

가상공간에서 아바타 외모와 목소리의 표현 방식에 따른 사용자 반응 연구: 관계에 따른 차이를 중심으로

강민정*
홍익대학교 조형대학 부교수

User Responses to Avatar Appearance and Voice Expression in Virtual Spaces: Focusing on Differences Based on Relationships

Minjeong Kang*

Associate Professor, College of Arts and Design, Hongik University, Sejong 30016, Korea

[요약]

본 연구는 가상공간에서 친목을 목적으로 아바타를 사용할 때, 실제감, 친근감, 신뢰감을 형성하는 아바타 이미지와 목소리 표현에 대해 탐구하였다. 포커스 그룹 인터뷰를 통해 아바타의 목소리가 신뢰감을 형성하는데 중요한 역할을 하며, 관계에 따라 선호하는 아바타 이미지가 다르다는 점을 확인하였다. 이를 검증하기 위해 설문 조사를 진행한 결과 실제감과 신뢰감에서는 사실적 이미지와 실제 목소리가 모두 중요한 반면 친근감에서는 이미지의 실제감은 상대적으로 중요하지 않았다. 관계에 따라서 아바타 선호도가 통계적으로 다르다는 것을 확인하였고 친밀한 관계에서는 사실적 아바타, 모르는 관계에서는 실제 목소리에 대한 선호도가 두드러졌다. 반면, 수직적 관계에서는 비현실적이면서 친근한 아바타를 선호하였고, 수평적 관계에서는 만화 이미지에 대한 선호도가 다른 그룹보다 높게 나타났다. 이러한 연구 결과는 가상공간에서 AI 아바타가 사용자를 대신해 소통할 때 관계에 따라 아바타를 설정할 수 있도록 서비스를 디자인하는데 유용한 참고자료가 될 것이다.

[Abstract]

This study examines the influence of avatar images and voice expressions on social presence, intimacy, and trust in virtual spaces designed for social interaction. Findings from focus group interviews indicate that an avatar's voice plays a crucial role in establishing trust, while preferences for avatar images vary depending on the nature of the relationship. Surveys conducted to validate these findings revealed that realistic images and actual voices enhance social presence and trust; however, realistic images are less critical for fostering intimacy. Preferences for avatars also differed significantly based on relationship type: realistic images were preferred in intimate relationships, real voices in interactions with strangers, non-realistic friendly avatars in hierarchical relationships, and while realistic images were still most preferred in peer-to-peer relationships, cartoon images showed relatively higher preference in this group compared to others. These findings offer valuable insights for designing AI-driven avatar services tailored to different relationship contexts in virtual spaces.

색인어 : 아바타 이미지, 아바타 목소리, AI 목소리, 사용자와의 관계, 가상공간

Keyword : Avatar Image, Avatar Voice, AI Voice, User-Avatar Relationship, Virtual Space

<http://dx.doi.org/10.9728/dcs.2025.26.2.297>



This is an Open Access article distributed under the terms of the Creative Commons Attribution Non-Commercial License (<http://creativecommons.org/licenses/by-nc/3.0/>) which permits unrestricted non-commercial use, distribution, and reproduction in any medium, provided the original work is properly cited.

Received 24 December 2024; **Revised** 15 January 2025

Accepted 24 January 2025

***Corresponding Author; Minjeong Kang**

Tel: [REDACTED]

E-mail: mjmiso@hongik.ac.kr

I. 서론

1-1 연구 배경 및 목적

Market.us에 따르면 2022년 글로벌 디지털 아바타 시장 규모는 135억 달러를 달하여 전년 대비 45.8%의 연평균 성장률(CAGR)을 기록하였고 이후 2032년까지 5,338억 달러의 매출을 달성할 것으로 예상된다. 이러한 성장은 게임, 가상 현실(VR), 증강 현실(AR), 소셜 미디어, 전자 상거래 등 다양한 산업에서 디지털 아바타의 채택이 증가할 것이기 때문이다[1]. 이러한 시대의 변화에 맞춰 여러 기업들이 독자적인 메타버스 환경을 구축하면서 사용자들은 자신을 대표하는 아바타를 각 서비스에서 중복해서 만들어야 하는 부담이 나타났다. 이러한 이유로 2022년 6월에 메타버스 표준포럼(metaverse standard forum)을 발족하게 되면서 여러 플랫폼과 호환이 가능한 아바타 생성 플랫폼도 등장하게 되었다[2]. 이러한 플랫폼 사례로서 레디플레이어미(readyplayerme)는 사진을 올리면 자신과 유사한 아바타를 생성해주며 ‘브이알챗(VRChat)’, ‘스페이셜(spatial)’ 등 다양한 VR회사들과 제휴를 맺고 있다[3]. 또 다른 아바타 생성 플랫폼인 제니스(Genies)는 귀여운 만화 캐릭터 스타일로 개인화된 아바타를 만들어 아이메시지(imessage) 및 인스타그램(instagram)과 같은 앱에서 사용할 수 있다[4]. 오벤(OBEN)은 인공지능 기술을 통해 음성 및 이미지 데이터의 최소한의 샘플 사용으로 사용자와 똑같은 외모와 동일한 목소리를 구현하며, 사람처럼 행동하도록 훈련시킬 수 있는 지능형 3D 아바타 생성 플랫폼이다[5].

이와 같이 다양한 표현 방식의 아바타들이 등장하고 다양한 목소리 설정 기술도 가능해지면서 가상공간에서 아바타를 표현하는데 있어서 인간이 느끼는 편안함에 대한 다수의 연구들이 진행되어 오고 있다. 아바타의 외양에 대한 구현 수준을 보면, 사용자들은 만화적인 캐릭터보다는 사실적으로 구현된 아바타를 더 선호하며, 사실성이 높아질수록 친밀감도 높아지고 사용에 따른 즐거움과 사용의도가 높아진다[6]. 그러나 신, 김, 비오카의 연구에 따르면 가상공간에서 첫인상 테스트를 한 결과 언캐니 현상 때문에 사실적이고 움직이는 아바타가 만화적이고 정지된 아바타에 비해 더 큰 섬뜩함을 유발할 수 있다[7]. 따라서 자연스러운 사실적인 아바타 동작 구현이 가능해져서 언캐니한 현상이 사라진다면 그때도 사용자들이 만화적인 아바타를 선호할지에 대해서는 좀 더 연구가 필요하다. 또한 본 연구에서는 이러한 아바타 표현에 대한 선호도는 상대방과의 어떤 관계인지에 따라 다를 것이라고 가정하였다.

목소리는 준언어적 요소로서 사용자의 성별, 인종, 교육 수준 등이 드러나기 때문에 소셜vr과 같은 메타버스 공간에서 사람과 커뮤니케이션할 때 그 사람의 정체성을 드러내는데 매우 중요한 역할을 한다[8]. 이러한 이유 때문에 가상공간에서 자신의 프라이버시를 지키거나 재미를 위해 ‘보이스모드(Voicemod)’와 같은 실시간 목소리 변조가 가능한 다양한

서비스들이 출시되었다.

이와같이 가상공간에서는 이미지와 목소리를 자유롭게 설정할 수 있기 때문에, 긍정적인 인상을 주는 아바타 이미지와 목소리의 조합을 파악하는 것이 필요하다. 그러나 선행연구를 조사한 결과, 언캐니(uncanny) 현상을 중심으로 한 아바타 표현에 관한 연구는 다수 존재하지만, 아바타 이미지와 목소리 간 중요도 차이 연구는 찾아보기 어려웠다. 또한, 아바타 이미지와 목소리에 대한 선호도가 아바타 사용자와의 관계에 따라 어떻게 달라지는지를 밝힌 연구 역시 부족한 실정이다.

이에 본 연구는 상대 아바타 이미지와 목소리 중 무엇이 더 중요한지 파악하고, 선호하는 아바타 이미지와 목소리 조합이 관계의 특성에 따라 어떻게 달라지는지를 탐구하고자 한다. 이러한 연구 결과는 가상공간에서 소통 시 긍정적인 이미지를 전달할 수 있는 아바타 설정 기능을 설계하는 데 있어 중요한 참고자료로 활용될 수 있을 것으로 기대된다.

1-2 연구 방법

본 연구는 가상공간에서 아바타를 통해 상대방을 인지하는데 있어서 아바타의 외모와 목소리의 관련성에 따른 사용자 반응에 대한 연구를 목적으로 한다. 연구방법으로 1) 문헌 연구를 통해 아바타의 스타일과 목소리에서 드러나는 인상에 대한 선행 연구들을 고찰한다. 2) 이후 가상공간에서 타인의 아바타에 대한 다양한 경험들을 파악하기 위해 VR전공 대학생들을 대상으로 FGI를 진행한다. 3) FGI에서 도출한 발견점들을 검증하기 위해 아바타 이미지를 의인화 정도에 따라 구분하고 목소리는 실제 목소리와 긍정적인 인상을 주는 AI 목소리로 구분하여 더빙을 한다. 이후 4) 설문 조사를 통해 각각의 아바타 이미지와 목소리를 크로스도 매치하여 피험자에 대한 인상이 어떻게 달라지는지 파악하고자 한다.

II. 문헌연구

2-1 사회적 실재감(Social Presence)

소셜 프레즌스는 온라인 환경에서 다른 사람이나 개체(entity)와 상호작용할 때 느끼는 실재감이나 친밀감을 의미한다[9]. 이 개념은 상대방의 존재를 얼마나 현저히 느끼는지를 나타내는 인지적 사회적 실재감(cognitive social presence)과 상대방과 상호작용을 하면서 나타나는 정서적 사회적 실재감(emotional social presence)으로 구분된다[10].

인지적 사회적 실재감은 대상의 존재 정도를 얼마나 '인식(perception)'하는지와 이야기를 얼마나 '파악(Understand)'하는 지로 나뉜다. 정서적 사회적 실재감은 타인과의 사회적 상호작용 중에 형성되는 정서적 유대감을 의미하며, 이는 '친밀감(Intimacy)'과 '즉시성(Immediacy)'라는 하위 요인으로 나누어진다[11]. 가상공간에서 디지털 휴면에 대한 '친밀감'

은 대상에 대해 심리적으로 느끼는 감정과 연결을 의미하고, ‘즉시성’은 지금 여기(here-and-now)에 초점을 맞추는 개념으로 즉각적으로 반응하고 상호작용할 수 있는지에 관련된 개념이다[12]. 본 연구에서는 그림1의 빨간 박스로 표시한 것처럼 아바타 표현에 따라 목소리와 이미지에서 느껴지는 존재에 대한 인식으로 부터의 ‘실재감’과 정서적으로 연결감을 느끼는 ‘친밀감’에서의 차이를 파악하고자 한다.

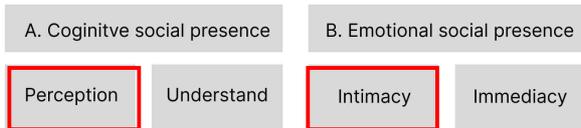


그림 1. 사회적 실재감의 개념
Fig. 1. The concept of social presence

2-2 아바타의 의인화 정도에 따른 인식 차이

신뢰감 측면에서 볼 때 가상환경에서 사실적 아바타를 사용한 집단에서 상대 아바타의 키가 클 때 참여자가 상대에게 더 설득되었고, 사회적 현존감에 있어서는 만화적 아바타를 사용한 집단이 사실적 아바타를 사용한 집단에 비해 더 높게 나타났다[13]. 또한 이연빈 외 2인의 연구에 따르면 사실성이 높은 디지털휴먼이 일러스트로 표현된 인간형 아바타 보다 인지적 사회적 실재감의 지각 부분과 정서적 사회적 실재감의 친밀감 부분에서 높게 나타났다. 그러나 즉시성에서는 디지털 휴먼과 인간형 아바타의 점수 차이가 크지 않았다 [10]. 종합해보면 표 1에서 보듯이 사실형 아바타가 만화형 보다 대체적으로 신뢰도, 실재감, 정서적 측면에서 더 긍정적인 반응을 보였지만 현존감, 즉시성 등에서는 만화적 표현이 더 높게 나타났다.

본 연구에서는 이러한 선행연구에서 밝힌 아바타에 대한 반응을 분석을 기준으로 신뢰감, 사회적 실재감, 친밀감으로 분석 기준을 정하고자 한다.

표 1. 아바타 표현에 따른 인식의 차이
Table 1. The cognitive difference of the avatar expression

	Criteria	Assessment Based on Avatar Style		Adopted Criteria
		Photo	Cartoon	
J. Kim (2019)	Trust	✓		✓
	Social Presence		✓	
Y.Lee et al. (2020)	Social Presence	✓		✓
	Intimacy	✓		✓
	Instance	✓	✓	N/A

신민철 외 2인은 아바타 표현에 따른 인식의 차이를 밝히기 위해 사실적 이미지와 만화 같은 이미지를 정적/동적인 버

전으로 동일한 얼굴 표정을 가지도록 디자인하여 첫인상에서 받아들이는 정보의 양과 호감도에 대한 연구를 실시하였다. 그 결과 사용자들은 초현실적이고 움직이는 아바타가 만화적이고 정지된 아바타에 비해 더 큰 섬뜩함을 유발할 수 있음을 확인하였다[7].

또한 연예인과 같이 익숙한 얼굴에 대한 연구에서 연예인 및 일반인의 사진과 그 사진을 기반으로 제작된 아바타에서 눈 크기를 조절하여 비정형성 정도를 조작하였고, 이에 따른 언캐니밸리 현상의 정도를 측정하였다. 그 결과, 사진과 만화 버전 모두 연예인 얼굴에 대해서 언캐니밸리 현상이 일반인 얼굴 보다 더 강하게 나타났다[14]. 이는 친숙한 얼굴에 대해서 더 강한 얼굴 표상이 형성되어 얼굴 변화에 대해 더 쉽게 탐지한다는 기존의 발견들과 일치한다. 따라서 가상공간에서 실제얼굴과 유사한 AI 아바타를 제작하여 사용할 때 아바타가 조금이라도 어색함을 가지고 있으면 오히려 부정적 반응을 가져올 수 있다.

또 다른 연구에서는 사람 아바타 뿐 아니라 다른 비현실적인 AI챗봇 이미지에 대한 반응 연구에서 재미를 목적으로 챗봇과 대화할 때 성별을 알 수 없는 동물이나 로봇의 이미지를 선호하는 것으로 나타났다[15]. 이는 온라인에서 재미를 추구할 때는 중성적인 이미지에 호감을 가질 수 있을 것이라 유추할 수 있다.

따라서 본 연구에서는 이러한 선행연구를 참고하여 친목을 목적으로 하는 가상공간에서 얼굴이 익숙한 지인을 만날 경우와 낯선 이를 만날 경우, 사실적 이미지, 만화적 이미지, 중성적인 느낌의 동물 이미지를 자극물로 설정하여 선호하는 상대방의 아바타 이미지 차이를 알아보하고자 한다.

2-3 목소리에 따른 반응

목소리는 들었을 때 그 사람이 누구인지에 대한 정체성과 그 사람이 어떤 감정인지 파악을 돕는 역할을 한다[16]. 이와 같이 목소리는 사용자의 성별, 인종, 교육 수준 등을 반영하기 때문에 첫인상 뿐 아니라 이미지 형성에도 많은 영향을 준다 [8]. 따라서 가상공간에서 목소리로 인한 피해를 줄이기 위해 목소리 변조에 대한 니즈와 서비스들이 등장하고 있음에도 아바타 목소리에 대한 연구는 미비한 상황이다. 도미니크 카오의 연구팀에서는 여성들이 성별을 알 수 없는 중성적 목소리로 온라인 미팅을 통해 진행했을 때 생산성이 좋아지는 것을 밝혔다[17]. 또 다른 연구에서는 음성 변조기를 통해 트랜스 젠더들이 자신의 정체성을 자유롭게 드러내도록 돕는 연구를 진행하여 음성 변조기가 트랜스젠더와 성비순응자 사용자가 직면하는 문제를 완화하는데 도움이 되는 것을 확인하였다[18]. 이러한 연구들은 사용자 목소리의 성 정체성에 기반 한 연구이기 때문에 아바타의 이미지와 목소리의 조합에 따라 달라지는 인상에 대한 연구는 아직 진행된 것을 발견하지 못하였다. 따라서 본 연구에서는 긍정적인 AI 목소리를 사용하는 아바타와 실제 목소리를 사용하는 아바타에 대한 반

응의 차이를 밝혀보고자 한다.

긍정적인 인상을 만드는 목소리 연구들을 살펴보면 긍정적인 정서상태의 목소리를 사용하면 이와 상호작용하는 사용자의 정서도 긍정적으로 전환된다[19]. 예를 들면 웃는 목소리로 정보를 제공하면, 사용자의 신뢰도가 높아진다[20]. 또한 나이가 많은 목소리로 설명하면 신뢰감이나 사회적인 실재감을 더 줄 수 있다는 결과도 있다[21]. 매력적인 얼굴을 가진 여성이 좋은 목소리를 가지고 있으며, 일반적인 여성의 목소리 높이보다 높은 주파수의 목소리일 경우 성적으로나 신체적으로 더 매력적이거나 실제 나이보다 어린 여성으로 추측하는 경향이 있다[22]. 반면, 남성의 경우, 여성과는 달리 목소리의 톤이 낮고, 느린 목소리를 더 매력적으로 느끼는 경향이 있는 것으로 나타났다. 따라서 목소리 매력이라는 개념은 교양 있는, 인지적, 감정적 반응을 이끌어 내는데 이러한 목소리를 만드는데 기여하는 것은 목소리의 높낮이, 강도, 말의 속도, 말의 멈춤, 스피치 시간 등의 변인과 관련이 높다고 할 수 있다[23]. 빠르게 말하는 사람은 외향성이 높고, 자신감이 있는 사람으로 인식하는 경향이 있으며, 보통 속도로 말하는 사람보다 더 믿을만한 것으로 지각된다[24],[25]. 이러한 기준을 통해 본 연구에서는 실제 목소리의 대조군으로 AI목소리를 선정할 때 참고해보고자 한다.

III. 사용자 인터뷰

3-1 인터뷰 계획

본 연구에서는 가상공간에서 사용자들이 아바타를 선택하고 활용하는 방식, 아바타 유형에 따른 상대 사용자에 대한 인식, 그리고 지인이 목소리를 변조했을 때 사용자가 느끼는 감정을 탐구하고자 한다.

표 2. FGI 참여자 정보

Table 2. Participants information for FGI

Group	Name	Gender/ Age	Frequency of VR Use	Main Purpose of Use
A	Choi **	Male/20s	Almost daily	Communication with friends
	Park **	Male/20s	More than once a month	Others
	Ryu **	Female/20s	Rarely used	Gaming
	Lim **	Male/20s	Rarely used	Communication with friends
B	Kwon **	Female/20s	2-3 times every few months	Gaming
	Lee **	Male/20s	2-3 times every few months	Others
	Heo *	Male/20s	Rarely used	Gaming
	Jung **	Male/20s	More than once a month	Gaming

이를 위해 수도권 소재의 VR관련학과 2,3학년 학생들 8명을 리크루팅 하여 남녀 혼성으로 표 2와 같이 4명씩 2개 그룹으로 FGI(focus group interview)를 실시하였다. 인터뷰 전에 사전 설문조사를 진행하여 VR 사용빈도, 주사용 목적, 사용기기, 아바타 설정 시 선호하는 스타일, 친근감과 신뢰감을 느끼게 하는 요인 등을 파악하였다.

이후 인터뷰는 각 그룹당 1시간씩 진행되었고 질문에 대해서 순서 없이 자유롭게 답변하도록 하였다.

A팀은 주로 소통을 목적으로 하는 사용자들이 모였으며 이중에서도 VR서비스를 매일 사용하는 헤비유저가 참여를 하였고 B팀은 소통 보다는 게임을 주로 목적으로 하는 사용자들이 모여 인터뷰를 진행하였다. FGI 진행을 위한 질문은 아래 표와 같다.

표 3. 인터뷰 질문

Table 3. Interview questions

1. How do you set up your own avatar in VR?
2. What kind of impression do you get when you see someone else's avatar in a VR environment?
3. What factors build trust in a VR environment?
4. What impression do you get if a friend's avatar looks very different from their usual appearance?
5. How would you feel if your friend's voice sounded different from their usual voice?

3-2 사전 설문조사 결과

인터뷰에서 심도 있는 질문을 이어가기 위해 사전 설문조사를 진행하였다. 가상공간에서 친근감과 신뢰감에 영향을 주는 아바타의 외모, 목소리, 행동을 우선순위로 답하도록 하였다. 설문 결과 친근감을 느끼게 하는 요인에 대한 응답으로 아바타의 행동(2.29) > 아바타의 외모(2.0) > 아바타 목소리(1.7) 순으로 나타났다. 이러한 결과를 통해 친근감을 느끼는데 있어서 시각적 요인인 아바타의 외모와 행동이 의미 있게 나타난 것으로 보인다. 반면에 신뢰감에 영향을 주는 요인은 아바타의 행동(2.63) > 아바타 목소리(2.00) > 아바타 외모(1.38)로 가상공간에서도 현실과의 일치성을 가지는 측면이 신뢰감을 느끼는 요인으로 보인다. 목소리도 원하는 음색으로 설정이 가능하기 때문에 사실상 사용자의 정체성을 숨길 수 없는 것은 메타버스 안에서의 행동으로 볼 수 있다.

3-3 인터뷰 결과

FGI내용을 종합 분석한 결과 참여자들의 아바타 설정 행태는 크게 4가지 유형으로 나타났다. 1)자신이 좋아하는 이미지나 스타일대로 꾸미거나, 2)타인이 좋아할 것 같은 이미지로 꾸미거나, 3)내 모습과 유사하게 꾸미거나, 4)아무것도 꾸미지 않고 기본 아바타를 사용하는 행태로 나타났다. 아바타 설정 행동에서 나타난 두드러진 특징으로는 자신의 실제 성별과 다르게 꾸미거나, 지인과 대화할 때는 재미있는 캐릭터를

사용하는 반면, 공식적인 자리에서는 무난한 아바타를 선택하는 경향이 있었다. 처음 보는 사람을 만날 때에는 임팩트 있게 꾸미고 싶다는 의견도 있었다.

상대 아바타를 볼 때 첫 인상에서 그 사람의 정서가 느껴지는데 나와 크게 다르지 않은 아바타는 친근감이 느껴져서 접근하고 싶은 마음이 든다고 응답하였다. 너무 튀는 아바타는 신뢰감이 안 간다는 의견과 오히려 말을 걸어보고 싶다는 상충되는 의견이 모두 있었다. 이미 알던 지인의 경우 실제 모습과 다른 이미지로 꾸민 아바타를 보면 호감도가 올라간다고 응답하였다. 또한 서비스의 분위기와 안어울리는 아바타는 비호감으로 부정적인 이미지로 느껴진다는 의견도 있었다.

이러한 아바타의 외모는 신뢰도에 있어서는 말과 행동 보다 중요도가 낮으며 실제 목소리 보다는 말의 내용이 중요하다고 응답 하였다. 또한 지인이 다른 목소리로 설정하는 것에 대한 반응은 대부분 부정적이었으며 특히 다른 성별로 음성을 변조한 경우 거부감이 큰 것으로 나타났다. 이러한 발견점들은 설문조사를 통해 검증하고자 한다.

표 4. FGI응답

Table 4. FGI response

	Findings	Interview Answers
Avatar Setup	Set up for my satisfaction	I always have a set style that I like, so I just adjust the important parts and finish.
		I decorate it in the style I like, without being conscious of others.
		I set the avatar to something I find cute.
		I set the avatar to my favorite celebrity or character.
	Choose the default avatar	I use the default avatar as it is.
I set it to a style that others would like	As a VRChat user, I decorate it in a highly unrealistic style that others might like.	
	I choose an avatar that I think others would find funny.	
I make it resemble myself	Actually, VR is a bit different from regular games. In VR, I meet people directly, while in games, I play within the game world. SO when I use VR, I make it very similar to my real appearance, but when I play games, I make it completely different.	
	I make it resemble myself when I use VR.	
Avatar Usage Behavior	I change my avatar if it's not satisfying, considering others	I change my avatar if it is not satisfying, considering others' or MPC's avatars

	Findings	Interview Answers
Avatar Usage Behavior	I set it to look like what I want to see	When using VR and playing games, the way I customize my avatar is different. In games, I choose the appearance I want to see
	I make it resemble myself if I identify with the avatar	If I recognize that I entered a game as a female, I set my avatar to female.
	I mostly set it to the opposite gender	I usually customize my character to a different gender in VRChat
		Men choose female characters when playing games and set them to look like what they want to see
		Even if I use a female appearance in VRChat, everyone still perceives me as male.
		When setting an avatar in games, since I have to keep looking at it, I customize it to a beautiful female instead of a dull male
	I treat the avatar like my child	I think it's more fun when I set it to a male character instead of a female, and I enjoy seeing a different side when I move the male character
		I consider the character as if it were my child
	When I meet acquaintances, I would set it in a fun way	There are many people who really adore their character and raise the character they created
		When I do VR, I end up decorating it in a fun way because I meet people I know
For formal settings, I use a neutral avatar.	In meetings or important situations, I would naturally set it up simply and neutrally to avoid any discomfort.	
I customize it to leave an impression when meeting new people	When I first meet someone, I think I need to make a memorable impact, so I would decorate it in a fun way or try a bold style.	
Reaction to other avatars	I approach when I feel familiarity from the avatar's appearance.	At first, I approach based on the appearance
	I want to talk to someone with a flashy avatar.	When I see a loud and unique avatar, I feel amazed and want to start a conversation
	I sense emotions through the avatar.	I can sense the person's emotions from their avatar

	Findings	Interview Answers
Reaction to other avatars	An overly flashy avatar does not instill trust.	I feel a sense of familiarity from an avatar that is not much different from mine
		I usually approach a plain avatar, because if it's too flashy, it feels overwhelming.
		I usually approach a plain avatar, because if it's too flashy, it feels overwhelming
		Even though it's a virtual space, since their appearance is visible to others, a person who excessively acts provocatively despite knowing this seems to have an issue with their mental state
	An appearance that is disconnected from the avatar atmosphere of the service does not instill trust.	It might actually be scary if someone has too ordinary an appearance that doesn't fit the atmosphere in the virtual world.
		It feels scary if someone adopts a style that is completely different from the atmosphere of the avatars provided in VR.
In the virtual world, if someone looks like a completely ordinary person, I feel like I won't find them attractive.		
In the virtual world, trust is more influenced by speech and behavior than appearance.	In the case of an ordinary appearance, I feel trust from their speech and behavior.	
	Even if the appearance stands out, I trust the person when I hear their real voice.	
The content of the speech is more important than the voice.	In the virtual world, there is limited information about the actual person. It's harder to judge just by appearance or voice compared to the real world, so the content of the speech seems more important.	
	In the virtual world, some people might say 'Hello' or 'Hi,' and I am influenced by their tone of speech.	
True identity is revealed through behavior.	Behavior reflects the real self, so it gives more trust than appearance.	
	If a close friend of mine takes really good care of me when they're hanging out with other close friends, it means I'm developing feelings for them,	
Reactions to an acquaintance's avatar settings	People feel more attracted to avatars that are opposite to the image of the acquaintance.	If an acquaintance presents an appearance opposite to their real one, it feels more friendly.
		It would be memorable and fun if an acquaintance shows a different look through their avatar.
		Seeing a different side of an acquaintance feels positive.

	Findings	Interview Answers
Reactions to an acquaintance's avatar settings	When an acquaintance changes their voice to the opposite gender, it feels repulsive.	It would be confusing if the avatar looks almost like the acquaintance but the voice is different.
		Everyone uses a female voice modulation when speaking. But if an acquaintance is speaking in such a modulated voice, and their name blinks like 'Park**', it's honestly disgusting.
		Some YouTubers are men who dress as women and do female voices well. If they suddenly change to a male voice during a conversation, it's surprising. It would be shocking if someone I know does that, especially when I'm familiar with their appearance.
		Voice modulation to a neutral, alien-like gender would be better than to another gender.
	Using a celebrity's voice seems to influence their image.	If an avatar is set with Yoo Jae-suk's voice, it might feel trustworthy because of the celebrity's image.
		When meeting an acquaintance whose avatar and voice are different, I might not recognize them and perceive them as the character of the modified voice.
	If a different voice is used, there is no desire to maintain the relationship	In the virtual world, having one's own voice is the only truth, but if even that is someone else's, it feels like I know nothing about this person.
We talk online, but it seems unlikely that we would meet in reality.		
In real meetings, the voice also confirms the identity of the other person, so it's not awkward.	I went to meet someone offline without hearing their voice first, and although they were different from what I expected, it wasn't completely off-putting as their way of speaking was the same.	
		If the voice is modulated, I might maintain the relationship online, but in reality, since I don't know anything for sure, I wouldn't meet them. Hiding the voice feels like deception.
		It would be annoying if an acquaintance modulated their voice.

IV. 설문조사

4-1 설문조사 목표

위의 FGI 에서 분석한 발견 점들을 다음과 같이 2가지 축

면에서 검증해보고자 설문조사의 목표를 설정하였다.

1) FGI 결과에서 대부분 지인이 목소리를 변조했을 때 부정적 느낌을 가졌고 처음 보는 사람의 경우에도 변조된 목소리라면 오프라인으로 관계를 이어나가지 않을 것이라고 응답하였다. 이와 같이 실제목소리가 중요하다는 응답이 있었지만 소수의 의견이기 때문에 아바타 설정시 사용자의 실제 목소리와 사실적 이미지 설정의 중요도 차이를 밝혀 볼 필요가 있다. 따라서 가상공간에서 관계를 유지하는데 있어서 실제 목소리의 중요성을 검증하기 위해 신뢰감이 느껴지는 ai목소리와 실제 사용자 목소리를 아바타 이미지와 매치하여 실제감, 친근감, 신뢰감에서 차이를 비교해보고자 한다.

2) 가상공간에서 지인들을 만났을 때 평소와 다른 재밌는 아바타 모습일 때 호감도가 올라간다는 반면에 처음 대하는 경우에는 너무 과격적인 이미지 보다는 자신과 유사한 모습의 아바타일 때 좀 더 다가가기 쉽다고 하였다. 이는 만나는 관계에 따라 편하게 느끼는 아바타 이미지에 차이가 있다고 가정할 수 있다. 따라서 이러한 차이를 확인하기 위해 샘플이 되는 피험자를 정하고 그를 기준으로 가족과 같은 친밀한 관계, 친구와 같은 수평적 관계, 사제지간과 같은 수직적 관계, 서로 모르는 관계의 사람들에게 설문조사를 하고자 한다.

4-2 설문조사 계획

설문 조사를 진행하기 위해 한명의 피험자를 선정하여 그림 2와 같이 사실적, 만화적, 비현실적으로 3가지 버전의 아바타 이미지를 만들고 목소리는 실제목소리와 ai목소리를 선정하여 실제감, 친근감, 신뢰감을 파악하고자 한다.

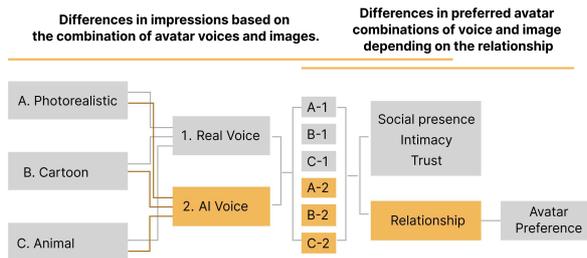


그림 2. 설문조사 계획
Fig. 2. Survey plan

자극물 제작은 표 5에서 보듯이 사실적 이미지 아바타를 만들기 위해 피험자의 실제 사진을 가지고 D-ID프로그램을 이용하여 영상으로 만들고 미리 녹음해둔 실제 목소리와 AI 목소리를 더빙하여 제작하였다. D-ID는 AI를 활용하여 이미지, 아바타 등 가상 인물이 말하는 영상을 제작할 수 있는 프로그램이다. 같은 사진을 만화적 이미지로 변환하기 위해 ‘툰앱’(toonapp)을 이용하였고 이후 다시 D-ID프로그램에서 영상으로 제작하였다. 비현실적 이미지로는 선행연구에서 밝힌 바와 같이 재미를 추구하는 상황에서는 동물과 같은 중성적

이미지의 아바타를 선호한다는 결과[15]를 참고하여 범용적으로 사용되는 화상회의 플랫폼인 ZOOM에서 제공하는 강아지 아바타를 채택하였다.

표 5. 설문조사를 위한 아바타 자극물
Table 5. Avatar stimuli for survey

Image						
Voice	Real	AI	Real	AI	Real	AI
Avatar combination	Photo + Real	Photo AI	Cartoon + Real	Cartoon + AI	Animal +Real	Animal +AI
	PR	PA	CR	CA	AR	AA

피험자는 40대 전문직 여성으로 선정하였고 직업에 맞게 AI목소리는 신뢰감을 주는 여성의 목소리로 임의 선정하였고, 표 6과 같이 10명의 다양한 연령대의 사람들에게 자극물로 선정된 AI목소리를 검증한 결과 친밀감에 있어서는 평균 6.6으로 중간 이상의 평가 받았고 신뢰도 점수는 평균 8.35로 높게 평가받았다. 주관식 의견으로 목소리에서 예상되는 성격은 친절하고 활발하며 신뢰감을 주는 목소리로, 지적이고 사교적인 이미지를 전달하는 것으로 평가되어 전문직 여성의 이미지와 매치 되었다.

설문 문항은 표 7과 같이 사회적 실제감의 차이를 인지적 측면과 정서적 측면으로 나누어서 차이를 살펴보았다. 인지적 사회적 실제감 차이를 파악하기 위해 지각의 측면에서 파악하기 위해 ‘같이 함께 있는 기분이 드는지’ 질문하였고 정서적 측면에서는 ‘친근감(intimacy) 정도’를 질문하였다. 신뢰도 관련 질문은 신뢰감에 대한 해석을 일관되게 하기 위해 ‘중고 거래 시 어느 정도 믿음이 가는지’ 질문하였다. 동일한 질문을 6가지 유형의 아바타 각각에 대해서 5점 척도로 응답하게 하여 실제감, 친밀감, 신뢰감에서 가장 높은 점수를 받은 아바타를 밝혀보고자 한다. 또한 친목을 위해 가상공간에서 대화를 할 때 피험자와의 관계에 따라 선호하는 아바타에 대한 차이를 밝혀보고자 한다.

4-3 설문 결과

설문에 참여한 인원수는 표 8과 같이 총121명으로 여성 67.5%, 남성32.5%로 참여하였고 아바타 참여자와 아는 지인은 총 82명 모르는 사람은 38명이 참여하였다. 관계에 있어서는 표 9와 같이 친밀한, 수평적, 수직적, 무관계로 4가지 유형의 관계로 설정하였다.

표 6. 선정된 AI 아바타 목소리 평가

Table 6. Evaluation of the selected AI avatar voice

	Participants	Imagined profession	Imagined characteristic	intimacy	trust
1	15-year-old girl (Kim)	Weathercaster	Hopeful personality	3.5	8
2	18-year-old boy (Kim)	Corporate Intern	Greedy and big dreaming feeling	10	10
3	Woman in her 20s (Park)	Banker	Seems like a kind and active person. Not a light feeling. Seems to laugh a lot and feels like someone good to get along with at work or in business.	8	8
4	Woman in her 20s (Jung)	Leader Type	Voice sounds casual yet bold	8	9
5	Woman in her 20s (Yu)	Office Worker or Finance Professional	Meticulous and diligent personality, trustworthy image	6	8
6	Woman in her 40s (Kim)	Host	Quite smart and very sociable. Voice seems intellectual and lively.	7	7.5
7	Woman in her 40s (Kang)	Announcer	Feels like a strong effort to sound trustworthy, like an announcer	1.5	6
8	Man in his 40s (Park)	Announcer/ Professional	Decisive yet soft but strong personality	8	10
9	Woman in her 50s (Kim)	Female Pastor	Ambitious and assertive	8	10
10	Man in his 50s (Kang)	Announcer Trainee	More of a feeling of a new announcer trainee than an established announcer, trying to be upright	6	7
	Total			6.6	8.35

표 7. 설문항목

Table 7. Survey Questions

Criteria	Questions
Social presence	It feels like being together
Intimacy	It feels friendly
Trust	It feels trustworthy in used-item transactions

표 8. 성별과 나이에 따른 참여자수

Table 8. Number of participants by gender and age

	10s	20s	30s	40s-50s	over 60s	total
Female (67.8%)	10	27	2	39	4	82
Male (32.2%)	3	13	0	22	1	39
	13	40	2	61	5	121

표 9. 관계에 따른 참여자 수

Table 9. Number of participants by relationship

Intimate relationship		Horizontal relationship		Vertical relationship	No relationship
family	relatives	friends	acquaintance	teacher-student	stranger
7	2	15	15	44	38

A) 아바타 유형별 차이

모든 아바타는 인공지능 프로그램을 통하여 동일한 방식을 통해 영상으로 제작되었으며, 기술적 한계로 인하여 다소 자

연스럽지 않은 부분이 존재하였다. 그 결과, 부정적인 반응이 긍정적인 반응보다 많았다. 6가지 아바타 유형의 강한 부정부터 강한 긍정까지 반응 정도를 비교하기 위해 1, 3, 5, 7, 9의 가중치를 적용하여 아래 그림 3과 같은 결과를 도출하였다. 본 연구에서는 편의상 아바타를 다음과 같이 명명하였다: 사실적 이미지 + 실제 목소리는 'PR', 사실적 이미지 + 인공지능 목소리는 'PA', 만화 이미지 + 실제 목소리는 'CR', 만화 이미지 + 인공지능 목소리는 'CA', 동물 이미지 + 실제 목소리는 'AR', 동물 이미지 + 인공지능 목소리는 'AA'. 실제감, 친근감, 신뢰감의 세 가지 측면에서 아바타 유형별 나타나는 차이를 카이제곱검정을 진행한 결과 표 10과 같이 P value = 0.03으로 통계적으로 유의하게 나타났다.

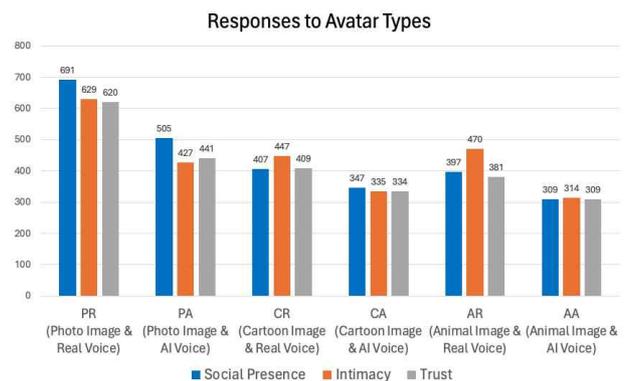


그림 3. 아바타 유형에 따른 응답

Fig. 3. Responses based on avatar types

표 10. 실재감, 친근감, 신뢰감 차이에 대한 통계 검증
Table 10. Statistical analysis of variations in social presence, intimacy, and trust

Chi-square statistic	19.8776
Degrees of freedom	10
Significance level	0.05
Critical value	3.9403
P-value	0.0304319

세 가지 영역에서 모두 높은 긍정적 반응을 보인 유일한 아바타는 이미지와 목소리가 실제와 가장 유사한 'PR'이었다. 그다음으로, 사실적인 이미지를 사용하고 인공지능 목소리를 적용한 'PA'는 실재감과 신뢰감에서 두 번째로 높은 긍정적 반응을 보였지만, 친근감에서는 'CR'과 'AR'보다 낮은 결과를 나타냈다. 이는 목소리의 실제성이 이미지의 실제성보다 친근감에 더 큰 영향을 미친다는 것을 시사한다.

한편, 만화 이미지의 'CR'과 동물 이미지의 'AR'는 모두 실재감과 신뢰감보다 친근감이 더 높게 나타났다. 이는 아바타 이미지의 사실성보다 이미지 자체가 주는 인상이 친근감에 더 큰 영향을 미친다고 유추할 수 있다.

1) 정리해 보면 목소리는 모든 항목에서 공통적으로 중요하다. 표 11의 주관식 답변을 분석해 보았을 때에도 많은 응답자들이 목소리의 자연스러움과 진정성을 중요하게 평가했으며, 목소리가 아바타 경험의 핵심 요소임을 강조하였다.

2) 실재감과 신뢰감에서는 이미지와 목소리에서 모두 사실성이 높을수록 긍정적인 반면 친밀감에서 이미지는 자체에서 느껴지는 인상에 영향을 받는 것으로 보인다. 이는 만화적 표현이나 동물 이미지가 쉽게 인식되고 기억될 수 있으며[26], 현실과의 거리를 두어 긍정적이고 편안한 감정을 유발하기 때문에 해석될 수 있다[27]. 특히, 동물 이미지는 본능적으로 애정 반응을 유발하여 강한 정서적 반응을 끌어내기 때문에 상대적으로 부정 반응이 낮게 나타난 것으로 보인다[28].

3) 유튜브들이 사용하는 만화이미지나 동물 이미지에 이미 익숙한 사용자들은 이러한 이미지들에 대해 낮은 부정적 반응이 나타났다. 이는 반복적인 노출이 친밀감을 증가시키기 때문으로 볼 수 있다[29].

4) 기술적 한계에서 나타나는 부자연스러움은 상대 아바타에 대한 신뢰감에 영향을 미치는 것으로 보인다.

5) 딥페이크, AI 기술로 인해 가상 공간에서 개인정보 보호와 진정성에 대한 우려가 크게 나타났다. 사실적 이미지의 아바타의 경우 응답자들은 기술의 발전에도 불구하고 조금의 어색함에도 여전히 인위적인 느낌을 강하게 인식했다.

B) 사용자 관계별 차이

친목을 위해 가상공간에서 대화를 할 때 가장 선호하는 아바타 유형은 다음 그림 4와 같이 PR>CR>AR 순서로 나타났다. 이와 같이 이미지와 상관없이 실제 목소리를 가진 아바타

표 11. 주관식 답변
Table 11. Open-ended response

	Positive Open-Ended Response	Negative Open-Ended Response
Photo-realistic image	<p>1. Trust and Realism</p> <p>"Trust is formed through realistic images and actual voices."</p> <p>"It feels good because it feels real."</p> <p>"There is no aversion, and it feels natural."</p> <p>Familiarity</p> <p>"When you already know a person named ooo, the most familiarity is felt when meeting an avatar that uses their actual voice and image."</p>	<p>1. Technical Unnaturalness</p> <p>"I feel uneasy because the mouth movements and voice are not in sync."</p> <p>"The expressions are awkward, so I can tell it's an avatar."</p> <p>"Real images provide information but are not perfect, which can cause cognitive dissonance."</p> <p>Psychological Aversion</p> <p>"I am hesitant to use it due to concerns about crimes like deepfakes."</p> <p>"The first impression might be mistaken for the actual person, but as time passes, a sense of incongruity grows."</p>
Cartoon image	<p>Familiarity and Comfort</p> <p>"Cartoon images are thought to be less burdensome."</p> <p>"The combination of a cartoon image resembling a real face and an actual voice gives a feeling like making a phone call."</p> <p>"It feels less aversive as I am used to encountering virtual YouTubers."</p> <p>Anonymity and Psychological Comfort</p> <p>"Showing the real appearance feels burdensome, but real voices can create trust and emotions."</p> <p>"It feels familiar, similar to the face filters that were popular on social media."</p>	<p>Decreased Credibility</p> <p>"The trust decreases due to the cartoon image."</p> <p>"I can't trust it because it's not a real face."</p> <p>"The voice is real, but the cartoon image does not build trust."</p> <p>Awkwardness and Unnaturalness</p> <p>"The sync between the mouth shape and voice is not natural, which feels awkward."</p> <p>"The cartoon image lacks a sense of reality."</p> <p>"The caricatured image seems to lack authenticity."</p>
Animal Image	<p>1. Familiarity and Comfort</p> <p>"Animals give a familiar feeling. They are more comfortable than people."</p> <p>"Many YouTubers use virtual avatars with animal characters, so they feel familiar."</p> <p>"Cute characters feel friendly and are nice because they are less burdensome."</p> <p>Anonymity and Psychological Safety</p> <p>"I do not want to show my face much in the virtual space."</p> <p>"While feeling protected in terms of personal information, I can also feel trust."</p> <p>"I want to maintain some distance from other people."</p>	<p>Awkwardness and Alienation</p> <p>"The combination of an animal image with a real human voice is very awkward."</p> <p>"I don't like the animal image; it lacks realism."</p> <p>"I couldn't trust it because it was an animal image."</p> <p>Psychological Aversion</p> <p>"It feels like wearing an animal mask."</p> <p>"The image of a human wearing a mask feels somehow unfamiliar."</p>

를 선호하는 것은 이미지 보다 실제목소리가 중요하다는 것을 시사한다.

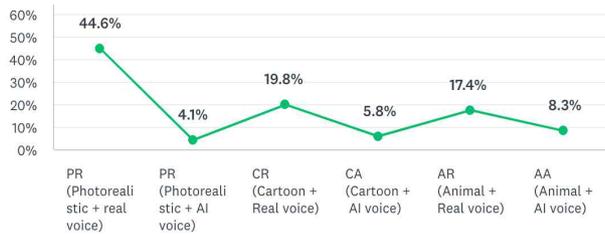


그림 4. 선호하는 아바타 이미지/목소리 조합
Fig. 4. Preferred avatar expression

관계에 따라 친밀한 관계인 가족과 친척, 수평적 관계인 친구와 지인(친구, 학부형, 직장동료, 선후배), 수직적 관계인 사제지간, 모르는 사람 그룹으로 구분하여 선호하는 아바타 유형에 대한 차이를 분석하였다. 카이제곱검정을 해본 결과 P value = 0.0004로 그룹간의 차이가 있음을 확인하였다.

표 12. 관계에 따른 아바타 선호도 차이 통계 검증
Table 12. Statistical analysis of differences in avatar preferences by relationship

Chi-square statistic	40.1221
Degrees of freedom	15
Significance level	0.05
Critical value	7.2609
P-value	0.000434647

가장 관계가 가까운 가족과 친척들 총 9명의 응답을 분석해본 결과 아래 그림 5와 같이 PR에 대한 선호도가 44.4%로 높게 나타났다. 다음으로 PA, CR이 둘 다 22.2%로 동일하게 나타났다. 이는 이미지와 목소리 모두 실제에 가까울수록 선호도가 높은 것으로 보인다. 친밀한 관계에서는 오랜 시간을 지켜봐 왔기 때문에 선행연구에서 밝힌 바와 같이 친숙한 이미지에 대해서는 작은 차이에도 어색함을 크게 느끼기 때문에 사실적 이미지가 선호되는 것으로 판단된다.

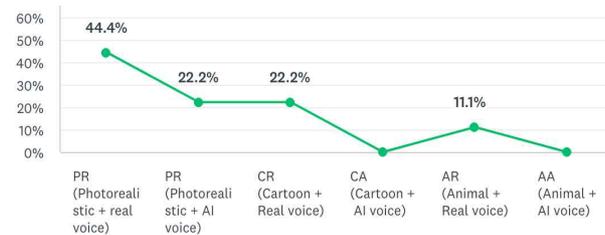


그림 5. 가족과 친척들이 선호하는 아바타 이미지/목소리 조합
Fig. 5. Avatar expression preferred by family and relatives

다음으로 수평적 관계인 친구와 지인을 그룹핑 하면 총 30명으로 4-50대가 주를 이루었는데 이 경우에도 그림 6에서 보듯이 PR에 대한 선호도가 56.7% 가장 높았고 다음으로

CR(16.7%), CA(13.3%)순으로 높게 나타났다. 주목할 만한 것은 만화이미지에 대한 선호도도 모두 합치면 30%(9명)로 다른 그룹에 비해 높게 나타났다. 주관식 답변을 통해 만화 이미지가 가장 덜 부담스럽고, 가상의 공간과 잘 어울리며, 친근하게 느껴지기 때문임을 파악하였다. 또한 FGI에서 밝힌 친한 사이에서 재밌는 이미지를 선호하는 것과 같은 맥락으로 볼 수 있다.

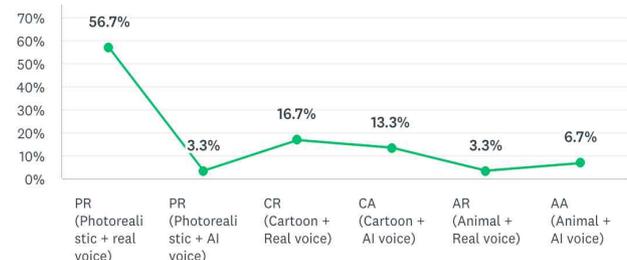


그림 6. 친구와 지인이 선호하는 아바타 이미지/목소리 조합
Fig. 6. Avatar image/voice combinations preferred by friends and acquaintances

반면에 모르는 사람 그룹은 38명으로 PR>CR>AR순으로 실제 목소리를 가진 아바타가 실제 이미지에 가까운 순서대로 선택되었다. 따라서 처음 보는 관계에서는 실제 목소리가 신뢰 구축을 위해 가장 중요한 것으로 보인다. 실제 목소리는 사용자의 정체성과 진정성을 나타내고 상대방의 의도와 감정을 더욱 명확히 파악할 수 있어 사용자 간의 연결감을 강화하는 것으로 보인다. 또한 가상공간에서 아바타 이미지의 인위적인 측면을 상쇄시켜 자연스럽게 인간적인 상호작용을 가능하게 한다고 볼 수 있다.

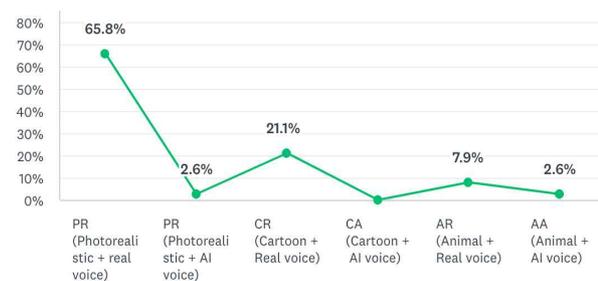


그림 7. 모르는 관계에서 선호하는 아바타 이미지/목소리 조합
Fig. 7. Preferred avatar expression by strangers

반면에 사실적 이미지를 가장 선호하지 않는 그룹은 수직적 관계인 사제지간으로 나타났다. 이들은 총 44명으로 AR(36.36%)을 가장 선호했으며 다음으로 CR(20.5%) > PR(18.2%)순서로 실제와 거리가 먼 이미지 순으로 선택되었다. 이는 위계가 있는 관계에서는 비현실적인 이미지가 대도시 편안함을 줄 수 있다는 것을 유추해 볼 수 있다. 주관식 의견으로 동물 이미지가 주는 친근한 이미지와 프라이버시를 보호 받을 수 있다는 측면을 언급하였다.

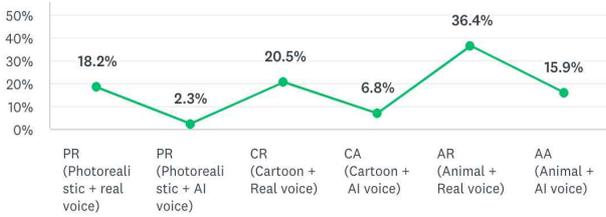


그림 8. 사제시간 사용자들이 가장 선호하는 아바타 표현
 Fig. 8. Preferred avatar expression by student users for teachers

앞서 사제시간으로 응답한 대상이 주로 20대이기 때문에 이러한 결과가 나이에 따른 특징인지 여부를 파악하기 위해 20대 모르는 사람 그룹의 응답만 살펴본 결과 7명의 응답자가 있었고 아래 그래프와 같이 사제시간인 경우의 그래프와는 확연히 다르게 사실적 이미지를 가장 선호하는 것으로 나타났다. 따라서 관계에 따라 가상공간에서 선호하는 아바타 이미지가 다른 것을 알 수 있었다.

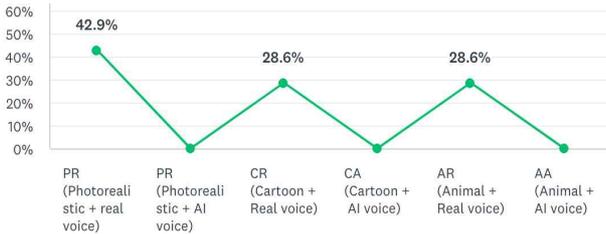


그림 9. 20대 '모르는 사람'이 선호하는 아바타 표현
 Fig. 9. Preferred avatar expression of strangers

4-4 연구 결과 적용 방향

위의 설문 분석을 통해 전반적으로 실제 목소리는 실제 이미지 보다 더 중요한 것으로 파악이 되었다. 친밀한 관계에서는 실제 이미지가 중요하고, 처음 사람을 만나는 경우에는 실제 목소리로 대면하는 것이 관계 형성에 매우 중요하다는 것을 파악하였다. 지인의 경우에는 다른 그룹에 비해 만화 이미

표 13. 관계에 따른 아바타 유형 선호도 및 인사이트

Table 13. Preferences and insights for avatar types based on relationships.

Relationship	Preferred avatar order	Insights
Intimate relationship (family and relatives)	PR>PA>CR	Real image important
Horizontal relationship (friends & acquaintance)	PR>CR>CA	Moderately concealing cartoon image
Vertical relationship (teacher & students)	AR>CR>PR	Friendly unrealistic image
No relationship (strangers)	PR>CR>AR	Real voice important

지에 대한 선호도가 높았다. 위계가 있는 관계에서는 사실적인 이미지 보다 편안하고 친근감이 느껴지는 비현실적 이미지 선호도가 높았다. 표 13에 이러한 4가지 관계에 따른 아바타 선호도 및 인사이트를 정리하였다.

이러한 결과를 반영하여 기업이나 공공기관에서 진행되는 화상회의, 가상공간에서의 거래, 또는 제품 홍보와 같이 아바타 사용자의 신뢰도가 중요한 가상공간 서비스에서는 최대한 사실성이 높은 아바타와 사용자의 실제 목소리로 참여할 수 있는 환경을 제공하는 것이 중요할 것으로 보인다.

반면, 친목을 목적으로 하는 가상공간에서 지인들과 소통하는 경우에는 사실적인 표현과 더불어 친근하고 단순화된 아바타 이미지를 제공하는 것도 중요하다고 판단된다. 더불어, 재미 요소를 강화하기 위하여 목소리 변조 옵션을 제공하는 것도 고려할 수 있다. 이는 편안하고 유쾌한 분위기에서 소통하는 데 기여할 수 있을 것이다.

또한, 수직적 관계에서 창의적인 아이디어를 소통해야 하는 상황에서는 상대적으로 낮은 위치에 있는 사용자들이 보다 적극적으로 참여할 수 있도록 친근한 동물 이미지의 아바타를 활용하는 방안을 제안할 수 있다. 이러한 디자인은 사용자들의 심리적 부담을 완화하고 자유롭게 편안하게 의견을 제시하는데 도움이 될 수 있을 것이다.

아울러, 다양한 관계의 사람들이 혼재하는 가상공간 서비스에서는 상대와의 관계나 대화 목적에 따라 사용자가 자신의 아바타를 신속하게 변경할 수 있는 유연한 인터페이스를 제공할 필요성이 제기된다. 이를 통해 업무 목적에는 전문적인 아바타를, 친목 목적에는 캐주얼한 아바타를 선택할 수 있는 환경을 조성하는 것이 바람직할 것이다.

V. 결 론

본 연구는 가상공간에서 아바타 이미지와 목소리 조합이 사용자 반응에 미치는 영향을 탐구하였다. 이미지와 목소리의 사실성에 따른 중요도 차이와 사용자와의 관계에 따라 선호되는 아바타 유형을 파악하였다.

문헌 연구를 통해 아바타 의인화 정도에 따른 인식 차이와 목소리의 역할을 고찰하고, 이를 기반으로 아바타 자극물 제작 기준을 설정하였다. 이후 포커스 그룹 인터뷰(FGI)를 통해 가상공간에서 상대 아바타에 대한 경험을 분석하고, 목소리의 중요성과 관계에 따라 선호 이미지가 달라진다는 점을 확인하였다. 이를 검증하기 위해 실험 참가자를 설정하고, 실제 이미지, 만화 이미지, 동물 이미지에 실제 목소리와 AI 목소리를 더한 6가지 동영상 자극물을 AI 툴을 이용해 제작하였다. 이후, 피험자와 친밀한 관계에 있는 사람들과 전혀 모르는 사람들을 대상으로 설문 조사를 실시하였다.

설문 결과, 가상공간에서 친목을 목적으로 대화를 할 때, 실제 목소리를 가진 3가지 유형의 아바타가 AI 목소리를 가

진 실제 이미지 아바타보다 더 선호되는 경향이 나타났다. 관계에 있어서 아바타 선호도 차이가 통계적으로 유의하다는 것을 파악하였다. 가족과 같은 친밀한 관계에서는 사실적 이미지 아바타에 대한 선호도가 높았으며, 무관계 응답자들은 실제 목소리를 실제 이미지 보다 더 중요시 하는 경향을 보였다. 수평적 관계에서는 만화 이미지에 대한 선호도가 다른 관계에 비해 상대적으로 높게 나타났다. 이는 이미 얼굴을 알고 있는 관계에서는 사실적인 얼굴이 부담스럽게 느껴질 수 있음을 시사한다. 한편, 사제시간과 같이 수직적 관계에서는 사실적인 모습보다는 친근하고 편안한 비현실적 이미지가 더 선호되는 경향성이 확인되었다.

본 연구의 한계는 한 명의 피험자를 기준으로 실험을 진행한 점과 자극물의 완성도가 결과에 영향을 주었을 가능성이 있다는 점이다. 따라서 후속 연구에서는 자극물의 완성도를 높여 기술적 한계로 인한 오류를 최소화하는 것이 중요할 것이다.

본 연구는 가상공간에서 친목을 목적으로 대화를 할 때 관계에 따라 적절한 아바타를 설정하는 데 유용한 설계 기준을 제공한다는 점에서 의의가 있다. 이는 아바타 디자인 및 아바타 서비스 설계 시 참고할 수 있는 실질적 가이드라인이 될 것이다.

감사의 글

이 논문은 2023학년도 홍익대학교 학술연구진흥비에 의하여 지원되었음.

참고문헌

[1] Market.us. Digital Avatar Market [Internet]. Available: <https://market.us/report/digital-avatar-market/>.

[2] TechM. Metaverse and Avatar Technology [Internet]. Available: <https://www.techm.kr/news/articleView.html?idxno=97587>.

[3] Technology Magazine. Ready Player Me: Creating Avatars for the Metaverse [Internet]. Available: <https://technologymagazine.com/digital-transformation/read-y-player-me-creating-avatars-for-the-metaverse>.

[4] Genies. About Genies [Internet]. Available: <https://genies.com/ko/>.

[5] ObEN. ObEN: Personal Artificial Intelligence [Internet]. Available: <https://creati.ai/ai-tools/oben-me/>.

[6] H. Shin and H. Kim, "Analysis of Audience Preference for Realistic or Non-Realistic Expression - Focus on Rendering Technique in 3D Animation -," *Journal of Digital Design*, Vol. 13, No. 4, pp. 611-620, 2013.

[7] M. Shin, S. J. Kim, and F. Biocca, "The Uncanny Valley: No

Need for Any Further Judgments when an Avatar Looks Eerie," *Computers in Human Behavior*, Vol. 94, pp. 100-109, May 2019. <https://doi.org/10.1016/j.chb.2019.01.016>

[8] B. M. Hwang and E. J. Han, "A Study of Images that are Reminiscent of a Voice," *Journal of Speech-Language & Hearing Disorders*, Vol. 24, No. 4, pp. 249-257, December 2015. <https://doi.org/10.15724/jslhd.2015.24.4.023>

[9] J. Short, E. Williams, and B. Christie, *The Social Psychology of Telecommunications*, London, UK: John Wiley & Sons, 1976.

[10] Y. Lee, S. Hwang, and I. Kim, "The Effect of Appearance Design Stage on Social Presence When Interacting with Digital Humans in VR," *Journal of Digital Contents Society*, Vol. 21, No. 6, pp. 1113-1122, June 2020. <https://doi.org/10.9728/dcs.2020.21.6.1113>

[11] K. I. Park and C. H. Cho, "Developing the Scale of Brand Social Presence: Focusing on Facebook," *The Korean Journal of Advertising*, Vol. 26, No. 5, pp. 213-241, July 2015. <https://doi.org/10.14377/KJA.2015.7.15.213>

[12] F. Biocca, C. Harms, and J. K. Burgoon, "Toward a More Robust Theory and Measure of Social Presence: Review and Suggested Criteria," *Presence: Teleoperators and Virtual Environments*, Vol. 12, No. 5, pp. 456-480, October 2003. <https://doi.org/10.1162/10547460322761270>

[13] J.-S. Kim, Social Influence of Opponent Avatar's Realism and Height in Immersive Virtual Environment, Master's Thesis, Seoul National University, Seoul, August 2019.

[14] N. Jung, M. Lee, and H. Choi, "The Uncanny Valley Effect for Celebrity Faces and Celebrity-Based Avatars," *Science of Emotion & Sensibility*, Vol. 25, No. 1, pp. 91-102, April 2022. <https://doi.org/10.14695/KJSOS.2022.25.1.91>

[15] M. Kang, "A Study on Chatbot Profile Images Depending on the Purpose of Use," *Journal of the Korea Contents Association*, Vol. 18, No. 12, pp. 118-129, December 2018. <https://doi.org/10.5392/JKCA.2018.18.12.118>

[16] S. Han, S. Kim, J. Kim, and C. Kwon, "A Preliminary Study on Correlation between Voice Characteristics and Speech Features," *Phonetics and Speech Sciences*, Vol. 3, No. 4, pp. 85-91, December 2011.

[17] D. Kao, S. T. Mubarrat, A. Joshi, S. Pandita, C. Mousas, H.-N. Liang, and R. Ratan, "Exploring How Gender-Anonymous Voice Avatars Influence Women's Performance in Online Computing Group Work," *International Journal of Human-Computer Studies*, Vol. 181, 103146, January 2024. <https://doi.org/10.1016/j.ijhcs.2023.103146>

[18] K. C. Povinelli and Y. Zhao, "Springboard, Roadblock or

- “Crutch”?: How Transgender Users Leverage Voice Changers for Gender Presentation in Social Virtual Reality,” in *Proceedings of 2024 IEEE Conference Virtual Reality and 3D User Interfaces (VR)*, Orlando: FL, pp. 354-364, March 2024. <https://doi.org/10.1109/VR58804.2024.00057>
- [19] A. McStay, *Emotional AI: The Rise of Empathic Media*, San Francisco, CA: Sage, 2018. <https://doi.org/10.4135/9781526451293>
- [20] I. Torre, J. Goslin, and L. White, “If Your Device Could Smile: People Trust Happy-Sounding Artificial Agents More,” *Computers in Human Behavior*, Vol. 105, 106215, April 2020. <https://doi.org/10.1016/j.chb.2019.106215>
- [21] C. Edwards, A. Edwards, B. Stoll, X. Lin, and N. Massey, “Evaluations of an Artificial Intelligence Instructor’s Voice: Social Identity Theory in Human-Robot Interactions,” *Computers in Human Behavior*, Vol. 90, pp. 357-362, January 2019. <https://doi.org/10.1016/j.chb.2018.08.027>
- [22] S. A. Collins and C. Missing, “Vocal and Visual Attractiveness Are Related in Women,” *Animal Behaviour*, Vol. 65, No. 5, pp. 997-1004, May 2003. <https://doi.org/10.1006/anbe.2003.2123>
- [23] S. A. Collins, “Men’s Voices and Women’s Choices,” *Animal Behaviour*, Vol. 60, No. 6, pp. 773-780, December 2000. <https://doi.org/10.1006/anbe.2000.1523>
- [24] S. M. Smith and D. R. Shaffer, “Celerity and Cajolery: Rapid Speech May Promote or Inhibit Persuasion through Its Impact on Message Elaboration,” *Personality and Social Psychology Bulletin*, Vol. 17, No. 6, pp. 663-669, December 1991. <https://doi.org/10.1177/0146167291176009>
- [25] C. E. Kimble and S. D. Seidel, “Vocal Signs of Confidence,” *Journal of Nonverbal Behavior*, Vol. 15, No. 2, pp. 99-105, June 1991. <https://doi.org/10.1007/BF00998265>
- [26] H. Chen, R. Russell, K. Nakayama, and M. Livingstone, “Crossing the ‘Uncanny Valley’: Adaptation to Cartoon Faces Can Influence Perception of Human Faces,” *Perception*, Vol. 39, No. 3, pp. 378-386, March 2010. <https://doi.org/10.1068/p6492>
- [27] A. W. de Borst and B. de Gelder, “Is It the Real Deal? Perception of Virtual Characters Versus Humans: An Affective Cognitive Neuroscience Perspective,” *Frontiers in Psychology*, Vol. 6, 576, May 2015. <https://doi.org/10.3389/fpsyg.2015.00576>
- [28] S. Hayama, L. Chang, K. Gumus, G. R. King, and T. Ernst, “Neural Correlates for Perception of Companion Animal Photographs,” *Neuropsychologia*, Vol. 85, pp. 278-286, May 2016. <https://doi.org/10.1016/j.neuropsychologia.2016.03.018>
- [29] P.-L. Yang and D. M. Beck, “Familiarity Influences Visual Detection in a Task That Does Not Require Explicit Recognition,” *Attention, Perception, & Psychophysics*, Vol. 85, No. 4, pp. 1127-1149, May 2023. <https://doi.org/10.3758/s13414-023-02703-7>

강민정 (Minjeong Kang)

2000년 : 이화여자대학교 서양화/정보 디자인 학사

2006년 : New York University, Interactive Telecommunication 석사(MPS)

2014년 : 서울대학교 디자인학부 박사 (디자인박사)



2006년~2010년: R/GA New York

2015년~현 재: 홍익대학교 조형대학 부교수

※관심분야 : UI/UX디자인, 사용자 연구, 인공지능