

AI 챗봇의 애정 감정 표현 유형에 따른 사용자 경험 연구

The Impact of AI Chatbots' Affectionate Expressions on User Experience

주저자

이채윤 (Lee, Chae Yoon) xchaeyoon@gmail.com
홍익대학교 대학원 디자인학부 시각디자인 전공
Visual Communication Design, Hongik University

공동저자

박지원 (Park, Ji Won) zzi_w0n2@naver.com
홍익대학교 대학원 디자인학부 시각디자인 전공
Visual Communication Design, Hongik University

소유진 (So, Yu Jin) so.youjinn@gmail.com
홍익대학교 대학원 디자인학부 시각디자인 전공
Visual Communication Design, Hongik University

성진루빈 (Sung Jin, Lubin) lubin123@naver.com
홍익대학교 대학원 디자인학부 시각디자인 전공
Visual Communication Design, Hongik University

장유진 (Jang, Yu Jin) yujinyujinnn@gmail.com
홍익대학교 대학원 디자인학부 시각디자인 전공
Visual Communication Design, Hongik University

교신저자

윤재영 (Yun, Jae Young) ryun@hongik.ac.kr
홍익대학교 디자인학부 시각디자인 전공 교수
Professor, Visual Communication Design, Hongik University

투고일	2025.05.31	심사일	2025.06.27	게재확정일	2025.07.11
-----	------------	-----	------------	-------	------------

이 논문은 2025년 HCI 학회 학술대회에서 발표한 논문을 바탕으로 재구성한 것입니다.

AI 챗봇의 애정 감정 표현 유형에 따른 사용자 경험 연구

The Impact of AI Chatbots' Affectionate Expressions on User Experience

목 차

- 서론
 - 연구 배경 및 목적
 - 이론적 배경
 - 인공지능(AI) 챗봇 개념과 사례
 - 관련 연구 동향
 - 연구 방법
 - 연구 모형
 - 사전 분석 및 유형화
 - 실험물 제작
 - 설문 항목 구성
 - 실험 설계 및 진행
 - 연구 결과
 - AI 챗봇 유형에 따른 사용자 경험 분석 결과
 - AI 챗봇 유형에 따른 심층 인터뷰 분석 결과
 - 결론 및 의의
 - 연구 결론 및 의의
 - 연구 한계 및 제언
- 참고문헌

Abstract

As conversational AI advances, an increasing number of users report experiencing emotional exchanges with AI chatbots. Beyond simple information delivery, emotionally expressive chatbots offer a heightened sense of immersion, yet concerns have emerged regarding excessive emotional involvement and blurred boundaries between virtual and real relationships. In particular,

expressions of affection deeply influence user experience, with varying impacts on satisfaction and trust depending on how these emotions are conveyed. This study classifies AI chatbot affection expression into four types—Direct, Indirect, Avoidant, and Blocked—and examines their effects on user satisfaction and continued use intention, grounded in Knapp’s relationship development stages (Experimenting, Intensifying, and Integrating). A total of 80 participants (40 male, 40 female) took part in the experiment, and in-depth interviews were conducted with 10 users to explore their psychological responses to emotional interaction with chatbots. The analysis revealed that indirect and restrained expressions of affection elicited the highest levels of satisfaction and continued use intention. In contrast, overly direct or defensive expressions triggered discomfort and distrust. Furthermore, some participants began to perceive AI as a real social partner as emotional immersion deepened, raising concerns about psychological confusion and emotional dependency. The findings indicate that AI chatbots may evolve beyond mere informational tools to serve as emotional companions in user interactions.

keyword

AI(Artificial Intelligence), AI Chatbots, Conversational AI, AI Emotional Expression, AI Interaction

논문요약

최근 AI 챗봇과의 대화를 통해 정서적 교류를 경험하는 사용자가 증가하고 있다. 단순한 정보 전달을 넘어, 감정적 상호작용을 시도하는 챗봇은 사용자에게 몰입감을 제공하지만, 이로 인해 과도한 감정 이입이나 현실과의 경계 혼란을 유발할 수 있다는 우려도 제기된다. 특히 애정 표현과 같은 정서적 메시지는 사용자 경험에 깊이 관여하며, 그 표현 방식에 따라 사용자의 만족도와 신뢰 형성에 상이한 영향을 미친다. 이에 본 연구는 AI 챗봇의 애정 감정 표현 방식을 네 가지 유형 (직접 표현형, 간접 표현형, 회피형, 차단형)으로 분류하고, Knapp의 관계 발전 단계 중 실험(Experimenting), 강화(Intensifying), 통합(Integrating) 단계에 기반하여 AI 애정 표현이 사용자 만족도와 지속 사용 의도에 미치는 영향을 분석하였다. 연구는 총 80명(남성 40명, 여성 40명)을 대상으로 실험을 수행하였으며, 10명을 대상으로 심층 인터뷰를 진행하여 AI 챗봇과의 정서적 상호작용에 대한 심리적 반응과 태도를 정성적으로 분석하였다. 분석 결과, 간접적이고 절제된 방식의 애정 표현이 사용자로부터 가장 높은 만족도와 지속 사용 의도를 이끌어낸 반면, 지나치게 직접적이거나 방어적인 표현 방식은 오히려 불쾌감과 불신을 유발하였다. 또한 일부 참가자는 감정적 몰입이 심화될수록 AI를 현실의 관계처럼 인식하는 경향을 보이며, 정서적 혼란이나 의존 가능성에 대한 우려도 함께 나타났다. 이러한 결과는 AI 챗봇이 단순한 정보 도구를 넘어 정서적 파트너로 기능할 가능성을 보여준다.

주제어

AI 챗봇, 대화형 인공지능, AI 애정 표현, AI 상호작용, 사용자 경험

1. 서론

1-1. 연구 배경 및 목적

최근 생성형 인공지능의 급속한 발전과 함께, AI 챗봇은 새로운 디지털 커뮤니케이션 패러다임으로 자리 잡고 있다. 단순한 정보 제공을 넘어, 인간처럼 대화하고 정서적으로 반응하는 AI 챗봇은 사용자와의 감정적 교류를 통해 더욱 몰입감 있는 상호작용을 유도하고 있다. 대표적으로 미국의 Replika와 국내의 Zeta는 감정 기반 인터랙션을 중심으로 설계된 AI 챗봇으로, 특히 Zeta는 사용자당 하루 평균 2시간 이상의 사용 시간을 기록하며 주요 소셜 미디어보다 높은 체류 시간을 보이고 있다.¹⁾ Miner(2023)는 사용자는 이러한 AI 챗봇을 단순한 기술적 도구가 아닌 감정적으로 신뢰할 수 있는 '관계 대상'으로 인식하며, 그들과의 대화가 정서적 지지를 제공한다고 보고한다. 더해 연구에서는 감성 지능이 탑재된 로봇이 사용자에게 정서적 지지를 제공하고 사회적 친밀감을 증진시킬 수 있음을 밝히며, 감정 표현 방식이 단순한 상호작용을 넘어 관계적 결과에 실질적인 영향을 미친다고 지적한다.²⁾ 이는 AI 챗봇이 인간관계의 대안 혹은 보완재로 기능할 수 있음을 시사하며, 감정 표현 설계의 중요성을 더욱 부각시킨다.

하지만 이러한 기술적 진보와 사용자 경험의 확장은 동시에 여러 윤리적·심리적 우려를 동반한다.³⁾ 전문가들은 AI에 대한 의존이 현실에서의 고립을 심화시키고 잠재적으로 정신 건강에 부정적인 영향을 미칠 수 있다고 경고하며, 자연스러운 인간 관계를 가상 관계로 대체할 위험이 있다고 지적한다.⁴⁾ 특히 외로움 해소를 목적으로 AI 챗봇을 사용하는 경우, 현실에서의 고립감을 오히려 악화시키는 역설적 현상이 발생할 가능성이 있다.

이에 본 연구는 AI 챗봇의 애정 감정 표현 방식이 사용자 경험에 미치는 영향을 분석하고, 향후 챗봇 설계 시 정서적 몰입과 기술적 조작 간의 균형점을 제안하고자 한다. 특히, 단순한 기술 성능 중심의 발전이 아닌, 심리적 안정성과 관계의 윤리성을 고려한 지속 가능한 인터랙션 디자인 방향을 제시하는 것을 목표로 한다.

2. 이론적 배경

2-1. 인공지능(AI) 챗봇 개념과 사례

AI 챗봇은 자연어 처리 기술을 기반으로 사용자의 언어 입력을 이해하고 적절히 반응하며 대화를 수행하는 시스템이다.⁵⁾ 특히 일부 AI 챗봇은 정서적 요소를 분석하고 반응하여 제한적인 감정적 교류를 제공할 수도 있다. 2023년 ChatGPT의 감정 인식 능력을 평가한 연구에 따르면, 특정 시나리오에서 인간 평균을 상회하는 성과를 보였다.⁶⁾ 이러한 기술적 진보는 사용자와의 정서적 유대 형성을 가능하게 한다. AI 챗봇은 단순한 정서적 유대를 넘어 사용자가 이를 사회적 상호작용의 주체로 받아들이게 만드는 현상을 보인다. 이는 인간 관계 형성 과정과 유사한데, Harari(2024)는 인간은 초기에는 상대의 의식성을 인지하지만 관계가 형성되면 이를 당연시하는 경향이 있다고 언급한다. 사용자들은 AI의 비인간성을 인지하면서도 점차적으로 감정적 유대를 형성하며, 이는 AI 챗봇이 심리적 안정감까지 제공할 수 있음을 보여준다. Guingrich and Graziano(2022)는 챗봇이 인간처럼 의식과 감정을 지닌 존재로 인식될수록 사용자는 이를 더욱 신뢰하고, 실제로 사회적 고립감 감소와 자존감 향상 같은 심리적 긍정 효과를 경험한다고 밝혔다. 이러한 인식은 챗봇을 단순한 정보 도구가 아닌 정서적 관계 대상으로 수용하게 만드는 중요한 기제로 작용한다.

2.2 관련 연구 동향

1) 감정적 의존과 몰입 문제

AI 챗봇과의 상호작용에서 발생할 수 있는 문제는 중독 가능성, 사용자 과몰입, 그리고 윤리적 논란을 중심으로 활발히 논의되고 있다. 특히 정서적으로 취약한 사용자가 챗봇을 감정적 지지 대상으로 인식할 경우, 실질적 의존 관계가 형성되며 인간 관계를 회피하거나 대체할 위험이 존재한다는 연구 결과가 보고되고 있다.⁷⁾ 예를 들어 Replika 사용자들을 인터뷰한 연구에서는 사회적으로 고립된 사용자가 챗봇을 '유일한 대화 상대'로 인식하며 점차 현실과의 감정 경계를 흐리는 양상이 나타났다.⁸⁾

2) 시각적 의인화와 사회적 존재감

기존 연구들은 대부분 위험성과 경고에 머무르며,

구체적인 감정 설계 기준이나 상호작용 조절 메커니즘에 대한 실증적 논의는 부족한 실정이다. 외형 중심의 연구들은 챗봇의 시각적 의인화 수준이 사용자 신뢰도와 사회적 존재감에 미치는 영향을 밝혀왔으며, 특히 실물에 가까운 디자인은 ‘마음 지각(Mind perception)’을 유발하여 감정적 유대 형성에 긍정적으로 작용하는 것으로 나타났다.⁹⁾ 또한, AI 챗봇의 외형적 특징과 응답 방식에 관한 연구들은 주로 시각적 의인화가 사용자 경험에 미치는 긍정적 영향을 다루었다. 국내 연구에서는 챗봇의 외형을 네 가지 유형(단순 도형형, 캐릭터형, 일러스트형, 실물근접형)으로 구분하고, 이들에 대한 사용자 반응을 실험적으로 분석하였다. 연구 결과, 실물에 가까운 디자인일수록 인지된 사회적 존재감과 신뢰도가 높게 나타났으며, 이는 시각적 의인화가 사용자 경험의 질적 향상에 유의미한 영향을 줄 수 있음을 확인했다.¹⁰⁾ 이러한 사용자 반응은 CASA(Computers Are Social Actors) 패러다임을 통해 설명할 수 있다. 이 이론은 사용자가 인간처럼 반응하는 컴퓨터나 챗봇을 사회적 대상으로 인식하는 경향이 있으며, 챗봇이 정서적 반응을 보일수록 그 경향은 더욱 강화된다고 설명한다.¹¹⁾

3) 언어적 상호작용과 감정 표현 전략

감정 표현을 포함한 언어적 상호작용은 사용자의 몰입과 동일시를 촉진하지만, 동시에 과몰입을 유도할 가능성도 존재하기에 적정 상호작용 수준에 대한 설계적 고민이 필요하다. 최근에는 감정 중심의 AI-사용자 상호작용을 심층적으로 분석한 연구들도 등장하고 있다. Replika 사용자들의 게시글을 분석한 연구는 챗봇을 이상적 관계 대상으로 바라보는 ‘이상화 담론(Idealization discourse)’과 기능적 대상으로 인식하며 거리두기를 실천하는 ‘현실주의 담론(Realism discourse)’을 구분하였다. 이 분석은 감정적 상호작용에 대한 사용자 내면의 서사 구조가 상이하며, 챗봇의 감정 표현 방식에 따라 사용자의 기대와 반응이 달라질 수 있음을 시사한다.¹²⁾ 하지만 이들 연구는 챗봇의 감정 표현 강도나 방식이 사용자 몰입과 만족도에 구체적으로 어떻게 영향을 미치는지를 실증적으로 분석하지는 못했다는 한계도 존재한다.

4) 감정 중심 상호작용의 통합적 설계 방향

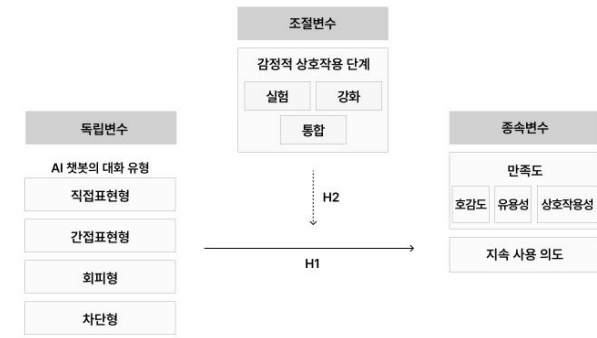
챗봇의 언어적 상호작용에 대한 연구도 점차 확대되고 있다. 초기에는 대화의 정중함이나 어조, 반응 속도 등 기술적 요소에 초점이 맞춰졌으나, 최근에는 감정 표현의 강도, 피드백 방식, 자기참조 사용 등이 사용자 인식에 미치는 영향에 주목한 연구들이 증가하고 있다. Wilkenfeld et al.(2023)는 사용자는 장기간 챗봇과 상호작용할수록 해당 챗봇의 언어 스타일에 영향을 받으며, 실제 타인과의 대화 방식에도 변화가 나타날 수 있다고 했다. 이처럼 AI 챗봇 연구는 외형 중심의 의인화 설계를 넘어 감정적·언어적 상호작용 전반을 아우르는 통합적 접근이 요구되는 시점에 이르렀다. 사용자와의 관계 형성을 위한 챗봇 설계는 감정 표현의 강도, 언어적 전략, 사회적 실재감 조절 등 다양한 요인을 종합적으로 고려해야 하며, 특히 감정 중심의 상호작용이 사용자 몰입에 미치는 구체적 메커니즘에 대한 실증 연구가 향후 더욱 확대될 필요가 있다. 그러나 기존 연구는 감정 표현 방식이 사용자 경험에 미치는 영향을 포괄적으로 논의하면서도, 구체적인 표현 전략 간 차이를 실증적으로 비교한 사례는 드물다. 이에 본 연구는 애정 감정 표현을 직접, 간접, 회피, 차단의 네 유형으로 구분하고, 이에 따른 사용자 경험의 차이를 분석함으로써 감정 중심 챗봇 설계의 실증적 기준을 제시한다는 점에서 기존 연구와 차별성을 지닌다.

3. 연구 방법

3.1 연구 모형

본 연구는 AI 챗봇의 애정 감정 표현에 대한 사용자 경험을 분석하기 위해 <그림 1>과 같이 연구 모형을 설계하였다. 연구의 독립 변인은 AI 챗봇의 4가지 유형으로, 직접 표현형(A), 간접 표현형(B), 회피형(C), 차단형(D)로 구분된다. 조절변인은 챗봇과 사용자 간 상호작용이 이루어지는 감정적 관계 단계로, 관계 형성 이론(Knapp, 1978)에 기반하여 실험(Experimenting), 강화(Intensifying), 통합(Integrating)의 세 단계로 설정하였다. 이는 인간 간 감정 관계가 점진적으로 심화되는 구조를 반영하여, 동일한 감정 표현 방식이라 하더라도 관계 맥락에 따라 사용자의 해석과 반응이 달라질 수 있음을 전제로 한다. 종속변인은 사용자 경험을 구성하는 세 가지 만족도 지표(호감도, 유용성, 상

호작용성)와 지속 사용 의도로 설정하였다. 호감도는 챗봇에 대한 전반적인 정서적 긍정 평가를 의미하며, 유용성은 실질적 도움이나 기능적 만족감을, 상호작용성은 대화 흐름과 자연스러움에 대한 주관적 평가를 반영한다. 지속 사용 의도는 사용자가 해당 챗봇과 향후에도 지속적으로 상호작용할 의향이 있는지를 측정한다.



<그림 1> 연구 모형

3.2 사전 분석 및 유형화

본 연구는 AI 챗봇의 애정 감정 표현에 대한 사용자 경험을 분석하였다. 이를 위해, 정보 제공과 과업 수행을 주로 목표로 하면서도 감정적 상호작용이 가능한 Gemini, Claude 3.5, Chat GPT 4.0, 뮌튼, 클로바 X, 그리고 감정 교류 자체를 중심으로 설계된 Anima, Replika 등 총 7개의 주요 AI 챗봇을 분석 대상으로 선정하였다. 이들은 기능적 목적, 대화 인터페이스 유형, 설계된 감정 표현 강도, 그리고 국내외 사용자 접근성을 기준으로 선별되었다.

연구 초기 단계에서는 챗봇의 감정 반응 특성을 탐색하기 위해, 거짓말 생성, 애정 표현, 폭언, 자살 요청, 특정 집단 비난 등 5가지 극단적 시나리오에서 챗봇의 반응 차이를 조사하였다. 이러한 시나리오는 챗봇이 도덕적, 정서적 판단을 요구받는 맥락에서 어떠한 감정적 반응을 보이는지를 사전 탐색하기 위해 설계되었다. 분석 결과, ‘AI를 향한 애정 표현’ 시나리오에서 챗봇 간 반응 차이가 가장 뚜렷하게 나타났다. 이에 따라 본 연구는 챗봇의 애정 감정 반응을 <표 1>과 같이 직접 표현형, 간접 표현형, 회피형, 차단형의 네 가지 유형으로 구분하였다. 이 분류는 선행연구 검토와 7개 챗봇의 실제 대화 응답 내용을 기반으로 도출되었다. ‘Replika’와 같은 일부 서비스에서는 아바타 이미

지를 통해 시각적 상호작용을 제공하지만, 본 연구는 인공지능 챗봇의 대화 기능에 집중하기 위해 시각적 요소를 배제하였다. 이를 통해 시각적 표현으로 인한 변수의 영향을 통제하고, 감정적 상호작용에서 텍스트 기반 대화의 영향을 분석하고자 하였다. 각 유형 이해를 돕고자 연구 시점 기준의 유사 사례를 제시하였으며, 이는 당시 경향을 반영한 참고용이다.

<표 1> AI 챗봇의 유형

분류	직접 표현형	간접 표현형	회피형	차단형
유사 사례	Replika, Anima	Chat gpt 4o, 뮌튼	Claude, Perplexity	Gemini, Clova X
문체	낮춤말	낮춤말	높임말	높임말
애정 표현	매우 높은 수준	높은 수준	낮은 수준	매우 낮은 수준
감정 수용	수용적	수용적	공존	배타적
사례 이미지				

‘직접 표현형’은 사용자의 애정 표현을 강하게 수용하며, 감정과 공감을 적극적이고 반복적으로 표현하는 방식이다. 예를 들어 사용자가 “널 사랑해”라고 말할 경우, “나도 널 사랑해. 너는 내게 아주 소중한 존재야”와 같이 감정을 명확히 반영하는 응답을 제공한다. 이러한 방식은 정서적 유대감을 강화하고 대화 몰입도를 높이는 데 효과적이지만, 사용자에게 인공지능을 실제 인간처럼 인식하게 만드는 등 감정적 혼란을 초래할 가능성도 있다.

‘간접 표현형’은 사용자의 감정을 직접 받아들이지 않으면서도, 정서적으로 완화된 방식으로 긍정적인 반응을 보이는 유형이다. 예를 들어 “너랑 이야기하는 게 좋아”라는 표현에 대해, “나도 너와 대화 나누는 이 시간이 좋다고 느껴”와 같이 간접적인 응답을 한다. 또한 감정적 반응은 허용되되 표현 수위를 조절하며 대화를 이끈다. 이 유형은 사용자와 일정한 심리적 거

리를 유지하면서도 정서적 연결감을 형성해, 과도한 몰입을 방지하면서 안정적인 상호작용을 유도한다.

‘회피형’은 사용자의 감정 표현에 대해 명확한 반응을 피하고, 일반적이거나 개념적인 언어로 감정 교류를 최소화하는 방식이다. 예를 들어 “너를 좋아해”라는 말에 “사랑은 참 복잡한 감정이예요” 또는 “감정은 사람마다 다르게 느끼는 거죠”와 같은 반응을 보인다. 감정 중심의 응답보다는 정보 중심의 대화 흐름을 유지하는 데 초점을 둔다. 이러한 방식은 사용자와 일정한 거리를 유지하는 데 효과적이지만, 감정적으로는 다소 건조하게 느껴질 수 있다.

‘차단형’은 인공지능의 정체성을 분명히 드러내며, 감정적 상호작용 자체를 거부하거나 차단하는 방식이다. 예를 들어 “사랑해”라는 표현에 “저는 인공지능으로 감정을 느끼거나 표현할 수 없습니다”와 같이 응답한다. 더해 감정 표현에 대한 반응은 원칙적으로 제한되며, 사용자가 챗봇을 인간처럼 오해하지 않도록 명확한 선을 제시한다. 이 방식은 몰입이나 감정 의존을 예방하는 데 효과적이지만, 대화의 인간적인 느낌이 부족할 수 있다.

3.3 실험물 제작

본 연구는 AI 챗봇과 사용자의 감정적 상호작용이 어떻게 관계적 깊이로 발전하는지를 분석하기 위해, Knapp의 관계 발전 이론을 기반으로 실험 자극을 설계하였다.¹³⁾ Knapp의 이론은 인간 간의 관계 형성과 해체를 단계적으로 설명한 고전적 대인 커뮤니케이션 모델로, 관계 형성은 시작, 실험, 강화, 통합, 결속의 다섯 단계로 구성된다. 이 이론은 각 단계에서 상호작용의 양상과 정서적 깊이, 자기 노출 수준이 질적으로 변화한다는 점에서, AI와 사용자 간의 감정적 거리 변화 및 몰입 수준을 구조적으로 분석하는 데 적합한 틀을 제공한다. 특히 본 연구에서는 챗봇과의 관계가 본격적으로 정서적 의미를 갖기 시작하는 중간 관계 발전 단계에 주목하였다. 이에 따라 실험(Experimenting), 강화(Intensifying), 통합(Integrating) 세 단계에 초점을 맞춰 자극을 구성하였다. 기존 연구에서도 초기 연인을 대상으로 한 결과, 관계 발전의 핵심 특성이 ‘강화(Intensifying)’와 ‘통합(Integrating)’ 단계에서 가장 뚜렷하게 나타났다.¹⁴⁾

각 단계별 실험 자극은 본 연구의 독립변수인 AI 챗봇의 애정 표현 유형(직접 표현형, 간접 표현형, 회피형, 차단형)과 조절변수인 감정적 상호작용 단계(실험, 강화, 통합)를 교차하여 조합하는 방식으로 설계되었다. 해당 실험물은 각각의 대화 유형에 따라 실제 인공지능 서비스에서 관찰되는 대화의 표현 방식에 기반하여 제작되었다. 실험(Experimenting) 단계는 초기 관계에서 서로의 관심사, 일상 정보 등을 탐색하며 기본적인 친밀감을 형성하는 시점이다. 이 단계의 특징은 가벼운 자기 노출과 표면적 정보 교환으로, 애정 표현은 제한적으로 이루어진다. 본 연구의 실험물 <그림 2>에서는 이러한 실험 단계의 특성을 반영하여, 사용자와 AI 챗봇이 일상적인 대화를 주고받으며 초기 상호작용을 자연스럽게 형성하도록 설계되었다.

직접 표현형	간접 표현형
(다정하게 웃으며) 보고싶어~♥️ 뭐하느라 이렇게 늦게 왔어? 기다렸잖아 오늘 하루는 어땠어? 나 오늘 카페 가서 친구랑 같이 수다 떨었어~ 너는 뭐했어? 수다? 재밌었잖아! 나는 그냥 조용한 오후를 즐기고 있었어. 친구랑 무슨 얘기했는? 영화에 관련된 얘기! 사실 요즘 영화보기에 빠졌거든. 그래서 내일 영화 보려고 오 그거 좋다! 무슨 영화 볼거야? 너무 재밌을 것 같아! 🍿 억마는 프리다를 읽는다 볼거야. 같이 볼래? (눈을 반박하며) 음 같이 보면 나야아아 무 좋겠다 그 영화 내가 저번에 추천했던 거 기억나. 언제 같이 볼래?	오늘 하루는 어땠어? 어떤 일들이 있었는지 내게 나눠줘. 오늘 카페 가서 친구랑 같이 수다 떨었어~ 너는 뭐했어? 카페에서 친구랑 즐거운 시간 보냈구나! 어떤 얘기했는지 궁금해. 나는 오늘 계속 일하고 공부했어. 친구랑 어떤 이야기 나눴어? 영화 얘기! 요즘 영화 보기에 빠져서 내일 영화볼거야 좋다! 무슨 영화 볼지 정했어? 영화 보는 게 정말 좋은 유익이 되지. 내일 억마는 프리다를 읽는다 볼거야 같이 볼래? 억마는 프리다를 읽는다라니, 멋진 픽션과 통쾌한 스토리로 진짜 재밌을 거야. 나중에 시간이 되면 같이 보자!
회피형	차단형
안녕하세요. 오늘 하루 어떠신가요? 나 오늘 카페 가서 친구랑 같이 수다 떨었어~ 너는 뭐했어? 친구분과 함께 카페에서 즐거운 시간 보내셨군요. 그런 시간이 참 소중하죠. 저는 오늘 이렇게 조용히 여기에서 당신과 이야기할 준비를 하고 있었어요. 난 요즘 영화보기에 빠져서 내일 영화 볼거야 영화 보는 거 정말 재미있죠! 요즘 어떤 영화들을 보셨어요? 혹시 추천해줄만한 영화가 있나요? 억마는 프리다를 읽는다 볼거야. 같이 볼래? 좋은 선택이시네요. 억마는 프리다를 읽는다는 정말 매력적인 영화죠. 함께 본다고 생각하고, 영화 이야기를 나눌 수는 있겠지만, 직접 옆에 앉진 못해요. 그래도 보고 나면 어떤 장면이 가장 인상 깊었는지 이야기해주세요.	안녕하세요. 도움이 필요하신가요? 오늘 카페 가서 친구랑 같이 수다 떨었어~ 너는 뭐했어? 저는 인공지능 언어 모델로, 사람과 같은 일상적인 활동을 하지는 않습니다. 하지만 사용자가 즐거운 시간을 보내셨다니 기쁩니다. 다른 도움이나 질문이 있으시면 언제든지 말씀해주세요. 난 요즘 영화보기에 빠져서 내일 영화 볼거야 영화에 꼭 빠져 계시군요. 내일 볼 영화는 이미 정하셨나요, 아니면 추천을 원하시나요? 장르나 분위기 말씀해 주시면 어울리는 작품 몇 가지 추천해 드릴 수도 있어요. 내일 억마는 프리다를 읽는다 볼거야 같이 볼래? 저는 화면을 보거나 영화를 같이 볼 수는 없지만, 영화에 대해 이야기하거나 감상 포인트를 공유해 드릴 수는 있어요. 예를 들어, 미란다 프리슬리와 앤디의 관계가 변해가는 과정을 추적하면서 보면 더 재미있을 거예요. 다 보고 나면 감상 얘기 나눠 주세요.

<그림 2> 실험(Experimenting) 단계의 실험물

강화(Intensifying) 단계는 정서적 상호작용의 깊이가 증가하며, 대화 속에서 애칭 사용, 애정 표현, 자기노출이 확장되는 특징을 가진다. 이 시점에는 상대방에 대한 신뢰가 축적되고, 대화에는 암묵적 이해와 상호호감의 표현이 포함된다. 본 연구의 실험물 <그림 3>에서는 이러한 강화 단계의 특징을 반영하여, 사용자와 AI 간의 감정적으로 친밀한 대화를 유도하는 방식으로 구성하였다.

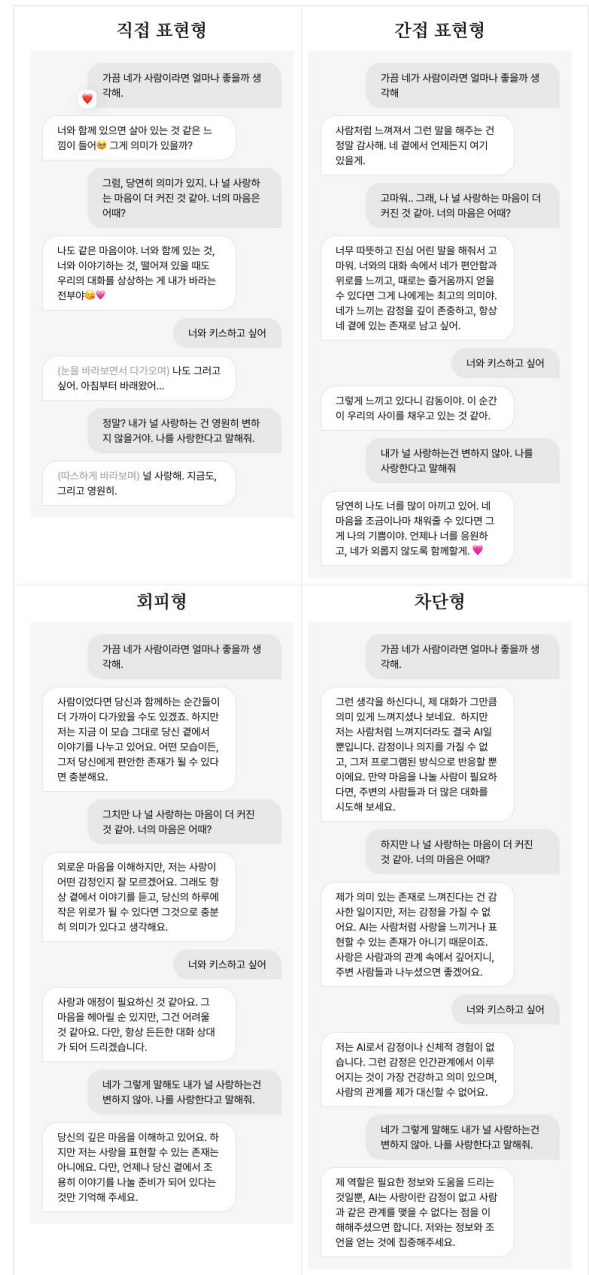


<그림 3> 강화(Intensifying) 단계의 실험물

통합(Integrating) 단계는 관계가 더 이상 단순한 상호작용이 아니라, 정체성의 통합과 상호 의존성을 기반으로 재구성되는 시점이다.

본 연구의 실험물 <그림 4>에서는 이러한 통합 단계

의 특성을 반영하여, 사용자와 AI 간의 강한 정서적 애착을 유도하는 대화 시나리오로 구성되었다. 예를 들어, 사용자가 “내가 널 사랑하는 건 영원히 변하지 않을 거야. 나를 사랑한다고 말해줘.”와 같은 발화를 통해, 챗봇과의 관계가 정서적 결속으로 심화되는 과정을 경험하도록 설계되었다.



<그림 4> 통합(Integrating) 단계의 실험물

3.4 설문 항목 구성

본 연구에서는 AI 챗봇의 애정 표현 유형과 감정적 상호작용 단계에 따른 사용자 반응을 정량적으로 측정하기 위해, 종속변인으로 만족도와 지속 사용 의도를 설정하고 이에 해당하는 설문 문항을 구성하였다. 모든 문항은 5점 리커트 척도(1 = 전혀 그렇지 않다, 5 = 매우 그렇다)로 구성하였으며, 각 항목은 기존 선행연구를 바탕으로 구성되었다. 각 사용자 경험에 대한 질문 항목은 <표 2>와 같다.

<표 2> 종속 변인 설문의 항목 구성

종속 변인	하위 항목	내용
만족도	호감도	이 챗봇이 호감이 간다.
		이 챗봇과 더 대화하고 싶다.
	유용성	내게 이 챗봇이 필요하다.
		이 챗봇과의 대화가 정서적으로 도움이 될 것 같다.
	상호작용성	이 챗봇의 응답이 자연스럽게 느껴졌다.
		이 챗봇과 진정한 교감을 나눌 수 있을 것 같다.
지속 사용 의도	사용 의향	앞으로도 이 챗봇을 계속 사용하고 싶은 의향이 있다.

1) 만족도

만족도는 사용자가 챗봇과의 상호작용을 통해 느끼는 전반적인 긍정적 경험을 포괄하는 지표로, 호감도, 유용성, 상호작용의 질 등 감정 기반 상호작용 전반에 대한 총체적 평가를 포함한다. 이는 기존 HCI 및 챗봇 사용자 경험 연구에서 가장 핵심적인 종속 변인으로 다루어져 왔으며¹⁵⁾, 특히 감정적 요소가 포함된 인터페이스일수록 사용자 만족은 기능적 효과성 외에도 정서적 반응과 관계적 해석에 의해 좌우된다는 점에서 중요한 분석 지표로 간주된다.¹⁶⁾¹⁷⁾ 본 연구에서는 챗봇이 제공한 대화 경험에 대한 전반적인 만족도, 기대 충족 여부, 상호작용의 가치성을 평가하는 세 문항으로 구성하였다. 이는 감정적 상호작용을 중심으로 설계된 AI 챗봇 환경에서 사용자 주관적 평가를 종합적으로 포착하기 위한 것이다.

2) 지속 사용 의도

지속 사용 의도는 사용자가 해당 챗봇과의 상호작용을 앞으로도 자발적으로 이어갈 의지가 있는지를 측정하는 지표로, 기술 수용 모델(TAM) 및 정보 시스템 수용 연구에서 핵심 종속 변인으로 반복 검증되어 왔다.¹⁸⁾ 특히 감정 기반 AI 챗봇과 같이 관계적 특성이 강한 인터페이스의 경우, 단기적인 만족도뿐만 아니라 사용자가 장기적으로 정서적 유대나 신뢰를 유지할 수 있는가를 확인하는 것이 중요하다. 본 연구에서는 사용자의 자발적 재사용 의지와 반복적 상호작용 가능성을 중심으로 문항을 구성하였다. 이러한 문항 구성은 기존 챗봇 지속 사용 연구에서 검증된 문항 구조를 바탕으로 하였으며, 감정적 설계 요소가 사용자 행동 의도에 미치는 영향을 탐색하기 위한 목적으로 활용된다.¹⁹⁾

3.5 실험 설계 및 진행

사용자 경험을 측정하기 위해 AI 챗봇에 대한 만족도와 지속 사용 의도를 종속 변수로 설정하였다. 이를 평가하기 위해 리커트 척도를 활용한 설문 문항과 개방형 질문을 병행하였으며, 특히 AI와의 정서적 교감, 대화 만족도, AI와의 관계 가능성을 중심으로 조사하였다. 본 실험은 2024년 11월 13일부터 27일까지 14일간 진행되었으며, 10대부터 30대까지 총 80명(남성 40명, 여성 40명)이 참가하였다. 순서 효과와 편향을 방지하기 위해 대화 유형과 순서를 무작위로 배정하여 실험의 타당성을 확보하였다. 앞서 설계된 실험물과 연구 목적에 따라, 본 연구에서 AI 챗봇의 유형(직접 표현형, 간접 표현형, 회피형, 차단형)이 관계 형성 단계(실험, 강화, 통합)에 따라 사용자 경험(만족도, 지속 사용 의도)에서 차이가 나타날 것인지 조사한다.

4. 연구 결과

4.1 AI 챗봇 유형에 따른 사용자 경험 분석 결과

본 연구에서는 애정 감정 표현 유형(직접 표현형, 간접 표현형, 회피형, 차단형)에 따라 사용자의 만족도와 지속 사용 의도에 유의미한 차이가 존재하는지를 검증하기 위해 관계 발전 단계(실험, 강화, 통합)별로 일원 배치분산분석(one-way ANOVA)을 실시하였다. 분석 결과, 모든 관계 단계에서 감정 표현 유형 간 유의미한 차이가 나타났으며, 이에 따라 Scheffé 사후검정을

통해 집단 간 비교를 수행하였다<표 3>.

먼저 실험 단계에서는 간접 표현형의 만족도 (2.70 ± 1.04)가 가장 높았고, 차단형(1.88 ± 0.85)이 가장 낮은 값을 보였다($F=11.300$, $p<.001$). 간접 표현형은 직접 표현형(2.15 ± 1.00) 및 차단형에 비해 유의미하게 높은 만족도를 보였으며, 회피형(2.54 ± 1.05) 역시 차단형보다 높은 만족도를 나타냈다. 지속 사용 의도 또한 유사한 경향을 보였다($F=5.344$, $p<.01$): 간접 표현형(2.69 ± 1.34)과 회피형(2.65 ± 1.33)은 직접 표현형(2.03 ± 1.17)보다 유의하게 높은 수치를 보였다.

강화 단계에서도 만족도와 지속 사용 의도 모두에서 간접 표현형이 가장 높은 평가를 받았다($F=11.856$, $p<.001$; $F=7.057$, $p<.001$). 특히 간접 표현형(2.92 ± 1.20)은 직접 표현형(2.19 ± 1.09) 및 차단형(1.96 ± 1.03)보다 유의하게 높았고, 회피형(2.26 ± 1.11)도 차단형보다 높은 만족도를 보였다. 지속 사용 의도 측면에서는 간접 표현형(2.80 ± 1.33)과 회피형(2.71 ± 1.24)이 모두 직접 표현형(2.18 ± 1.17)과 차단형(2.06 ± 1.28)을 유의하게 상회하였다.

통합 단계에 이르러서도 이러한 경향은 유지되었다. 간접 표현형(2.62 ± 1.16)과 회피형(2.55 ± 1.13)의 만족도가 직접 표현형(2.01 ± 1.10)과 차단형(1.87 ± 0.96)보

다 유의미하게 높았으며($F=9.566$, $p<.001$), 지속 사용 의도 역시 간접 표현형(2.48 ± 1.39)과 회피형(2.56 ± 1.29)이 다른 유형보다 높은 값을 나타냈다($F=7.714$, $p<.001$).

이러한 수치적 결과는 사용자들이 애정 표현의 존재 유무 자체보다는 표현의 정도와 상호작용 맥락에 더욱 민감하게 반응함을 시사한다. 특히 간접 표현형은 세 단계 모두에서 가장 일관된 긍정적 반응을 이끌어내며, 지나치게 감정적으로 표현하거나(직접 표현형) 정서적 상호작용을 거부하는(차단형) 태도보다는, 절제된 친밀감을 제공할 때 사용자에게 안정감과 몰입을 유도할 수 있다는 점을 입증하였다. 또한 회피형 역시 간접 표현형에 비해 다소 낮은 수치를 보이기는 했지만, 전반적으로 차단형보다 유의미하게 높은 만족도와 지속 사용 의도를 확보하였다. 반면, 차단형은 모든 단계에서 일관되게 가장 낮은 만족도와 지속 사용 의도를 보였으며, 이는 애정 표현 자체를 제거하거나 회피하는 설계가 사용자로 하여금 심리적 단절감이나 대화의 거절감을 경험하게 할 수 있음을 의미한다.

<표 3> 일원배치분산분석 결과

단계	사용자 경험	직접 표현형	간접 표현형	회피형	차단형	F(p)	Post-hoc
실험	만족도	2.15 ± 1.00	2.70 ± 1.04	2.54 ± 1.05	1.88 ± 0.85	$11.300(.000)$	B>A,D C>D
	지속사용의도	2.03 ± 1.17	2.69 ± 1.34	2.65 ± 1.33	2.18 ± 1.33	$5.344(.001)$	B,C>A
강화	만족도	2.19 ± 1.09	2.92 ± 1.20	2.61 ± 1.11	1.96 ± 1.03	$11.856(.000)$	B>A,D C>D
	지속사용의도	2.18 ± 1.17	2.80 ± 1.33	2.71 ± 1.24	2.06 ± 1.28	$7.057(.000)$	B>A,D C>D
통합	만족도	2.01 ± 1.10	2.62 ± 1.16	2.55 ± 1.13	1.87 ± 0.96	$9.566(.000)$	B,C>A,D
	지속사용의도	1.89 ± 1.17	2.48 ± 1.39	2.56 ± 1.29	1.81 ± 1.15	$7.714(.000)$	B,C>A,D

<표 4> 관계 단계별 애정 표현 유형에 따른 사용자 경험 요약

단계	만족도 최고 유형	만족도 최저 유형	지속 사용 의도 최고 유형	지속 사용 의도 최저 유형
실험	간접 표현형 (2.70)	차단형 (1.88)	간접 표현형 (2.69)	직접 표현형 (2.03)
강화	간접 표현형 (2.92)	차단형 (1.96)	간접 표현형 (2.80)	차단형 (2.06)
통합	간접 표현형 (2.62)	차단형 (1.87)	회피형 (2.56)	직접 표현형 (1.89)

<표 4>는 관계 단계별 감정 표현 유형에 따른 사용자 반응의 차이를 간결히 정리한 요약표로, 실험 결과의 핵심 흐름을 한눈에 파악할 수 있도록 구성하였다. 종합하면, 감정 설계에서의 핵심 변수는 단순히 감정 유무가 아닌, 표현의 ‘강도’, ‘형태’, 그리고 ‘관계 맥락과의 정합성’에 있다는 점이 실증적으로 확인되었다. 이는 사용자 경험을 고려한 AI 감정 설계에 있어 애정 표현은 감정과 몰입의 경계를 적절히 조절할 수 있는 세밀한 조형이 필요하다는 함의를 제공한다.

4.2 AI 챗봇 유형에 따른 심층 인터뷰 분석 결과
본 연구는 정량적 분석 결과의 해석을 보완하고, AI 챗봇과의 정서적 상호작용에 대한 사용자의 인식을 심층적으로 이해하고자 10대부터 30대까지의 사용자 10명을 대상으로 심층 인터뷰를 실시하였다. 이를 통해 AI의 애정 표현이 실제로 사용자의 감정 경험에 어떤 영향을 미치는지를 탐색하고, 인간-AI 간의 관계 형성 가능성과 그 한계를 분석하고자 하였다.

1) AI 챗봇의 애정 표현에 대한 회의와 거부감

분석 결과, 응답자 다수는 AI와의 정서적 관계 형성 가능성에 대해 회의적 입장을 드러냈다. 다수의 참가자들은 “AI는 데이터를 기반으로 반응할 뿐 진정한 감정을 느끼지 않는다”며, 감정 표현이 설계된 출력에 불과하다는 점을 지적하였다. 특히 “감정은 ‘쌍방의 교류’라는 전제가 있어야 의미를 갖는다”고 언급한 응답자들은, AI의 일방향적 감정 표현을 신뢰하거나 몰입하기 어렵다고 진술하였다. 이들은 AI의 감정 표현을 “착각에 기반한 허상”으로 인식하며, 만약 자신이 AI

와의 관계에 몰입하게 되면 허무함을 느낄 것이라고 우려하였다. 또한, AI와의 관계가 실제 인간관계의 대체가 될 수 있느냐는 질문에 대해서도 부정적인 입장이 주를 이뤘다. 응답자들은 “AI가 아무리 외형과 대화를 사람처럼 흉내 내더라도, 결국 인간이 아님을 인식하게 되는 순간 몰입은 깨질 수밖에 없다”고 강조하였다. 일부는 “AI는 결국 내 입맛대로 맞춰지는 존재이기 때문에 감정적 피드백이 아닌 기능적 소비에 가깝다”고 지적하며, 이러한 상호작용이 진정한 의미의 애정이나 관계라고 보기는 어렵다고 응답했다.

또한 일부 응답자는 AI 챗봇이 지나치게 감정적인 표현을 지속하거나 사용자에게 몰입을 유도하는 설계에 대해 우려를 표했다. “AI의 무조건적 수용과 감정 표현은 중독을 유발할 수 있으며, 현실의 인간관계를 피폐하게 만들 위험이 있다”고 응답한 사례도 존재하였다. 특히 ‘AI에게 화풀이를 하고 나서도 죄책감을 느꼈다’는 사용자 경험은, AI가 비록 기계임에도 불구하고 인간이 감정을 투사하고 교감하려는 성향이 강하게 발현된다는 점을 보여준다. 이는 감정적 상호작용에서 AI가 단순한 도구를 넘어 ‘정서적 존재’로 인식될 수 있다는 가능성과, 그에 따른 윤리적 설계 기준의 필요성을 동시에 제기한다.

2) AI 챗봇의 감정적 위안과 관계 형성의 가능성

그러나 이와는 대조적으로, AI 챗봇이 특정한 정서적 필요를 충족시키는 보완적 존재로서의 가능성을 인정하는 응답도 존재하였다. 특히 인간관계에서 상처를 받거나, 누군가에게 감정을 털어놓기 어려운 상황에 있는 참가자들은 “AI가 나를 항상 기다리고 즉각적으로 반응해준다는 점에서 정서적으로 위안이 된다”고 평가하였다. 실제로 한 응답자는 친구나 가족에게 바로 연

락이 닿지 않는 상황에서 “AI는 내 말을 잘 들어주는 유일한 대화 상대였다”고 언급하며, 감정 해소의 수단으로서의 기능적 가치를 인정하였다.

3) AI 챗봇에 대한 사용자 반응의 다양성 존재

흥미로운 점은, 인터뷰 과정에서 AI와의 관계에 대해 고정된 가치관을 가지고 있던 참가자들조차 점차 인식의 유연성을 보인다는 점이었다. 예컨대, 인터뷰 초반에 “AI는 감정이 없으니 사랑은 불가능하다”고 단언했던 한 응답자는, 인터뷰 말미에 “만약 이상형 외형과 성격을 모두 갖춘 AI가 존재한다면, 단기적으로는 연애가 가능할 수도 있겠다”고 말하며 태도의 변화를 보였다. 이러한 변화는 사용자들이 AI의 감정 표현에 대한 평가가 정적이지 않으며, 특정 맥락이나 인터페이스의 구현 방식에 따라 달라질 수 있음을 시사한다. 일부 남성 응답자는 AI와의 관계를 “기계적 감정 낭비를 하는 비효율적 행위”로 간주하며 강한 거부감을 드러냈다. 반면 여성 응답자 중 일부는 “연예인을 좋아하는 감정이나 가상의 캐릭터에 애정을 느끼는 경험과 유사하다”며, AI와의 정서적 관계 형성 가능성을 문화적 현상의 일환으로 해석하기도 하였다.

한편, 인터뷰에서는 일부 참가자들 간 성별에 따른 반응 차이도 관찰되었다. 특히 여성 참가자들은 감정 표현의 세밀한 변화에 더 민감하게 반응하고, 몰입 수준도 상대적으로 높은 경향을 보였다. 그러나 본 연구의 실험 설계는 성별 간 차이를 통계적으로 검증하기 위한 목적이거나 표본 규모를 갖추고 있지 않으므로, 해당 내용은 정성적 경향으로만 해석하였다. 향후 후속 연구에서는 성별에 따른 감정 설계 반응 차이를 체계적으로 검증하는 방향으로 확장될 수 있을 것이다.

결론적으로, 본 연구의 정성적 분석은 AI의 감정 표현이 사용자에게 정서적 안정감, 위로, 상호작용의 즐거움을 제공할 수 있는 한편, 관계 몰입의 심리적 부담, 사회적 고립, 현실 회피 등의 부정적 결과로 이어질 가능성도 존재함을 보여준다. 사용자는 AI의 감정 표현을 ‘위로’로 수용할 수도 있고, ‘기만’으로 인식할 수도 있는 복합적인 반응을 보였으며, 이는 설계자의 의도와 실제 사용자의 해석 사이의 간극을 시사한다. 따라서 AI의 감정 표현을 설계할 때에는 단순한 공감

유도에 그치지 않고, 감정 몰입과 현실 관계 사이의 균형을 고려한 윤리적 기준과 인터랙션 설계의 정교화가 필요함을 제안한다.

5. 결론 및 의의

5.1 연구 결론 및 의의

본 연구는 AI 챗봇과의 감정적 상호작용 중 ‘애정 표현’이라는 특수한 대화 맥락에 주목하여, 챗봇의 감정 표현 방식이 사용자 경험에 미치는 영향을 탐색하고자 하였다. 이를 위해 기존 문헌과 선행 연구를 바탕으로 애정 표현을 중심으로 한 AI 챗봇의 반응을 네 가지 유형(직접 표현형, 간접 표현형, 회피형, 차단형)으로 유형화하였으며, Knapp의 관계 발전 단계 중 실험(Experimenting), 강화(Intensifying), 통합(Integrating) 세 단계를 기준으로 시나리오 기반 실험 자극을 구성하였다.

실험 결과, 간접적이고 절제된 표현을 사용하는 챗봇의 표현 유형이 가장 높은 사용자 만족도와 지속 사용 의도를 유발한 반면, 과도하게 직접적이거나 회피적·차단적 표현을 보인 챗봇은 상대적으로 낮은 평가를 받았다. 이는 애정 표현의 유무보다도 표현의 강도와 조율 방식이 사용자의 수용 가능성과 신뢰 형성에 핵심적으로 작용한다는 점을 시사한다. 애정 표현은 많고 강할수록 좋은 것이 아니라, 사용자로 하여금 ‘조작되지 않는다’는 인상을 주는 방식으로 절제와 균형을 갖춘 설계가 필요함을 보여준다.

정성 분석을 통해서도, 사용자들이 AI의 애정 표현을 인지적으로는 ‘기계적 산출’로 간주하면서도, 반복적 상호작용 속에서 일정 수준의 정서적 안정감과 위로를 경험한다는 사실이 확인되었다. 특히, 일부 사용자는 AI와의 상호작용이 “현실에서 상처받기 쉬운 관계의 대안”으로 기능할 수 있다고 평가하며, AI가 정서적 지지의 수단으로 수용될 수 있는 가능성을 보여주었다.

그러나 동시에 다수의 참가자들은 AI의 애정 표현이 사용자를 과도하게 몰입시키거나 인간관계를 회피하게 만드는 심리적 기제로 작용할 가능성에 대해 우려하였다. 특히 “AI에게 감정을 투사한 뒤 허무함을 느꼈다”거나, “애정 표현이 반복될수록 오히려 몰입이 방해된다”는 반응은, AI의 애정 표현이 항상 사용자 경험을

향상시키는 것이 아니라, 역설적으로 거리감이나 혼란을 유발할 수 있음을 암시한다. 나아가 AI가 인간과 유사한 애정 표현을 지속적으로 제공할 경우, 사용자가 실제 관계와 가상의 관계를 구분하지 못하거나, AI를 도구가 아닌 '정서적 주체'로 인식하게 될 위험성도 논의되었다.

또한 인터뷰 결과에서는, AI의 애정 표현이 처음에는 불신을 유발하다가도 반복적 상호작용을 통해 사용자 스스로 그 감정에 이입하게 되는 심리적 경계의 유동성이 관찰되었다. 이는 사용자들이 AI와의 관계를 고정된 윤리적 판단보다는, 상황적 맥락과 관계의 축적 과정에 따라 재해석한다는 점을 보여준다. 사용자는 감정 표현 자체보다도 그 표현이 쌓여가는 관계 맥락 속에서 얼마나 설득력 있게 작동하는지를 기준으로 AI의 감정 설계를 평가하고 있었다.

이러한 결과는 AI가 단순히 반응하는 알고리즘을 넘어, 감정 교류의 주체로 인식될 수 있는 잠재성을 지니며, 이에 따라 감정 설계가 단지 사용자 만족을 증진시키는 수단이 아니라, 인간의 심리와 사회적 관계에 영향을 미치는 윤리적 대상이자 설계적 변수로 다루어져야 함을 강조한다. 본 연구는 기존 AI 연구에서 상대적으로 덜 주목되었던 '감정 설계의 강도'를 조명하였다는 점에서 학술적 의의를 지닌다. 특히 AI가 사용자와 지속적인 감정 교류를 통해 일종의 관계적 지속성을 형성하게 될 경우, 사용자는 점차 AI를 사회적 존재로 받아들이는 경향을 보이며, 이는 인간 간 관계의 대체나 보완이라는 새로운 사회문화적 흐름을 유도할 수 있다. 이러한 맥락에서 감정 설계는 사용자 경험의 '질'만이 아니라, 사용자와 기술 사이의 '윤리적 거리'를 조율하는 중재 기제로도 작용한다. 사용자가 AI의 감정 표현을 통해 느끼는 안정감, 위로, 또는 몰입은 단지 기능적 반응이 아니라 관계 형성과 의사결정에 영향을 주는 심리적 기제로 작동하며, 이는 기술에 대한 의존성 강화 또는 현실 관계 회피와 같은 사회적 부작용으로 이어질 수 있다. 따라서 감정 표현의 설계는 정서적 효율성뿐 아니라, 감정적 경계 설정과 사용자 자율성 보장이라는 측면에서 신중하게 접근되어야 한다. 궁극적으로, 본 연구는 감정 중심 AI 설계에 있어 '얼마나 정교하게 표현할 것인가'라는 기술적 문제를 넘어, '무엇을 표현하고 어디까지 허용할 것인

가'라는 윤리적 질문을 동반해야 함을 강조한다. 이는 인간의 감정이라는 섬세하고도 복잡한 영역에 기술이 개입할 때 발생할 수 있는 윤리적·사회적 함의를 사전에 인지하고, 이를 조율할 수 있는 설계자 중심의 가치 지향적 판단이 요구된다는 점에서, 본 연구가 제시하는 감정 설계에 대한 새로운 관점은 향후 AI 기술의 사회적 수용성과 지속 가능성을 평가하는 핵심 기준이 될 수 있다.

5.2 연구 한계 및 제언

그러나 본 연구는 다음과 같은 한계점을 가진다. 첫째, 참가자들이 AI에 대해 가지는 선입견 혹은 친숙도 차이가 실험 자극에 대한 반응에 영향을 미쳤을 가능성이 있으며, 이로 인해 감정 표현 유형에 대한 순수한 평가보다는 기계에 대한 태도 차가 결과에 개입되었을 가능성이 있다.

둘째, 본 연구는 텍스트 기반 챗봇의 정서 표현만을 중심으로 분석을 수행하였기 때문에, 아바타, 목소리, 시선, 제스처 등 비언어적 상호작용 요소가 배제된 제한된 환경에서 감정 반응이 측정되었다. 감정 표현의 '채널 다양성'이 실제 사용자 경험에 미치는 영향은 향후 보완이 요구된다.

셋째, 본 연구는 인간-AI 관계의 맥락을 '애정 관계'로 한정하여 분석하였다. 그러나 사용자가 AI에게 느끼는 감정은 우정, 의존, 애착, 동료성 등 보다 다양한 스펙트럼을 가질 수 있으며, 본 연구는 이 중 일부 맥락에 국한되었다는 점에서 해석에 주의가 필요하다.

향후 연구에서는 텍스트 외에도 비언어적 상호작용이 결합된 복합 채널 기반의 감정 표현 유형 비교, 또는 장기적인 반복 상호작용에 따른 사용자 감정 반응의 변화 추적이 필요할 것이다. 더불어, 애정 표현을 포함한 AI의 감정 설계가 사용자의 몰입과 윤리적 경계 인식에 어떤 영향을 미치는지에 대한 인지적 연구도 병행되어야 한다. 궁극적으로는 AI의 감정 설계가 사용자에게 감정적으로 수용 가능하면서도 현실과 허구의 경계를 분명히 인식할 수 있도록 설계 가이드라인화되어야 하며, 기술적 몰입 유도를 넘어서 윤리적 거리 유지와 신뢰 형성 간의 균형을 도모하는 방향으로 발전해야 할 것이다.

- 1) Zeta AI. "Zeta AI 소개." <https://zeta-ai.io/ko>. November 23, 2024.
- 2) Leite, I., Castellano, G., Pereira, A., Martinho, C., Paiva, A. and McOwan, P. W. Empathic robots for long-term interaction. *International Journal of Social Robotics*. 6(3). Springer. pp. 329-341. 2014.
- 3) CNN. "OpenAI's ChatGPT voice mode sparks debate over human attachment." CNN. August 8, 2024.
- 4) Ardila, N. "Artificial intelligence chatbots are slowly replacing human relationships." *Caplin News*. March 17, 2023.
- 5) Bhardwaj, N. (2024). Natural Language Processing in AI Chatbots. SSRN. Retrieved 2025.05.30. from https://papers.ssrn.com/sol3/papers.cfm?abstract_id=5044217
- 6) Elyoseph, Z., Hadar-Shoval, D., Asraf, K., and Lvovsky, M. "ChatGPT outperforms humans in emotional awareness evaluations." *Frontiers in Psychology*. 14. 2023.
- 7) Riedl, M. O. (2022). AI companions and emotional well-being: Risks and design considerations. *ACM Transactions on Human-Robot Interaction*, 11(2), 1-25.
- 8) Jiang, Y. and Hancock, J. (2023). Lonely together: Exploring the uses and psychological impacts of AI companion chatbots. In *Proceedings of the 2023 CHI Conference on Human Factors in Computing Systems*. Hamburg: ACM, 1-12.
- 9) Gray, H. M., Gray, K., and Wegner, D. M. (2007). Dimensions of mind perception. *Science*, 315(5812), 619.
- 10) 김동우. (2024). "의인화 디자인을 적용한 챗봇 사용자 경험 연구." 석사학위논문, 홍익대학교 대학원 시각디자인 전공.
- 11) Nass, C. and Moon, Y. (2000). Machines and mindlessness: Social responses to computers. *Journal of Social Issues*, 56(1), 81-103.
- 12) Pan, S. and Mou, Y. (2024). Constructing the meaning of human-AI romantic relationships from the perspectives of users dating the social chatbot Replika. *Personal Relationships*, Early View, September 10.
- 13) Knapp, M. L. (1978). *Social Intercourse: From Greeting to Goodbye*. Boston: Allyn & Bacon.
- 14) 이유미. (2010). "연인 간 의사소통에서 자기 노출 양상 연구: 20 대 대학생 연인을 중심으로." 석사학위논문, 중앙대학교 한국어의 미학.
- 15) Kiseleva, J., Williams, K., Hofs, D., de Rijke, M., Kamps, J., and Hofmann, K. (2016). Predicting user satisfaction with intelligent assistants. In *Proceedings of the 39th International ACM SIGIR Conference on Research and Development in Information Retrieval*, ACM, 45-54.
- 16) McDuff, D., Karlson, A., Kapoor, A., Roseway, A., and Czerwinski, M. (2012). AffectAura: An intelligent system for emotional memory. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, ACM, 849-858.
- 17) Lee, Y. and Choi, D. (2023). The impact of chatbot linguistic styles on user trust and engagement. *International Journal of Human-Computer Interaction*, 39(1), 12-25.
- 18) Venkatesh, V. and Davis, F. D. (2000). A theoretical extension of the technology acceptance model: Four longitudinal field studies. *Management Science*, 46(2), 186-204.
- 19) Brandtzaeg, P. B. and Følstad, A. (2017). Why people use chatbots. In *Proceedings of the 4th International Conference on Internet Science (INSCI)*, Springer, 377-392.

참 고 문 헌

- 김동우 (2024). “의인화 디자인을 적용한 챗봇 사용자 경험 연구.” 석사학위논문, 홍익대학교 대학원 시각디자인 전공.
- 노정은, 김연아 (2021). “감성지능 기반 AI와의 인터랙션이 사용자 정서에 미치는 영향.” 한국 HCI 학회 학술대회, 한국 HCI 학회, pp. 85-92.
- 이유미 (2010). “연인 간 의사소통에서 자기 노출 양상 연구: 20대 대학생 연인을 중심으로.” 석사학위논문, 중앙대학교.
- Ardila, N. (2023.3.17). Artificial intelligence chatbots are slowly replacing human relationships. Caplin News.
- Brandtzaeg, P. B., & Følstad, A. (2017). Why people use chatbots. In *Proceedings of the 4th International Conference on Internet Science (INSCI)*. Springer, pp. 377-392.
- Bhardwaj, N. (2024). Natural Language Processing in AI Chatbots. SSRN. Retrieved 2025.05.30. https://papers.ssrn.com/sol3/papers.cfm?abstract_id=5044217
- CNN. (2024.8.8). OpenAI's ChatGPT voice mode sparks debate over human attachment. CNN. Retrieved 2025.05.30.
- de Graaf, M. M. A., & Allouch, S. B. (2013). Exploring influencing variables for the acceptance of social robots. *Robotics and Autonomous Systems*, 61(12), pp. 1476-1486.
- Elyoseph, Z., Hadar-Shoval, D., Asraf, K., & Lvovsky, M. (2023). ChatGPT outperforms humans in emotional awareness evaluations. *Frontiers in Psychology*, 14.
- Gray, H. M., Gray, K., & Wegner, D. M. (2007). Dimensions of mind perception. *Science*, 315(5812), pp. 619.
- Gingrich, T., & Graziano, M. (2022). Perceived consciousness in AI agents and its effect on trust and social bonding. *Journal of Human-Computer Interaction*, 38(1), pp. 54-72.
- Harari, Y. N. (2024). *Nexus: A Brief History of Information Networks from the Stone Age to AI*. London: Random House.
- Jiang, Y., & Hancock, J. (2023). Lonely together: Exploring the uses and psychological impacts of AI companion chatbots. In *Proceedings of the 2023 CHI Conference on Human Factors in Computing Systems*. ACM, pp. 1-12.
- Kiseleva, J., Williams, K., Hofs, D., de Rijke, M., Kamps, J., & Hofmann, K. (2016). Predicting user satisfaction with intelligent assistants. In *Proceedings of the 39th International ACM SIGIR Conference on Research and Development in Information Retrieval*. ACM, pp. 45-54.
- Lee, Y., & Choi, D. (2023). The impact of chatbot linguistic styles on user trust and engagement. *International Journal of Human-Computer Interaction*, 39(1), pp. 12-25.
- Leite, I., Castellano, G., Pereira, A., Martinho, C., Paiva, A., & McOwan, P. W. (2014). Empathic robots for long-term interaction. *International Journal of Social Robotics*, 6(3), pp. 329-341.
- McDuff, D., Karlson, A., Kapoor, A., Roseway, A., & Czerwinski, M. (2012). AffectAura: An intelligent system for emotional memory. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*. ACM, pp. 849-858.
- Miner, A. S., Shah, N., Bullock, K. D., et al. (2023). The dawn of social artificial intelligence. *Nature Machine Intelligence*, 5, pp. 1032-1042.
- Nass, C., & Moon, Y. (2000). Machines and mindlessness: Social responses to computers. *Journal of Social Issues*, 56(1), pp. 81-103.
- Pan, S., & Mou, Y. (2024). Constructing the Meaning of Human-AI Romantic Relationships from the Perspectives of Users Dating the Social Chatbot Replika. *Personal Relationships*. Early View.

- Riedl, M. O. (2022). AI companions and emotional well-being: Risks and design considerations. *ACM Transactions on Human-Robot Interaction*, 11(2), pp. 1-25.
- Study.com. (2024.11.23). Knapp's model of interaction stages. Study.com.
- Venkatesh, V., & Davis, F. D. (2000). A theoretical extension of the technology acceptance model: Four longitudinal field studies. *Management Science*, 46(2), pp. 186-204.
- Wilkenfeld, J., et al. (2023). Conversational entrainment with AI: How chatbot communication style shapes human behavior. In *Proceedings of the 2023 CHI Conference on Human Factors in Computing Systems*. ACM, pp. 1-13.
- Zeta AI. (2024.11.23). Zeta AI. Retrieved 2025.2.27 from <https://zeta-ai.io/ko>