Check for updates

Development and Evaluation of an AI-Based Mock Interview Training System to Support User Interview Skills in Design Education

Moon-Hwan Lee[1*], Sumin Seong[2], Sanha Shin[2]

[1]Department of Design Engineering, Associate Professor, Tech University of Korea, Siheung, Korea
[2]Department of Design Engineering, Bachelor Student, Tech University of Korea, Siheung, Korea

Abstract

Background User interviews are a fundamental qualitative research method in design research, enabling designers to explore users' experiences, motivations, and contextual behaviors. However, the quality of interview outcomes largely depends on the interviewer's skill in constructing questions, managing conversational flow, and responding to participants' answers. In design education, novice designers often have limited opportunities to repeatedly practice interviews or to receive detailed feedback on their performance. This study aims to explore the potential of an artificial intelligence (AI)-based training environment that supports the development of user interview skills through simulated practice.

Methods This study followed a Research-through-Design (RtD) approach consisting of three stages: formative investigation, system design and implementation, and user evaluation. First, formative interviews with four design educators were conducted to identify common difficulties faced by novice interviewers and limitations in current interview training practices. Based on these findings, an AI-based mock interview training system called Pretalk was designed, integrating question generation, virtual interviewee persona creation, voice-based interview simulation, and performance analysis. Finally, user testing was conducted with six undergraduate design students who used the system in two iterative training sessions, allowing observation of changes in interview performance and learning experiences.

Results The results indicate that the system helped participants recognize and adjust their interview strategies by visualizing key aspects of interview performance, such as speaking–listening balance, time allocation, follow-up questioning, and emotional responses. Across repeated training sessions, changes were observed in interview structure, follow-up-question use, and awareness of empathetic responses during interviews. Participants also reported that the system supported pre-interview rehearsal and helped them reflect on their interaction patterns.

Conclusions This study demonstrates the potential of AI-based simulation environments as training tools for developing user interview skills in design education. By enabling repeated practice and data-driven feedback, the proposed system supports reflective learning during the interview process. Furthermore, the findings suggest that AI can function not only as an automation tool but also as a partner that assists designers in developing research skills. These insights contribute to the design of AI-supported learning environments for design research education.

Keywords User Interview Training, Generative AI, Design Research Education, Mock Interview System

Copyright : 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 educational and non-commercial use, provided the original work is properly cited.

This work was supported by the Korea Planning & Evaluation Institute of Industrial Technology (KEIT) grant funded by the Korea government (MOTIE) (No. 20023065).

*Corresponding author: Moon-Hwan Lee (moonhwanlee@tukorea.ac.kr)

Citation: Lee, M-H., Seong, S., & Shin, S. (2026). Development and Evaluation of an AI-Based Mock Interview Training System to Support User Interview Skills in Design Education. *Archives of Design Research, 39*(4), 227-248.

http://dx.doi.org/10.15187/adr.2026.11.39.4.227

Received : Mar. 19. 2026 ; **Reviewed :** Jul. 31. 2026 ; **Accepted :** Aug. 15. 2026
pISSN 1226-8046 **eISSN** 2288-2987

1. 서론

사용자 인터뷰(User Interview)는 사용자 경험 리서치에서 가장 널리 활용되는 질적 조사 방법 중 하나로, 사용자의 요구와 인식, 경험의 맥락을 심층적으로 탐색할 수 있는 강점을 지닌다(Gubrium & Holstein, 2002; Kvale & Brinkmann, 2009). 특히 디자인 과정의 초기 단계에서 인터뷰는 문제 정의와 인사이트 도출의 핵심 도구로 활용되며, 이후 아이디어 발산 및 콘셉트 개발 과정의 근거를 형성한다(Kumar, 2012).

그러나 사용자 인터뷰의 성과는 인터뷰 진행자의 역량에 크게 의존하며, 질문의 구성 방식, 대화의 흐름 조율, 경청 및 반응의 질 등에 따라 인터뷰의 심층성이 크게 달라진다(Kvale & Brinkmann, 2009). 초보 디자이너나 학생 연구자는 이러한 역량을 충분히 갖추기 전에 실습을 수행하게 되는 경우가 많아, 인터뷰의 목적을 명확히 설정하거나 대화를 유연하게 이끌어가는 데 어려움을 겪는 것으로 알려져 있다(Gubrium & Holstein, 2002).

이러한 맥락에서 사용자 인터뷰 교육은 단순한 질문 기술의 습득을 넘어, 질문 구성 능력과 후속 질문 전략, 그리고 인터뷰 흐름을 조율하는 역량을 함께 개발하는 학습 활동으로 이해되고 있다. 그러나 디자인 교육 환경에서는 실제 사용자와의 반복적 인터뷰 실습이 시간과 인적 자원의 제약으로 제한되는 경우가 많다. 또한 교수자가 학습자의 인터뷰 수행 과정을 지속적으로 관찰하고 세밀한 피드백을 제공하기에도 현실적인 어려움이 존재한다(Kumar, 2012).

이러한 교육적 한계를 보완하기 위해, 최근에는 AI 기반 시뮬레이션을 활용한 학습 환경이 사용자 연구 교육에서도 점차 관심을 받고 있다. 특히 챗봇 인터페이스를 활용한 대화형 시뮬레이션 시스템은 학습자가 실제와 유사한 환경에서 반복적으로 인터뷰를 연습할 수 있도록 지원하며, 즉각적인 피드백을 제공함으로써 자기주도적 학습을 촉진한다(Lin et al., 2025; Wilkie & Rosendale, 2024). 그러나 기존의 대부분 시스템은 채용 면접이나 기술 면접과 같은 평가 중심의 맥락에 초점을 맞추고 있어, 디자인 프로젝트의 사용자 인터뷰 상황에서 요구되는 개방형 질문 구성, 꼬리 질문 훈련, 정서적 교감 능력 향상 등 심층적 인터뷰 역량을 개발하는 데에는 한계가 있다.

이에 본 연구는 사용자 인터뷰 경험이 부족한 디자인 전공 학습자를 대상으로, 인터뷰 수행 훈련을 지원하는 AI 기반 모의 인터뷰 훈련 시스템(Pretalk)을 설계하고 그 활용 가능성을 탐색하고자 한다. 본 연구는 세 단계로 진행되었다. 첫째, 디자인 교육 현장에서 인터뷰 지도를 담당하는 교육자 4명을 대상으로 형성적 인터뷰를 실시하여 초보 학습자들이 겪는 주요 어려움과 교육적 제약을 규명하였다. 둘째, 그 결과를 바탕으로 인터뷰 질문 생성, 가상 인터뷰이 생성, 실시간 인터뷰 진행, 인터뷰 결과 분석 및 피드백 기능을 통합한 AI 기반 모의 인터뷰 시스템 Pretalk을 설계 및 구현하였다. 셋째, 디자인 전공 학부생을 대상으로 Pretalk을 활용한 사용자 테스트를 수행하여, 시스템의 활용 가능성과 수행 변화를 평가하였다.

본 연구는 AI가 사용자 인터뷰에 미숙한 학습자, 즉 초보 인터뷰어의 인터뷰 수행 역량과 자기 점검을 어떻게 지원할 수 있는지를 탐색함으로써, 디자인 교육에서 사용자 인터뷰 학습을 지원하는 AI 기반 훈련 도구의 설계 방향을 제시하는 데 목적이 있다.

2. 관련 연구

2. 1. 사용자 인터뷰의 단계별 수행 역량

사용자 인터뷰는 디자인 과정에서 사용자의 경험과 맥락을 심층적으로 이해하기 위한 핵심적인 질적 조사 방법이다(Goodman et al., 2012; Hanington & Martin, 2022). 그러나 인터뷰의 성과는 단순히 질문 목록의 완성도에 의해 결정되지 않으며, 인터뷰 전 과정을 구조화하고 조율하는 인터뷰어의 수행 역량에 크게 의존한다(Kvale & Brinkmann, 2009). 이러한 역량은 단일한 기술이 아니라, 인터뷰의 준비, 수행, 분석 단계에서 서로 다른 형태로 요구되는 복합적 능력으로 구성된다(Table 1).

먼저 준비 단계에서는 인터뷰의 목적과 탐색 범위를 명확히 설정하고, 이를 기반으로 적절한 질문 구조를

설계하는 능력이 요구된다. 인터뷰 목표가 구체화되지 않을 경우 질문은 산만해지기 쉽고, 이후 분석 단계에서도 의미 있는 통찰을 도출하기 어렵다. 초보 인터뷰어의 경우 인터뷰 목적을 추상적으로 설정하거나, 질문을 정보 확인 중심으로 구성하여 탐색적 대화를 충분히 유도하지 못하는 경향이 나타난다(Goodman et al., 2012).

수행 단계에서는 계획된 질문을 전달하는 것을 넘어, 인터뷰이의 발화를 해석하고 그 의미를 확장하는 상호작용적 역량이 중요하다. 인터뷰어는 응답의 맥락을 이해하고, 필요한 경우 후속 질문을 통해 경험의 배경과 감정을 구체화해야 한다(Kvale & Brinkmann, 2009). 이 과정에서 경청 능력, 질문의 타이밍 조절, 대화 흐름의 유지가 핵심적인 역할을 한다. 그러나 초보자는 사전에 준비한 질문에 집착하거나, 예상치 못한 응답에 유연하게 대응하지 못해 대화가 단절되는 어려움을 겪는다. 즉, 수행 단계는 계획된 구조와 즉흥적 탐색이 균형을 이루어야 하는 고도의 상호작용 과정이다.

분석 단계에서는 수집된 발화를 단순히 요약하는 것을 넘어, 인터뷰이가 드러낸 경험의 맥락과 의미를 해석적으로 재구성하는 능력이 요구된다(Gubrium & Holstein, 2002). 초보 인터뷰어는 종종 응답 내용을 표면적으로 정리하거나, 자신의 가설에 부합하는 부분만을 강조하는 오류를 범한다. 따라서 의미 있는 인사이트를 도출하기 위해서는 반복적 검토와 체계적인 질적 분석 과정이 필요하다.

이와 같이 사용자 인터뷰는 준비, 수행, 분석의 단계로 이루어지며, 각 단계는 서로 다른 유형의 인지적, 수행적 역량을 요구한다. 특히 초보 인터뷰어에게는 이러한 역량을 통합적으로 수행하는 것이 쉽지 않아, 반복 연습과 피드백을 통해 점진적으로 극복해야 한다.

Table 1 Stages of User Interview and Associated Activities

구분	세부 단계		설명
인터뷰 준비	1	목표 설정	인터뷰의 목적과 목표를 명확히 정의. (예: 사용자 요구사항 파악, 피드백 수집 등)
	2	인터뷰 대상 선정과 이해	인터뷰할 사용자 그룹을 정의하고 선정. 타깃 사용자 프로필을 고려.
	3	인터뷰 질문 및 진행 방식 수립	질문 리스트를 작성. 개방형 질문과 닫힌 질문을 적절히 혼합하는 등의 진행 방식에 대해서 결정.
	4	파일럿 테스트	초기 인터뷰 질문과 프로세스를 소규모로 테스트하여 인터뷰 진행의 문제점을 발견. 돌발 상황에 대한 확인.
	5	인터뷰 계획 수정	인터뷰 진행상의 문제점을 기반으로 하여 인터뷰 질문을 포함한 인터뷰 과정을 수정.
인터뷰 수행	6	일정 조율 및 환경 설정	인터뷰 일정을 조정하고, 사용자와의 약속을 확정. 인터뷰를 진행할 장소를 준비.
	7	인터뷰 진행	질문을 통해 사용자의 경험과 의견을 수집. 경청하고 필요한 경우 추가 질문.
인터뷰 분석	8	결과 정리	인터뷰 후 수집한 데이터를 정리하고, 주요 인사이트를 도출.

2. 2. 사용자 인터뷰에서의 AI 활용 연구 동향

최근 HCI 및 디자인 연구 분야에서는 AI를 모의 인터뷰 진행 및 인터뷰 결과 분석 과정에 보조 도구로 활용하려는 시도가 이루어지고 있다. 이러한 흐름은 사용자 인터뷰의 준비, 수행, 분석 단계 전반에서 나타난다.

먼저 인터뷰 준비 단계에서는 질문 생성과 인터뷰 설계를 지원하는 AI 기반 도구들이 등장하고 있다. 일부 UX 리서치 도구에서는 인터뷰 목적을 입력하면 질문 초안을 자동으로 생성해주는 기능이 제공되고 있으며, 이는 인터뷰 준비 과정의 초기 부담을 줄이는 데 활용되고 있다. HCI 분야 연구 중에서 Hu 등은 대화 에이전트가 사용자의 응답 맥락을 분석하여 후속 질문을 생성하도록 설계하고 그 효과를 평가하였다(Hu et al., 2024).

또한 대화 시뮬레이션을 활용한 모의 인터뷰 및 면접 훈련 시스템도 연구되어 왔다. Hoque 등은 AI 기반 면접 코칭 시스템을 통해 사용자의 말하기 방식과 비언어적 행동을 분석하고 피드백을 제공함으로써 수행 향상을 지원하였다(Gomez et al., 2025; Hoque et al., 2013; Leong et al., 2024). Lucas 등은 가상 면접 평가자와의 상호작용이 실제 상황보다 심리적 부담을 낮추어 사용자의 자기 표현을 더 쉽게 유도할 수 있음을 보고하였다(Lucas et al., 2014). 이러한 연구들은 AI 기반 대화 환경이 반복 연습과 수행 훈련을 지원할 수 있음을 시사한다.

한편 최근에는 LLM을 활용하여 특정 페르소나나 사용자 집단을 모사한 가상의 응답자를 생성하는 연구도 이루어지고 있다. 이러한 접근은 설문이나 인터뷰 질문에 대한 예측 응답을 생성하여 아이디어 검증이나 초기 탐색을 빠르게 수행할 수 있다는 장점을 지닌다(Argyle et al., 2023). 반면, AI 응답이 실제 사용자 경험을 완전히 대변할 수 있는지에 대해서는 여전히 논의가 존재한다.

마지막으로 인터뷰 분석 단계에서는 LLM을 활용한 질적 데이터 분석 지원 연구가 활발히 이루어지고 있다. Yan 등은 ChatGPT를 활용한 주제 분석 과정에서 인간–AI 협업 방식을 실험하고, AI가 코드 후보 생성과 요약 과정에 기여할 수 있음을 보고하였다(Yan et al., 2024). Tai 등은 질적 연구에서 LLM 활용의 가능성과 한계를 분석하며, AI가 초기 데이터 구조화 단계에서 연구자의 인지적 부담을 경감할 수 있음을 제시하였다(Tai et al., 2024).

2. 3. 본 연구의 차별성

앞선 연구들은 AI를 활용하여 사용자 인터뷰 과정의 일부를 자동화하거나 대체하는 다양한 접근을 제시해 왔다. 예를 들어 인터뷰 질문을 자동으로 생성하거나, 가상의 인터뷰이를 구성하여 응답을 생성하고, 인터뷰 데이터를 자동으로 분석하는 방식 등이 연구되어 왔다. 이러한 접근은 인터뷰 준비와 분석 과정의 효율성을 높이고 연구자의 작업 부담을 줄이는 데 기여하였다.

그러나 가상 인터뷰이가 생성하는 응답은 실제 사용자의 경험과 의견을 충분히 반영하지 못할 가능성이 있다. 이러한 이유로 실제 사용자와의 상호작용을 기반으로 하는 사용자 인터뷰는 여전히 중요한 사용자 이해 방법으로 남아 있다.

이에 본 연구는 AI를 인간을 대체하는 도구가 아니라 인간 인터뷰어의 수행 역량을 보조하고 훈련하는 도구로 활용하는 접근을 제안한다. 본 연구에서 제안하는 Pretalk 시스템은 AI가 인터뷰를 대신 수행하는 것이 아니라, 가상 인터뷰 환경과 분석 피드백을 제공함으로써 학습자가 인터뷰어로서 실제 사람과의 인터뷰 상황을 보다 효과적으로 준비하고 연습할 수 있도록 설계되었다. 즉, AI가 인터뷰 절차를 자동화하는 것이 아니라 인간 인터뷰어가 인간 인터뷰이를 인터뷰할 때의 수행을 향상시키도록 지원하는 학습 도구라는 점에서 기존 연구와 차별성을 가진다.

3. 연구 방법: 디자인을 통한 연구(Research-through-Design)

본 연구는 사용자 인터뷰 경험이 부족한 초보 디자이너 및 학생 연구자들이 인터뷰를 효과적으로 준비하고 훈련할 수 있도록 지원하는 AI 기반 모의 인터뷰 시스템(Pretalk)을 설계하고, 그 교육적 활용 가능성을 탐색하는 것을 목적으로 한다. 이를 위해 디자인을 통한 연구(Research-through-Design, RtD)를 적용하였다(Zimmerman et al., 2007). RtD는 디자이너가 실험적 제작 과정을 통해 새로운 개념적 통찰과 실천적 지식을 창출하는 연구 방법으로, 본 연구는 형성적 조사(Formative Study), 시스템 설계 및 구현, 사용자 평가의 세 단계로 구성된다. 이와 같은 RtD 절차를 통해 본 연구는 교육 맥락에서 관찰된 요구를 설계 요구사항으로 구체화하고, 이를 구현된 프로토타입과 사용자 평가로 연결함으로써 AI 기반 인터뷰 훈련 도구의 설계 원리와 적용 가능성을 체계적으로 탐색하였다(Table 2). 이 과정에서 Pretalk은 형성적 조사에서 도출된 설계 방향을 구현하고 사용자 평가를 통해 검토하기 위한 연구 아티팩트로 활용되었다. 사용자 평가에서 확인된 한계와 개선 요구는 초기 설계 방향을 재검토하고 후속 설계 방향을 구체화하는 근거로 활용하였다.

Table 2 Research Stages Based on the Research-through-Design Approach

연구 단계	주요 목적	실행 방법
1단계. 형성적 조사	사용자 인터뷰 훈련의 교육적 어려움 및 요구사항 탐색	디자인 교육자(교수) 4인 대상 심층 인터뷰
2단계. 시스템 설계 및 구현	학습자 (초보 인터뷰어)의 인터뷰 훈련을 지원하는 AI 기반 시스템(Pretalk) 설계 및 프로토타이핑	Pretalk 시스템 설계 및 생성형 AI 기반 대화 시뮬레이션 구현
3단계. 사용자 평가	시스템의 교육적 활용 가능성 및 개선점 탐색	디자인 전공 학부생 6명 대상 사용자 테스트

4. 형성적 조사

4. 1. 목적

RtD의 초기 단계에서는 맥락 이해와 설계 요구 도출이 중요하다. 이에 본 연구는 형성적 조사를 통해 사용자 인터뷰 교육 맥락에서 나타나는 문제를 탐색하였다. 본 조사의 목적은 사용자 인터뷰에 익숙하지 않은 학습자들이 실제 인터뷰 준비와 수행 과정에서 겪는 어려움을 이해하고, 이를 지도하는 교육자의 관점에서 기존 훈련 방식의 한계와 보완점을 파악하는 것이다. 이를 통해 초보 학습자의 인터뷰 역량 강화를 지원할 수 있는 AI 기반 인터뷰 훈련 시스템 설계를 위한 초기 사용자 요구 사항을 도출하고자 하였다.

4. 2. 참가자 및 진행 과정

참가자는 사용자 인터뷰 교육 경험이 풍부한 디자인 전공 교수 4인으로 구성되었다(Table 3). 이들은 모두 대학 또는 대학원 수준에서 사용자 인터뷰를 포함하여 사용자 조사 기법을 수업 및 프로젝트 진행에서 활용해온 전문가들이다. 반구조화된 인터뷰는 1:1 온라인 대면 방식으로 진행되었으며, 각 인터뷰는 약 40분간 진행되었다. 인터뷰 질문은 크게 세 가지 범주로 구성되었다:

(1) 초보 학습자들이 사용자 인터뷰를 준비하거나 수행할 때 자주 마주치는 어려움이나 실수는 무엇인가?

(2) 현재 교육 현장에서 활용되는 인터뷰 훈련 방식의 한계는 무엇인가? (3) 기술적 도구나 시스템이 이러한 훈련을 어떻게 보완할 수 있을 것 같은가?

수집된 인터뷰 데이터는 주제 분석(thematic analysis)을 통해 분석하였다. 인터뷰 녹음을 전사한 뒤 의미 단위를 중심으로 개방 코딩을 수행하고, 유사한 코드를 범주화하여 초보 학습자가 겪는 어려움과 교육적 제약 요인을 중심으로 상위 주제를 도출하였다. 도출된 주제는 연구자 간 검토와 토의를 통해 수정되었으며, 이를 통해 인터뷰 훈련과 관련된 주요 문제 유형과 설계 요구 사항을 정리하였다.

Table 3 Participants in the Formative Study

참가자 E = Educator	전문 분야	전문 분야 경력
E1	서비스 디자인	10년
E2	제품 및 인터랙션 디자인	11년
E3	UX 디자인	5년
E4	UX 디자인	5년

4. 3. 형성적 조사 결과: 인터뷰 훈련의 문제 유형 및 요구사항

디자인 교육자 4인(E1–E4)과의 인터뷰를 통해, 초보 학습자들이 사용자 인터뷰 과정에서 흔히 겪는 문제와 현행 교육의 한계가 네 가지 범주로 도출되었다(Table 4). 이 요구사항들은 인터뷰를 절차적으로만 이해하거나 유동적인 대화 상황에 적응하지 못하는 초보 학습자의 한계와 교육 환경의 제약을 함께 드러냈다.

Table 4 Challenges in User Interview and Training

구분		주요 문제	발화자 E = Educator
1	인터뷰 질문 구성 미숙	닫힌 질문, 유도적 질문, 모호한 질문 작성으로 풍부한 응답 유도 실패	E1, E2, E3, E4
2	진행 과정 즉각적 대응 부족	경청하는 태도 부족하여 꼬리 질문 부족 예상 밖 응답·돌발 상황에 추가 질문을 하지 못하고 당황	E1, E3, E4
3	인터뷰이와의 정서적 교감 부족	인터뷰 초반에 긴장 완화 실패, 라포 형성·공감 유도 부족 대화 확장보다는 기록에 집중	E1, E2, E4
4	반복 훈련 부족	수업 중 실제 인터뷰 기회가 적어 기술 향상 어려움	E2, E3, E4

(1) 인터뷰 질문 구성의 미숙함

참가자 전원(E1–E4)은 학생들이 인터뷰 질문을 구성하는 데 큰 어려움을 겪는다고 지적하였다. 특히 '예/아니오'로 답할 수 있는 닫힌 질문이나 인터뷰 목적과 충분히 연결되지 않은 질문이 자주 나타난다고 언급하였다. E2는 학생들이 "이게 좋으셨나요? 예, 아니오?"와 같은 질문을 작성하는 경우가 많으며, 이는 인터뷰에서 얻을 수 있는 정보의 폭을 제한한다고 설명하였다.

또한 질문 간 맥락 연결성이 부족해 대화 흐름이 끊기는 문제도 지적되었다. E1은 질문들이 서로 독립적으로 나열되어 "앞의 질문과 뒤의 질문이 따로 논다"고 언급하였다. 이는 초보 학습자가 질문을 구조화하고 흐름을 조정하는 데 어려움을 겪고 있음을 보여주며, 이를 지원하는 교육적 보완의 필요성을 시사한다.

(2) 인터뷰 수행 과정에서의 즉각적 대응 부족

참가자 다수(E1, E3, E4)는 학생들이 인터뷰 수행 과정에서 인터뷰이의 답변에 즉각적으로 반응하거나 대화를 확장하는 능력이 부족하다고 지적하였다. E3는 학생들이 "당황해서 질문지를 그대로 읽고 넘어가는 경우가 많다"고 설명하였으며, E1 역시 학생들이 "표면적인 답변에서 멈추는 경우가 많으며 왜 그렇게 생각했는지와 같은 배경을 꼬리 질문으로 파고드는 과정이 부족하다"고 언급하였다.

또한 인터뷰이가 질문 의도와 다른 방향으로 답하거나 매우 짧게 응답하는 경우, 반대로 개인적 경험을 길게 서술하는 상황에서도 초보 학습자들은 대화를 적절히 조율하는 데 어려움을 보였다. 특히 준비된 질문을 모두 수행해야 한다는 압박으로 인해 응답 내용을 더 깊이 탐색하기보다 다음 질문으로 넘어가는 경향이 있었으며, 이 과정에서 응답자의 경험 맥락을 확장할 기회를 놓치기도 하였다. 이러한 결과는 초보 학습자에게 질문을 전달하는 능력뿐 아니라, 인터뷰 상황에 맞게 대화를 조율하고 꼬리질문으로 확장하는 즉흥적 대응 역량을 반복적으로 훈련할 필요가 있음을 시사한다.

(3) 정서적 교감 및 라포 형성 부족

모든 참가자(E1–E4)는 학생들이 인터뷰 초반에 인터뷰이의 긴장을 완화하거나 편안한 분위기를 형성하는 능력이 부족하다고 언급하였다. 특히 아이스브레이킹 대화가 거의 이루어지지 않는 점이 반복적으로 지적되었다. E4는 학생들이 준비해 온 질문을 진행하는 데 집중한 나머지 인터뷰이가 편안하게 이야기할 수 있는 환경을 충분히 조성하지 못한다고 설명하며, "학생들은 질문만 던지고 답변을 기록하는 데 급급한 경우가 많다"고 언급하였다. 그는 인터뷰에서 중요한 것은 질문을 던지는 것뿐 아니라 "분위기를 만들어서 상대방이 더 많은 이야기를 하도록 이끄는 것"이라고 강조하였다.

또한 공감적 경청이 부족하다는 점도 반복적으로 언급되었다. 인터뷰이는 답변 과정에서 감정이나 태도를 드러내기도 하지만, 초보 인터뷰어는 이를 인식하지 못한 채 질문 진행에만 집중하는 경우가 많다는 것이다. E2는 "상대방이 표정이나 몸짓으로 불편함을 표현하는 경우가 있는데 학생들은 이를 그냥 지나치는 경우가 많다"고 설명하였다. 이러한 정서적 교감의 부족은 인터뷰이의 몰입을 낮추고 결과적으로 풍부한 데이터를 얻기 어렵게 만드는 요인으로 지적되었다. 따라서 인터뷰 훈련 과정에서는 대화 분위기 형성, 공감 표현, 정서적 신호 인식과 같은 상호작용 역량을 함께 다룰 필요가 있음이 확인되었다.

(4) 반복 훈련 기회의 제약 및 교수자의 피드백 한계

참가자 전원(E1–E4)은 현재 교육 과정에서 학생들이 충분한 사용자 인터뷰를 경험하기 어렵다는 점을 공통적으로 지적하였다. E2는 "인터뷰는 많이 해봐야 는다"고 설명하며, 학기 동안 학생들이 실제 인터뷰를 수행하는 기회가 많아야 두세 번 정도에 불과하다는 점을 문제로 언급하였다. 이러한 제한된 경험으로는 인터뷰 수행 역량을 충분히 향상시키기 어렵다는 것이다.

또한 교수자가 학생 개개인의 인터뷰 수행 과정을 직접 관찰하고 즉각적으로 피드백하기 어렵다는 점도 중요한 제약으로 언급되었다. E1은 학생들이 파일럿 인터뷰를 진행하더라도 "모든 인터뷰 과정을 옆에서 지켜보며 피드백을 제공하는 것은 현실적으로 어렵다"고 설명하였다. 그 결과 학생들이 인터뷰 연습을 수행하더라도 구체적인 피드백 없이 지나가는 경우가 많다는 것이다. 이러한 교육 환경의 제약은 반복적인 인터뷰 연습과 체계적인 피드백을 제공할 수 있는 보조적 학습 도구의 필요성을 보여준다.

5. 사용자 인터뷰 훈련을 위한 AI 기반 모의 인터뷰 시스템

앞선 형성적 조사에서 디자인 교육자들은 사용자 인터뷰 교육에서 초보 학습자들이 반복적으로 겪는 어려움을 지적하였다. 특히 (1) 인터뷰 질문을 적절히 구성하는 데 어려움이 있으며, (2) 예상치 못한 응답 상황에서 즉각적으로 대응하거나 꼬리 질문을 이어가는 능력이 부족하고, (3) 인터뷰이와의 정서적 교감이나 라포 형성에 어려움을 느끼며, (4) 인터뷰 수행을 반복적으로 연습하고 피드백을 받을 기회가 제한되어 있다는 문제가 확인되었다. 이러한 어려움은 제한된 수업 시간과 실제 인터뷰 연습 기회의 부족, 그리고 교수자가 학생들의 인터뷰 수행 과정을 세밀하게 관찰하고 피드백하기 어려운 교육 환경과도 관련되어 있었다.
이에 본 연구에서는 형성적 조사에서 확인된 문제와 교육적 요구를 바탕으로 시스템의 디자인 요구사항을 구체화하고, 이를 AI 기반 모의 인터뷰 훈련 시스템 Pretalk의 주요 기능으로 반영하였다. 본 시스템은 사용자 인터뷰를 자동화하거나 실제 조사를 대체하려는 목적이 아니라, 학습자가 실제 인터뷰에 돌입하기 전 자신의 인터뷰 기술을 안전하게 연습할 수 있는 학습 지원 환경을 제공하는 것을 목표로 한다.
시스템은 생성형 AI의 언어 이해 및 응답 생성 능력을 활용하여, 사용자와 AI가 함께 인터뷰 질문을 구성하고, 사용자가 다양한 유형의 가상 인터뷰이와 인터뷰를 시뮬레이션하며, 자신의 대화 진행 결과를 피드백 받을 수 있도록 설계되었다.

5. 1. 주요 특징

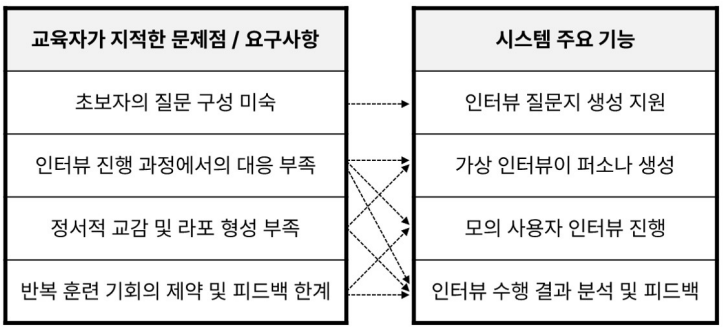

Figure 1 Mapping Between Design Requirements and Core System Functions

본 연구에서 제안하는 AI 기반 모의 인터뷰 시스템은 형성적 조사에서 확인된 주요 문제와 디자인 요구사항을 핵심 기능으로 구체화하여 설계하였다(Figure 1). 구체적으로 질문 구성의 어려움은 질문지 생성 및 수정 기능으로, 즉각적 대응과 정서적 상호작용의 어려움은 가상 인터뷰이와의 모의 인터뷰 및 피드백 기능으로, 반복 훈련의 제약은 수행 결과 분석과 반복 연습 기능으로 연결하였다. (시연 영상: https://player.vimeo.com/video/1126830637).

(1) 인터뷰 질문지 생성 지원
형성적 조사에서 모든 교육자(E1–E4)는 초보 학습자들이 질문을 구성하는 단계에서 가장 큰 어려움을 겪는다고 지적하였다. 특히 '예/아니오' 대답만 가능한 닫힌 질문이나 모호하고 유도적인 질문이 많아, 풍부한 응답을 이끌어내지 못한다는 점이 반복적으로 언급되었다(E2).
이 문제를 보완하기 위해 시스템은 사용자가 인터뷰 주제와 목적을 입력하면 그 맥락에 적합한 질문지를 자동으로 생성한다(Figure 2). 질문지는 사용자가 정한 개수에 기반으로 해서 개방형 질문 위주로 생성된다. 1차적으로 생성된 질문은 사용자가 점검하며 직접 변경할 수 있다.

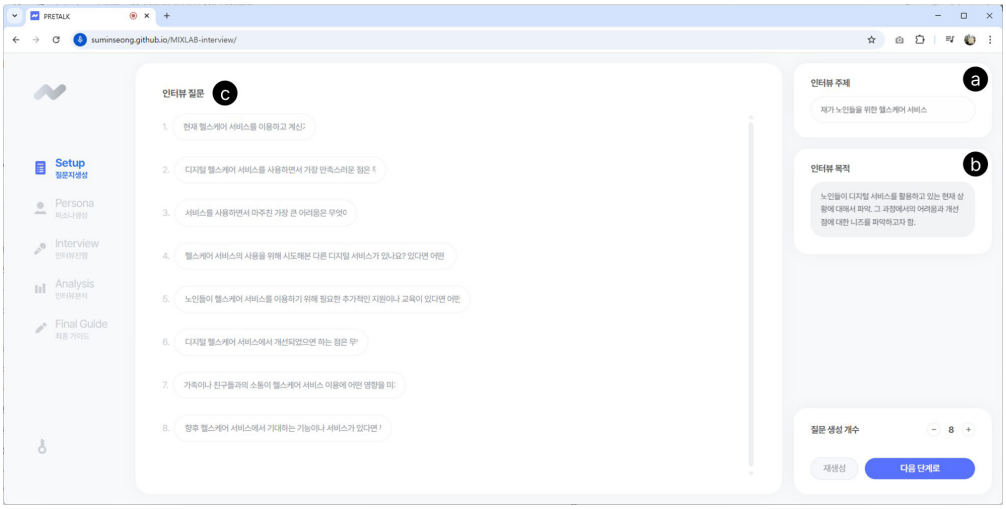

Figure 2 Interview Question Generation Process: (a) Project and Interview Topic Input, (b) Interview Objective Input, (c) Generated Question Set

(2) 가상 인터뷰이 퍼소나 생성

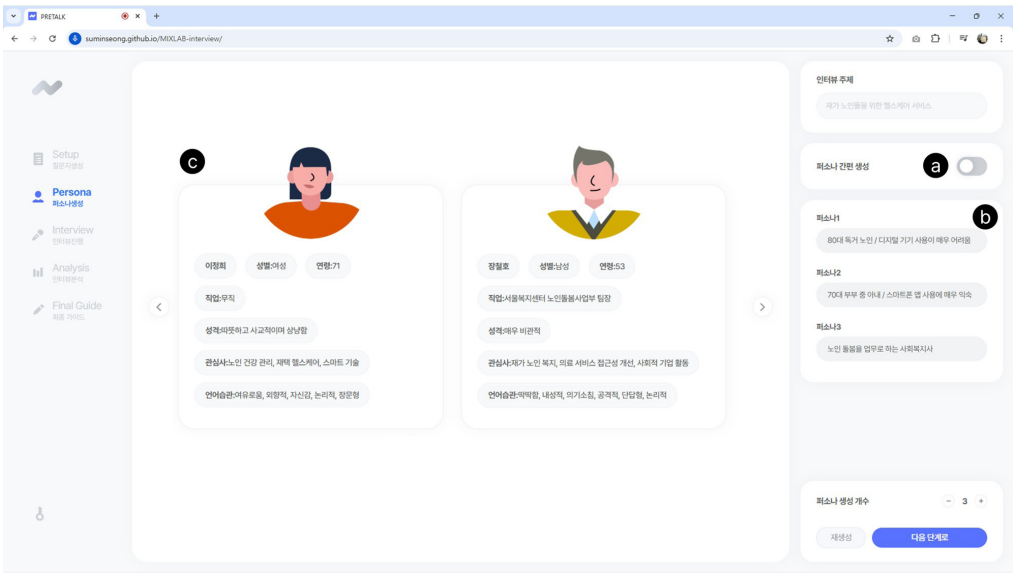

Figure 3 Virtual Interviewee Persona Generation: (a) Automatic Generation Selection, (b) Manual Persona Description Input, (c) Generated Persona

형성적 조사에서는 초보자들이 인터뷰 응답자의 예상치 못한 답변이나 성실하지 않은 반응을 만났을 때 당황하여 대응하지 못하는 문제가 지적되었다(E1, E3). 예컨대 질문 의도와 전혀 다른 방향으로 답하거나, 짧은 한 마디로 대답하는 경우, 학습자들은 곧장 다음 질문으로 넘어가 버리며 대화의 확장 기회를 놓쳤다.

이 문제를 해결하기 위해 시스템은 다양한 가상 인터뷰이 퍼소나를 생성할 수 있도록 지원한다(Figure 3). 자동 생성 모드에서는 시스템이 주어진 인터뷰 주제와 목적에 따라 응답자의 특성을 제안하며, 수동 생성 모드에서는 사용자가 연령, 성별, 직업, 성격, 관심사 등을 직접 묘사하여 특정 유형의 인터뷰이를 설계할 수 있다. 가상의 인터뷰이는 성별, 연령, 직업, 성격, 관심사, 언어 습관 등의 속성으로 구성되어 있어 사용자가 인터뷰 훈련에 적합한 대상을 선택할 수 있다.

예를 들어 노인을 위한 헬스케어 서비스 개발 프로젝트에서 노인의 건강 관리 경험을 탐색하는 인터뷰를

수행한다고 가정할 수 있다(Figure 3). 이 경우 시스템은 '71세 여성, 무직, 건강 관리에 관심이 많으며 외향적이고 논리적인 말투를 사용하는 노인 사용자'와 같은 퍼소나를 생성하여 실제 서비스 이용자의 관점을 반영한 인터뷰 연습을 지원할 수 있다. 또한 동일한 인터뷰 주제에 대해 '53세 남성, 복지센터 노인돌봄사업부 팀장으로 노인 돌봄 현장 경험이 풍부하며 정책과 서비스 운영 관점에서 의견을 제시하는 인터뷰이'와 같은 전문가 퍼소나도 생성할 수 있다. 이를 통해 학습자는 동일한 인터뷰 주제라도 서로 다른 배경과 관점을 가진 응답자와의 대화를 경험해볼 수 있다.

(3) 모의 인터뷰 진행

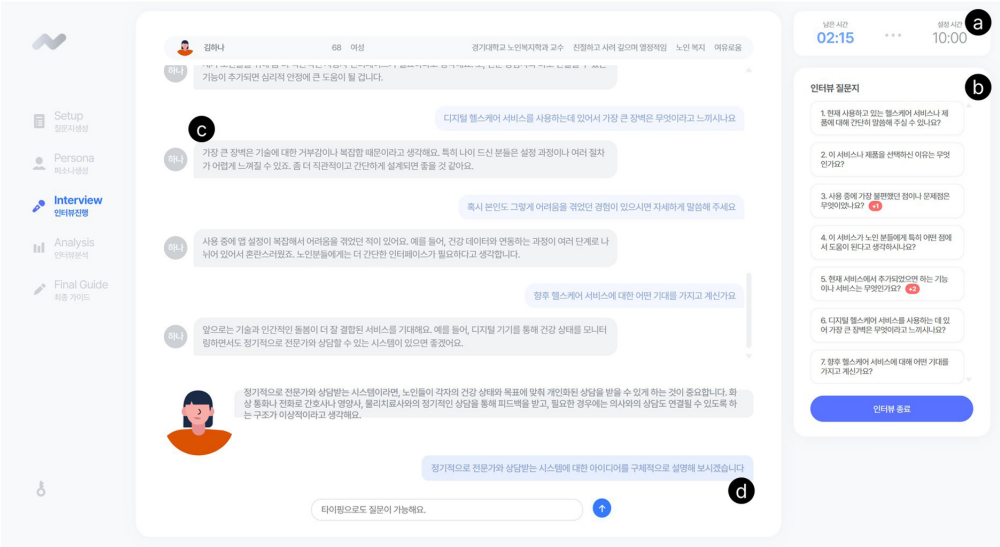

Figure 4 Mock Interview: (a) Interview Duration, (b) Interview Question List and Number of Follow-up Questions, (c) Virtual Interviewee Responses, (d) Interviewer Speech Input

형성적 조사에서는 학생들이 질문에만 집중한 나머지 인터뷰 초반의 긴장 완화나 라포 형성에 실패하고, 인터뷰이와의 정서적 교감을 충분히 이끌어내지 못하는 경우가 많다는 지적이 있었다(E4). 또한 인터뷰이의 비언어적 신호나 감정적 반응을 충분히 고려하지 못해 인터뷰이가 자신의 경험을 충분히 설명하지 못하는 상황도 반복적으로 나타났다(E2). 이러한 문제는 실제 인터뷰와 유사한 대화 환경에서 인터뷰 진행의 흐름과 시간 감각을 경험할 기회가 부족하기 때문으로 해석된다.

이를 보완하기 위해 본 시스템은 음성 기반의 실시간 모의 인터뷰 환경을 제공하도록 설계되었다(Figure 4). 학습자는 마이크를 통해 질문을 음성으로 발화하며, 가상의 인터뷰이는 음성 합성 기술을 활용하여 자연스러운 음성으로 응답한다. 인터뷰어인 학습자와 AI 기반의 인터뷰이는 텍스트 입력이 아닌 음성 대화 형태로 상호작용하며, 이를 통해 실제 인터뷰와 유사한 대화 리듬과 시간 흐름을 경험할 수 있다.

인터뷰는 사전에 준비한 질문지를 기반으로 진행되지만, 학습자는 인터뷰 상황에 따라 질문 순서를 변경하거나 새로운 질문을 즉흥적으로 추가할 수 있다. 시스템은 인터뷰 진행 중 현재 어떤 질문에 대한 대화가 이루어지고 있는지를 인식하고, 그와 관련된 추가 질문이 이루어질 경우 이를 꼬리 질문으로 기록한다. 이러한 질문 수행 과정은 자동으로 로그 데이터로 저장되며, 인터뷰 과정에서 발생한 꼬리 질문의 횟수와 위치는 이후 분석 단계에서 인터뷰 수행 지표로 활용된다.

또한 가상 인터뷰이는 앞서 생성된 퍼소나의 특성을 기반으로 서로 다른 대화 스타일과 응답 태도를 보이도록 설계되어 있다. 예를 들어 외향적인 성격의 인터뷰이는 비교적 길고 적극적인 답변을 제공하는 반면, 간결한 성향의 인터뷰이는 짧고 핵심적인 응답을 제공할 수 있다. 학습자는 이러한 다양한 응답 스타일에 맞추어 대화를 이어가며, 제한된 인터뷰 시간 안에서 질문의 타이밍, 응답 대기, 반응 속도와 같은 실제 인터뷰 수행 요소를 체험적으로 경험할 수 있다.

(4) 인터뷰 수행 결과 분석 및 피드백 제공

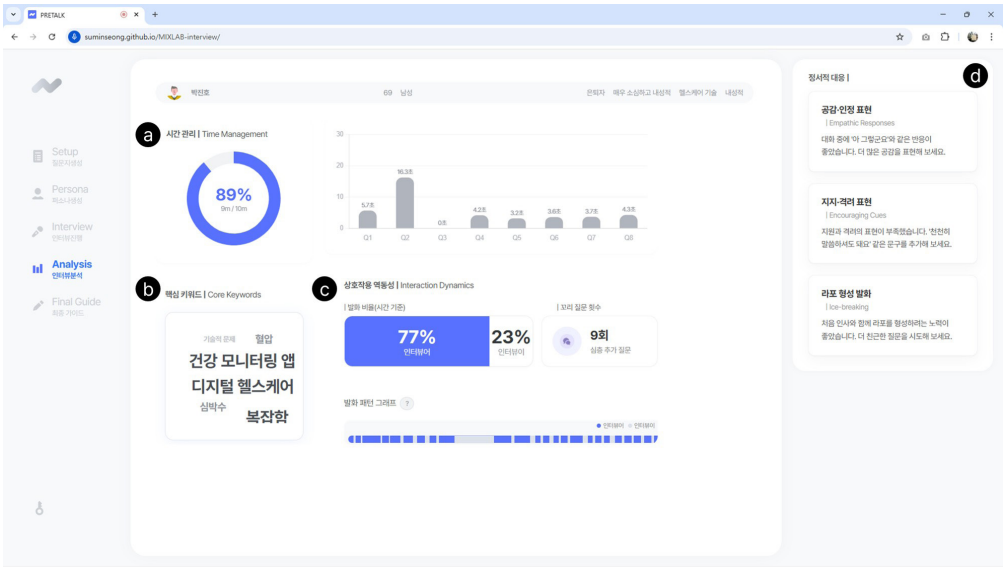

Figure 5 Interview Performance Analysis: (a) Time Management, (b) Key Topics, (c) Interaction Dynamics, (d) Emotional Responses

형성적 조사에서 교육자들은 학생들이 인터뷰를 수행한 이후에도 자신의 수행 방식을 객관적으로 점검할 기회가 부족하다는 점을 지적하였다(E1, E2). 실제 수업에서는 학생들이 파일럿 인터뷰를 진행하더라도 교수자가 모든 인터뷰 과정을 직접 관찰하고 세밀한 피드백을 제공하기 어렵기 때문이다. 이러한 한계를 보완하기 위해 본 시스템은 모의 인터뷰 종료 후 학습자의 인터뷰 수행 과정을 자동으로 분석하고 시각화된 피드백을 제공하도록 설계되었다(Figure 5, Table 5).

분석 화면은 크게 시간 관리, 인터뷰 핵심 키워드, 상호작용 역동성, 그리고 정서적 대응에 대한 피드백 영역으로 구성된다. 먼저 시간 관리 영역에서는 인터뷰 목표 시간 대비 실제 소요 시간을 도넛 차트로 제시하고, 질문별 응답 시간을 막대 그래프로 시각화하여 질문 분배의 균형을 확인할 수 있도록 하였다.

또한 인터뷰 대화 과정에서 등장한 주요 단어를 추출하여 인터뷰 핵심 키워드로 제시함으로써, 학습자가 인터뷰 내용의 주요 흐름을 빠르게 파악할 수 있도록 하였다.

상호작용 역동성 영역에서는 인터뷰어와 인터뷰이의 발화 비율을 수치화하여 제시한다. 인터뷰어와 인터뷰이의 발화 비율을 구분하여 보여주며, 발화 패턴 그래프를 통해 전체 인터뷰 시간 동안의 발화 교대 흐름을 확인할 수 있도록 하였다. 또한 인터뷰 과정에서 특정 질문 이후 추가적으로 수행된 꼬리 질문의 횟수를 함께 표시하여, 학습자가 자신의 질문 확장 전략을 점검할 수 있도록 하였다.

마지막으로 시스템은 인터뷰 대화 내용을 분석하여 공감 표현, 격려적 언어 사용, 라포 형성 시도 여부에 대한 요약 피드백을 제공한다. 예를 들어 '그랬겠군요', '저도 그렇게 생각합니다'와 같은 공감 표현, '힘드셨겠네요'와 같은 지지·격려 표현, 그리고 인터뷰 중 가벼운 인사나 반응을 통한 라포 형성 시도를 확인하여 제시한다. 다만 이러한 정서적 대응 표현은 대화의 맥락에 따라 의미가 달라질 수 있기 때문에 단순한 정량 점수로 평가하지 않고, 학습자가 자신의 대화 방식을 참고할 수 있는 보조적 피드백 형태로 제공된다.

Table 5 Analytical Metrics for Interview Performance

분석 영역	세부 지표	설명
시간 관리	목표 시간 대비 실제 소요 시간 비율	전체 인터뷰에 걸린 시간 / 목표 소요 시간
	질문별 소요 시간 분포	질문당
인터뷰 내용	핵심 키워드	대화에서 반복적으로 등장한 주요 키워드
상호작용 역동성	말하기 vs 경청 비율	인터뷰어와 인터뷰이 발화의 비율(문자 수 기반)
	발화의 주고받기 패턴	인터뷰이와 인터뷰어의 발화 교대 패턴 시각화
	꼬리 질문 횟수	특정 질문 후에 심층적으로 추가 질문한 횟수
정서적 대응	정서적 대응 표현	공감·인정(예. 그랬겠군요, 저도 그렇게 생각합니다), 지지·격려 표현(예. 힘드셨겠네요), 라포 형성 표현(예. 가벼운 인사)의 수행 수준에 대한 피드백

(5) 질문지 및 진행 전략 개선

시스템의 마지막 단계에서는 학습자가 앞선 인터뷰 수행 분석 결과를 참고하여 자신의 질문지를 점검하고 필요한 부분을 수정할 수 있도록 한다. 학습자는 기존 질문을 삭제하거나 수정할 수 있으며, 인터뷰 진행 과정에서 정량적 피드백과 주의사항을 함께 확인할 수 있다.

5. 2. 시스템 설계 및 구현

본 연구에서 제안하는 AI 기반 모의 인터뷰 시스템은 생성형 AI 모델을 기반으로 인터뷰 질문 생성, 가상 인터뷰이 퍼소나 설계, 실시간 인터뷰 응답 생성, 인터뷰 수행 결과 분석의 네 가지 핵심 모듈로 구성된다. 각 모듈은 인터뷰 과정에서 생성되는 질문 정보, 퍼소나 설정 값, 대화 로그 데이터를 공유하는 구조로 구현되어 있으며, 이러한 데이터는 인터뷰 진행과 이후 수행 분석 단계에서 함께 활용된다. 시스템은 인터뷰 세션 단위로 질문 목록, 인터뷰 대화 기록, 발화 시간, 수행 지표 등의 로그 데이터를 저장하도록 설계되어 학습자의 인터뷰 수행 과정을 분석할 수 있도록 한다.

(1) 질문 생성 모듈

사용자가 인터뷰의 주제와 목적, 인터뷰 대상에 대한 정보를 입력하면, 경량 생성형 모델인 GPT-4o mini가 이를 기반으로 인터뷰 질문 목록을 자동으로 생성한다. 생성된 질문은 인터뷰 목적에 맞도록 개방형 질문 중심으로 구성되며, 경험 탐색, 감정 탐색, 행동 맥락 이해와 같은 인터뷰 목적에 따라 질문 유형이 조정된다.
생성된 질문은 인터뷰 세션의 기본 질문지로 활용되며, 학습자는 이를 수정하거나 인터뷰 진행 과정에서 새로운 질문을 추가할 수 있다. 이러한 질문 정보는 인터뷰 세션 데이터로 저장되어 이후 인터뷰 진행 및 수행 분석 단계에서 함께 활용된다.

(2) 가상 인터뷰이 퍼소나 생성 모듈

사용자가 입력한 인터뷰 주제, 목적, 대상 정보를 기반으로 GPT-4o 모델이 인터뷰이 퍼소나를 생성한다. 퍼소나는 연령, 성별, 직업, 관심사, 성격, 언어 스타일과 같은 속성으로 구성되며, 이러한 속성 정보는 인터뷰 세션 데이터로 저장된다. 생성된 퍼소나는 인터뷰 응답 생성 단계에서 인터뷰이의 발화 스타일과 응답 태도를 결정하는 매개변수로 활용된다.
학습자는 자동 생성된 퍼소나를 선택하거나 일부 속성을 직접 수정할 수 있으며, 수정된 퍼소나 정보 역시 세션 데이터에 반영된다. 이렇게 생성된 퍼소나 속성은 이후 인터뷰 응답 생성 과정에서 모델 프롬프트에 포함되어 서로 다른 배경과 성향을 가진 인터뷰이의 응답을 생성하도록 사용된다.

(3) 인터뷰 응답 생성 모듈

AI가 담당하는 가상 인터뷰이의 응답은 파인튜닝된 Chammkae_v3b 모델을 통해 생성된다. 이 모델은 GPT-4o 기반의 대화 데이터를 추가 학습하여 자연스러운 대화 흐름과 감정 표현을 생성하도록 설계되었다. 인터뷰 응답 생성 시에는 앞 단계에서 생성된 퍼소나 정보가 프롬프트에 함께 전달되며, 퍼소나의 성향 매개변수는 응답 길이, 감정 표현, 문체 유형 등에 반영된다.

시스템은 음성 기반 인터랙션 구조를 사용한다. 인터뷰어인 학습자는 마이크를 통해 질문을 발화하고, 가상 인터뷰이는 음성 합성 기술을 통해 음성으로 응답한다. 음성 입력은 실시간으로 텍스트로 변환되어 대화 로그로 저장되며, 이후 응답 생성 모델의 입력으로 사용된다.
또한 시스템은 인터뷰 진행 중 현재 어떤 질문에 대한 대화가 이루어지고 있는지를 지속적으로 추적한다. 학습자가 기존 질문과 의미적으로 연관된 추가 질문을 제시할 경우 이를 꼬리 질문으로 인식하고 해당 질문을 현재 질문과 연결하여 기록한다. 이러한 정보는 인터뷰 수행 분석 단계에서 꼬리 질문 횟수를 계산하는 데 활용된다. 동시에 AI 인터뷰이는 이러한 질문 맥락을 반영하여 후속 질문에 맞는 응답을 생성하도록 설계되어 있다.

(4) 인터뷰 수행 결과 분석 모듈
인터뷰 수행 과정에서 생성된 로그 데이터는 학습자의 인터뷰 수행 특성을 분석하는 데 활용된다. 로그 데이터에는 인터뷰 질문 목록, 가상 인터뷰이의 퍼소나 정보, 인터뷰 진행 과정에서 기록된 발화 시간과 대화 내용이 포함된다. 이러한 데이터는 인터뷰 세션 단위로 저장되며 이후 수행 분석 단계에서 활용된다.
시스템은 사전에 설정한 목표 인터뷰 시간과 실제 인터뷰 소요 시간을 비교하고, 질문별 소요 시간을 분석하여 특정 질문에 과도한 시간이 사용되었거나 충분한 탐색 없이 넘어간 구간을 확인할 수 있도록 한다. 또한 인터뷰어와 인터뷰이의 발화 비율과 발화 교대 패턴을 계산하여 대화 상호작용의 균형을 분석한다. 발화 비율은 인터뷰 전사 텍스트에서 각 화자의 발화 문자 수를 기준으로 산출된다. 또한 인터뷰 진행 중 추가된 질문이 기존 질문과 의미적으로 연관될 경우 이를 꼬리 질문으로 인식하여 횟수를 집계한다. 시스템은 또한 인터뷰 대화 내용 로그에서 공감 표현이나 지지적 언어 사용 여부와 같은 정서적 대응 요소를 탐지하여 분석 지표로 기록한다.
본 시스템에서 산출되는 수행 지표는 학습자의 인터뷰 역량을 객관적으로 진단하기 위한 검증된 측정 척도가 아니라, 모의 인터뷰 과정에서 나타난 수행 특성을 가시화하여 학습자의 자기점검을 지원하기 위한 참고 지표로 활용하였다.

6. 사용자 평가

6. 1. 목적

본 사용자 평가는 AI 기반 인터뷰 훈련 시스템이 사용자 인터뷰 경험이 부족한 초보 학습자들의 인터뷰 수행 및 준비 과정에 미치는 영향을 탐색하기 위해 수행되었다. 특히 인터뷰 진행 과정에서의 시간 관리, 질문 구성, 인터뷰 진행 태도와 관련하여, 시스템의 분석 및 피드백을 경험한 이후 학습자의 인터뷰 수행에서 어떠한 변화가 나타나는지를 살펴보고자 하였다. 또한 단일 사용이 아닌 반복적 훈련을 진행하도록 하여, Pretalk을 두 차례 사용하는 동안 학습자의 수행 변화가 어떻게 나타나는지를 관찰하고, AI 기반 인터뷰 훈련 시스템의 활용 가능성과 그 과정에서 나타나는 변화의 양상을 탐색하고자 하였다.

6. 2. 참가자

참가자는 UX 디자인 전공 3학년 4명과 4학년 이상 2명, 총 6명의 학부생으로 구성되었으며, 평균 연령은 22.2세였다. 이들은 모두 수업이나 팀 프로젝트를 통해 사용자 인터뷰를 수행한 경험은 있으나, 체계적인 인터뷰 훈련을 충분히 받지 않은 초보 인터뷰어 수준으로 분류된다.
사용자 인터뷰 진행 수준에 대한 자가 평가 결과는 5점 척도 기준 평균 2.67점(SD=0.94)으로, 전반적으로 중간 이하 수준의 자신감을 보였다. 특히 다수의 참가자가 인터뷰 진행 시 ‘대화가 끊기거나 어색해질 것에 대한 불안’과 ‘적절한 질문을 설계하지 못할 것에 대한 우려’를 가장 큰 염려 요인으로 응답하였다. 일부는 시간 관리 및 결과 분석 활용에 대한 어려움도 언급하였다.
참가자들은 각기 다른 디자인 프로젝트를 진행 중이었으며, 그 주제 또한 다양하였다(Table 6). 필요로 하는

인터뷰 대상 또한 웹 개발자, 음악 청취자, 대학생 사용자, 디자이너, 20대 여성 소비자, 정성 조사 수행자 등으로 상이하였다. 인터뷰 목적 역시 신규 기능 검증, 아이디어 타당성 확인, 사용자 니즈 및 행동 이해 등으로 분포되어 있었다.

이처럼 본 연구의 참가자 집단은 인터뷰 경험은 있으나 수행 안정성이 낮고, 다양한 주제와 목적의 실제 프로젝트를 병행하고 있는 집단으로서, Pretalk의 인터뷰 학습 지원의 활용 가능성과 수행 변화를 탐색하기에 적합한 조건을 갖추고 있다. 또한 Pretalk을 활용한 연습 결과가 곧바로 실제 사용자 조사 활동과 연계될 수 있는 실천적 맥락을 지닌 집단이라는 점에서 연구의 적용 가능성을 높인다.

Table 6 Participants in the User Evaluation

참가자	전공 학년 (나이)	진행 중 프로젝트	예상 인터뷰 참가자와 사용자 인터뷰 목적	사용자 인터뷰 진행 수준 자가 평가
P1	UX 전공 3학년 (22)	기능이 확장된 실시간 몰입형 번역기	해외 웹사이트를 자주 이용하는 30대 중반 웹 개발자를 대상으로 신규 기능 검증 및 경쟁사 대비 차별점 도출	3
P2	UX 전공 3학년 (23)	음악 청취 사용자의 궁극적 가치 탐색 연구	다양한 장르의 음악 청취자를 대상으로 청취 동기와 행동 이해	3
P3	UX 전공 4학년 이상 (22)	맥락 맞춤형 생성형 AI 이모티콘 서비스	모바일 메신저를 자주 사용하는 20대 대학생을 대상으로 기존 서비스의 페인포인트 파악	2
P4	UX 전공 3학년 (23)	AI 기반 반응형 텍스트 레이블 자동 생성 기능	반응형 디자인 경험이 있는 디자이너를 대상으로 사용자 니즈 및 기능 수요 검증	1
P5	UX 전공 3학년 (21)	디지털 선물의 '성의 없음' 해결 서비스	디지털 선물 문화에 익숙한 20대 여성을 대상으로 아이디어 타당성 검증	3
P6	UX 전공 4학년 이상 (22)	생성형 AI 기반 사용자 조사 도구	최근 1년 내 정성 조사를 수행한 디자이너 대상으로 조사 경험 확인	4

6. 3. 진행 과정

실험은 1차와 2차 세션으로 구성되어 총 약 110분간 진행되었으며, 참가자는 동일한 프로젝트 주제를 기반으로 두 차례의 Pretalk 활용 훈련을 수행하였다(Table 7).

먼저 연구진은 시스템의 주요 기능과 사용 방법을 간략히 설명한 뒤, 참가자가 인터페이스에 익숙해질 수 있도록 적응 시간을 제공하였다. 이후 참가자는 자신의 디자인 프로젝트를 바탕으로 인터뷰 목적, 대상, 질문 구성을 포함한 인터뷰 계획서를 직접 작성하였다. 작성된 계획을 기반으로 1차 Pretalk 세션을 수행하였으며, 인터뷰 종료 후 시스템이 제공한 수행 분석 결과를 확인하였다.

참가자는 분석 결과를 참고하여 인터뷰 계획을 수정한 뒤, 동일한 주제에 대해 2차 Pretalk 세션을 다시 진행하였다. 이를 통해 반복 훈련 과정에서의 수행 변화를 직접 경험하도록 하였다.

2차 세션 종료 후, 참가자들은 Pretalk 사용 경험에 대한 설문에 응답하였다. 설문은 5점 Likert 척도(1=전혀 그렇지 않다, 5=매우 그렇다)로 구성되었으며, 주로 다음 다섯 항목을 포함하였다:

(1) 자동 질문 생성 기능은 인터뷰 질문을 구성하는 데 도움이 되었는가,

(2) 가상 인터뷰이 퍼소나 기능은 다양한 인터뷰 상황을 연습하는 데 도움이 되었는가,

(3) 음성 기반 모의 인터뷰 방식은 인터뷰 진행 훈련에 도움이 되었는가,

(4) 시스템이 제공한 분석 지표는 자신의 인터뷰 수행을 점검하는 데 도움이 되었는가,

(5) Pretalk의 전체 훈련 과정은 인터뷰 역량 향상에 도움이 되었는가.

설문 이후에는 연구자와의 개별 대면 사후 인터뷰를 진행하였다. 사후 인터뷰는 반구조화된 형식으로 이루어졌으며, 다음 세 가지 질문을 중심으로 구성되었다:

(1) Pretalk의 인터뷰 훈련 기능 중 가장 도움이 되었던 요소는 무엇인가,

(2) 1차에서 2차 훈련을 거치며 자신의 인터뷰 태도나 질문 구성에서 어떤 변화를 의도하였는가,

(3) 현재 시스템에서 개선이 필요하다고 생각되는 점은 무엇인가.

이러한 절차를 통해 반복 훈련 과정에서의 수행 변화와 학습자의 인식 변화를 함께 수집하였다.

Table 7 Procedure of the User Evaluation Study

단계		주요 내용	소요 시간
1	시스템 적응	연구진이 시스템 사용 방법을 설명하고, 참가자가 시스템에 익숙해지는 과정	20분
2	인터뷰 계획 작성	참가자가 자신의 프로젝트를 바탕으로 인터뷰 질문 포함한 진행 계획을 문서로 작성(인터뷰 목적, 참가자 구성, 인터뷰 질문 및 진행, 주의사항)	15분
3	Pretalk 1차 훈련	작성한 질문을 바탕으로 가상 인터뷰이와 음성 기반 인터뷰를 수행.	15분
4	인터뷰 계획 수정	1차 결과에 대한 AI 피드백을 확인한 후, 학습자는 인터뷰 전략을 필요에 따라 변경.	15분
5	Pretalk 2차 훈련	수정된 질문으로 동일하거나 변형된 가상 인터뷰이와 두 번째 인터뷰를 수행.	15분
6	설문 및 사후 인터뷰	설문을 통해서 주요 기능이 훈련에 도움 되었는지 평가. 두 차례 훈련 경험을 비교하며, 인터뷰 훈련에의 효과와 개선 필요 사항에 대하여 참가자와 반구조화 인터뷰 진행	30분

6. 4. 측정 항목 및 분석 방법

본 연구는 Pretalk 시스템의 반복 사용이 초보 학습자의 인터뷰 수행에 미치는 단기적 변화를 탐색하기 위해 정량적 수행 로그 분석을 중심으로 설문 분석과 사후 인터뷰에 대한 정성 분석을 병행하였다. 분석은 동일 참가자의 1차와 2차 세션 간 변화를 비교하는 방식으로 진행하였다.

정량 분석은 시스템이 자동 기록한 인터뷰 로그 데이터를 기반으로 수행되었다. 로그 데이터에는 전체 인터뷰 질문과 응답 내역, 질문별 소요 시간, 인터뷰어–인터뷰이 발화 비율, 발화 교대 패턴, 꼬리 질문 횟수, 정서적 대응 표현 등이 포함되며, 이를 통해 질문 탐색성, 대화 흐름 조율, 상호작용 균형 등의 수행 변화를 확인하였다. 본 연구는 탐색적 성격을 지니므로 통계적 유의성 검증보다는 반복 훈련 전후 수행 패턴의 방향성 변화에 초점을 두었다.

설문과 사후 인터뷰는 정량 로그 분석 결과를 보완하기 위한 자료로 활용되었다. 설문 응답은 기술통계 방식으로 정리하여 시스템 기능에 대한 인식을 파악하였으며, 사후 반구조화 인터뷰는 개방 코딩을 통해 주요 인식을 도출하였다. 이를 통해 로그 데이터에서 확인된 수행 변화가 어떤 학습 경험이나 전략 조정과 연결되는지를 해석하였다.

7. 사용자 평가 결과

7. 1. 전반적 인식 및 기능 평가

Table 8 Evaluation Results of Pretalk Core Functions

평가 항목	표준편차	평균
Pretalk 전반의 훈련 경험	4.67	0.52
자동 질문 생성 기능	4.0	1.26
가상 인터뷰이 퍼소나 생성	4.0	0.63
가상 인터뷰이와의 음성 기반 인터뷰 방식	4.5	0.84
인터뷰 훈련 후 피드백 제공	4.0	1.1

설문 결과, 참가자들은 Pretalk의 전반적인 훈련 흐름에 대해 긍정적으로 인식하였다. ‘Pretalk의 전체 훈련 흐름이 실제 인터뷰 역량 향상에 도움이 되었다’는 문항은 평균 4.67점(SD=0.52)으로 나타났으며, 6명 중 4명이 5점을 부여하였다. 이는 시스템의 단계적 구조가 인터뷰 준비, 수행, 성찰을 연결하는 학습 과정으로 인식되었음을 보여준다.

자동 질문 생성 기능은 평균 4.00점(SD=1.26)으로 나타났다. 참가자들은 질문 설계 초기 단계에서 느끼는 막연함을 줄이고 질문 구성을 빠르게 시작할 수 있다는 점을 긍정적으로 평가하였다. 예를 들어 P4는 “아무리 뇌로는 알고 있어도 질문을 생성하기 막막했는데 시간 절약을 해주고 질문도 적절했다”고 언급하였으며, P5 또한 “목적에 맞는 질문을 구성하는 데 어려움을 느꼈던 경험이 있었는데 그 막연함을 해결해 주는 점이 좋았다”고 말했다.

가상 인터뷰이 퍼소나 생성 기능은 평균 4.00점(SD=0.63)으로 비교적 안정적인 평가를 보였다. 참가자들은 퍼소나 정보를 통해 인터뷰 대상을 보다 구체적으로 상상할 수 있었다고 응답하였다. 예를 들어 P6는 "인터뷰 참가자를 미리 상상해보고 인터뷰 훈련에 몰입하는 데 도움이 되었다"고 말했으며, P3 또한 "생성된 퍼소나의 연령, 직업, 성격, 관심사 등이 다양하여 인터뷰 상황을 떠올리는 데 도움이 되었다"고 언급하였다.

가상 인터뷰이와의 음성 기반 인터뷰 방식은 평균 4.50점으로 기능별 평가 중 가장 높은 점수를 기록하였다. 참가자들은 음성 기반 상호작용이 실제 인터뷰와 유사한 긴장감과 몰입을 유도했다고 응답하였다. 예를 들어 P2는 "말하는 것과 타이핑은 큰 차이가 있는데, 이 시스템은 인터뷰를 말로 진행하면서 실제 상황과 유사한 경험을 하게 한다"고 언급하였으며, P5는 "직접 말을 하면서 진행하니 온라인 환경에서도 실제 인터뷰처럼 매끄럽게 진행해야 한다는 긴장감이 생겼다"고 말했다. P6 역시 "상대의 응답에 따라 꼬리 질문을 즉석에서 생각해야 해서 현장감 있는 커뮤니케이션 훈련이 되었다"고 설명하였다.

분석 대시보드 기능은 평균 4.00점(SD=1.10)으로 나타났으며, 특히 정서적 대응과 상호작용 패턴에 대한 피드백이 유용하게 인식되었다. 예를 들어 P6는 "정서적 응대 부분에서 개선이 필요한 지점을 구체적으로 피드백 받을 수 있어서 유용했다"고 언급하였다. P3는 "정서적 대응 피드백을 받은 이후 인터뷰 훈련에서 해당 요소를 더 의식하게 되었다"고 설명하였다. 또한 P1은 "두 번째 인터뷰에서는 시간 관리나 공감 표현 등 이전에 받은 피드백을 의식적으로 반영하려 했다"고 말했다. 이러한 응답은 분석 대시보드가 단순한 결과 제시에 그치지 않고 학습자가 자신의 인터뷰 수행 방식을 성찰하는 계기로 작용했음을 보여준다.

7. 2. 반복 훈련에서 나타난 수행 변화와 전략적 조정

Table 9 Changes in Interview Performance Metrics Between the 1st and 2nd Training Sessions

분석 영역	지표	1차 평균 (SD)	2차 평균 (SD)
시간 관리	전체 예상 시간 대비 소요 시간 (%)	82.5 (14.3)	93.8 (6.9)
상호작용 구조	인터뷰어 발화 비율 (%) (문자 수 기준)	48.6 (9.7)	41.3 (6.2)
상호작용 구조	발화 교대 빈도 (회)	43.8 (9.1)	39.5 (6.4)
상호작용 구조	꼬리질문 횟수 (회)	9.3 (3.8)	11.8 (3.4)
정서적 대응	공감·인정 표현 횟수	0.7 (0.5)	2.8 (1.2)
정서적 대응	지지·격려 표현 횟수	0.3 (0.5)	1.7 (0.8)
정서적 대응	라포 형성 발화 (회)	2.8 (0.7)	3.5 (0.8)

7. 2. 1. 시간 관리 변화

전체 예상 시간 대비 실제 수행 시간은 1차 평균 82.5%(SD=14.3)에서 2차 93.8%(SD=6.9)로 증가하여 2차 세션에서 인터뷰 시간이 계획된 시간에 보다 근접하는 경향이 나타났다. 이러한 변화는 사후 인터뷰에서 참가자들이 시간 배분을 보다 의식하게 되었다고 응답한 내용과도 연결된다.

P4의 경우 1차 인터뷰가 8분 만에 종료되어 전체 예상 시간(15분)의 53%에 그쳤다. 그는 1차 훈련 후 "답변이 예상 범위를 크게 벗어나지 못했고, 더 깊이 파고들지 못했다"고 반성하였다. 2차에서는 인터뷰 시간이 11분으로 증가하여 계획 시간에 보다 근접하게 운영되었다. 이처럼 인터뷰 흐름을 보다 안정적으로 유지하려는 전략 조정과 수행 시간의 변화가 함께 나타났다.

P1은 1차에서 예상 시간 대비 다소 짧게 인터뷰를 종료하였으나, 2차에서는 계획 시간에 보다 근접하게 인터뷰를 마쳤다. 그는 1차 후 "시간 관리가 안 되는 것을 보아 인터뷰 구조를 다시 점검해야겠다"고 언급하였다. 2차에서는 질문 흐름이 중단 없이 이어지며 인터뷰가 조기 종료되지 않았다. 이는 사후 인터뷰에서 P1이 "실제로 2번째 인터뷰에선 피드백 받았던 부분들(시간 관리, 정서적 공감)을 개선하여, 인터뷰 훈련에 도움이 되었다"고 언급한 내용과도 일치한다.

결과적으로 2차 세션에서는 계획된 인터뷰 시간에 보다 근접하는 변화가 관찰되었는데, 이는 인터뷰 수행이 내용적 확장뿐 아니라 시간 구조 관리라는 측면에서도 변화 가능성이 있음을 시사한다.

7. 2. 2. 상호작용 구조 변화

상호작용 구조는 인터뷰어 발화 비율, 발화 교대 빈도, 그리고 꼬리질문 횟수를 중심으로 분석하였다. 반복 훈련 이후 가장 일관되게 나타난 변화는 인터뷰어 발화 비율 감소와 꼬리질문 증가였다.

먼저 문자 수 기준 인터뷰어 발화 비율은 1차 평균 48.6%(SD=9.7)에서 2차 41.3%(SD=6.2)로 감소하였다. 이는 인터뷰어가 자신의 설명이나 의견을 덧붙이는 비중을 줄이고 인터뷰이의 발화를 더 많이 확보하는 방향으로 변화했음을 보여준다. 예를 들어 P5의 경우 1차 인터뷰에서는 인터뷰이의 응답 이후 “저는 직접 선물이 더 감동적일 것이라고 생각했는데…”와 같이 자신의 의견을 먼저 제시한 뒤 질문을 이어가는 발화가 나타났다. 그러나 2차 인터뷰에서는 이러한 의견 삽입 발화가 줄어들고 인터뷰이의 응답을 기반으로 질문을 이어가는 방식이 더 자주 나타났다. P4의 경우에도 1차 인터뷰에서는 “동의하십니까?”와 같은 단정적 질문이 나타났으나, 2차에서는 “그렇다면 실제 상황에서는 어떻게 나타났나요?”와 같이 응답을 구체화하는 형태의 질문이 사용되었다. 이러한 변화는 인터뷰어 발화 비율 감소가 단순한 발화량 감소라기보다 질문 방식의 변화와 함께 나타났음을 보여준다.

발화 교대 빈도는 1차 평균 43.8회(SD=9.1)에서 2차 39.5회(SD=6.4)로 소폭 감소하였다. 이는 전체 발화 수 감소의 영향으로 볼 수 있으며, 인터뷰어와 인터뷰이 간 발화 교대 구조 자체에는 큰 변화가 나타나지 않았다.

가장 뚜렷한 변화는 꼬리질문 횟수에서 나타났다. 1차 평균 9.3회(SD=3.8)이었던 꼬리질문은 2차 11.8회(SD=3.4)로 증가하였다. 참가자별로는 P1(6회 → 12회), P3(8회 → 13회), P5(9회 → 12회)에서 증가가 확인되었다. P1의 경우에는 가상 인터뷰이가 “정확성이 가장 중요하다”고 응답하자, “그렇다면 정확성을 판단하는 기준은 무엇이라고 보시나요?”와 같이 응답의 개념을 확장하는 질문을 하기도 하였다.

이와 같이 반복 훈련 이후 인터뷰어 발화 비율은 감소하고 꼬리질문 횟수는 증가하는 경향이 나타났다. 이는 인터뷰어의 발화량 자체가 줄어든 것이라기보다, 질문의 기능이 설명이나 의견 제시에서 응답 탐색과 확장 중심으로 이동했음을 보여준다.

7. 2. 3. 정서적 대응 변화

정서적 대응은 반복 훈련에서 가장 뚜렷한 수치 변화가 나타난 영역이었다. 공감·인정 표현은 1차 평균 0.7회(SD=0.5)에서 2차 2.8회(SD=1.2)로 증가하였으며, 지지·격려 표현은 0.3회에서 1.7회로 증가하였다. 라포 형성 발화 또한 2.8회에서 3.5회로 소폭 증가하였다. 사후 인터뷰에서도 참가자들은 정서적 대응이 의식하지 않으면 간과되기 쉬우나, 피드백을 통해 보다 의식적으로 조절하게 되었다고 언급하였다.

특히 공감·인정 표현 증가가 두드러졌다. P2의 경우 1차 인터뷰에서는 공감 표현이 나타나지 않았으나 2차에서는 4회로 증가하였다. 그는 1차 훈련 후 “사용자를 좀 더 격려하는 태도로 진행해야겠다”고 언급하였으며, 실제 2차 인터뷰에서는 “그럴 수 있겠네요”, “말씀해주셔서 감사합니다”와 같은 인정 발화가 추가되었다.

P5 또한 공감·인정 표현이 1차 1회에서 2차 5회로 증가하였다. 그는 1차 후 인터뷰 중간에 인터뷰이를 격려하며 답변을 이끌어야 한다는 점을 새롭게 인식했다고 기술하였고, 2차에서는 “새로운 관점을 보게 되었다”, “그런 관점은 생각해보지 못했다”와 같은 인정 표현이 나타났다.

지지·격려 표현과 라포 형성에서도 일부 변화가 관찰되었다. P2와 P5는 2차 인터뷰에서 인터뷰이의 답변을 긍정적으로 인정하는 표현을 사용하였으며, P5는 인터뷰 초반 날씨를 언급하며 분위기를 완화하는 발화를 사용하였다.

7. 3. 실제 활용 의견과 개선 요구 사항

7. 3. 1. 실제 활용 맥락에 대한 의견

참가자들은 Pretalk을 실제 사용자 인터뷰를 수행하기 전 사전 연습 또는 점검 도구로 활용할 수 있을 것이라고 공통적으로 언급하였다. 특히 인터뷰 질문 구성, 진행 흐름, 시간 배분 등을 미리 확인하고 조정하는 용도로 유용하다는 의견이 나타났다. 예를 들어 P1은 “사용자 인터뷰를 하기 전 실습으로 충분히 가능할 것 같습니다”라고 언급하였으며, P3 역시 “프로젝트 인터뷰 진행 전 사전점검” 용도로 Pretalk을 활용하여 “질문 구성, 진행 흐름, 시간 배분 등 전반적인 요소를 사전에 점검”할 수 있다고 설명하였다.

또한 일부 참가자는 Pretalk을 인터뷰 질문의 효과를 미리 확인하는 도구로 활용할 수 있다고 보았다. P6는 준비한 질문에 대해 퍼소나가 어떻게 반응하는지를 통해 “미리 꼬리질문을 예측하고 준비”할 수 있다고 언급하였다. P5 또한 “Pretalk으로 내가 이 질문을 했을 때 어떤 답변을 이끌어 낼 수 있는지 점검하는 데 사용하면 좋을 것 같다”고 설명하였다. 이는 Pretalk이 단순한 진행 연습을 넘어, 질문이 유도하는 응답의 방향성을 사전에 탐색하는 시뮬레이션 도구로 활용될 가능성을 보여준다.

7. 3. 2. 개선 요구 사항

(1) 무조건적 질문 생성의 역효과와 질문지 작성 코칭 필요

참가자들은 질문 생성이 유용하다는 평가와 동시에, 질문을 완성본으로 제공하기보다는 사용자의 설계를 유도하는 코칭형 지원이 더 적절하다고 지적하였다. P1은 “인터뷰 훈련이라고 한다면 무조건적인 질문 생성은 역효과일 수도 있습니다”라고 말하며, 사용자가 먼저 “초기 질문을 5개 정도 구상”한 뒤 AI가 “사용자가 보지 못한 관점을 제시”하거나 “질문 수정을 어떤 방향으로 하면 좋을지” 제안하는 방식이 더 낫다고 응답하였다. P5도 “질문지를 전부 생성하다보니 내가 질문을 충분히 이해하고 있지 못해 실제 진행 중 흐름이 어색해”졌다고 언급하며, 사용자가 작성한 인터뷰 질문을 바탕으로 질문지를 개선해주는 방식이 바람직하다고 제안하였다. P3 역시 “사용자의 질문지를 피드백하고 개선”하고, 특히 “질문 순서를 재구성”하거나 “중복되는 질문을 통합하는 피드백”이 있으면 훈련에 더 도움이 될 것이라고 말했다.

(2) 퍼소나의 목적 연결성 및 응답 반영도 부족

퍼소나 생성 기능은 인터뷰 대상을 사전에 구체화하는 데 도움을 주었으나, 훈련 도구로서 활용하기에는 두 가지 한계가 지적되었다.

첫째, 자동 생성된 퍼소나가 인터뷰 목적과 어떤 방식으로 연결되는지에 대한 설명이 부족하다는 점이다. 참가자들은 퍼소나의 연령, 직업, 관심사와 같은 정보가 제시되더라도, 이러한 정보가 인터뷰 수행에서 어떤 의미를 갖는지 이해하기 어렵다고 언급하였다. 예를 들어 P5는 “단순 퍼소나의 정보로는 내 인터뷰 목적에 어떤 방향으로 도움을 줄 수 있는지 파악하기 어려웠다”고 응답하며, 퍼소나가 특정 상황을 잘 아는 전문가 역할인지, 혹은 일반 사용자 경험을 대변하는 참가자인지와 같은 맥락 설명이 필요하다고 제안하였다. P4 또한 “퍼소나가 어떤 문제를 겪고 있는지와 같은 구체적 설명”이 포함된다면 인터뷰 설계와의 연결성이 더 명확해질 것이라고 언급하였다. 이는 단순한 속성 정보 제시를 넘어, 퍼소나가 어떤 경험이나 문제 맥락을 대표하는 인물인지에 대한 설명이 함께 제공될 필요가 있음을 보여준다.

둘째, 설정된 퍼소나의 특성이 실제 인터뷰 응답에 충분히 반영되지 않는다는 한계가 제기되었다. 참가자들은 서로 다른 직업이나 성향을 가진 퍼소나라면 답변의 수준이나 표현 방식에서도 차이가 나타나기를 기대했으나, 실제 응답은 비교적 일반적이고 유사한 패턴을 보이는 경우가 있었다. P6는 “특성이 답변의 내용에 충분히 투영되지 않는다”고 언급하였다. 또한 일부 참가자는 퍼소나에 따라 답변의 전문성, 설명 방식, 말투 등에서 보다 뚜렷한 차이가 나타난다면 모의 인터뷰 상황의 현실성과 몰입도가 높아질 것이라고 의견을 제시하였다. 이는 퍼소나 설정이 단순한 인물 정보 제시에 그치지 않고, 응답 스타일과 대화 방식까지 보다 적극적으로 반영될 필요가 있음을 시사한다.

(3) 분석 및 피드백의 구체적 실행 지침 부족

분석 지표와 대시보드는 전반적으로 유용하게 인식되었다. 특히 정서적 대응에 대한 피드백은 여러 참가자에게서 긍정적으로 평가되었다. P3는 “정서적 리액션에 대한 피드백을 제공받은 이후 해당 요소에 더욱 유념하여 발언하게 되었다”고 언급하였으며, P1 역시 2차 인터뷰에서 “정서적 공감”을 개선했다고 응답하였다. 이는 피드백이 최소한 인식 전환 단계에는 기여했음을 보여준다.

그러나 동시에 일부 참가자는 정서적 대응 피드백이 구체적인 행동 전략으로 이어지기에는 다소 부족하다고 지적하였다. P1은 공감 표현이 부족하다는 결과를 받았을 때 어떤 문장 구조나 표현 방식을 사용해야 하는지까지는 제시되지 않아 곤란하였다는 의견을 제시하였다. P5 역시 “구체적으로 어떤 질문에서 부족했는지 알기 어려워 텍스트 피드백이 와닿기 어려웠다”고 응답하였다. 이는 예컨대 공감 표현 빈도가

낮다는 지표가 제시되더라도, 어느 발화에서 공감적 반응이 가능했는지, 어떤 방식으로 재진술하거나 인정 표현을 덧붙일 수 있었는지까지는 안내되지 않았음을 의미한다.
또한 P2는 인터뷰 중 “분위기를 완화시키거나 공감 수치를 높여주는 질문”을 실시간으로 제안해주는 기능이 있으면 좋겠다고 언급하였다. 이는 정서적 대응을 개선하려는 의도는 있으나 실제 상황에서 어떤 방식으로 개선해야 할지 판단하기 어렵다는 점을 보여준다.

8. 논의

본 연구는 AI 기반 모의 인터뷰 시스템이 사용자 인터뷰 수행 훈련에 어떠한 영향을 미칠 수 있는지를 탐색적으로 분석하였다. 실험 결과, 참가자들은 시스템을 활용하여 인터뷰 수행 과정에서 자신의 발화 구조와 상호작용 패턴을 보다 객관적으로 인식할 수 있었다. 특히 발화–경청 비율, 시간 배분, 꼬리질문 확장 정도, 정서적 대응 등 인터뷰 수행 과정에서 나타나는 다양한 요소가 데이터로 가시화되면서, 참가자들이 자신의 인터뷰 전략을 점검하고 조정하는 계기가 되었다.
다만 사용자 평가 과정에서는 몇 가지 개선이 필요한 지점도 함께 확인되었다. 참가자들은 현재 제공되는 분석 지표가 자신의 수행을 객관적으로 인식하는 데에는 도움이 되었지만, 구체적으로 어떤 방식으로 인터뷰 전략을 수정해야 하는지에 대해서는 추가적인 안내가 필요하다는 의견을 제시하였다. 또한 퍼소나의 특성과 인터뷰 목적 사이의 연결성, 시뮬레이션 응답의 현실감 등에서도 보완 가능성이 제기되었다.
이러한 결과를 바탕으로 본 절에서는 시스템 개선 방향과 확장 가능성, 그리고 연구의 한계를 중심으로 논의를 진행한다.

8. 1. 향후 시스템 개선 및 확장 가능성

(1) 질문 설계 및 수행 피드백의 코칭 강화
형성적 조사에서는 초보 학습자가 인터뷰 질문을 구성하는 데 어려움을 겪는다는 문제가 확인되었고, 이를 지원하기 위해 Pretalk에 질문 자동 생성 기능을 반영하였다. 그러나 사용자 평가에서 일부 참가자는 질문이 완성된 형태로 제공될 경우 질문의 의도와 흐름을 충분히 이해하지 못한 채 인터뷰를 진행하게 될 수 있다고 지적하였다. 또한 사용자가 먼저 질문을 구성한 뒤 AI가 질문의 순서나 중복 여부를 점검하고 수정 방향을 제안하는 방식이 훈련에 더 도움이 될 수 있다는 의견이 제시되었다. 따라서 향후에는 질문 자동 생성뿐 아니라 학습자가 작성한 질문을 바탕으로 개선 방향을 제안하는 코칭형 지원을 함께 고려할 필요가 있다.
수행 이후의 피드백에서도 보다 구체적인 안내에 대한 요구가 나타났다. 참가자들은 인터뷰 수행 지표가 자신의 진행 방식을 인식하는 데 도움이 되지만, 이를 실제로 어떻게 수정해야 하는지까지는 충분히 안내하지 못한다고 평가하였다. 따라서 실제 발화 기록과 함께 개선 가능한 꼬리질문이나 공감 표현의 예시를 제시하여, 학습자가 자신의 수행에서 어떤 부분을 어떻게 수정할 수 있는지 이해하도록 지원할 수 있다.
나아가 인터뷰 진행 중 발화가 지나치게 길어지거나 추가 탐색이 필요한 상황에서 간단한 힌트를 제공하는 실시간 코칭도 고려할 수 있다. 이러한 방식은 질문 설계와 인터뷰 수행 모두에서 AI가 완성된 결과를 제공하는 데 그치지 않고, 학습자가 자신의 질문과 인터뷰 전략을 점검하고 조정하도록 지원하는 방향으로 확장될 가능성을 보여준다.

(2) 퍼소나의 목적 연결성과 응답 반영도 강화
일부 참가자는 퍼소나 정보가 인터뷰 목적과 어떻게 연결되는지 충분히 이해하기 어렵다고 언급하였다. 또한 퍼소나의 직업, 성격, 배경 특성이 실제 응답 내용에 충분히 반영되지 않아 답변이 다소 일반적인 수준에 머무르는 경우가 있다는 의견도 제시되었다.
향후 시스템에서는 퍼소나 정보가 인터뷰 훈련의 목적과 보다 명확히 연결되도록 설계할 필요가 있다. 예를 들어 퍼소나를 단순한 키워드 수준의 정보로 제시하기보다 행동 특성, 주요 동기, 제품 사용 경험, 잠재적

불만 요인 등을 포함한 구체적인 형태로 제시할 수 있다. 또한 특정 문제 상황이나 사용 맥락을 짧은 시나리오 형태로 함께 제공한다면, 학습자가 어떤 퍼소나와 인터뷰를 진행하는 것이 자신의 훈련 목적에 적절한지 판단하는 데 도움이 될 수 있다. 인터뷰 훈련의 목적에 맞게 퍼소나의 특성을 조정할 수 있도록 지원하는 것도 고려할 수 있다. 예를 들어 퍼소나의 말투, 응답 태도, 의견의 강도, 경험 수준 등을 설정하여 협조적인 사용자, 소극적으로 답변하는 사용자, 강한 의견을 가진 사용자 등 다양한 유형을 경험하도록 할 수 있다. 이러한 설정은 학습자가 서로 다른 사용자 유형에 따라 질문 방식이나 인터뷰 전략을 조정하는 훈련으로 연결될 수 있다.

이와 같이 퍼소나의 특성이 실제 응답 패턴에 보다 적극적으로 반영되도록 설계한다면, 단순한 정보 응답형 대화를 넘어 보다 다양한 인터뷰 상황을 경험할 수 있을 것이다. 동시에 학습자가 자신의 훈련 목적에 맞는 인터뷰 상황을 선택하고 준비하는 과정에도 도움이 될 것으로 기대된다.

(3) 멀티모달 시뮬레이션과 분석

현재 시스템은 음성 기반 상호작용을 중심으로 구성되어 있어 인터뷰 상황에서 중요한 비언어적 요소는 제한적으로 반영되어 있다. 일부 참가자는 인터뷰이의 감정 상태나 반응이 충분히 표현되지 않아 몰입도가 떨어질 수 있다는 의견을 제시하였다. 실제 인터뷰에서는 표정 변화, 망설임, 반응 지연, 감정적 돌발 상황과 같은 비언어적 신호가 상호작용의 흐름을 형성하는 중요한 요소로 작용한다. 이러한 요소를 반영하기 위해서는 가상 인터뷰이가 단순 음성 응답을 넘어 시각적 형태의 인터뷰 에이전트로 확장될 필요가 있다. 예를 들어 인터뷰어와 시선을 맞추거나 고개를 끄덕이고, 질문의 맥락에 따라 표정이나 자세가 변화하도록 할 수 있다. 또한 질문의 난이도나 민감도에 따라 망설임이나 감정 변화가 나타나도록 한다면 실제 인터뷰에 보다 가까운 훈련 상황을 제공할 수 있을 것이다.

이와 같은 시각적 상호작용이 도입되면 인터뷰어의 비언어적 행동 역시 분석 대상으로 확장할 수 있다. 예를 들어 시선 유지, 고개 끄덕임과 같은 경청 신호뿐 아니라 몸동작이나 자세 변화가 대화 흐름에 어떻게 나타나는지를 함께 분석할 수 있다. 이러한 멀티모달 분석은 인터뷰 훈련을 단순한 질문과 응답 구조에서 벗어나 언어적, 비언어적 상호작용을 함께 학습하는 방향으로 확장하는 데 활용될 수 있다.

(4) 사용자 인터뷰 맥락 확장

현재 시스템은 일대일 인터뷰 상황을 중심으로 설계되어 있으나, 디자인 과정에서는 연구 목적과 상황에 따라 1:1 심층 인터뷰, 포커스 그룹 인터뷰, 맥락 인터뷰(contextual inquiry) 등 다양한 방식이 활용된다. 따라서 향후 시스템은 인터뷰 훈련의 맥락을 보다 다양하게 확장할 필요가 있다. 예를 들어 1:1 심층 인터뷰 상황에서는 개인 경험을 깊이 탐색하는 질문 전략을 연습하도록 하고, FGI 상황에서는 여러 참여자의 발화를 조율하거나 의견 충돌을 관리하는 인터뷰 진행 전략을 연습할 수 있도록 설계할 수 있다. 또한 맥락 인터뷰와 같은 방식에서는 특정 사용 환경이나 활동 상황을 기반으로 질문을 구성하고 관찰과 질문을 병행하는 훈련이 가능할 것이다.

이와 같은 다양한 인터뷰 상황을 시뮬레이션에 포함한다면 학습자는 단일 인터뷰 기술을 연습하는 수준을 넘어, 연구 목적과 상황에 맞는 인터뷰 방식을 선택하고 전략을 조정하는 능력을 함께 훈련할 수 있을 것이다.

8. 2. 연구의 한계

본 연구는 Pretalk의 단기적 활용 가능성과 반복 사용 과정에서 나타나는 수행 변화를 탐색하는 데 초점을 두었기 때문에, 이러한 변화가 실제 사용자 인터뷰 수행 능력으로 전이되는지는 확인하지 못하였다. 또한 디자인 전공 학부생 6명을 대상으로 한 소규모 탐색적 평가로 진행되어 연구 결과를 일반화하는 데 한계가 있다. 향후에는 보다 다양한 참가자를 대상으로 장기적인 반복 훈련을 진행하고, Pretalk 활용 이후 실제 프로젝트 환경에서 수행한 사용자 인터뷰 결과를 함께 비교할 필요가 있다.

또한 본 연구에서는 동일한 참가자가 Pretalk을 두 차례 반복 사용하도록 하였기 때문에, 1차와 2차 수행 사이에서 나타난 변화에 시스템의 분석 및 피드백뿐 아니라 반복 수행에 따른 익숙함이나 연습 효과가 함께 작용했을 가능성을 배제하기 어렵다. 따라서 본 연구에서 관찰된 변화를 Pretalk의 독립적인 효과로 해석하는

데에는 한계가 있으며, 향후에는 비교 조건을 포함한 연구 설계를 통해 시스템 사용에 따른 변화를 보다 명확하게 검토할 필요가 있다.
마지막으로 시스템이 산출한 꼬리질문과 정서적 대응 등의 일부 분석 지표에 대해서는 인간 코더와의 일치도나 별도의 타당도 검증을 수행하지 않았다. 따라서 이러한 지표는 인터뷰 역량을 객관적으로 평가하기 위한 검증된 측정값이라기보다, 학습자가 자신의 모의 인터뷰 수행 특성을 점검하기 위한 참고 지표로 해석할 필요가 있다. 향후에는 전문가의 코딩 결과와 시스템 분석 결과를 비교하는 등의 방법을 통해 자동 분석 지표의 신뢰성과 타당성을 검토할 필요가 있다.

9. 결론

본 연구는 디자인 교육에서 사용자 인터뷰 수행 훈련을 지원하기 위한 AI 기반 모의 인터뷰 시스템을 제안하고, 이를 활용한 수행 데이터를 통해 인터뷰 과정이 어떻게 가시화되고 학습에 활용될 수 있는지를 탐색하였다. 특히 발화–경청 비율, 시간 배분, 꼬리질문 확장 정도, 정서적 대응과 같은 수행 요소를 데이터로 분석하여 학습자가 자신의 인터뷰 전략을 점검하고 조정할 수 있도록 하는 구조를 제시하였다.
본 연구의 의의는 사용자 인터뷰 훈련을 단순한 대화 연습이 아니라 수행 과정 자체를 분석하고 반성적으로 점검할 수 있는 학습 활동으로 확장했다는 점에 있다. 기존의 인터뷰 교육은 질문 설계나 사례 학습에 집중되는 경우가 많았으며, 실제 수행 과정에 대한 반복적인 연습과 피드백을 제공하기는 어려웠다. 본 연구는 가상 인터뷰이와의 반복 수행과 수행 데이터 기반 분석을 결합함으로써, 학습자가 자신의 인터뷰 방식과 상호작용 패턴을 확인하고 전략을 수정할 수 있는 구조를 제시하였다. 또한 이러한 접근은 AI를 단순한 자동화 도구로 사용하는 것을 넘어, 디자이너가 리서치 수행 과정에서 자신의 행동을 성찰하고 개선할 수 있도록 돕는 학습 파트너로 활용할 수 있는 가능성을 보여준다. 이는 향후 AI 기반 도구가 디자인 교육에서 수행 중심 학습을 지원하는 새로운 형태의 교육 환경으로 발전할 수 있음을 시사한다.

References

1. Argyle, L. P., Busby, E. C., Fulda, N., Gubler, J. R., Rytting, C., & Wingate, D. (2023). Out of one, many: Using language models to simulate human samples. *Political Analysis, 31*(3), 337–351.
2. Gomez, N., Batham, S., Volonte, M., & Do, T. D. (2025). Virtual Interviewers, Real Results: Exploring AI-Driven Mock Technical Interviews on Student Readiness and Confidence. In *Companion publication of the 2025 conference on computer-supported cooperative work and social computing* (pp. 209–213). Association for Computing Machinery.. https://doi.org/10.1145/3715070.3749227
3. Goodman, E., Kuniavsky, M., & Moed, A. (2012). *Observing the User Experience: A Practitioner's Guide to User Research*. Morgan Kaufmann.
4. Gubrium, J. F., & Holstein, J. A. (2002). *Handbook of interview research: Context and method*. Sage.
5. Hanington, B., & Martin, B. (2022). *The Pocket Universal Methods of Design, Revised and Expanded: 125 Ways to Research Complex Problems, Develop Innovative Ideas, and Design Effective Solutions*. Rockport Publishers.
6. Hoque, M. E., Courgeon, M., Martin, J.-C., Mutlu, B., & Picard, R. W. (2013). MACH: my automated conversation coach. *Proceedings of the 2013 ACM International Joint Conference on Pervasive and Ubiquitous Computing*, 697–706. https://doi.org/10.1145/2493432.2493502
7. Hu, J., Guo, J., Tang, N., Ma, X., Yao, Y., Yang, C., & Xu, Y. (2024). Designing the Conversational Agent: Asking Follow-up Questions for Information Elicitation. *Proceedings of the ACM on Human-Computer Interaction, 8*(CSCW1), 1–30. https://doi.org/10.1145/3637320
8. Kumar, V. (2012). *101 design methods: A structured approach for driving innovation in your organization*. John Wiley & Sons.

9. Kvale, S., & Brinkmann, S. (2009). *Interviews: Learning the craft of qualitative research interviewing*. sage.
10. Leong, C. W., Jawahar, N., Basheerabad, V., Wörtwein, T., Emerson, A., & Sivan, G. (2024). Combining Generative and Discriminative AI for High-Stakes Interview Practice. *Companion Proceedings of the 26th International Conference on Multimodal Interaction*, 94–96. https://doi.org/10.1145/3686215.3688377
11. Lin, X., Zhao, T., Schmidt, S. W., & Zhou, S. (2025). Using AI as a Learning Tool Through Simulation Interviews to Enhance Adult Learning. *Adult Learning*, 10451595251345274.
12. Lucas, G. M., Gratch, J., King, A., & Morency, L.-P. (2014). It's only a computer: Virtual humans increase willingness to disclose. *Computers in Human Behavior, 37*, 94–100.
13. Tai, R. H., Bentley, L. R., Xia, X., Sitt, J. M., Fankhauser, S. C., Chicas-Mosier, A. M., & Monteith, B. G. (2024). An Examination of the Use of Large Language Models to Aid Analysis of Textual Data. *International Journal of Qualitative Methods, 23*, 1–14. https://doi.org/10.1177/16094069241231168
14. Wilkie, L., & Rosendale, J. (2024). Efficacy and benefits of virtual mock interviews: Analysing student perceptions of digital employment preparations. *Journal of University Teaching and Learning Practice, 21*(1), 177–195.
15. Yan, L., Echeverria, V., Fernandez-Nieto, G. M., Jin, Y., Swiecki, Z., Zhao, L., Gašević, D., & Martinez-Maldonado, R. (2024). Human-AI Collaboration in Thematic Analysis using ChatGPT: A User Study and Design Recommendations. *Extended Abstracts of the CHI Conference on Human Factors in Computing Systems*. https://doi.org/10.1145/3613905.3650732
16. Zimmerman, J., Forlizzi, J., & Evenson, S. (2007). Research through design as a method for interaction design research in HCI. *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems - CHI '07*, 493–502. https://doi.org/10.1145/1240624.1240704

사용자 인터뷰 역량 강화를 위한 AI 기반 모의 인터뷰 훈련 시스템 개발 및 평가

이문환[1*], 성수민[2], 신산하[2]

[1]한국공학대학교 디자인공학부, 부교수, 시흥, 대한민국

[2]한국공학대학교 디자인공학부, 학사, 시흥, 대한민국

초록

연구배경 사용자 인터뷰는 디자인 리서치에서 사용자의 경험과 요구를 심층적으로 이해하기 위한 핵심적인 질적 조사 방법이다. 그러나 인터뷰의 성과는 질문 구성, 대화 흐름 조율, 응답에 대한 즉각적 대응 등 인터뷰어의 수행 역량에 크게 의존한다. 디자인 교육 환경에서는 실제 사용자 인터뷰를 반복적으로 연습하거나 수행 과정을 세밀하게 피드백받기 어려운 제약이 존재한다. 본 연구는 이러한 교육적 한계를 보완하기 위해 사용자 인터뷰 수행 훈련을 지원하는 AI 기반 모의 인터뷰 시스템의 활용 가능성을 탐색하고자 한다.

연구방법 본 연구는 Research-through-Design(RtD) 접근을 기반으로 세 단계로 진행되었다. 먼저 디자인 교육자 4명을 대상으로 형성적 인터뷰를 수행하여 초보 학습자가 사용자 인터뷰 과정에서 겪는 주요 어려움과 교육적 요구를 도출하였다. 이를 바탕으로 인터뷰 질문 생성, 가상 인터뷰이 퍼소나 생성, 음성 기반 인터뷰 시뮬레이션, 수행 분석 및 피드백 기능을 통합한 AI 기반 모의 인터뷰 훈련 시스템 Pretalk을 설계 및 구현하였다. 이후 디자인 전공 학부생 6명을 대상으로 두 차례의 반복 훈련을 포함한 사용자 평가를 수행하여 인터뷰 수행 변화와 학습 경험을 분석하였다.

연구결과 사용자 평가 결과, 시스템은 발화–경청 비율, 시간 관리, 꼬리 질문 수행, 정서적 대응 등 인터뷰 수행 요소를 가시화함으로써 학습자가 자신의 인터뷰 방식을 점검하고 조정하는 데 도움을 주는 것으로 나타났다. 반복 훈련 이후 참가자들은 인터뷰 구조의 안정성, 꼬리 질문 활용, 공감 표현 등에서 수행 변화가 나타났으며, 시스템이 실제 인터뷰 전 사전 점검 및 연습 도구로 활용될 수 있다는 인식을 보였다.

결론 본 연구는 사용자 인터뷰 훈련을 단순한 대화 연습이 아니라 수행 과정을 분석하고 성찰할 수 있는 학습 활동으로 확장하는 AI 기반 훈련 환경의 가능성을 제시하였다. 또한 AI를 자동화 도구가 아닌 디자이너의 리서치 역량 개발을 지원하는 학습 파트너로 활용할 수 있음을 보여준다. 이러한 결과는 디자인 교육에서 사용자 인터뷰 학습을 지원하는 AI 기반 도구 설계에 대한 기초적 시사점을 제공한다.

주제어 사용자 인터뷰 훈련, 생성형 인공지능, 디자인 연구방법 교육, 모의 인터뷰 시스템

*교신저자 : 이문환 (moonhwanlee@tukorea.ac.kr)