]. It is a fundamental technology for animating Digital Humans realistically and responsively. MoCap systems can capture a wide range of human behaviour — from facial expressions and hand gestures to full-body movements — and mapping them onto virtual models in real time or for offline animation [43].

There are several types of MoCap systems. Optical MoCap uses multiple cameras to track markers placed on a performer's body, providing highly accurate spatial data. These systems are typically used in professional studios due to their precision but require controlled environments [44]. Inertial MoCap, on the other hand, relies on wearable sensors embedded in suits or devices that detect acceleration and orientation. These systems are more portable and affordable but may lack the granularity of optical systems [45].

Recently, markerless facial capture using machine learning algorithms and standard camera input has gained popularity. This approach enables facial animation using readily available hardware, such as webcams or smartphones, and has become particularly useful for applications where cost and accessibility are important [46].

In the context of this research, MoCap plays a vital role in transferring real human expressions and movements to the MetaHuman character. This ensures that the digital human responds naturally and convincingly, enhancing immersion and emotional connection in digital interactions.

2.5. Video Conferencing Platforms and Privacy Considerations

Video conferencing platforms have become ubiquitous tools for communication in personal, educational, and professional settings. Their utility was particularly highlighted during the COVID-19 pandemic, which necessitated widespread remote collaboration [47]. Platforms like Zoom [48], Microsoft Teams [49], Google Meet [50], and Cisco Webex [51] provide a suite of functionalities, including screen sharing, real-time chat, background filters, and collaborative whiteboards, designed to support efficient communication across geographical boundaries.

However, despite their widespread use, these platforms are often limited in conveying the full spectrum of human interaction. Flat video feeds can hinder non-verbal communication, leading to issues like misinterpretation, reduced engagement, and "Zoom fatigue" [52]. Research attributes "Zoom fatigue" to nonverbal overload — our brains working harder due to prolonged eye contact, self-monitoring via self-view, hyper-gaze, and heightened cognitive demand from interpreting limited non-verbal cues [53], [54].

Integrating Digital Humans into video conferencing platforms could address many of these challenges by offering customizable avatars that reflect a user's real-time expressions, gestures and emotions. Similar technologies are already being explored in other domains, such as AI-generated news anchors [55], [56], which replicate human presenters using real-time or pre-rendered animation. While these systems display the potential of lifelike digital humans in broadcasting, they also highlight certain limitations — such as delays in real-time rendering and constraints in replicating spontaneous interaction — that could similarly impact live video communication.

However, such integration introduces new concerns around privacy and data security. Real-time facial tracking and emotion analysis require access to biometric data, raising questions about consent, data storage, and misuse.

To mitigate these risks, several measures are being adopted, including end-to-end encryption, anonymized data processing, and opt-in tracking features. Nonetheless, the ethical implications of using AI to analyse and reproduce human behaviour in digital spaces remain a topic of ongoing debate. Any deployment of Digital Humans in video communication systems must consider both the technical and social dimensions of user privacy, trust, and control.

3. Related Work

In the context of applying Digital Humans for communication, either for personal or professional purposes, establishing an empathic connection with the avatar is crucial for enabling meaningful and realistic engagements. This is particularly relevant for custom MetaHumans, where authenticity in emotional expression directly influences user trust and engagement. Research in this field has explored both the emotional and cognitive components of empathy, examining factors such as emotion recognition, facial mimicry, voice realism, and social presence.

3.1. Studies with user testing

Kroes et al. [57] (2022) evaluated empathic responses towards a Virtual Character (VC) using a Post-Experiment Survey adapted from the TEQ. Participants first identified the emotion displayed by the VC (sadness, in all cases) then rated their agreement with statements on emotional contagion (EmCon), emotion understanding (EmUnd), sensitivity (Sens), sympathetic physiological arousal (SympPhy), and altruism (AltEmp). While curiosity about the cause of the VC’s emotion and willingness to help were consistently high, emotional contagion and physiological arousal varied between participants. The authors also investigated the effect of “personification” stories on empathy, finding mixed results, and explored whether awareness that the VC could not feel emotions influenced responses.

Loveys et al. [58] (2022) examined empathy towards Digital Humans through user tests assessing affective, cognitive, and somatic components. Participants, recruited via Amazon Mechanical Turk and university advertisements, completed the Interpersonal Reactivity Index (IRI) at baseline to measure trait empathy. They then viewed videos of Digital Humans either displaying emotional expressions or remaining neutral, rating empathy and of appropriate responses via visual analogue scales and Likert items. The study highlighted the role of multimodal cues — speech, gestures, facial expressions — in eliciting empathic responses and suggested that these cues should be carefully modelled in custom MetaHumans.

Higgins et al. [59] (2021) explored the influence of synthetic voice realism on affinity, social presence, and empathy in photorealistic virtual humans. Over 200 participants took part in VR-based experiments, viewing characters whose voices either matched or mismatched their visual realism. Emotional responses (e.g., concern, calmness), comfort levels, and affinity were measured via questionnaires. Results showed that incongruent realism between voice and appearance negatively impacted perceived empathy and trust, underscoring the need for natural-sounding voices in immersive Digital Human applications.

McQuiggan and Lester [60] (2007) evaluated automated empathy systems in virtual game characters, comparing them to human-controlled expressions. Using behavioural observation and user questionnaires, the authors found that automated empathy responses could be as effective as those provided by human operators. This finding points towards scalable approaches for implementing empathic behaviours in Digital Humans without requiring continuous human intervention.

3.2. Studies without user testing

Holland et al. [61] (2020) conducted a meta-analysis of correlations between facial mimicry, empathy, and emotion recognition across numerous studies. Facial mimicry — the automatic imitation of another's facial expression — was found to be a key driver of both affective empathy (sharing an emotional state) and cognitive empathy (understanding without sharing the state). For Digital Humans, these findings imply that mimicking users' expressions can foster more authentic and engaging interactions.

Kreijns et al. [62] (2021) reviewed and redefined the concept of social presence, separating it from technological determinism and linking it to sociability (an attribute of the medium) and social space (an attribute of the group). Their framework, while theoretical, has practical implications for empathy modelling by stressing the role of perceived authenticity in virtual communication — critical for fostering trust in MetaHumans used in customer service or collaborative settings.

Jiang et al. [63] (2023) proposed mFERMeta++, a robust multi-view facial expression recognition (FER) system using MetaHuman models and meta-learning. Although no human participants were involved, the system was evaluated on the Karolinska Directed Emotional Face (KDEF) dataset to classify seven basic emotions. The authors compared traditional handcrafted feature extraction methods such as Gabor texture analysis [64] and Local Binary Patterns (LBP) [65] with deep learning approaches (CNNs), noting that while the latter excel in controlled settings, performance in real-world conditions remains a challenge due to variations in race, gender, age, and head pose [66], [67].

3.3.Technological trends and challenges

Across these studies, the evaluation methods range from controlled laboratory experiments with immersive VR [59] to large-scale online surveys [58] and secondary statistical meta-analyses [61]. Technologies span deep learning for FER [63], multimodal perception systems incorporating speech and gesture recognition [58], and emotion contagion modelling through behavioural rules [60]. Common measurement tools include the TEQ [57], IRI [58], and bespoke empathy rating scales [59].

While promising results have been achieved — such as the ability for automated systems to match human-level empathy expression [60] — several gaps remain. Most notably, there is limited investigation into somatic empathy (physiological responses) in Digital Human interaction [68], [69], [70] and challenges persist in adapting high-accuracy emotion recognition models to unconstrained, real-world contexts. Addressing these gaps will be essential for developing MetaHumans capable of consistent, authentic, and context-appropriate empathic engagement.

3.4.Tools to evaluate empathy

Several standardized instruments are frequently used to measure empathy in human–computer interaction studies, each with different emphases and application contexts. In the reviewed works, the main tools were the TEQ, Post-Experiment Surveys adapted from TEQ, and the IRI.

- **Toronto Empathy Questionnaire**

The TEQ, developed by Spreng et al. (2009), is a unidimensional self-report instrument designed to assess an individual's general empathic tendencies. It comprises 16 items rated on a 5-point Likert scale, producing a total score between 0 and 64, with higher scores indicating greater empathy. While it does not separate empathy into subcomponents, its items indirectly capture emotional contagion, emotion understanding, sensitivity, and sympathetic physiological arousal.

In the reviewed literature, Kroes et al. [57] (2022) adapted the TEQ as part of their experimental design to establish a baseline for empathic disposition and to inform the structure of their Post-Experiment Survey. In previous studies, general population averages tend to cluster between 40 and 45, providing a benchmark for interpreting participant scores.

- **Post-Experiment Survey**

In [57], the TEQ was adapted into a task-specific post-experiment survey to assess state empathy, the empathic response triggered by interaction with a Virtual Character (VC). The survey began by asking participants to identify the emotion expressed by the VC, followed by Likert-scale items targeting six key components: emotional contagion (EmCon), emotion understanding (EmUnd), sensitivity (Sens), sympathetic physiological arousal (SympPhy), altruism, and higher-order empathic behaviour (AltEmp).

This approach allowed researchers to distinguish between the recognition of an emotion and the depth of empathic engagement, as well as to explore moderating factors such as the VC's perceived emotional authenticity. It also offered a fine-grained analysis of how different empathy dimensions are elicited in controlled digital interactions.

- **Interpersonal Reactivity Index (IRI)**

The IRI, developed by Davis (1980), is a 28-item self-report questionnaire that measures trait empathy across four subscales: Perspective Taking (PT), Empathic Concern (EC), Fantasy (FS), and Personal Distress (PD). Each subscale targets a distinct cognitive or affective dimension, enabling a multidimensional empathy profile.

In Loveys et al. [58] (2022), the IRI was administered as part of a baseline questionnaire before participants engaged with Digital Human videos. By correlating IRI scores with experimental outcomes, the authors were able to examine how pre-existing empathic tendencies influenced perceived empathy, appropriateness of responses, and emotional engagement. This combination of trait measurement (IRI) and state evaluation (visual analogue scales, Likert items) provided a richer understanding of participant reactions than either measure could alone.

4. Methodology

This study aims to investigate the role of emotionally responsive Digital Humans in fostering greater user engagement and empathy. The methodology followed a structured five-phase approach, moving from literature review to proposal and requirements, through development, user testing, and statistical analysis. An overview of the workflow is presented in Figure 7.

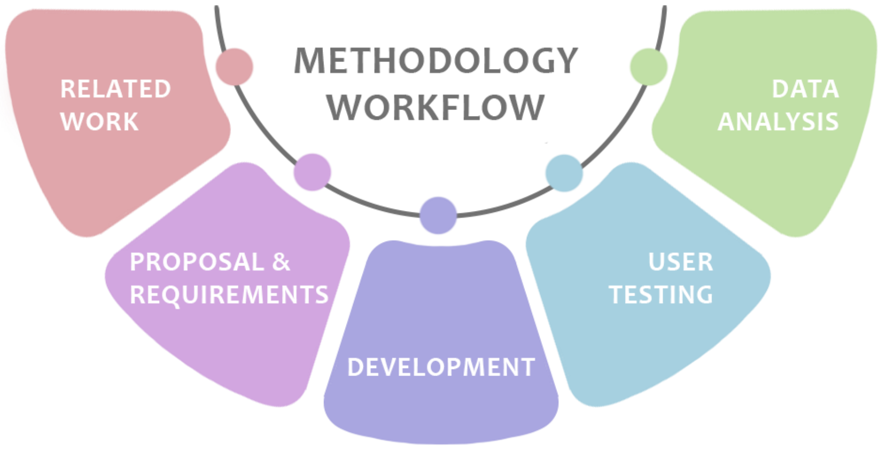


Figure 7 – Methodology Workflow

4.1. Phase 1 — Related Work Review

The first phase involved a targeted literature review to identify existing research, technologies, and evaluation methods relevant to emotionally expressive Digital Humans.

- **Search Strategy:** Articles were collected using specific keywords: Avatars (2 articles), Digital Humans (5), Virtual Humans (1), Photorealistic Humans (3), Virtual Reality (4), FACS (1), Facial Emotion Recognition (3), Uncanny Valley (2), Metaverse (2), and Virtual Environments (4).
- **Data Extraction:** Each study was recorded in an Excel database divided into three categories:
    - **Content** (Figure 8)**:** year, important topics, similarities with this project, relevance.
    - **Technology** (Figure 9)**:** engine used, capture technologies, emotional customization, motion files/databases and real-time capture.

- **Evaluation** (Figure 10)**:** evaluation parameters, methods, user tests, tools and results.
- **Selection Criteria:** Priority was given to studies including user tests, especially those incorporating emotional states (few met this criterion).

The review identified a lack of empirical research measuring empathy in interactions with Digital Humans, particularly when comparing default and customized emotional expressions.

Content					
Order	**Year**	**Title**	**Important Topics**	**Similarities**	**Why it is relevant**
1	2021	Sympathy for the digital: Influence of synthetic voice on affinity, social presence and empathy for photorealistic virtual humans	o Explores emotional responses and social presence of users towards characters expressing strong emotions. o Provides insights into the impact of synthetic voices on perception and empathy towards virtual characters.	o Focuses on understanding how perceptual cues affect user interaction with virtual characters. o Evaluates emotional responses and social presence, which are crucial aspects of creating believable expressions for metahumans. o Examines the influence of voice realism, akin to the importance of lifelike expressions in metahuman customization.	o Offers valuable insights into optimizing realism and expressiveness in virtual characters. o Highlights the significance of maintaining consistency between appearance and voice to enhance user engagement. o Suggests avenues for improving synthetic voice synthesis to match the visual fidelity of metahumans.

Figure 8 – Related Work Review excel with "Content" category.

Technology				
Engine (if applicable)	**Capture Technologies**	**Emotion costumization (yes or no)**	**Used motion files /database (type, origin, ...)**	**Realtime capture (yes or no)**
o Not mentioned	o Facial tracking markers o Sensors in VR headsets o Monocular RGB cameras o Monocular RGB+D cameras o Other RGB+D sensors (e.g., Intel Realsense, Structure Sensor, Microsoft Kinect Azure) o Multi-view capture in HMCs	o Investigates emotional responses to virtual characters expressing different emotions. o No specific mention of emotion customization features.	o Data from social media and video conferencing for facial capture and development of Codec Avatar models	o Not mentioned

Figure 9 – Related Work Review excel with "Technology" category.

Evaluation				
What is evaluated? (parameters)	**How is it evaluated? (methods)**	**User Tests**	**Tool**	**Results**
o Emotional responses (e.g., Concern, Excited, Afraid, Calm). o Social presence. o Comfort level and affinity towards characters. o Voice realism and its impact on perception.	o Conducted experiments with over two hundred participants in Virtual Reality (VR). o Utilized ANOVA analysis for evaluating subjective responses. o Post-hoc tests (e.g., Tukey's HSD) for further analysis of results.	Yes	o Text-To-Speech system for generating synthetic voice clips. o Participants were exposed to different scenarios involving virtual characters with manipulated voice conditions. o The scenarios likely varied in emotional content (e.g., sad, friendly, unfriendly) to elicit different emotional responses from participants. o Two experiments were conducted: one in immersive VR and another in a screen-based environment. o Questionnaires were administered to collect data on participants' responses to the scenarios and characters.	o Synthetic voice lowered social presence but did not significantly impact empathy or concern levels. o Voice realism influenced emotional responses and appeal of characters. o No significant effects of voice or scenario on discomfort were found. o Appearance realism of characters had medium ratings, with behavior realism slightly higher. o Synthetic voices were considered unappealing with photorealistic characters, affecting likeability. o Expressiveness in synthetic voices could enhance social

Figure 10 – Related Work Review excel with "Evaluation" category.

4.2. Phase 2 — Proposal and Requirements Analysis

From the literature review, key requirements for the study were defined:

1. **Empathy Measurement:** Adoption of the TEQ, chosen for its reliability and suitability for short-form empathy assessment.
2. **Emotional Expression Framework:** Initially planned to follow the Facial Action Coding System (FACS) with seven emotions but reduced to two positive emotions — happiness and surprise — due to scenario constraints (narration by the MetaHuman).
3. **Evaluation Structure:** Comparative testing between default Unreal Engine facial animations and customized facial animations based on real human expressions; and between real human body movement and MetaHuman body movement replicating the same actions.
4. **Mixed-Method Approach:** Combining quantitative data (questionnaire scores) with qualitative feedback to gain a holistic understanding of user responses.

4.3. Phase 3 — Development Process

The development process followed two distinct but related pipelines — one for facial animation (User Test 1) and one for body animation (User Test 2).

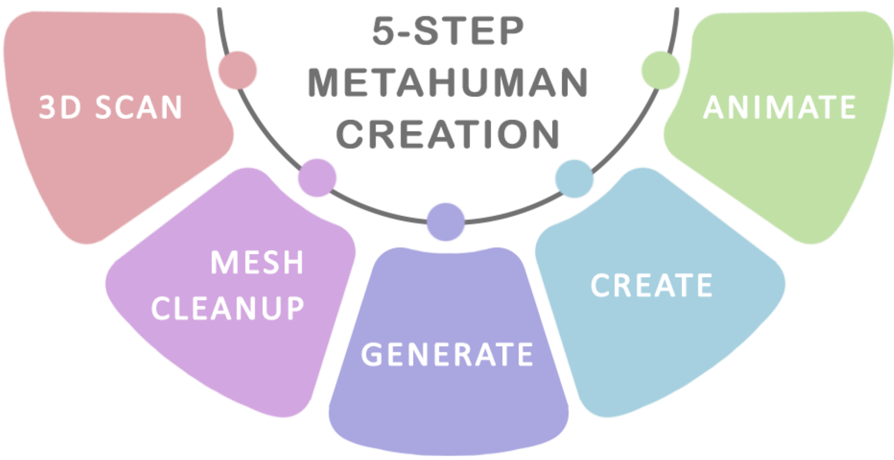


Figure 11 – Steps to MetaHuman creation.

The steps involved in creating a MetaHuman can be summarized in 5 steps, as depicted in Figure 11:

1. **3D Scan:** The participant's head was scanned using the Polycam mobile app (iPhone 11).
2. **Cleanup:** Captured meshes were refined in Blender to remove unwanted elements and improve topology.
3. **Generate:** In Unreal Engine, using the MetaHuman plugin, use the "MetaHuman Identity" to create a MetaHuman from the model.
4. **Create:** Visit the "MetaHuman Creator" website to customize the model previously generated.
5. **Animate:** Facial and body movements recorded using the Live Link Face app and Rokoko Studio.

Two customized MetaHumans (one for each scenario) were developed, each based on a different real person.

4.4. Phase 4 — User Testing

Two independent controlled user tests were conducted, each with 20 participants (40 total), with no overlap between groups.

Procedure:

1. **Demographic Survey:** Participants provided basic personal information (age, gender, academic background) and stated the frequency of interacting with virtual environments.
2. **Pre-Interaction Empathy Assessment:** The TEQ was administered using a 5-point Likert scale.
3. **Stimulus Presentation:**
    a. **User Test 1**: Group A viewed a MetaHuman with default Unreal Engine expressions; Group B viewed a MetaHuman with customized expressions.
    b. **User Test 2**: Group A viewed a real human speaking; Group B viewed the MetaHuman replicating the same face and body movements and narration.
4. **Post-Interaction Survey:** The PIS was designed based on the same logic established by authors K Kroes, I Saccardi and J Masthoff, adapting the TEQ to the study in question [57] to explore how users' levels of empathy may change after the

visualization of the MetaHuman video. While the original TEQ contains 16 questions, this adapted version includes only 5, selected to represent key elements of empathy. These questions are scored on a 5-point Likert scale, similar to the TEQ.

To maintain balance, 2 questions are written in a positive tense (where responses are scored from 0 to 4), while 3 questions are written in a negative tense (where responses are scored from 4 to 0). This structure aligns with the TEQ's logic, enabling a maximum possible score of 20 for the adapted survey. However, it is important to note that this adaptation has not been formally validated, and as such, precise conclusions akin to those derived from the original TEQ cannot be drawn.

The results of this survey are best used in comparison to participants' initial TEQ scores, collected at the beginning of the user tests. This approach helps us understand how empathy levels may shift after viewing a MetaHuman video. Specifically, it enables us to evaluate whether users demonstrate greater empathy toward a MetaHuman character with custom facial expressions compared to one with default facial expressions.

While the TEQ provides benchmarks and averages (e.g., 40-45 out of 64, representing moderate empathy in general populations), such comparisons are not feasible with the adapted survey due to its reduced scope and unverified status. Instead, this tool serves as an exploratory measure to observe relative changes in empathy between pre- and post-test scenarios, offering insights into the impact of facial expression customization on empathetic responses.

All questionnaires and surveys were administered digitally via Google Forms.

4.5.Phase 5 — Data Analysis

The collected data were analysed using mixed quantitative methods:

- **Shapiro–Wilk Test:** Used to assess the normality of the PIS score distribution within each experimental group.
- **Pearson Correlation:** Measured the strength and direction of the relationship between empathy levels TEQ and the PIS within each group.
- **Linear Regression with ANOVA output:** Performed within each group to examine whether empathy levels (TEQ) significantly predict perceived interaction quality (PIS). The regression ANOVA table reports an F-test that evaluates whether the

model explains a considerable proportion of variance in the dependent variable compared to a model with no predictors. Although similar in outcome to the Pearson correlation for a single predictor, regression additionally provides the slope (B) and R^2 values, offering interpretable effect sizes in the original units.

- **One-way ANOVA:** Conducted separately for each user test to compare the mean PIS scores between the two experimental groups (Default vs. Custom facial expressions; Human vs. MetaHuman body movement). With only two groups per test, this analysis is equivalent to an independent samples t-test and determines whether the differences in perceived interaction quality between conditions are statistically significant.

All statistical analyses were performed in SPSS, with an alpha level of 0.05 used to determine significance.

5. Development

This section outlines the methodologies utilized in the creation of a highly interactive and emotionally responsive Digital Human, leveraging advanced tools and techniques to address the emotional limitations of traditional automated systems. By employing technologies such as Unreal Engine, MetaHuman Creator, and Live Link Face, this study focused on designing Digital Humans capable of delivering nuanced emotional expressions and fostering deeper user engagement. The development process was driven by the hypothesis that customized facial animations, tailored to reflect authentic human emotions, could significantly enhance empathy and trust in HCI.

5.1. Equipment

To achieve the objectives of this case study, a wide range of tools and software were used to design and evaluate Digital Human characters. This involved utilizing cutting-edge equipment and sophisticated software applications to ensure that the development and testing process is comprehensive and effective.

Hardware:

- **Laptop:** The primary device used for the development and animation of Digital Humans was an MSI laptop featuring an NVIDIA GeForce RTX 3060 GPU, an 11th Generation Intel Core i7-11800H processor, 32GB of RAM and 1TB NVMe SSD storage. This high-performance laptop is chosen to ensure seamless handling of resource-intensive tasks, especially real-time animation and rendering. Its powerful specifications enable efficient and effective creation and manipulation of Digital Humans.
- **Webcam:** Integrated MSI laptop webcam, used in conjunction with the Rokoko web-based MoCap system to capture body movement through two-camera tracking during an experimental phase of the project.
- **Smartphone:** In addition, an iPhone 11 with iOS 17.4 was used to capture facial motion. The iPhone 11's TrueDepth front camera enables the capture of intricate facial expressions, which played a crucial role in creating lifelike and emotionally expressive Digital Humans.

Software:

- **Polycam:** Used to capture 360-degree photographic scans, which were exported as 3D models for use in the project.
- **Blender:** To clean and refine the geometry of 3D models generated in *Polycam*, removing unwanted elements and preparing assets for integration into Unreal Engine.
- **Unreal Engine:** Unreal Engine 5.3 was the primary one for creating and animating MetaHuman characters. This version of Unreal Engine provides advanced tools for real-time rendering, facial animation, and integrating MoCap data.
- **MetaHuman Creator:** Integrated within Unreal Engine 5.3, MetaHuman Creator was also used, allowing the design and customization of highly realistic Digital Humans. This tool was used to fine-tune the facial features and expressions of the MetaHumans to ensure they can effectively communicate emotions.
- **Live Link Face:** This application, in conjunction with the iPhone 11, provided the capability to record videos that specifically captured facial expressions. These videos were then transmitted to Unreal Engine, where the facial movements were synced with the MetaHuman character. This process ensured that the character's animations were not only accurate but also realistic, resulting in a lifelike portrayal.
- **Rokoko Studio (Web):** Used in two ways: (1) experimentally with the integrated laptop webcam for real-time multi-camera body tracking, and (2) to process pre-recorded video footage into body animation for integration with the MetaHuman characters.

This combination of tools and workflows enabled the capture of both facial and body movement data, with varying approaches explored to assess their effectiveness in creating realistic and emotionally expressive Digital Humans.

5.2. Digital Human Creation

To create a custom Digital Human based on a real person, we began by creating a 3D model of the individual using the *Polycam* application [71], as shown in Figure 12. This application uses sensors to scan objects and capture full 360-degree photos which can be exported in various file formats.

Figure 12 – 3D model created from the *Polycam* application.

After exporting the 3D model, the next step was to refine it using 3D modelling software. During this stage, it was important to clean up the geometry by removing any unwanted parts of the model, keeping only the frontal part of the face. Additionally, any parts of the model containing hair were removed, as shown in Figure 13, as they would be recreated later.

Figure 13 – 3D model before and after cleanup on Blender.

To start creating the Digital Human, we first needed to select an Unreal Engine version that supported the "MetaHumans" plugin. Once the version is chosen, the model was imported into the project, and a "MetaHuman Identity" component was created for it, as shown in Figure 14 – MetaHuman Identity component with the 3D model of the face cleaned

. This feature served as an important link between the imported 3D model and the MetaHuman system in the Unreal Engine. It allowed them to manipulate MetaHumans' characteristics and characteristics such as visual appearance, actions, and interactions.

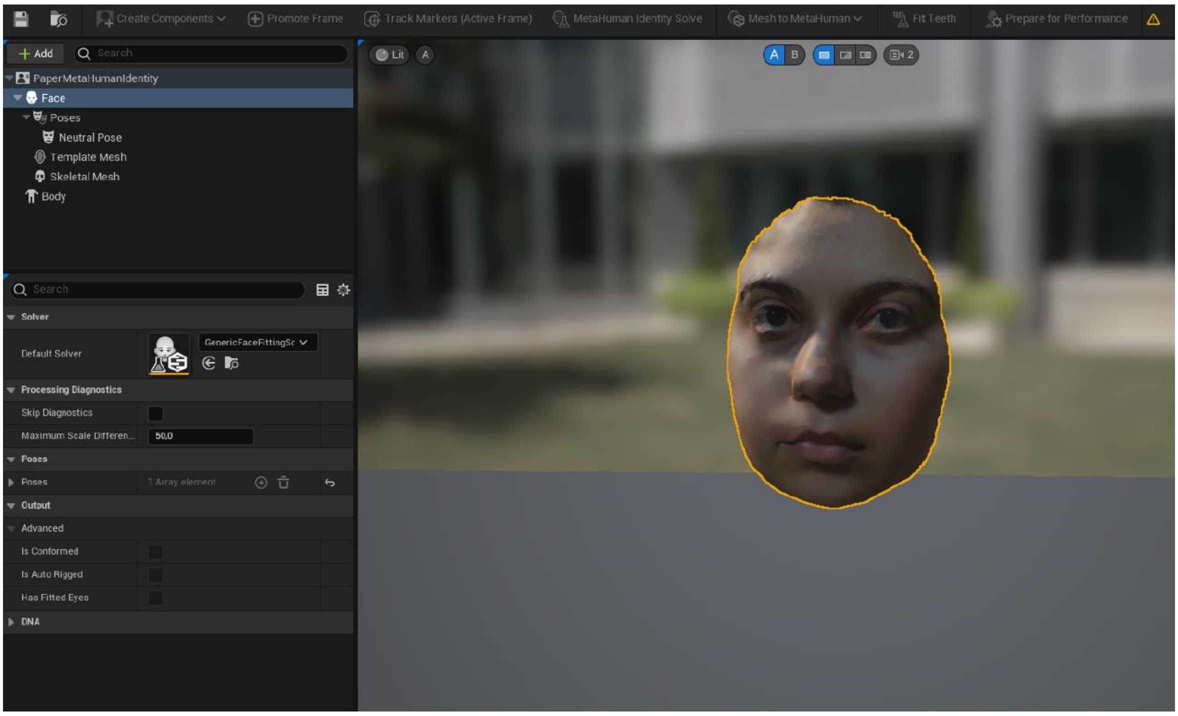


Figure 14 – MetaHuman Identity component with the 3D model of the face cleaned up.

This component served as the foundation for building the MetaHuman and mapping its various elements. From this point, the engine used the 3D model provided to display facial markers and create a visual representation of how the MetaHuman would look like, shown in Figure 15 – Visual representation of the MetaHuman .

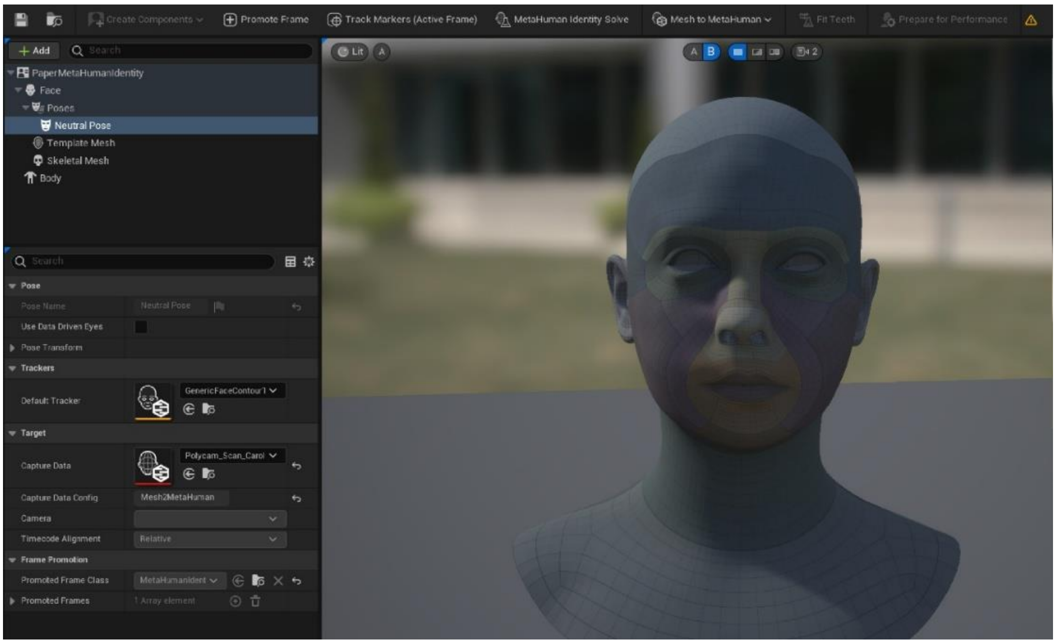


Figure 15 – Visual representation of the MetaHuman mesh.

It was important to evaluate how similar the created mesh was to the provided 3D model by overlapping the two. As shown in Figure 16 – MetaHuman Identity mesh , the best way to do this was by comparing the side profiles of both the mesh and the 3D scan, paying close attention to the contours and carefully assessing the accuracy of the created mesh in relation to the provided 3D model. If there were significant differences, adjustments to the facial markers needed to be made, and a new mesh should have been generated.

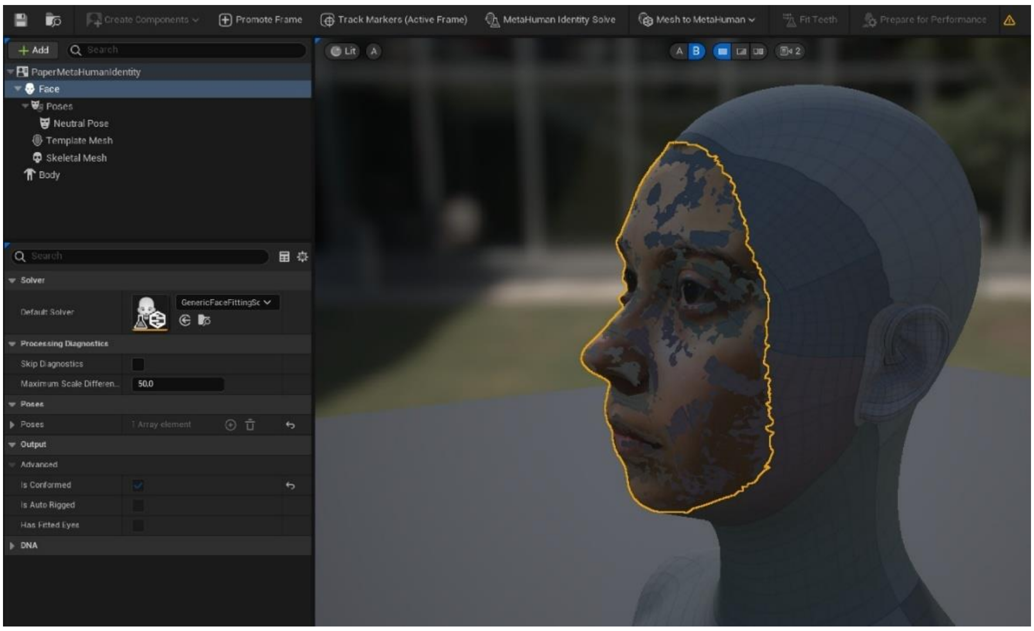


Figure 16 – MetaHuman Identity mesh comparison.

Finally, to complete the process of creating a MetaHuman, we needed to visit the "MetaHuman Creator" website. In Figure 17 – MetaHuman on the "MetaHuman Creator" , it is possible to see that this website provided us with the base MetaHuman created in Unreal Engine, which can then be customized to match the user's characteristics. Once the final model was complete, it was exported from the website and imported into Unreal Engine.

Figure 17 – MetaHuman on the "MetaHuman Creator" website.

5.3.Facial Animation

Upon completion of the customization and setup of the MetaHuman, it was confirmed that the rigging controllers were precisely aligned with the FACS. During the examination, it was possible to conclude discovery that the rigging was purposefully and meticulously developed with the FACS in mind. This intentional design ensured that the range of facial expressions was deliberately limited to prevent any potential distortion of the avatar's features, thus maintaining a prominent level of fidelity and realism in the facial animations.

To thoroughly test the suitability of the rigging, as shown in Figure 18 – MetaHuman displaying FACS , a menu was designed with seven buttons, each representing a different FACS emotion (anger, fear, disgust, contempt, joy, sadness and surprise). When a button was pressed, it triggered a transition between the current facial expression and the selected one, utilizing the facial expression database provided by Unreal Engine. This comprehensive test allowed us to evaluate the MetaHuman's ability to convey realistic emotions effectively. We then made the decision to create custom facial expressions to further evaluate the MetaHuman, and we also experimented with the transition between animations and adjusted its exponential curve.

Figure 18 – MetaHuman displaying FACS emotions.

To capture more nuanced and realistic facial expressions, the Live Link Face application was used to record custom expressions. This approach resulted in a significantly more authentic and natural appearance. Additionally, a fine-tune calibration was made within Unreal Engine to enhance the realism, bringing the virtual characteristics of the MetaHuman closer to real-life qualities. This calibration allowed the data transmitted from the "Live Link Face" application to be integrated more seamlessly and realistically within Unreal Engine. Furthermore, we incorporated body and eye animations to further augment the naturalness of the MetaHuman's overall performance.

5.4.Lip Sync Animation

To make a MetaHuman say a specific phrase, the "Live Link Face" application was utilized, which enabled the precise capture of facial movements and audio synchronization as someone spoke the exact words. This process ensured that the facial animation and audio were seamlessly integrated, resulting in an authentic outcome.

After that, the project focused on improving the MetaHuman custom animation to better match the conversation context. This involved adjusting various parameters, such as the transition between animations (particularly the fade-in and fade-out) and the exponential curve of the animation to customize the intensity and speed of specific moments. The objective was to guarantee a smooth and natural integration of the animations with the conversation flow.

5.5.Body Animation

To enhance the realism and expressiveness of the MetaHuman beyond facial animation, full-body MoCap was implemented using the Rokoko MoCap system. The setup involved the use of two cameras, strategically placed to accurately track and record the movements of a human actor. This method enabled the capture of natural and fluid body gestures, which are essential for conveying believable emotional cues and improving the overall empathetic impact of the digital performance.

Once the motion data was captured, it was imported into Unreal Engine, where a skeleton retargeting process was required to apply the animations to the MetaHuman. Initially, this retargeting was performed manually using Unreal Engine’s built-in tools. However, due to structural differences between the Rokoko actor skeleton and the Unreal Engine mannequin, the results were imprecise; movements lacked natural alignment and some body parts showed minor dislocation.

To address these issues, an automated retargeting feature was subsequently employed. While this approach caused a slight scaling adjustment to the MetaHuman mesh, it significantly improved the accuracy and fidelity of the movement. The captured gestures were better preserved and more effectively translated onto the digital character, resulting in smoother and more realistic body animations.

This integration of MoCap data, alongside calibrated facial expressions, played a vital role in enhancing the overall naturalness and credibility of the MetaHuman, especially in emotionally driven narrative sequences.

5.6.Development of the Case Study

The primary objective of this case study is to develop and evaluate customized Digital Humans equipped with a range of facial and bodily expressions, aiming to enable more emotionally resonant interactions with users. The study investigates the potential of Digital Humans to simulate human-like empathy and emotional responses to address the common challenges of user detachment and lack of trust in automated services.

To assess the impact of different animation approaches on user perception and empathy, two distinct user tests were conducted:

- **Facial Expression Test**

A narrated video interaction was created featuring a Digital Human displayed from the chest up. The video presented a short emotional narrative while participants focused on the Digital Human's facial expressions. Two versions of the animation were compared: one using Unreal Engine’s default motion-capture-based facial expressions, and another using customized facial animations designed to enhance emotional clarity and authenticity. This setup allowed for a controlled assessment of how the style and fidelity of facial animation influence user empathy and engagement.

- **Body Movement Test**

To explore the role of full-body expressiveness, a second test evaluated two conditions: one using a real human actor performing expressive body movements, and another using a MetaHuman animated with high-fidelity MoCap data replicating similar movements. Both Digital Human and real-human performances were shown in video format, with participants assessing the emotional connection and interaction quality. This test aimed to determine how realistic body movement contributes to perceived empathy in digital interactions.

Throughout both experiments, user responses were measured using standardized empathy questionnaires and qualitative feedback to evaluate the effectiveness of the Digital Human interactions.

6. Tests and Results

After completing the case study, we conducted two separate user tests to evaluate the impact of different animation techniques on user perception and empathy. Each test involved recruiting participants, dividing them into two groups, and presenting each group with a customized Google Forms containing four structured stages. The first test focused on facial expressions, comparing default Unreal Engine animations with custom-designed facial animations. One of the groups was presented with the test presented in the attachment named "UserTest 1 - Custom Facial Expressions - Google Forms" while the other group was presented with the test presented in the attachment named "UserTest 1 - Default Facial Expressions - Google Forms".

The second test examined body movement, comparing real human motion with motion-captured MetaHuman motion. The results from both tests were then collected, compared, and analysed to assess the emotional effectiveness and realism of the Digital Human interactions. One of the groups was presented with the test presented in the attachment named "UserTest 2 - Human Body Movement - Google Forms" while the other group was presented with the test presented in the attachment named "UserTest 2 - Metahuman Body Movement - Google Forms".

6.1. Participants

Two separate user tests were conducted as part of this study, each focusing on a different aspect of Digital Human expressiveness: facial expressions and body movement. Each test involved a distinct group of participants recruited via convenience sampling. All participants were Portuguese, ensuring language consistency throughout the study.

6.1.1. User Test 1 – Facial Expressions

For the first user test, which focused on facial expressions, a total of 20 participants (16 men and 4 women) were enrolled, with ages ranging from 21 to 59 years. Regarding their interaction with virtual environments, none reported never interacting with such environments, 31% indicated they rarely did, 6% sometimes, 37% frequently, and 25% reported always interacting with virtual environments. Participants were randomly assigned to one of two groups of 10. One group was shown a MetaHuman with default emotional facial animations provided by Unreal Engine, while the other group viewed a version of the same character with custom emotional facial expressions, captured through motion tracking.

6.1.2. User Test 2 – Body Movement

For the second user test, which examined the impact of body movement, another group of 20 participants was recruited. Their ages ranged from 20 to 35 years. In terms of interaction with virtual environments, none of the participants reported never engaging with them, while 20% said they rarely interacted, 5% sometimes, 47% frequently, and 28% always interacted with virtual environments. As with the first test, participants were divided into two equal groups. One group watched a MetaHuman using digitally generated MetaHuman body animation, while the other group experienced the same narration featuring body movement captured from a real human actor using MoCap technology.

6.2. Measures

This section outlines the user tests conducted to assess participants' empathetic responses to various facial animations displayed by Digital Humans. The tests were designed to evaluate how participants' empathy levels were influenced by viewing both default and customized facial animations from Unreal Engine. Empathy was measured by comparing participants' scores on the TEQ, administered before watching the video, with results from an adapted version of the same questionnaire given after watching the video.

6.2.1. Toronto Empathy Questionnaire (TEQ)

The TEQ is a validated assessment tool utilized to evaluate the empathic tendencies of individuals quantitatively. A higher score on the TEQ is indicative of a heightened level of empathy.

6.2.2. Post-Interaction Survey

The post-interaction was used to evaluate participants' empathetic responses to the MetaHuman character. Participants were tasked with identifying the emotions exhibited by the character and assessing their level of agreement with empathy-related statements using a 5-point Likert scale. Additional inquiries delved into the influence of the MetaHuman's emotional expressions on participants' empathetic engagement.

The integration of the TEQ and the adapted PIS yielded a comprehensive measure of empathy both before and after exposure to the narration, facilitating an intricate analysis of the effects of various emotional expression styles on user empathy. Qualitative feedback underwent thematic analysis to identify prevalent themes and insights pertaining to user perceptions and preferences.

Data analysis encompassed the comparison of pre- and post-interaction TEQ scores to discern alterations in empathy levels. Statistical methodologies were employed to identify significant disparities between the pre- and post-interaction scores, ensuring the production of robust and reflective findings representative of genuine user experiences.

6.3. Procedure

The participants involved in this research case study were asked to follow a four step Google Forms questionnaire, which was intended to assess their empathetic responses to MetaHumans displaying different facial animations. The study involved a comparison between the default facial animations from Unreal Engine and customized facial animations. To quantitatively measure empathy, the TEQ was utilized.

It's worth noting that the entire procedure was conducted in Portuguese to ensure clear communication and precise data collection. Participants in this research case study underwent a structured process designed to evaluate their empathetic responses to MetaHumans with different facial animations. The study compared default facial animations from Unreal Engine with customized facial animations captured from the real person used as reference for the MetaHuman, employing the TEQ to measure empathy quantitatively.

- **Stage 1: Preliminary Questions**

In the first phase, participants completed a brief questionnaire via Google Forms to gather demographic information and assess their prior experience with virtual characters and intelligent agents. This included questions about their age, gender, and familiarity with virtual characters.

- **Stage 2: TEQ Pre-Test**

To establish a baseline measure of empathy, participants completed the TEQ, a 16-item, 5-point Likert scale questionnaire that assesses various dimensions of empathy, including emotional contagion, emotion understanding, and sympathetic physiological arousal.

- **Stage 3: MetaHuman Narration**

During the third and most critical phase of the study, participants were asked to watch a MetaHuman narration on a laptop, using either provided or personal headphones to ensure consistent audio quality. This narration used in the study was as follows (original Portuguese version):

"Um homem do País de Gales descobriu que o rato de estimação lhe arrumava os items dispersos no barracão todas as noites, relatou a BBC. Rodney Holbrook, um carteiro reformado de 75 anos, instalou uma câmara de visão noturna por estranhar ver tudo arrumado ao acordar.
'No início percebi que alguns alimentos que eu deixava para os pássaros acabavam em sapatos velhos que eu guardava no barracão,' explicou Rodney Holbrook."

And for the English version, the translation is as follows:

"A man from Wales discovered that his pet rat was tidying up the items scattered in his shed every night, reported the BBC. Rodney Holbrook, a 75-year-old retired postman, installed a night-vision camera after noticing everything was organized when he woke up.
'At first, I noticed that some of the food I left out for the birds ended up in old shoes I kept in the shed,' explained Rodney Holbrook."

In the first user test, which focused on facial expressions, two separate groups of participants watched a 35-second MetaHuman narration. One group viewed a version using Unreal Engine's default facial animations, while the other group saw a version with custom-tailored facial expressions. During the narration, the MetaHuman told a story and displayed emotional reactions to it, allowing for a comparison between the emotional impact of default versus custom expressions. The specific expressions displayed are summarized in Table 1.

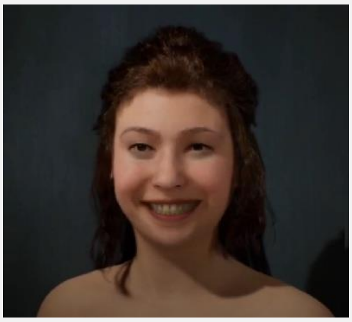
	Default Expression	Custom Expression
Surprise		
Happiness		

Table 1 – Comparison between default and custom expressions

In the second user test, the focus shifted to body movement. Again, two groups were formed. One group watched a real human performing the narration, while the other viewed the same narration enhanced with full-body motion captured using Rokoko technology, applied to a MetaHuman. In both cases, the narration was delivered with synchronized facial and body expressions designed to elicit empathetic responses from participants, as illustrated in Table 2.

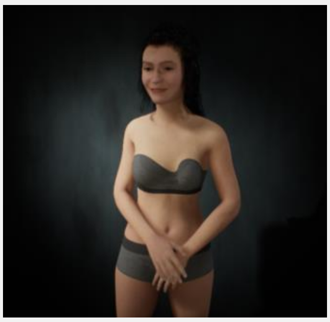
Human Body Movement	MetaHuman Body Movement
	

Table 2 – Comparison between Human and MetaHuman body movement

- **Stage 4: Post-Interaction Survey**

Following the video interaction, all participants were invited to complete a survey derived from the TEQ to evaluate potential changes in their empathy levels. This survey aimed to specifically measure the participants' empathetic responses after viewing the video, focusing on the emotions demonstrated in the video. Out of the sixteen original questions in the TEQ, only five were included, as the remaining questions were unrelated to the emotional dimensions under observation.

Table 3 below illustrates the selected TEQ questions, their corresponding adaptations for the PIS, and translations where necessary. The adapted questions were carefully designed to ensure relevance to the MetaHuman's behaviours and emotions, allowing researchers to assess how closely participants felt connected to and in tune with the character. This comparative analysis highlights the extent to which the MetaHuman's display of emotions influenced participants' empathetic alignment and emotional engagement.

TEQ Question	Post-Interaction TEQ Question	TEQ question translated	Post-Interaction TEQ question translated
1. Quando a pessoa se sente animada, tenho tendência a sentir-me animado(a) também.	Quando o MetaHuman está animado, sinto-me animado(a) também.	When someone else is feeling excited, I tend to get excited too.	When the MetaHuman is feeling excited, I tend to get excited too
4. Permaneço indiferente quando alguém que me é próximo está feliz.	Permaneço indiferente quando o MetaHuman está feliz.	I remain unaffected when someone close to me is happy.	I remain unaffected when the MetaHuman is happy.
9. Consigo estar “sintonizado(a)” com o estado de ânimo das outras pessoas.	Consigo estar “sintonizado(a)” com o estado de ânimo do MetaHuman.	I find that I am “in tune” with other people’s moods.	I find that I am “in tune” with the MetaHuman mood.
10. Não sinto simpatia por pessoas que causam as suas próprias doenças graves.	Não sinto simpatia pelo MetaHuman.	I do not feel sympathy for people who cause their own serious illnesses.	I do not feel for the MetaHuman.
12. Não me interesso realmente pela forma como as outras pessoas se sentem.	Não me interesso realmente pela forma como o MetaHuman se sente.	I am not really interested in how other people feel.	I am not really interested in how the MetaHuman feels.

Table 3 – PIS adapted from TEQ comparison.

6.4. Provisional Test

To ensure the clarity and effectiveness of the test formulation, provisional tests were conducted with four users. These preliminary tests aimed to evaluate the comprehensibility and organization of the questionnaire. The feedback indicated that while the users appreciated the organization of the questionnaire, they encountered difficulties relating to and understanding the last few questions. The primary issue was that the emotions displayed by the MetaHuman were mainly positive, while the questions asked if the participants could identify negative emotions. This discrepancy highlighted the need for adjustment in the questions to better align with the emotional content presented by the MetaHuman.

6.5.Results

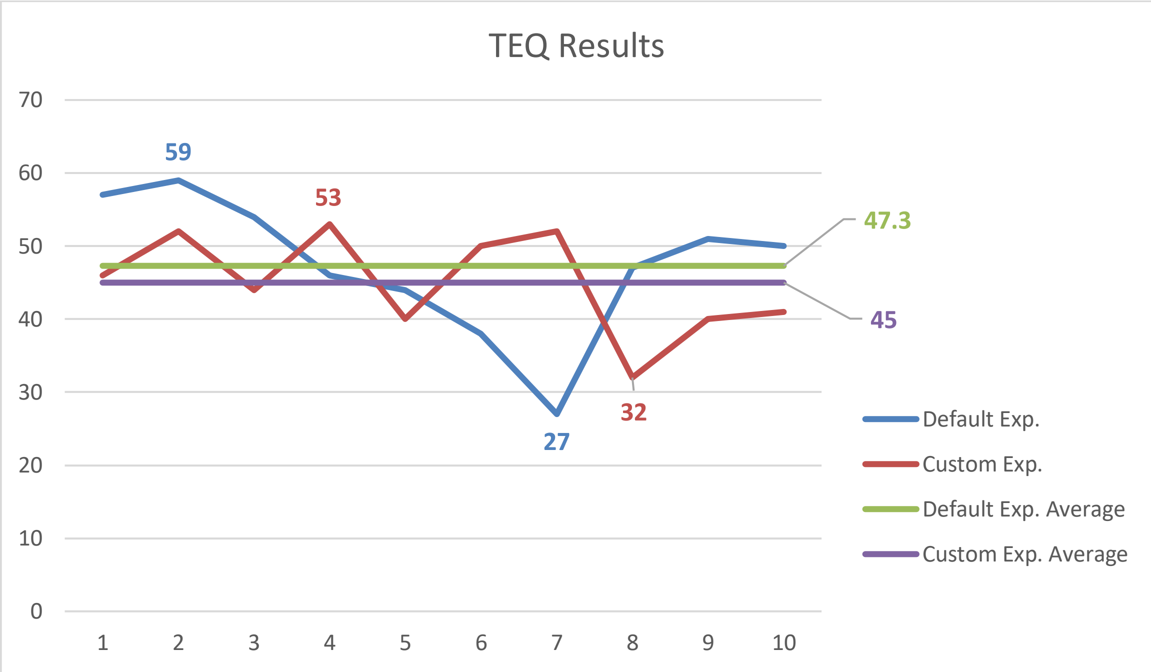
The results of the TEQ in this study were compared to reference values reported in previous validation studies, as shown in Table 4. In the Korean version validation, mean TEQ scores were 44.6, with a Standard Deviation (SD) of 7.36 and a range from 20 to 60 [72]. In the Arabic version validation, mean scores were 47.8 with a SD of 5.8 [73]. Among Saudi medical students, mean scores were 42.31 with a SD of 7.86 [74]. The original TEQ development study reported typical scores in the moderate range, approximately 40–45 [75]. Additionally, a Greek validation study confirmed the reliability of the TEQ (Cronbach's α = 0.73) but did not report descriptive statistics [76]. Across these studies, mean TEQ scores ranged from 42.31 to 47.8, with a mean of means of approximately 44.9. SD ranged from 5.8 to 7.9, and reported score ranges spanned from 20 to 60, indicating a moderate level of empathy.

Sample	Mean	SD	Range	Notes
Original TEQ development	~40–45	—	—	Descriptive stats not fully reported
Korean medical students	44.6	7.36	20–60	Full descriptive statistics reported
Arabic version, Saudi medical students	47.8	5.8	—	SD reported, range not reported
Saudi medical students	42.31	7.86	—	SD reported, range not reported
Greek medical students	—	—	—	Only reliability reported, no descriptive stats

Table 4 – Summary of TEQ scores from different validation studies

Regarding the first User Test, which compared default expressions with custom expressions, the group exposed to the default Unreal Engine emotional expressions had an average TEQ score of 47.375, as shown in Figure 19. This value lies within the upper range of scores reported in previous studies, suggesting that participants in this group displayed slightly higher-than-average empathic tendencies. For the group experiencing custom emotional expressions, the average TEQ score was 45, as shown in Figure 19, which is at the upper bound of the typical range, indicating a moderate level of empathy consistent with the general population.

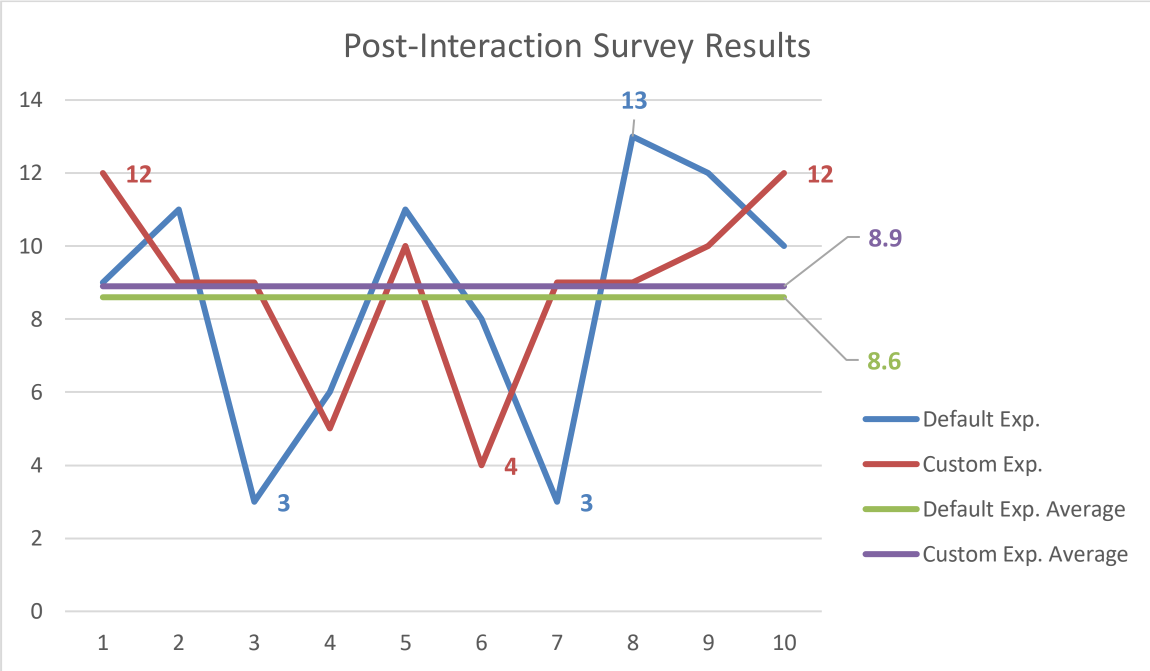
Given that both sets of average TEQ scores (47.375 for the default expressions and 45 for the custom expressions) fall within the range reported in the literature, it is unlikely that the main study results were influenced by pre-existing differences in participant empathy levels. This supports the validity of the findings and suggests that any observed changes in empathy are more plausibly attributed to the different facial animations of the MetaHumans rather than to baseline disparities.



	1	2	3	4	5	6	7	8	9	10
Default Exp.	57	59	54	46	44	38	27	47	51	50
Custom Exp.	46	52	44	53	40	50	52	32	40	41

Figure 19 – TEQ results for the first user tests and respective averages

The results of the PIS, summarized in Figure 20, show slight differences in empathy levels between the two conditions. Participants who viewed the MetaHuman with default expressions had an average PIS score of 8.6, while those who viewed the MetaHuman with custom expressions scored slightly higher at 8.9. Although this suggests a marginal increase in empathy for the custom expressions condition, the difference is not substantial enough to draw definitive conclusions. Furthermore, as the PIS is an unvalidated adaptation of the TEQ, these results should be interpreted with caution and considered only as a relative measure for comparison with the baseline TEQ scores.

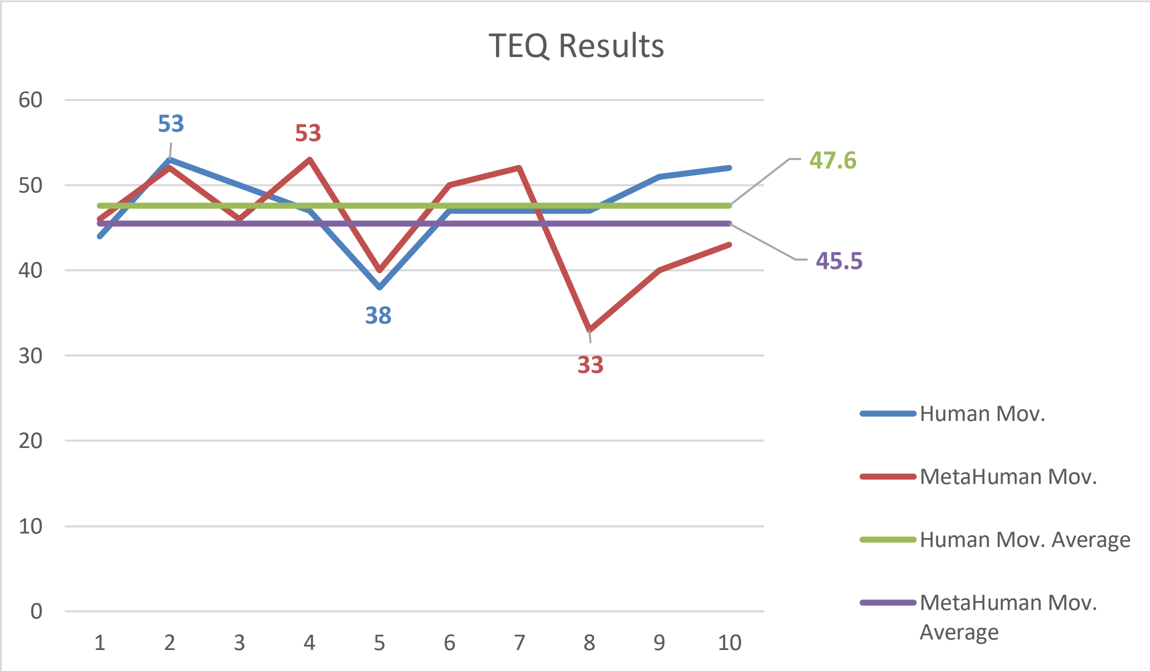


	1	2	3	4	5	6	7	8	9	10
Default Exp.	9	11	3	6	11	8	3	13	12	10
Custom Exp.	12	9	9	5	10	4	9	9	10	12

Figure 20 – PIS results for the first user tests and respective averages

For the second User Test, which compared Human Body Movement with MetaHuman Body Movement, the group exposed to Human Body Movement had an average TEQ score of 47.6, as shown in Figure 21. This value is within the upper range of scores reported in the literature, suggesting slightly higher-than-average empathic tendencies. The group experiencing MetaHuman Body Movement had an average TEQ score of 45.5, placing it near the upper bound of the typical range and indicating a moderate level of empathy consistent with the general population.

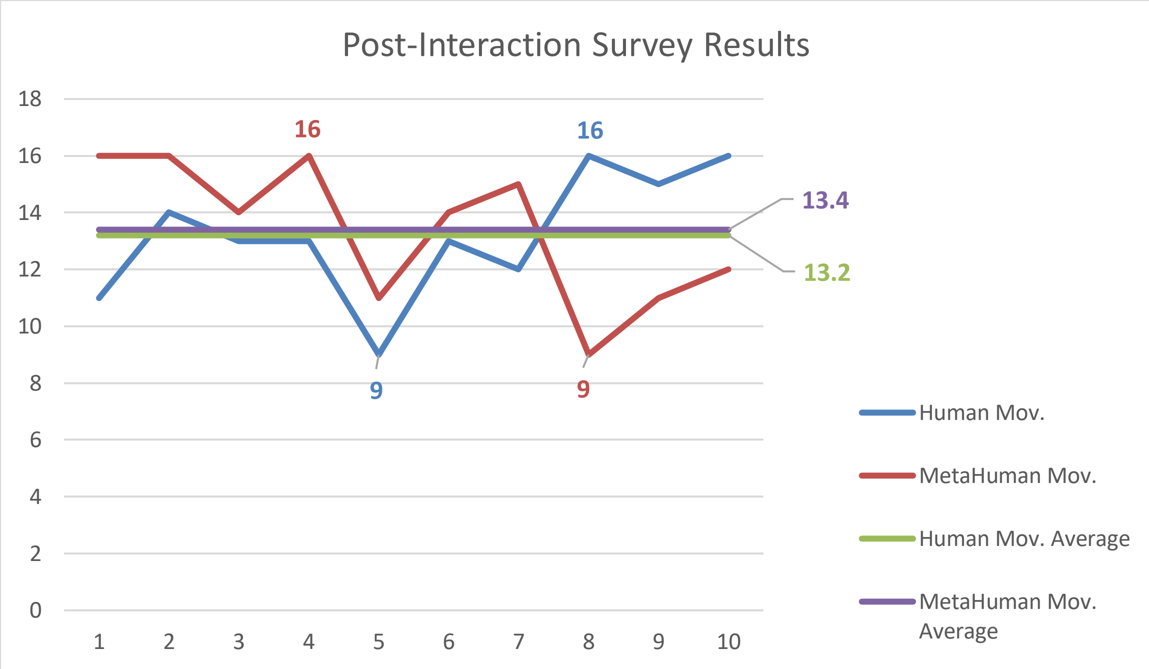
Since both average TEQ scores (47.6 for Human Body Movement and 45.5 for MetaHuman Body Movement) fall within the range reported in previous studies, it is unlikely that differences in baseline empathy influenced the main study outcomes. This strengthens the validity of the results and suggests that any observed differences in empathy can be attributed to the type of body movement animation rather than pre-existing empathy disparities.



	1	2	3	4	5	6	7	8	9	10
Human Mov.	44	53	50	47	38	47	47	47	51	52
MetaHuman Mov.	46	52	46	53	40	50	52	33	40	43

Figure 21 – TEQ results for the second user tests and respective averages

The results of the PIS, summarized in Figure 22, show minimal differences between the two conditions. Participants in the Human Body Movement condition had an average PIS score of 13.2, while those in the MetaHuman Body Movement condition had an average of 13.4. Although this represents a marginal increase in the MetaHuman Body Movement condition, the difference is too small to support firm conclusions. Moreover, as the PIS is an unvalidated adaptation of the TEQ, these findings should be interpreted with caution and viewed only as a relative measure alongside the baseline TEQ scores.



	1	2	3	4	5	6	7	8	9	10
Human Mov.	11	14	13	13	9	13	12	16	15	16
MetaHuman Mov.	16	16	14	16	11	14	15	9	11	12

Figure 22 – PIS results for the second user tests and respective averages

7. Analysis and Discussion

This section presents the results of the empirical study and examines their implications in relation to the research questions. The aim is not only to describe the outcomes of the user tests but also to critically interpret them within the broader context of digital human communication.

7.1. User Test 1 – Facial Expressions

The data analysis for this user test focuses on evaluating the quantitative and qualitative measures collected regarding facial expressions. Statistical methods, including Linear Regression, Shapiro-Wilk normality tests, Pearson correlation, and one-way ANOVA analysis, were employed to explore the relationships between TEQ scores and the PIS, as well as to compare the two animation approaches (custom vs. default facial expressions) within the same group.

7.1.1. Linear Regression

Linear regression is a statistical method that helps explore the relationship between an independent variable (also known as a predictor) and a dependent variable (also known as an outcome). In this case, the TEQ scores were used as the independent variable, while the PIS served as the dependent variable.

By conducting a linear regression analysis, it would be possible to determine whether TEQ scores significantly predict PIS scores, thereby offering insights into the relationship between emotional intelligence and interaction outcomes.

Linear regression provides key metrics, such as the strength and direction of the relationship (measured by the R-squared value and standardized coefficients), the overall fit of the model, and the statistical significance of the predictor. These outputs allow researchers to assess whether the independent variable contributes meaningfully to explaining the variance in the dependent variable.

The first analysis focused on the Custom Expressions dataset. The results, as shown in Table 5 – Linear Regression model for the Custom Expressions, showed an R-squared value of 0.224, indicating that 22.4 percent of the variance in the PIS Score is explained by TEQ scores. However, the adjusted R-squared was 0.127, reflecting a limited model fit when considering the sample size and number of predictors.

Model

Model	R	R squared	Adjusted R squared	Error
1	0,473 [a]	0,224	0,127	2,431

a. Predictor: (Constant), TEQ Score

Table 5 – Linear Regression model for the Custom Expressions

As part of the regression analysis, the Analysis of Variance (ANOVA) test was used to assess the overall significance of the model. In this context, the ANOVA F-test examines whether the independent variable(s) collectively explain a statistically significant proportion of the variance in the dependent variable, compared to a model with no predictors. This output, generated by SPSS alongside the regression coefficients and R-squared values, complements the interpretation of model fit and predictive power.

The ANOVA results from this model are shown in Table 6 – ANOVA results for the Custom Expressions, where the Z-statistic was 2.305 with a significance value (p) of 0.167, suggesting that the overall model was not statistically significant. This implies that TEQ scores, as a predictor, do not significantly explain the variation in PIS scores in this dataset.

ANOVA [a]

Model		Sum of the squares	Df	Sum of the squares	Z	Sig
1	Regression	13,623	1	13,623	2,305	0,167 [b]
1	Residue	42,277	8	5,91		
1	Total	60,9	9			

a. Dependent Variable: PIS Score
b. Predictor: (Constant), TEQ Score

Table 6 – ANOVA results for the Custom Expressions

In the coefficients table in Table 7 – Coefficient results for the Custom Expressions, the unstandardized beta coefficient for TEQ scores was -0.179 (p = 0.167). While this suggests a weak negative relationship between TEQ and PIS scores, the lack of statistical significance prevents us from making strong inferences about the effect.

Coefficients [a]

Model		Unstandardized Coefficients B	Unstandardized Coefficients Error	Beta Standardized Coefficients	t	Sig
1	(Constant)	16,966	5,368		3,161	0,013
1	TEQ Score	-0,179	0,118	-0,473	-1,518	0,167

a. Dependent Variable: PIS Score

Table 7 – Coefficient results for the Custom Expressions

The second analysis examined the Default Expressions dataset. As shown in Table 8 – Linear Regression model for the Default Expressions, the R-squared value was 0.167, meaning that 16.7 percent of the variance in PIS Score is explained by TEQ scores. The adjusted R-squared was lower at 0.063, again reflecting a modest explanatory power.

Model

Model	R	R squared	Adjusted R squared	Error
1	0,408 [a]	0,167	0,063	3,452

a. Predictor: (Constant), TEQ Score

Table 8 – Linear Regression model for the Default Expressions

The ANOVA table in Table 9 – ANOVA results for the Default Expressions showed an F-statistic of 1.602 with a significance value of 0.241, indicating that the regression model was not statistically significant. Similar to the Custom Expressions dataset, TEQ scores did not significantly predict PIS scores in this case.

ANOVA [a]

Model		Sum of the squares	Df	Sum of the squares	Z	Sig
1	Regression	19,989	1	19,989	1,602	0,241 [b]
1	Residue	95,311	8	11,914		
1	Total	114,4	9			

a. Dependent Variable: PIS Score
b. Predictor: (Constant), TEQ Score

Table 9 – ANOVA results for the Default Expressions

In the coefficients table in Table 10 – Coefficient results for the Default Expressions, the unstandardized beta coefficient for TEQ scores was 0.154 (p = 0.241). This suggests a weak positive relationship between TEQ and PIS scores, but, like in the previous dataset, the lack of significance prevents us from drawing definitive conclusions.

Coefficients [a]

Model		Unstandardized Coefficients		Beta Standardized Coefficients	t	Sig
		B	Error			
1	(Constant)	1,33	5,846		0,228	0,826
1	TEQ Score	0,154	0,121	0,408	0,1266	0,241

a. Dependent Variable: PIS Score

Table 10 – Coefficient results for the Default Expressions

Overall, the linear regression analysis for both Custom and Default Expressions datasets revealed no statistically significant relationship between TEQ scores and PIS scores. While slight variations were observed in the direction and magnitude of the relationships, these were not meaningful in explaining the variance in the dependent variable.

7.1.2. Shapiro-Wilk Normality

The Shapiro-Wilk test is a statistical test specifically designed to assess the normality of a dataset. Normality is a crucial assumption for many parametric statistical tests, as it ensures the validity of the results and interpretations. In this study, the test was conducted to verify whether the residuals of the model, with the TEQ as the independent variable and the PIS as the dependent variable, followed a normal distribution. By analysing the normality, we ensure that the assumptions for subsequent analyses, such as regression or hypothesis testing, are adequately met.

Conducting the Shapiro-Wilk test allows us to quantitatively assess the deviation of the data from a normal distribution. If the p-value associated with the test is greater than the chosen significance level (e.g., 0.05), the null hypothesis of normality cannot be rejected, indicating that the data is approximately normal. On the other hand, a p-value below the threshold suggests significant deviations from normality.

For the Custom Expressions, the results shown in Table 11 – Shapiro-Wilk Normality results for the Custom Expressions and Figure 23 – Shapiro-Wilk Normal Q-Q Plot of Unstandardized Residual for the Custom Expressions of the Shapiro-Wilk test indicates the following:

- Statistic: $W = 0.955$
- p-value: $p = 0.723$

Since the p-value is greater than 0.05, we fail to reject the null hypothesis of normality. This indicates that the residuals for the custom expressions are normally distributed. Supporting this conclusion, the Q-Q plot shows data points closely aligning with the reference line, further confirming the normality of the residuals.

Case Processing Summary

	Valid		Omitted Cases		Total	
	N	Percentage	N	Percentage	N	Percentage
Unstandardized Residual	10	100%	0	0%	10	100%

Descriptive

			Statistic	Standard test Statistics
Unstandardized Residual	Average		0	0,72477858
	95% Confidence Interval for Mean	Lower Limit	-1,639563	
		Upper Limit	1,639563	
	5% of the trimmed mean		0,0402516	
	Median		0,2037736	
	Variance		5,253	
	Standard Error		2,2919511	
	Minimum		-4,00377	
	Maximum		3,27925	
	Amplitude		7,28302	
	Interquartil Amplitude		3,90094	
	Asymmetry		-0,421	0,687
	Kurtosis		-0,596	1,334

Normality TESTS

	Kolmogorov-Smirnov [a]			Shapiro-Wilk		
	Statistic	gl	Sig	Statistic	gl	Sig
Unstandardized Residual	0,186	10	0,200*	0,955	10	0,723

*. This is the lower limit of true significance

a. Lilliefors Significance Correlation

Table 11 – Shapiro-Wilk Normality results for the Custom Expressions

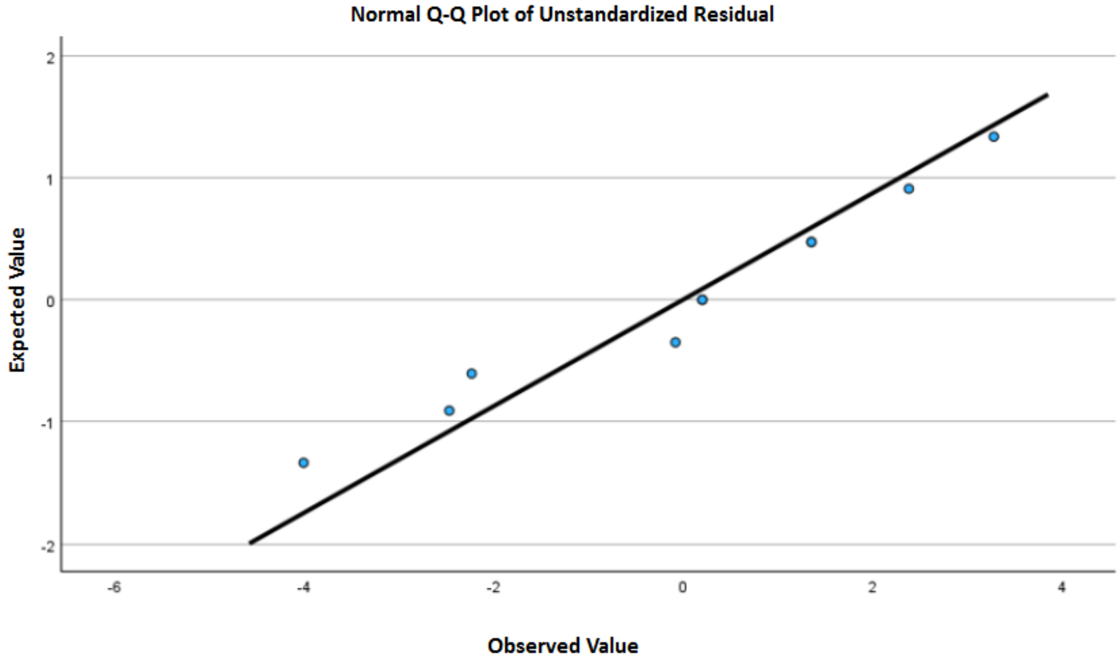


Figure 23 – Shapiro-Wilk Normal Q-Q Plot of Unstandardized Residual for the Custom Expressions

For the Default Expressions, the results shown in Table 12 – Shapiro-Wilk Normality results for the Default Expressions and Figure 24 – Shapiro-Wilk Normal Q-Q Plot of Unstandardized Residual for the Default Expressions of the Shapiro-Wilk test are as follows:

- Statistic: W = 0.949
- p-value: p = 0.660

Similarly, the p-value is above 0.05, leading us to conclude that the residuals for the default expressions also follow a normal distribution. The Q-Q plot for this case further validates this, with only minor deviations from the reference line.

Case Processing Summary

	Valid		Omitted Cases		Total	
	N	Percentage	N	Percentage	N	Percentage
Unstandardized Residual	10	100%	0	0%	10	100%

Descriptive

			Statistic	Standard test Statistics
Unstandardized Residual	Average		0	1,02908386
	95% Confidence Interval for Mean	Lower Limit	-2,3279494	
		Upper Limit	2,3279494	
	5% of the trimmed mean		0,1213134	
	Median		0,7155674	
	Variance		10,59	
	Standard Error		3,25424891	
	Minimum		-6,62975	
	Maximum		4,44611	
	Amplitude		11,07586	
	Interquartil Amplitude		5,27045	
	Asymmetry		-0,745	0,687
	Kurtosis		0,589	1,334

Normality TESTS

	Kolmogorov-Smirnov [a]			Shapiro-Wilk		
	Statistic	gl	Sig	Statistic	gl	Sig
Unstandardized Residual	0,173	10	0,200*	0,949	10	0,66

*. This is the lower limit of true significance

a. Lilliefors Significance Correlation

Table 12 – Shapiro-Wilk Normality results for the Default Expressions

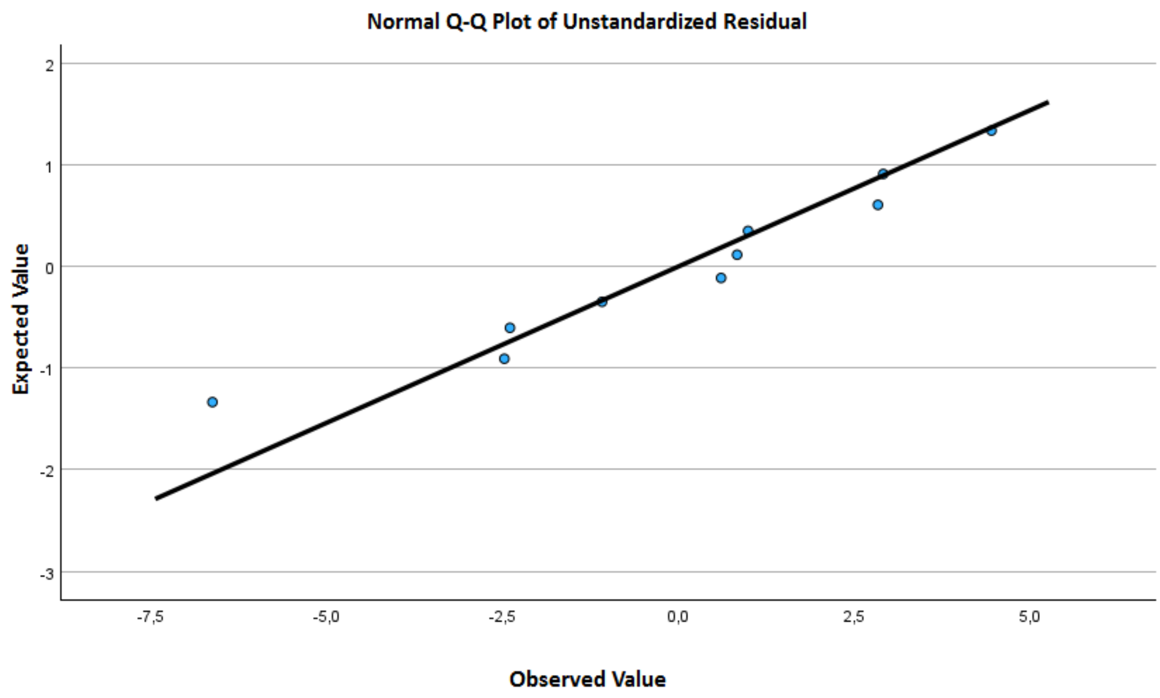


Figure 24 – Shapiro-Wilk Normal Q-Q Plot of Unstandardized Residual for the Default Expressions

Both custom and default expressions demonstrate residuals that are normally distributed, as evidenced by the Shapiro-Wilk test results and supporting visualizations. This confirms the suitability of the data for parametric analyses involving the TEQ and the PIS, ensuring the robustness of the results and their interpretations.

7.1.3. Pearson Correlation

The Pearson correlation test assesses the strength and direction of the linear relationship between two variables. It analyses the relationship between the TEQ Score and the PIS Score. This analysis was conducted for both custom and default expressions as well as the metahuman and human body movement to determine how closely these variables are associated.

A correlation coefficient (r) range from -1 to 1:

- **Positive Correlation:** As one variable increases, the other also increases;
- **Negative Correlation:** As one variable increases, the other decreases;
- **No correlation:** A coefficient near 0 indicated no linear relationship.

If the p-value is below the significance level (e.g. 0,05), the correlation is considered statistically significant.

For the Custom Expressions, the Pearson correlation results are shown in Table 13 – Pearson Correlation results for the Custom Expressions and indicate the following:

- **Correlation coefficient (r):** -0,473
- **p-value:** 0,167

Since the p-value is greater than 0.05, the correlation between the TEQ Score and the PIS Score is not statistically significant. Although there is a moderate negative correlation (r = -0.473), this relationship is not strong enough to be considered significant. This suggests that changes in one score do not reliably predict changes in the other score for the custom expressions.

Correlations

		PIS Score	TEQ Score
PIS Score	Pearson Correlation	1	-0,473
	Sig. (2 ends)		0,167
	N	10	10
TEQ Score	Pearson Correlation	-0,473	1
	Sig. (2 ends)	0,167	
	N	10	10

Table 13 – Pearson Correlation results for the Custom Expressions

For the Default Expressions, the Pearson correlation results shown in Table 14 – Pearson Correlation results for the Default Expressions are as follows:

- **Correlation coefficient (r):** 0,408
- **p-value:** 0,241

Similar to the custom expressions, the p-value is above 0.05, indicating that the correlation is not statistically significant. There is a moderate positive correlation (r = 0.408), but this relationship is not strong enough to be deemed significant. Therefore, for the default expressions, the TEQ Score and the PIS Score also do not show a statistically reliable linear relationship.

Correlations

		PIS Score	TEQ Score
PIS Score	Pearson Correlation	1	0,408
	Sig. (2 ends)		0,247
	N	10	10
TEQ Score	Pearson Correlation	0,408	1
	Sig. (2 ends)	0,241	
	N	10	10

Table 14 – Pearson Correlation results for the Default Expressions

Both custom and default expressions exhibit moderate correlations, but neither relationship is statistically significant. This suggests that, while some linear association exists between the TEQ Score and the PIS Score, it is not strong enough to conclude a consistent predictive relationship.

7.1.4. One-Way ANOVA

To evaluate whether the type of facial expressions (Default vs. Custom) and body movement (Real Human vs. MetaHuman) used had an impact on participants' evaluations in the Post Interaction Survey, a one-way ANOVA was conducted. This test allows the comparison of mean scores between two independent groups and determines whether any observed difference is statistically significant.

Descriptives

	N	Mean	Std. Deviation	Std. Error	95% Confidence Interval for Mean		Minimum	Maximum
					Lower Bound	Upper Bound		
Default	10	8.60	3.565	1.127	6.05	11.15	3	13
Custom	10	8.90	2.601	0.823	7.04	10.76	4	12
Total	20	8.75	3.041	0.680	7.33	10.17	3	13

Table 15 – One-Way ANOVA descriptives for the First User Test

The descriptive statistics, as shown in Table 15, indicated that participants in the default facial expressions condition (Mean = 8.60, Std. Deviation = 3.57) reported scores very similar to those in the custom facial expressions condition (Mean = 8.90, Std. Deviation = 2.60). This suggests that participants evaluated both conditions similarly (difference = 0.30).

Levene's Test (Test of Homogeneity of Variances)

		Levene Statistic	df1	df2	Sig.
PIS Score	Based on mean	1.839	1	18	0.192
	Based on Median	1.429	1	18	0.247
	Based on Median and with adjusted df	1.429	1	17.564	0.248
	Based on trimmed mean	1.903	1	18	0.185

Table 16 – One-Way ANOVA Levene's Test for the First User Test

In Table 16, the test of homogeneity of variances was not significant (p = .192, which is > .05), indicating that the assumption of equal variances between groups was met. This supports the use of the standard ANOVA results without the need for adjustments.

ANOVA

	Sum of Squares	df	Mean Square	F	Sig.
Between Groups	0.450	1	0.450	0.46	0.832
Within Groups	175.300	18	9.739		
Total	175.750	19			

Table 17 – ANOVA for the First User Test

The one-way ANOVA in Table 17 revealed no significant effect of expression type on Post Interaction Survey scores, $F(1, 18) = 0.046$, $p = .832$. This shows that the difference between the default and custom facial expression groups was not statistically meaningful since $p > .05$.

ANOVA Effect Sizes [a, b]

		Point Estimate	95% Confidence Interval Lower	Upper
PIS Score	Eta-squared	0.003	0.00	0.162
	Epsilon-squared	-0.053	-0.056	0.116
	Omega-squared Fixed-effect	-0.050	-0.053	0.110
	Omega-squared Random-effect	-0.050	-0.053	0.110

a. Eta-squared and Epsilon-squared are estimated based on the fixed-effect model.

b. Negative but less biased estimates are retained, not rounded to zero.

Table 18 – ANOVA Effect Sizes for the First User Test

The effect size shown in Table 18 was extremely small ($\eta^2 = .003$), indicating that less than 1% of the variance in Post Interaction Survey scores could be explained by the type of facial expressions. The 95% confidence interval included zero, further supporting the absence of a practical effect.

Overall, these results suggest that the use of default or custom facial expressions in the MetaHuman did not influence participants' perceptions as measured by the Post Interaction Survey. Both approaches were evaluated similarly, indicating that the customization of expressions, at least in this experimental setup, did not lead to measurable differences in user experience.

7.2. User Test 2 – Body Movement

The data analysis for this user test focuses on evaluating the quantitative and qualitative measures collected regarding body movement. Statistical methods, including Linear Regression, Shapiro-Wilk normality tests, Pearson correlation, and one-way ANOVA analysis, were employed to examine the relationships between TEQ scores and the PIS, as well as to compare the two animation approaches (digital human vs. real human body movement) within the same group.

7.2.1. Linear Regression

For the MetaHuman Body Movement condition, as shown in Table 19 – Linear Regression model for the MetaHuman Body Movement, the R-squared value was 0.853, indicating that 85.3 percent of the variance in SQ Score is explained by TEQ scores. The adjusted R-squared was similarly high at 0.835, confirming the strength and reliability of the model in explaining the relationship between empathy levels and the perception of social quality during MetaHuman Body Movement interactions.

Model

Model	R	R squared	Adjusted R squared	Error
1	0,924 [a]	0,853	0,853	1,018

a. Predictor: (Constant), TEQ Score

Table 19 – Linear Regression model for the MetaHuman Body Movement

The ANOVA table in Table 20 – ANOVA results for the MetaHuman Body Movement shows an F-statistic of 46.388 with a significance value less than 0.001, strongly indicating that the regression model is statistically significant. This means that TEQ scores are a highly meaningful predictor of SQ scores in the MetaHuman Body Movement condition.

ANOVA [a]

Model		Sum of the squares	Df	Sum of the squares	Z	Sig
1	Regression	48,104	1	48,104	46,388	< 0,001 [b]
1	Residue	8,296	8	1,037		
1	Total	56,400	9			

a. Dependent Variable: PIS Score
b. Predictor: (Constant), TEQ Score

Table 20 – ANOVA results for the MetaHuman Body Movement

The coefficients in Table 21 – Coefficient results for the MetaHuman Body Movement further support the robustness of the model. The unstandardized beta coefficient for TEQ scores was 0.354, and the standardized beta was 0.924, both indicating a strong positive relationship between empathy and perceived social quality. The p-value was less than 0.001, confirming that the predictor is highly statistically significant. This implies that participants with higher empathy levels rated their interaction quality significantly higher when the agent displayed MetaHuman Body Movement.

Coefficients [a]

Model		Unstandardized Coefficients B	Error	Beta Standardized Coefficients	t	Sig
1	(Constant)	-2,694	2,385		-1,130	0,291
1	TEQ Score	0,354	0,052	0,924	6,811	< 0,001

a. Dependent Variable: PIS Score

Table 21 – Coefficient results for the MetaHuman Body Movement

Overall, the regression results for the MetaHuman Body Movement condition showed the strongest predictive relationship between empathy (TEQ) and perceived social quality (SQ Score) out of all conditions analysed. These results suggest that high-fidelity, realistic body movements significantly enhance the social perception of digital humans, especially among users with higher levels of empathy.

For the Human Body Movement dataset, as shown in Table 22 – Linear Regression model for the Human Body Movement, the R-squared value was 0.648, meaning that 64.8 percent of the variance in PIS Score is explained by TEQ scores. The adjusted R-squared was slightly lower at 0.604, still reflecting a strong explanatory power and a well-fitting model.

Model

Model	R	R squared	Adjusted R squared	Error
1	0,805 [a]	0,648	0,604	1,385

a. Predictor: (Constant), TEQ Score

Table 22 – Linear Regression model for the Human Body Movement

The ANOVA table in Table 23 – ANOVA results for the Human Body Movement showed an F-statistic of 14.738 with a significance value of 0.005, indicating that the regression model is statistically significant. This result supports the conclusion that TEQ scores can meaningfully predict PIS scores when Human Body Movement is involved in interaction.

ANOVA [a]

Model		Sum of the squares	Df	Sum of the squares	Z	Sig
1	Regression	28,260	1	28,260	14,738	0,005 [b]
1	Residue	15,340	8	1,917		
1	Total	43,600	9			

a. Dependent Variable: PIS Score
b. Predictor: (Constant), TEQ Score

Table 23 – ANOVA results for the Human Body Movement

In the coefficients table, Table 24 – Coefficient results for the Human Body Movement, the unstandardized beta coefficient for TEQ scores was 0.405 with a p-value of 0.005, confirming that the relationship is both positive and statistically significant. The standardized beta value was 0.805, indicating a strong effect of TEQ on PIS Score. This suggests that higher empathy levels (as measured by TEQ) are strongly associated with higher perceived interaction smoothness in scenarios involving human body movement.

Coefficients [a]

Model		Unstandardized Coefficients		Beta Standardized Coefficients	t	Sig
		B	Error			
1	(Constant)	-6,072	5,039		-1,205	0,263
1	TEQ Score	0,405	0,105	0,805	3,839	0,005

a. Dependent Variable: PIS Score

Table 24 – Coefficient results for the Human Body Movement

Overall, the linear regression analysis for the Human Body Movement dataset revealed a strong and statistically significant relationship between TEQ scores and PIS scores. This contrasts with the Default Expressions and Custom Expressions datasets, where the relationship was weak and not statistically significant. These findings suggest that when human-like body movement is integrated into interactions, users with higher empathy levels perceive the experience as significantly smoother and more natural.

7.2.2. Shapiro-Wilk Normality

For the MetaHuman Body Movement, the results shown in Table 25 – Shapiro-Wilk Normality results for the MetaHuman Body Movement and Figure 25 – Shapiro-Wilk Normal Q-Q Plot of Unstandardized Residual for the MetaHuman Body Movement of the Shapiro-Wilk test are as follows:

- **Statistic:** W = 0.894
- **p-value:** p = 0.188

The p-value is greater than 0.05, indicating that the null hypothesis of normality cannot be rejected. This means that the residuals for the MetaHuman Body Movement are also normally distributed. The Q-Q plot supports this interpretation, displaying points that generally align with the reference line, with only minor deviations.

Case Processing Summary

	Valid		Omitted Cases		Total	
	N	Percentage	N	Percentage	N	Percentage
Unstandardized Residual	10	100%	0	0%	10	100%

Descriptive

			Statistic	Standard test Statistics
Unstandardized Residual	Average		13,40	0,792
	95% Confidence Interval for Mean	Lower Limit	11,61	
		Upper Limit	15,19	
	5% of the trimmed mean		13,50	
	Median		14,00	
	Variance		6,267	
	Standard Error		2,503	
	Minimum		9	
	Maximum		16	
	Amplitude		7	
	Interquartil Amplitude		5	
	Asymmetry		-0,516	0,687
	Kurtosis		-1,066	1,334

Normality TESTS

	Kolmogorov-Smirnov [a]			Shapiro-Wilk		
	Statistic	gl	Sig	Statistic	gl	Sig
Unstandardized Residual	0,195	10	0,200*	0,894	10	0,188

*. This is the lower limit of true significance

a. Lilliefors Significance Correlation

Table 25 – Shapiro-Wilk Normality results for the MetaHuman Body Movement

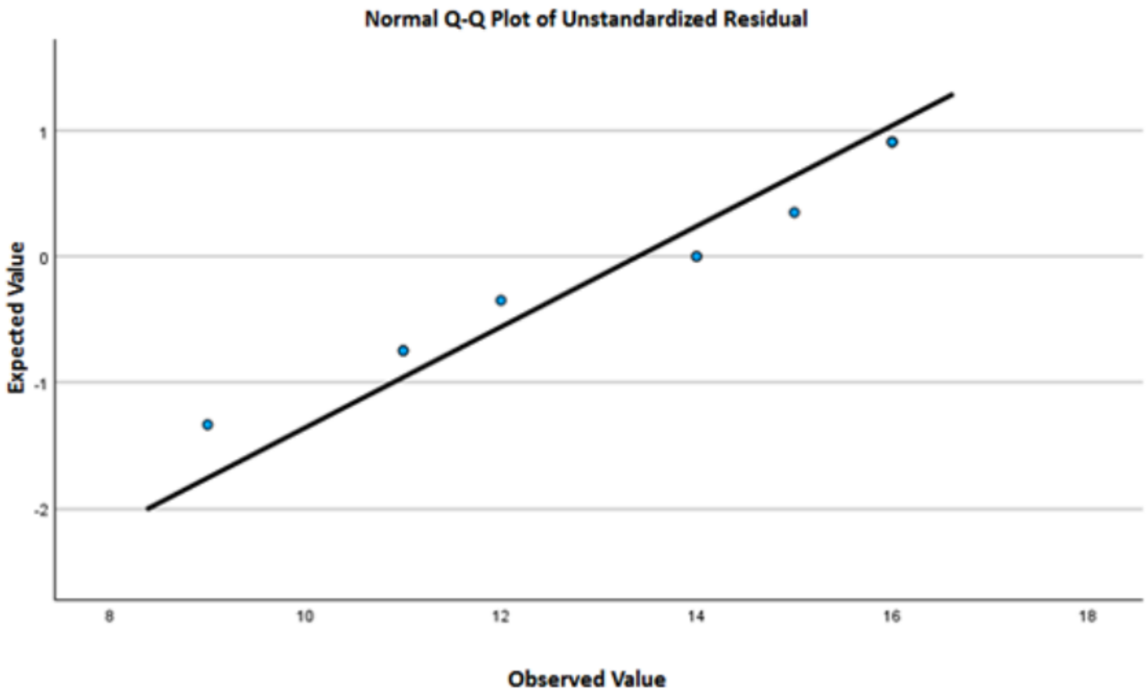


Figure 25 – Shapiro-Wilk Normal Q-Q Plot of Unstandardized Residual for the MetaHuman Body Movement

Together with the results for the default expressions, these findings suggest that both sets of data meet the assumption of normality. This supports the use of parametric statistical methods in analysing the responses to the TEQ and the PIS, thereby strengthening the validity and interpretability of the study's conclusions.

For the Human Body Movement, the results shown in Table 26 – Shapiro-Wilk Normality results for the Human Body Movement and Figure 26 – Shapiro-Wilk Normal Q-Q Plot of Unstandardized Residual for the Human Body Movement of the Shapiro-Wilk test are as follows:

- **Statistic:** W = 0.947
- **p-value:** p = 0.637

As with the previous cases, the p-value exceeds the 0.05 threshold, indicating that the residuals are normally distributed. The Q-Q plot further supports this conclusion, with the plotted points closely following the reference line and exhibiting only slight deviations.

Case Processing Summary

	Valid		Omitted Cases		Total	
	N	Percentage	N	Percentage	N	Percentage
Unstandardized Residual	10	100%	0	0%	10	100%

Descriptive

			Statistic	Standard test Statistics
Unstandardized Residual	Average		13,20	0,696
	95% Confidence Interval for Mean	Lower Limit	11,63	
		Upper Limit	14,77	
	5% of the trimmed mean		13,28	
	Median		12,00	
	Variance		4,844	
	Standard Error		2,201	
	Minimum		9	
	Maximum		16	
	Amplitude		7	
	Interquartil Amplitude		4	
	Asymmetry		-0,472	0,687
	Kurtosis		0,036	1,334

Normality TESTS

	Kolmogorov-Smirnov [a]			Shapiro-Wilk		
	Statistic	gl	Sig	Statistic	gl	Sig
Unstandardized Residual	0,164	10	0,200*	0,947	10	0,637

*. This is the lower limit of true significance

a. Lilliefors Significance Correlation

Table 26 – Shapiro-Wilk Normality results for the Human Body Movement

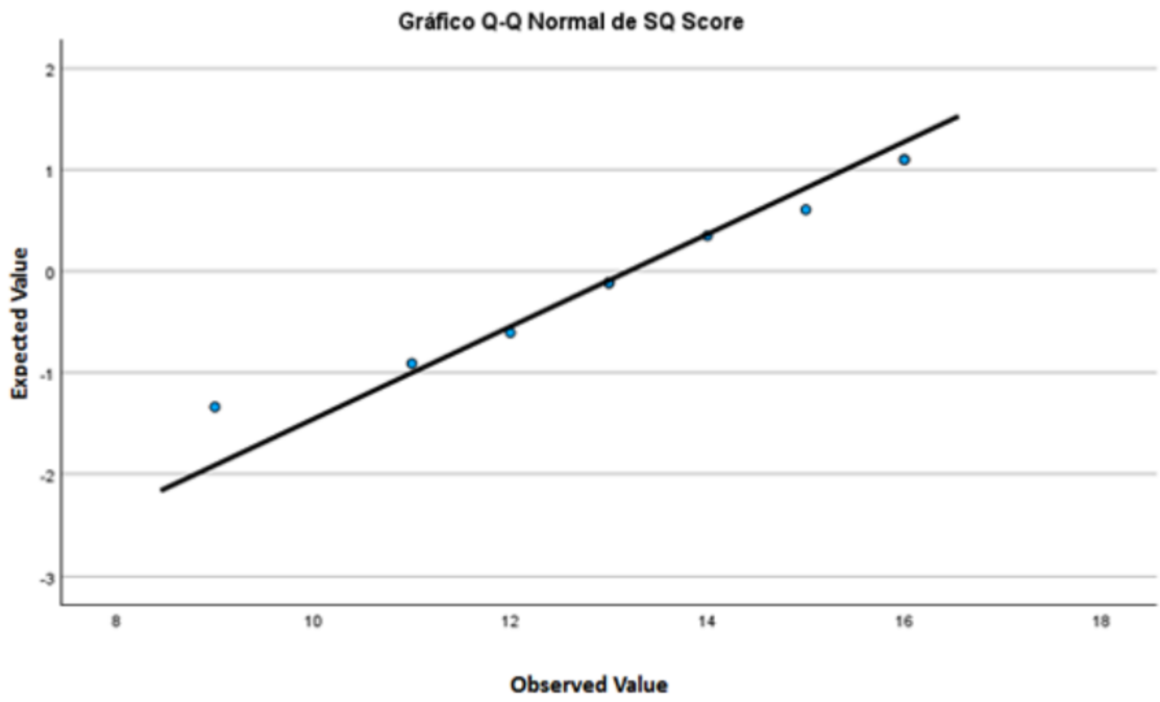


Figure 26 – Shapiro-Wilk Normal Q-Q Plot of Unstandardized Residual for the Human Body Movement

Collectively, the results affirm the assumption of normality. This consistency across datasets justifies the application of parametric tests in evaluating outcomes from TEQ and the PIS, enhancing the statistical reliability and interpretive clarity of the research.

7.2.3. Pearson Correlation

For the MetaHuman Body Movement condition, the Pearson correlation results shown in Table 27 are as follows:

- **Correlation coefficient (r):** 0.924
- **p-value:** < 0.001

Unlike the results obtained for both custom and default expressions, the correlation between the TEQ Score and the PIS Score in the MetaHuman Body Movement condition is statistically significant. The p-value is well below the 0.05 threshold, indicating a highly significant linear relationship. Additionally, the correlation coefficient ($r = 0.924$) reveals a strong positive correlation, suggesting that participants who scored higher in perceived interpersonal sensitivity (PIS) also tended to score higher in empathy (TEQ) when interacting with the MetaHuman Body Movement.

This strong and statistically significant relationship contrasts sharply with the weaker, non-significant correlations observed in previous conditions. It suggests that the inclusion of realistic body movements in digital human interactions may enhance the emotional resonance or empathetic connection perceived by users, reinforcing the potential of body dynamics in fostering meaningful HCI.

Correlations

		PIS Score	TEQ Score
PIS Score	Pearson Correlation	1	0,924
	Sig. (2 ends)		< 0,001
	N	10	10
TEQ Score	Pearson Correlation	0,924	1
	Sig. (2 ends)	< 0,001	
	N	10	10

Table 27 – Pearson Correlation results for the MetaHuman Body Movement

For the Human Body Movement condition, the Pearson correlation results shown in Table 28 are as follows:

- **Correlation coefficient (r):** 0.805
- **p-value:** 0.005

The results indicate a strong positive correlation between the TEQ Score and the PIS Score. With a correlation coefficient of 0.805, this relationship suggests that participants who scored higher in empathy also tended to score higher in perceived interpersonal sensitivity. Furthermore, the p-value is below the 0.05 threshold, confirming that this correlation is statistically significant.

This contrasts with the non-significant findings observed for both the default and custom expressions, where moderate correlations did not reach statistical significance. The significant relationship found in the Human Body Movement condition highlights the importance of realistic physical movement in human interactions, suggesting that such embodied behaviours enhance participants' emotional engagement and perception of interpersonal sensitivity.

Correlations

		PIS Score	TEQ Score
PIS Score	Pearson Correlation	1	0,805
	Sig. (2 ends)		0,005
	N	10	10
TEQ Score	Pearson Correlation	0,805	1
	Sig. (2 ends)	0,005	
	N	10	10

Table 28 – Pearson Correlation results for the Human Body Movement

These findings align with those obtained from the MetaHuman Body Movement condition, reinforcing the idea that movement, whether digitally replicated or human, is a key contributor to empathy and social perception in HCI scenarios.

7.2.4. One-Way ANOVA

Following the analysis of facial expressions in User Test 1, a second test was conducted to assess whether the type of body movement and narration (performed either by a real human or by a MetaHuman) had an impact on participants' evaluations in the Post Interaction Survey.

Descriptives

					95% Confidence Interval for Mean			
	N	Mean	Std. Deviation	Std. Error	Lower Bound	Upper Bound	Minimum	Maximum
Human	10	13.20	2.201	0.696	11.63	14.77	9	16
MetaHuman	10	13.40	2.503	0.792	11.61	15.19	9	16
Total	20	13.30	2.296	0.514	12.23	14.37	9	16

Table 29 – One-Way ANOVA descriptives for the Second User Test

The descriptive statistics in Table 29, showed that participants in the real human body movement condition (Mean = 13.20, Std. Deviation = 2.20) reported scores very similar to those in the MetaHuman body movement condition (Mean = 13.40, Std. Deviation = 2.50). This suggests that participants evaluated both conditions similarly (difference = 0.20).

Levene's Test (Test of Homogeneity of Variances)

		Levene Statistic	df1	df2	Sig.
PIS Score	Based on mean	0.736	1	18	0.402
	Based on Median	0.375	1	18	0.548
	Based on Median and with adjusted df	0.375	1	17.969	0.548
	Based on trimmed mean	0.620	1	18	0.441

Table 30 – One-Way ANOVA Levene's Test for the Second User Test

In Table 30, the test of homogeneity of variances was not significant (p = .402, which is > .05), indicating that the assumption of equal variances between groups was met. This validates the interpretation of the standard ANOVA results without the need for corrections.

ANOVA

	Sum of Squares	df	Mean Square	F	Sig.
Between Groups	0.200	1	0.200	0.036	0.852
Within Groups	100.000	18	5.556		
Total	100.200	19			

Table 31 – ANOVA for the Second User Test

The one-way ANOVA shown in Table 31 revealed no significant effect of body movement type on Post Interaction Survey scores, F(1, 18) = 0.036, p = .852. This shows that the difference between the real human and MetaHuman body movement groups was not statistically meaningful since p > .05.

ANOVA Effect Sizes [a, b]

			95% Confidence Interval	
		Point Estimate	Lower	Upper
PIS Score	Eta-squared	0.002	0.000	0.153
	Epsilon-squared	-0.053	-0.056	0.106
	Omega-squared Fixed-effect	-0.051	-0.053	0.101
	Omega-squared Random-effect	-0.051	-0.053	0.101

a. Eta-squared and Epsilon-squared are estimated based on the fixed-effect model.

b. Negative but less biased estimates are retained, not rounded to zero.

Table 32 – ANOVA Effect Sizes for the Second User Test

The effect size shown in Table 32 was extremely small ($\eta^2 = .002$), indicating that less than 1% of the variance in Post Interaction Survey scores could be explained by the type of body movement. The 95% confidence interval included zero, reinforcing the absence of a meaningful effect.

Overall, these results suggest that the use of real human or MetaHuman body movement and narration did not influence participants' perceptions as measured by the Post Interaction Survey. Both conditions were evaluated similarly, indicating that, within the scope of this study, MetaHumans were able to perform on par with real humans in terms of body movement and narration.

7.3. Discussion

The findings of this study offer significant insights into the impact of customized MetaHuman facial animations on user empathy, as well as the perceived effectiveness of digital humans in simulating human-like interactions. This section will address the research questions posed at the beginning of the study, drawing conclusions from the results of the user tests.

7.3.1. RQ1: How does user perception of a Digital Human compare to that of a real person in terms of realism and emotional expressiveness?

The One-Way ANOVA comparing user evaluations of real human versus MetaHuman conditions revealed no significant difference in perceived realism and emotional expressiveness. This indicates that participants judged the two conditions as roughly equivalent in terms of how authentic and expressive the agent appeared. Notably, despite the intuitive expectation that humans would always be perceived as more expressive, the results suggest that high-fidelity MetaHumans can effectively emulate human behaviours to a perceptually convincing degree. This is supported by regression and correlation analyses showing strong associations between empathy (TEQ) and perceived interaction quality (PIS) in both conditions (MetaHuman: $r = 0.924$; Human: $r = 0.805$). These findings suggest that empathetic engagement may amplify perceptions of expressiveness, and when body movements and animations are of sufficient quality, users may respond to MetaHumans almost as strongly as they would to real humans. This aligns with theories of social presence and embodiment, which posit that visual and behavioural realism can trigger authentic emotional responses, even in artificial agents.

7.3.2. RQ2: How do users perceive the realism and emotional expressiveness of a Digital Human's body movements compared to those of a real person performing the same actions?

Analysis using the One-Way ANOVA showed no significant difference between real human and MetaHuman body movements, suggesting that users perceive digital body movements with high-fidelity animation as comparably realistic to actual human movements. Regression analyses provide additional nuance: the predictive power of empathy on PIS scores was higher for MetaHuman interactions ($R^2 = 0.853$, $\beta = 0.924$) than for real human interactions ($R^2 = 0.648$, $\beta = 0.805$). This indicates that, while the overall average perception between groups is similar, individuals with higher empathy may find the MetaHuman particularly engaging, potentially because the animations are optimized for clarity and emphasis in a controlled digital environment. The results suggest that digital humans can leverage idealized expressiveness, where movements are exaggerated or refined in ways that enhance perceptual clarity without breaking realism. The combination of ANOVA and regression results implies that high-quality animation can equal real human performance in eliciting emotional responses among highly empathetic users, highlighting the critical role of movement fidelity in digital human design.

7.3.3. RQ3: Can personalized facial animation of emotion increase the feeling of empathy compared to motion capture-based facial animation?

The One-Way ANOVA results for default versus custom facial animations indicated no significant difference in PIS scores, suggesting that personalization alone does not significantly enhance perceived empathy. Both default and custom facial animations elicited similar user responses. Interestingly, regression and correlation analyses reveal subtle trends: the TEQ–PIS correlation was moderately negative for custom expressions ($r = -0.473$) and moderately positive for default expressions ($r = 0.408$), though neither reached significance. This might suggest that overly personalized animations, while visually distinctive, do not automatically align with users' expectations of emotional authenticity and may introduce minor inconsistencies in perception. In contrast, default animations benefit from industry-standard optimization and may produce more predictable emotional cues. These findings underscore that technical fidelity and user familiarity may be as important, or more so, than customization in influencing empathetic engagement. In practical terms, this indicates that for tasks aiming to foster empathy, investing resources into high-quality default animations could be more efficient than developing fully personalized facial expressions, unless personalization is accompanied by careful emotional calibration and contextual integration.

7.3.4. RQ4: How do users perceive the emotional authenticity and effectiveness of default facial expressions compared to custom-created facial expressions in a Digital Human?

The One-Way ANOVA again showed no significant effect of facial expression type on perceived emotional authenticity and effectiveness. Participants did not systematically perceive custom expressions as more authentic than default ones. This is further supported by the weak and non-significant correlation patterns (TEQ vs PIS: $r = -0.473$ for custom; $r = 0.408$ for default), which indicate that empathy did not strongly influence perception in either condition. These results suggest that facial expressions alone are insufficient to drive perceived emotional authenticity; rather, the integration of facial cues with body movement and contextual content likely plays a more substantial role. The minimal effect size emphasizes that customization, in isolation, may be visually noticeable but does not translate into measurable improvements in social perception. This has important implications for the design of Digital Humans: while personalization may enhance visual distinctiveness and user engagement, designers should prioritize holistic behavioural realism, ensuring that facial animations, gestures, and body language are coherently aligned to maximize emotional authenticity.

7.3.5. RQ5: What are the technical challenges and limitations in developing Digital Humans with emotional intelligence?

Developing emotionally expressive Digital Humans remains technically challenging. Translating real-time MoCap into high-fidelity MetaHuman animations requires significant computational power and careful calibration to avoid artificiality or detachment. Even minor discrepancies between intended emotional content and participant interpretation can reduce the perceived authenticity of the interaction. While current technologies, such as Unreal Engine and MetaHuman Creator, allow for detailed animation and behavioural control, achieving robust emotional intelligence requires integration of context-aware behavioural adjustments. The combination of technical constraints, computational requirements, and the need for precise alignment between visual, behavioural, and narrative cues underscores that producing emotionally convincing Digital Humans is a multi-dimensional problem. However, the study demonstrates that when these elements are optimized, particularly in the case of high-fidelity body movements, Digital Humans can evoke engagement and empathetic responses comparable to real human interactions, revealing both the potential and the limitations of current technology in supporting social and affective HCI.

8. Conclusion and Future Work

In today's increasingly digital world, there is a growing reliance on services that can mimic human interactions while prioritizing customer satisfaction and engagement. This study explored the potential of Digital Humans to produce empathetic multimodal expressions to enhance HCI. By personalizing user experiences, Digital Humans aim to reduce the emotional distance between users and automated services, offering a solution to the traditionally impersonal nature of such systems. Through a combination of Unreal Engine and advanced facial and body animation technology, the study developed Digital Humans capable of displaying human-like expressions and movements.

The study's findings revealed that Digital Humans are indeed capable of simulating emotional realism, providing an engaging and authentic user experience. While customized facial animations did not significantly outperform default motion-capture-based animations in eliciting empathy, both methods were equally effective in generating empathetic responses. This suggests that while personalizing facial animations enhances realism, other factors, such as narrative context and emotional appropriateness, play a critical role in fostering user empathy.

Importantly, the user tests demonstrated that body movement plays a critical role in shaping user perception. Both real human motion and digitally replicated MetaHuman movement exhibited strong, statistically significant correlations with empathy scores and perceived interaction quality, indicating that realistic motion — whether human or digital — enhances emotional engagement. These results highlight that, beyond facial expressions, full-body expressiveness is essential for creating emotionally engaging and believable Digital Human experiences.

This study makes several key contributions to the fields of HCI and digital communication. First, it demonstrates the practical viability of integrating facial and body animations to create emotionally resonant Digital Humans capable of enhancing empathy in user interactions. Second, it underscores the critical role of movement realism in shaping emotional perception, providing new insights into how motion design influences the overall user experience. Third, by comparing default and customized facial animations, the study reveals that personalization alone may not be sufficient to foster empathetic engagement; instead, contextual and narrative factors must also be considered.

Finally, this work builds upon broader concerns related to trust, security, and interaction quality in digital environments, as reflected in two scientific articles developed within the scope of the research grant and the work carried out for this dissertation. The first, “Security and Privacy in Physical–Digital Environments: Trends and Opportunities” [77], co-authored by the author and supervisors, explores issues of security, privacy, and trust in hybrid environments. The second, “Digital Avatars: Developing and Applying Customised MetaHumans” has been accepted for publication on the conference ICGI (International Conference on Graphics and Interaction) 2025. This latter contribution focuses on the technical development and application of customized MetaHumans, directly supporting the experimental component of the thesis. Together, these works highlight the dual emphasis of the project: addressing ethical and trust-related dimensions while advancing the practical creation and deployment of Digital Humans in service contexts.

Looking forward, future research could focus on integrating MetaHumans into real-world applications, including Metaverse spaces and customer service platforms, to evaluate their performance in practical, emotionally demanding settings. One promising direction is to combine Digital Humans with AI-driven chatbots capable of real-time emotional expressions and verbal interaction, enabling more complex, adaptive, and contextually appropriate emotional exchanges.

Additionally, the development of a custom editor within Unreal Engine, leveraging the MetaHuman framework, could simplify the workflow for creating and customizing Digital Humans, making the design process more accessible to developers and artists without technical expertise in animation or coding. Building on this, an automated tool could be created to allow any user to generate a personalized MetaHuman without requiring prior knowledge of the involved software.

Further directions include enabling the export of 3D MetaHuman models for use in other platforms and environments, as well as integrating real-time Digital Humans into video conferencing systems to enhance remote communication with richer emotional cues. Future studies might also explore ways to broaden emotional responsiveness, including handling negative emotions and conflict resolution scenarios, which are frequent in service contexts. Moreover, investigating the impact of cultural differences, age groups, and accessibility needs will be important to ensure these systems are inclusive and effective for diverse user populations.

Bibliography

[1] L. Atzori, A. Iera, and G. Morabito, “The Internet of Things: A survey,” *Computer Networks*, vol. 54, no. 15, pp. 2787–2805, Oct. 2010, doi: 10.1016/j.comnet.2010.05.010.

[2] M. Rosenlund, U.-M. Kinnunen, and K. Saranto, “The Use of Digital Health Services Among Patients and Citizens Living at Home: Scoping Review,” *J Med Internet Res*, vol. 25, p. e44711, Mar. 2023, doi: 10.2196/44711.

[3] C. Yang, M. Gu, and K. Albitar, “Government in the digital age: Exploring the impact of digital transformation on governmental efficiency,” *Technol Forecast Soc Change*, vol. 208, p. 123722, Nov. 2024, doi: 10.1016/J.TECHFORE.2024.123722.

[4] P. Dourish, V. B. Rank, and X. Europarc, “Awareness and Coordination in Shared Workspaces,” 1992.

[5] A. J. Sellen, “Human-Computer Interaction □ Remote Conversations: The Effects of Mediating Talk With Technology,” 2009.

[6] Rani Molla, “The pandemic was great for Zoom. What happens when there’s a vaccine?,” Vox. [Online]. Available: https://www.vox.com/recode/2020/12/4/21710457/zoom-covid-19-vaccine-users

[7] E. DeFilippis, S. M. Impink, M. Singell, J. T. Polzer, and R. Sadun, “The impact of COVID-19 on digital communication patterns,” *Humanit Soc Sci Commun*, vol. 9, no. 1, p. 180, May 2022, doi: 10.1057/s41599-022-01190-9.

[8] J. N. Bailenson, “Nonverbal Overload: A Theoretical Argument for the Causes of Zoom Fatigue,” *Technology, Mind, and Behavior*, vol. 2, no. 1, 2021, doi: 10.1037/tmb0000030.

[9] L. Laranjo *et al.*, “Conversational agents in healthcare: a systematic review,” *Journal of the American Medical Informatics Association*, vol. 25, no. 9, pp. 1248–1258, Sep. 2018, doi: 10.1093/jamia/ocy072.

[10] B. Schreiweis, M. Pobiruchin, V. Strotbaum, J. Suleder, M. Wiesner, and B. Bergh, “Barriers and Facilitators to the Implementation of eHealth Services: Systematic Literature Analysis,” *J Med Internet Res*, vol. 21, no. 11, p. e14197, Nov. 2019, doi: 10.2196/14197.

[11] S. Kokolakis, “Privacy attitudes and privacy behaviour: A review of current research on the privacy paradox phenomenon,” *Comput Secur*, vol. 64, pp. 122–134, Jan. 2017, doi: 10.1016/j.cose.2015.07.002.

[12] M. Abomhara and G. M. Kien, “Cyber Security and the Internet of Things: Vulnerabilities, Threats, Intruders and Attacks,” *Journal of Cyber Security and Mobility*, vol. 4, no. 1, pp. 65–88, 2015, doi: 10.13052/jcsm2245-1439.414.

[13] A. Caliskan, J. J. Bryson, and A. Narayanan, “Semantics derived automatically from language corpora contain human-like biases,” *Science (1979)*, vol. 356, no. 6334, pp. 183–186, Apr. 2017, doi: 10.1126/science.aal4230.

[14] J. Hill, W. Randolph Ford, and I. G. Farreras, “Real conversations with artificial intelligence: A comparison between human–human online conversations and human–chatbot conversations,” *Comput Human Behav*, vol. 49, pp. 245–250, Aug. 2015, doi: 10.1016/j.chb.2015.02.026.

[15] David H. Eberly, *3D game engine design: a practical approach to real-time computer graphics*. CRC Press, 2000.

[16] Epic Games, “MetaHuman Documentation.” [Online]. Available: https://dev.epicgames.com/documentation/en-us/metahuman/metahuman-documentation

[17] T. Antoniades, “Hellblade: Senua’s Sacrifice,” 2017, *Ninja Theory / Xbox Game Studios*: PlayStation 5 / Windows / Xbox Series X/S game.

[18] T. Antoniades, “Senua’s Saga: Hellblade II,” 2024, *Ninja Theory / Xbox Game Studios*: PlayStation 5 / Windows / Xbox Series X/S game.

[19] Kojima Productions, “OD,” Dec. 08, 2023, *Xbox Game Studios*: Unreal Engine 5. [Online]. Available: https://www.kojimaproductions.jp/en/od

[20] Kojima Productions, *OD – Official Reveal Trailer*, (Dec. 08, 2023). [Online Video]. Available: https://www.youtube.com/watch?v=Yi-5xTlWiiM

[21] M. Korban and X. Li, “A Survey on Applications of Digital Human Avatars toward Virtual Co-presence,” Jan. 2022, [Online]. Available: http://arxiv.org/abs/2201.04168

[22] X. Yang, Y. Fang, L. Wang, and R. C. Raga, “Digital Human Intelligent Interaction System Based on Multimodal Pre-training Mode,” *Applied Artificial Intelligence*, vol. 38, no. 1, 2024, doi: 10.1080/08839514.2024.2405953.

[23] M. Gleicher, “Motion capture and its applications,” *ACM Trans Graph*, vol. 37, no. 6, 2018.

[24] Tobias. Höllerer, Victoria. Interrante, Anatole. Lecuyer, and Evan. Suma, *2016 IEEE Virtual Reality (VR) : proceedings : Greenville, South Carolina, USA, 19-23 March 2016*. IEEE, 2016.

[25] M. Slater and S. Wilbur, “A Framework for Immersive Virtual Environments (FIVE): Speculations on the Role of Presence in Virtual Environments,” *Presence:*

Teleoperators and Virtual Environments, vol. 6, no. 6, pp. 603–616, Dec. 1997, doi: 10.1162/pres.1997.6.6.603.

[26] Epic Games, “Epic Games launches MetaHuman Creator.” [Online]. Available: https://www.cgchannel.com/2021/04/epic-games-announces-metahuman-creator/

[27] Epic Games, “Unreal Engine.” [Online]. Available: https://www.unrealengine.com

[28] Unity Technologies, “Unity.” [Online]. Available: https://www.unity.com

[29] CD Projekt Red, “Cyberpunk 2077.” [Online]. Available: https://www.cyberpunk.net

[30] Capcom, “Resident Evil Village.” [Online]. Available: https://www.residentevil.com/village/

[31] “Monster Hunter Wilds becomes fastest-selling RE engine title in Capcom’s history,” *VentureBeat*, 2025, [Online]. Available: https://venturebeat.com/games/monster-hunter-wilds-becomes-fastest-selling-re-engine-title-in-capcoms-history/

[32] Guerrilla Games, “Horizon Zero Dawn.” [Online]. Available: https://www.playstation.com/en-us/games/horizon-zero-dawn/

[33] Wccftech, “Death Stranding Uses Guerrilla’s Decima Engine; Kojima Setting Up Satellite Studio In Amsterdam,” 2016, [Online]. Available: https://wccftech.com/death-stranding-decima-engine-guerrilla/

[34] Epic Games, “Nanite Virtualized Geometry in Unreal Engine.” [Online]. Available: https://dev.epicgames.com/documentation/en-us/unreal-engine/nanite-virtualized-geometry-in-unreal-engine

[35] Epic Games, “Lumen Technical Details — Global Illumination and Reflections.” [Online]. Available: https://dev.epicgames.com/documentation/en-us/unreal-engine/lumen-global-illumination-and-reflections-in-unreal-engine

[36] Epic Games, “Balenciaga and Unreal Engine - Metaverse Trends.” [Online]. Available: https://www.unrealengine.com/en-US/blog/real-time-round-up-the-metaverse-and-emerging-2022-trends

[37] Unity Technologies, “Introducing the Metaverse Minute.” [Online]. Available: https://unity.com/blog/industry/introducing-the-metaverse-minute

[38] Guerrilla Games, “Horizon Forbidden West.” [Online]. Available: https://www.playstation.com/en-us/games/horizon-forbidden-west/

[39] Eidos Montreal, “Deus Ex: Mankinds Divided.” [Online]. Available: https://deusex.square-enix-games.com/

[40] Epic Games, “MetaHuman Creator is Now Available in Early Access.” [Online]. Available: https://www.epicgames.com/site/en-US/news/metahuman-creator-is-now-available-in-early-access

[41] Epic Games, “Unreal Engine 5.6.” [Online]. Available: https://www.unrealengine.com/en-US/news/unreal-engine-5-6-is-now-available

[42] S. García-de-Villa, D. Casillas-Pérez, A. Jiménez-Martín, and J. J. García-Domínguez, “Inertial Sensors for Human Motion Analysis: A Comprehensive Review,” Jan. 2024, doi: 10.1109/TIM.2023.3276528.

[43] M. Gutierrez *et al.*, “Comparing Optical and Custom IoT Inertial Motion Capture Systems for Manual Material Handling Risk Assessment Using the NIOSH Lifting Index,” *Technologies (Basel)*, vol. 12, no. 10, p. 180, Sep. 2024, doi: 10.3390/technologies12100180.

[44] C. Gu, W. Lin, X. He, L. Zhang, and M. Zhang, “IMU-based motion capture system for rehabilitation applications: A systematic review,” *Biomimetic Intelligence and Robotics*, vol. 3, no. 2, p. 100097, Jun. 2023, doi: 10.1016/j.birob.2023.100097.

[45] G. Fauville, M. Luo, A. C. M. Queiroz, A. Lee, J. N. Bailenson, and J. Hancock, “Video-conferencing usage dynamics and nonverbal mechanisms exacerbate Zoom Fatigue, particularly for women,” *Computers in Human Behavior Reports*, vol. 10, p. 100271, May 2023, doi: 10.1016/j.chbr.2023.100271.

[46] J. N. Bailenson, “Nonverbal overload: A theoretical argument for the causes of Zoom fatigue.,” *Technology, Mind, and Behavior*, vol. 2, no. 1, Feb. 2021, doi: 10.1037/tmb0000030.

[47] J. M. Basch, P. Albus, and T. Seufert, “Fighting Zoom fatigue: Evidence-based approaches in university online education,” *Sci Rep*, vol. 15, no. 1, p. 7091, Feb. 2025, doi: 10.1038/s41598-025-90973-6.

[48] Zoom, “About Zoom – We Deliver Happiness.” [Online]. Available: https://www.zoom.com/en/about/

[49] Microsoft, “Microsoft Teams service description.” [Online]. Available: https://learn.microsoft.com/en-us/office365/servicedescriptions/teams-service-description

[50] Google Workspace, “Google Meet: Online Web and Video Conferencing Calls.” [Online]. Available: https://workspace.google.com/products/meet/

[51] Cisco, “Webex – Collaboration & Meeting Platform Suite.” [Online]. Available: https://www.webex.com/

[52] A. C. M. Queiroz, A. Y. Lee, M. Luo, G. Fauville, J. T. Hancock, and J. N. Bailenson, “Too tired to connect: Understanding the associations between video-

conferencing, social connection and well-being through the lens of zoom fatigue," *Comput Human Behav*, vol. 149, p. 107968, Dec. 2023, doi: 10.1016/j.chb.2023.107968.

[53] M. Webb, "Zoom Fatigue and How to Prevent It.," *J Registry Manag*, vol. 48, no. 4, pp. 181–182, 2021.

[54] "How to deal with Zoom calls in 2025: in smaller groups with static backgrounds," *The Guardian*, Jan. 2025, [Online]. Available: https://www.theguardian.com/technology/2025/jan/04/how-to-deal-with-zoom-calls-in-2025-in-smaller-groups-with-static-backgrounds

[55] V. Woolaston, "AI news anchors: Why audiences might find digitally generated TV presenters hard to trust," *BBC Future*, 2024. [Online]. Available: https://www.bbc.com/future/article/20240126-ai-news-anchors-why-audiences-might-find-digitally-generated-tv-presenters-hard-to-trust

[56] Y. Wang, "Research on the Application of AI Anchors in TV Media under the Background of Artificial Intelligence," in *2021 2nd International Conference on Artificial Intelligence and Information Systems*, New York, NY, USA: ACM, May 2021, pp. 1–4. doi: 10.1145/3469213.3470671.

[57] K. Kroes, I. Saccardi, and J. Masthoff, "Empathizing with virtual agents: the effect of personification and general empathic tendencies," in *2022 IEEE International Conference on Artificial Intelligence and Virtual Reality (AIVR)*, IEEE, Dec. 2022, pp. 73–81. doi: 10.1109/AIVR56993.2022.00017.

[58] K. Loveys, M. Sagar, M. Billinghurst, N. Saffaryazdi, and E. Broadbent, "Exploring Empathy with Digital Humans," in *2022 IEEE Conference on Virtual Reality and 3D User Interfaces Abstracts and Workshops (VRW)*, IEEE, Mar. 2022, pp. 233–237. doi: 10.1109/VRW55335.2022.00055.

[59] D. Higgins, K. Zibrek, J. Cabral, D. Egan, and R. McDonnell, "Sympathy for the digital: Influence of synthetic voice on affinity, social presence and empathy for photorealistic virtual humans," *Comput Graph*, vol. 104, pp. 116–128, May 2022, doi: 10.1016/j.cag.2022.03.009.

[60] S. W. McQuiggan and J. C. Lester, "Modeling and evaluating empathy in embodied companion agents," *Int J Hum Comput Stud*, vol. 65, no. 4, pp. 348–360, Apr. 2007, doi: 10.1016/j.ijhcs.2006.11.015.

[61] A. C. Holland, G. O'Connell, and I. Dziobek, "Facial mimicry, empathy, and emotion recognition: a meta-analysis of correlations," *Cogn Emot*, vol. 35, no. 1, pp. 150–168, Jan. 2021, doi: 10.1080/02699931.2020.1815655.

[62] K. Kreijns, K. Xu, and J. Weidlich, "Social Presence: Conceptualization and Measurement," Mar. 01, 2022, *Springer*. doi: 10.1007/s10648-021-09623-8.

[63] S. Jiang, J. Liu, Y. Zhou, Z. Wang, and B. He, “mFERMeta++ : Robust Multiview Facial Expression Recognition Based on Metahuman and Metalearning,” *Advanced Intelligent Systems*, vol. 5, no. 10, Oct. 2023, doi: 10.1002/aisy.202300210.

[64] C. Liu and H. Wechsler, “Gabor Feature Based Classification Using the Enhanced Fisher Linear Discriminant Model for Face Recognition,” 2002.

[65] C. Shan, S. Gong, and P. W. McOwan, “Facial expression recognition based on Local Binary Patterns: A comprehensive study,” *Image Vis Comput*, vol. 27, no. 6, pp. 803–816, May 2009, doi: 10.1016/j.imavis.2008.08.005.

[66] S. Li and W. Deng, “Deep Facial Expression Recognition: A Survey,” *IEEE Trans Affect Comput*, vol. 13, no. 3, pp. 1195–1215, 2022, doi: 10.1109/TAFFC.2020.2981446.

[67] S. Jiang, P. Kang, X. Song, B. Lo, and P. Shull, “Emerging Wearable Interfaces and Algorithms for Hand Gesture Recognition: A Survey,” 2022, *Institute of Electrical and Electronics Engineers Inc.* doi: 10.1109/RBME.2021.3078190.

[68] A. Paiva *et al.*, “LEARNING BY FEELING: EVOKING EMPATHY WITH SYNTHETIC CHARACTERS,” *Applied Artificial Intelligence*, vol. 19, no. 3–4, pp. 235–266, Mar. 2005, doi: 10.1080/08839510590910165.

[69] A. Paiva, I. Leite, H. Boukricha, and I. Wachsmuth, “Empathy in Virtual Agents and Robots,” *ACM Trans Interact Intell Syst*, vol. 7, no. 3, pp. 1–40, Sep. 2017, doi: 10.1145/2912150.

[70] A. M. Rosenthal-von der Pütten, N. C. Krämer, L. Hoffmann, S. Sobieraj, and S. C. Eimler, “An Experimental Study on Emotional Reactions Towards a Robot,” *Int J Soc Robot*, vol. 5, no. 1, pp. 17–34, Jan. 2013, doi: 10.1007/s12369-012-0173-8.

[71] Polycam Inc., “Polycam,” 2025, *App Store*: 5.1.1. [Online]. Available: https://apps.apple.com/us/app/polycam-3d-scanner-lidar-360/id1532482376?utm_source=chatgpt.com

[72] S. Yeo and K.-J. Kim, “A validation study of the Korean version of the Toronto empathy questionnaire for the measurement of medical students’ empathy,” *BMC Med Educ*, vol. 21, no. 1, p. 119, Dec. 2021, doi: 10.1186/s12909-021-02561-7.

[73] A. AL Lihabi, “Validity and reliability of the Arabic version of the Toronto Empathy Questionnaire (TEQ) among Saudi medical students: a cross-sectional study,” *Ment Health Relig Cult*, vol. 27, no. 2, pp. 135–149, Feb. 2024, doi: 10.1080/13674676.2023.2264199.

[74] K. A. Bin Abdulrahman *et al.*, “Empathy level among Saudi medical students using the Toronto empathy scale,” *World Family Medicine Journal /Middle East Journal of Family Medicine*, vol. 20, no. 11, Nov. 2022, doi: 10.5742/mewfm.2022.95251392.

[75] R. N. Spreng*, M. C. McKinnon*, R. A. Mar, and B. Levine, “The Toronto Empathy Questionnaire: Scale Development and Initial Validation of a Factor-Analytic Solution to Multiple Empathy Measures,” *J Pers Assess*, vol. 91, no. 1, pp. 62–71, Jan. 2009, doi: 10.1080/00223890802484381.

[76] P. Voultsos, F. Chatzinikolaou, A. Papana, and A. Deliligka, “Reliability of Greek version of the Toronto empathy questionnaire in medical students and associations with sociodemographic and lifestyle factors,” *BMC Psychol*, vol. 10, no. 1, p. 113, May 2022, doi: 10.1186/s40359-022-00824-6.

[77] C. Pereira *et al.*, “Security and Privacy in Physical–Digital Environments: Trends and Opportunities,” *Future Internet*, vol. 17, no. 2, p. 83, Feb. 2025, doi: 10.3390/fi17020083.