



# Journal of Virtual Convergence Research

Volume 2

Number 2

Apr. 2026

Received: 28 January 2026. Accepted: 22 April 2026

© The Author(s) 2026. Published by International Metaverse Association. All rights reserved. For commercial reuse and other permissions please contact [hdq.ima@gmail.com](mailto:hdq.ima@gmail.com) for reprints and translation rights for reprints.

## 아바타 기반 온라인 강의의 상호작용 효과 연구:

### 실시간 질의응답, 응답 속도 및 언어 스타일의 기반으로

이태환

서강대학교 메타버스 전문대학원

luisfynn1@sogang.ac.kr

#### Abstract

온라인 강의란 인터넷을 활용하여 교수자가 사전 제작한 녹화 영상이나 실시간 영상을 학습자에게 전달하는 원격 교육의 한 형태이다. 이 중 녹화 영상의 경우 학습자가 학습 도중 중요한 개념을 제대로 이해 하지 못하는 경우 학습자가 강의 전체 내용을 제대로 이해하지 못하거나, 중도 포기하는 일이 발생할 수 있다.

이 연구는 이러한 문제점을 개선하기 위해 아바타 기반 온라인 강의 환경에서 학습자와의 상호작용을 촉진하기 위해 설계한 요소들이 학습자의 인지적 및 정의적 경험에 어떠한 영향을 미치는지를 탐색적으로 분석하는데 목적이 있다. 온라인 강의 중 실시간 질의응답 제공을 가능하게 하고, 학습자의 질문에 대한 아바타의 응답 속도(즉시 응답 vs 지연된 응답), 그리고 아바타의 언어 스타일(대화체 vs 격식체)을 조작 변인으로 설정함으로써, 이들 변인이 학습자의 학습 성과, 몰입, 사회적 실재감, 인지부하에 미치는 영향을 실증적으로 검증하고자 하였다. 이를 검증하기 위해 본 저자는 Unity 를 활용하여 실험용 온라인 AI 아바타 강의 시스에 Google STT(Speech-to-Text), LLM, ElevenLabs TTS 를 연동하여 학습자가 상호작용할 수 있도록 환경을 구성하였으며, 총 52명의 참가자를 모집하여, 2x2 요인 설계 실험을 수행하였으며, 사전·사후 테스트, 설문, 구조화된 인터뷰를 통해 자료를 수집하여 정량·정성 통합 분석을 실시하였다.

첫째, 데이터 분석한 결과 학습자가 온라인 강의 학습 중 실시간으로 질의하고,

아바타가 질문에 대해 답변하는 실시간 질의응답 기능은 학습 성과의 평균을 유의하게 증가시키지는 않았으나, 성취 점수의 분산을 유의하게 감소시키는 안정화 효과를 보였다.

둘째, 아바타가 학습자의 질문에 대해 응답하는 속도(응답 속도)와 아바타의 온라인 강의 언어 스타일의 주효과는 대부분의 종속변수에서 통계적으로 유의하지 않았으나, 사회적 실재감과 인지부하에서는 두 변인의 조합에 따른 교차적 평균 패턴과 중간 수준의 효과 크기가 관찰되었다. 특히 강의에 대화체를 사용한 경우 학습자의 질문에 빠르게 응답 시 사회적 실재감이 높고 인지부하가 낮은 반면, 격식체를 사용시 학습자의 질문에 대해 느리게 응답할 경우 사회적 실재감이 더 높게 나타나는 경향을 확인하였다.

이 연구의 실험군 인원이 10~11명으로, 수집된 자료를 통계 검정 결과 유의하지 않은 결과를 얻었지만 변인들의 효과 크기, 평균 패턴의 구조적 일관성, 정성 의견과의 정합성을 종합하여 설명하면, 아바타 기반 상호작용 설계 변인의 잠재적 효과를 뒷받침하는 예비 근거를 제공한다는 점에서 학술적 의의가 있다고 본다. 또한 대규모 실험을 위한 설계 가설과 변수 조합 전략을 제시하는 탐색적 파일럿 연구로서의 성격을 가지며, AI 아바타 기반 교육 시스템 설계 구축에 도움이 될 수 있기를 기대한다.

**Keywords :** AI Avatar, Real-Time Interaction, Response Latency, Linguistic Style, Learning Outcomes, Flow, Social Presence, Cognitive Load Theory

**아바타 기반 온라인 강의의 상호작용 효과 연구:  
실시간 질의응답, 응답 속도 및 언어 스타일의 기반으로**

## 1. 서론

20년 ~ 23년에 걸친 코로나 19 팬데믹을 기점으로 온라인 강의는 보편적인 교육 형태로 자리 잡았다. 또한 22년말 ChatGPT 출시 이후 급속도로 발전하고 있는 인공지능(AI) 기술과 결합하여 학습자 맞춤형 교육을 제공하는 방향으로 교육 방식이 진화하고 있다. 온라인 강의는 시간과 장소의 제약 없이 언제 어디서나 학습할 수 있다는 강점이 있으나, 교수자와 학습자 간의 상호작용이 제한된 단방향성(One-way communication)이라는 본질적인 한계를 갖고 있다. 이러한 단방향성은 학습자가 온라인 강의에서 언급하는 특정 개념 또는 용어를 학습자의 수준으로 이해하지 못한다면, 이어지는 뒷부분의 강의 내용을 온전히 파악할 수 없을 가능성이 높다. 또한 학습 중 발생하는 의문점을 해소하기 위해 학습자가 강의를 일시 중단하고 인터넷을 사용하여 외부 자료를 검색할 경우, 검색 시간 및 진위 여부 확인으로 인한 학습의 흐름을 단절(Disruption of Flow)시키고 외생적 인지 부하(Extraneous Cognitive Load)를 증가시켜, 결과적으로 학습 동기 저하 및 중도 포기를 유발하는 주요 원인이 된다.

이 연구는 기존의 온라인 학습 환경에 AI 기반 아바타를 도입하여, 온라인 학습 중 발생하는 의문점을 실시간으로 해소시켜 주기 위한 실시간 질의응답 시스템을 제안하고자 한다. 이는 학습자의 궁금증을 즉각적으로 해소하여 학습 몰입을 유지하고 인지 부하를 경감시키는 한편, 아바타와의 상호작용을 통해 사회적 실재감(Social Presence)을 고양함으로써 궁극적으로 온라인 학습의 성과를 제고하는 데 그 목적이 있다.

더 나아가 실험을 통해 관찰된 통계적 유의성 검증에 그치지 않고, 각 실험 집단 간 평균 차이의 상대적 크기를 나타내는 효과의 규모에 주목하고, 상호작용 설계 변인이 학습자 경험에 미치는 영향의 실제적 크기와 방향성을 정량적으로 탐색하고자 한다. 이를 통해 차후 대규모 실험에서 표본 수 산정, 변인 조합 최적화, 상호작용 조건 세분화에 필요한 기초 정보를 마련하는 것을 목적으로 한다. 최종적으로 이 연구는 AI 아바타 기반 교육 시스템의 상호작용 설계 원칙을 경험적으로 구체화하기 위한 연구로서의 출발점을 제공하고자 한다.

## 2. 이론적 배경

### 2.1. 온라인 교육 환경의 발전과 AI 아바타의 등장 배경

온라인 교육은 지난 수십 년간 고등 교육 및 직업 훈련 분야에서 보편적인 학습 방식으로 자리 잡았으며, 특히 최근 인공지능(AI) 기술의 급속한 발전과 AI 아바타의 접목을 통해 새로운 혁신 국면에 접어들었다 (Islam & Wang, 2025). 이러한 기술적 진보는 교육 분야를 포함한 글로벌 AI 시장 성장을 주도하고 있으며, 연평균 30% 이상의 성장률을 보일 것으로 예측된다 (Islam & Wang, 2025). AI 아바타가 최근 여러 디지털 플랫폼에서 주목받고 있다. 이를 테면, 교수나 강사가 직접 출연하지 않아도 온라인 강의를 제작할 수 있고, 연예인을 섭외하지 않고도 광고 영상을 만들 수 있으며, 뉴스 앵커 없이 뉴스를 송출하는 것도 가능하다. 미디어 산업 전반에서 이런 기술이 확산되는 추세다. 영국의 Synthesia 플랫폼을 살펴보면, 텍스트만 입력하고 아바타를 고르면 음성 합성과 립싱크로 영상이 완성된다. 교육 콘텐츠 제작 분야에서도 이런 솔루션이 도입되고 있다.

### 2.2. 온라인 기반 학습의 단점

온라인 학습의 가장 큰 장점은 시간과 장소에 구애받지 않는다는 것이다. 지하철이나 버스에서 스마트폰으로 강의를 듣는 사람들을 보는 것은 더 이상 낯선 풍경이 아니다. 출퇴근 시간이나 점심시간 같은 자투리 시간을 활용하여 학습을 하는 것은 이제 일상이 되었다.

하지만 온라인 강의를 제공하는 접근성과 유연성에도 불구하고, 대부분의 강의 시스템은 단순히 정보를 전달하는 데 초점을 맞추고 있어, 학습자의 고립감을 심화시키고 몰입도를 저하시킨다는 구조적인 한계를 갖고 있다(Briz-Ponce et al., 2017). 특히, 학습 과정에서 발생하는 질문이 즉시 해결되지 못할 경우, 학습 흐름이 끊겨 학습 효율 저하와 중도 포기의 주요 원인으로 작용한다 (Briz-Ponce et al., 2017).

따라서 AI 아바타를 활용하여 학습자와 상호작용할 수 있는 시스템을 구축한다면, 학습자의 학습 효과를 극대화하는 데 결정적인 요소가 될 수 있을 것으로 기대할 수 있다.

### 2.3. 연구의 필요성 및 핵심 연구 공백

최근 연구들은 아바타가 온라인 학습의 질을 향상시키며, 주의력과 긍정적 감정을 증진시킨다는 점을 보고하며 아바타의 잠재적 가치를 입증하였다 (Xia et al., 2024). Kodani et al. (2025)의 연구는 아바타의 체화된 의인화가 교육적 효과를 개선함을 보여주었다 (Kodani et al., 2025). 그러나 이러한 선행 연구들은 주로 아바타의 시각적 속성이나 전반적인 사회적 실재감에 초점을 맞추었을 뿐, 상호작용의 질을 구성하는 미시적인 두 가지 핵심 요소인 응답 속도(Latency)와 언어스타일(Tone)이 학습자의

인지적, 정서적 상태 및 최종 학습 성과에 미치는 영향을 통합적으로 검증하지 못했다 (Ukenova et al., 2025).

이 연구는 이러한 이론적 공백을 메우고자, 실시간 질의응답 제공이라는 기본 조건 하에, 응답 속도(빠름 vs. 느림)와 언어 스타일(대화체 vs. 격식체)을 독립 변인으로 조작하여, 상호작용의 시간적 차원과 언어적 차원이 학습자의 인지적·정의적 경험에 미치는 영향을 실증적으로 분석하는 것을 목적으로 한다.

구체적으로, 응답 속도는 아바타가 학습자의 질문을 인식한 시점부터 실제 답변을 출력하기까지의 지연 시간으로 조작하였으며, 언어 스타일은 동일한 의미 내용을 전달하되 문장 종결 어미와 표현 방식만을 체계적으로 달리하는 방식으로 구현하였다. 이 두 변인이 학습 성과에 미치는 영향은 직접적인 경로뿐 아니라, Flow, 사회적 실재감, 인지부하라는 세 가지 심리적 매개 변인을 경유하는 간접 경로를 통해서도 작동할 것으로 가정하였다. 이를 통해 단순히 어떤 조건이 더 효과적인가를 확인하는 수준을 넘어, 상호작용 설계 변인이 학습자 경험에 영향을 미치는 내적 메커니즘을 구조적으로 규명하고자 하였다.

### 3. 선행연구 검토

이 연구는 세 가지 핵심 이론을 채택하여 통합 프레임으로 구성하였다. 각 이론이 독립변인(응답 속도, 언어 스타일, Q&A 제공 여부)과 종속·매개변인을 연결하는 역할을 다음과 같이 담당한다.

첫째, Flow 이론(Csikszentmihalyi, 1991)은 즉각적 피드백이 학습 몰입을 유지 및 강화하는 메커니즘을 설명하며, 이 연구에서 응답 속도가 몰입과 학습 성과에 미치는 경로의 이론적 근거를 제공한다.

둘째, 인지부하 이론(Sweller, 1988) 및 Mayer(2021)의 개인화 원리는 언어 스타일이 외재적 인지부하를 증감시키는 메커니즘을 설명하며, 언어 스타일이 인지부하를 매개로 학습 성과에 영향을 주는 경로를 이론적으로 뒷받침한다.

셋째, 사회적 실재감 이론(Biocca et al., 2003)은 상호작용 방식이 학습자가 아바타를 사회적 존재로 인식하는 정도를 결정하며, 이것이 학습 참여도와 성취로 이어짐을 설명한다.

세 이론의 관계를 정리하면, 응답 속도는 주로 Flow 와 사회적 실재감 경로를, 언어 스타일은 주로 인지부하와 사회적 실재감 경로를, Q&A 제공 여부는 세 매개변인 모두를 통해 학습 성과에 영향을 미치는 구조로 설정된다(Figure 2 참조).

### 3.1. Flow 이론과 상호작용의 즉시성 (Immediacy)

Flow 이론은 Csikszentmihalyi (1991)에 의해 정립되었으며, 개인이 활동에 완전히 몰입하여 최적의 경험을 느끼는 심리적 상태를 설명한다 (Csikszentmihalyi, 1991). Flow 상태를 유도하는 핵심 조건 중 하나는 명확한 목표(Clear Goals)와 함께 즉각적이고 명확한 피드백(Immediate and Unambiguous Feedback)의 제공이다 (Csikszentmihalyi, 1991; Sherry, 2004). 이러한 피드백은 학습자가 도전과 기술 간의 균형을 유지하고, 현재 활동에 완전히 집중하도록 돕는다 (Csikszentmihalyi, 1991). Sherry (2004)는 미디어 환경에서 Flow 가 향유(Enjoyment)와 밀접하게 연관됨을 보여주었으며 (Sherry, 2004), 이는 몰입 경험이 긍정적인 학습 동기로 이어진다는 것을 시사한다. 온라인 실시간 질의응답 환경에서 즉각적인 피드백 구현을 위해서 AI 응답을 파이썬의 re package 를 이용하여 문장 단위로 분할하고, 각 문장을 실시간으로 TTS 처리하였다. 응답 지연이 길어질 경우, 학습자는 피드백의 즉시성을 상실하게 되며, 이는 Flow 상태의 구성 요소인 '과제 집중' (Csikszentmihalyi, 1991)을 방해하고 학습의 흐름을 끊는다. 이는 학습자의 주관적인 시간적 압박 및 좌절감을 높여 Flow 경험을 저해하는 요인으로 작용한다 (Sherry, 2004). Georgiou 와 Kyza (2018)는 위치 기반 증강 현실(AR) 학습 환경에서 몰입 수준이 개념적 학습 성과와 긍정적인 관계를 가짐을 실증하였다 (Georgiou & Kyza, 2018). Flow 이론에서 강조하는 즉각적 피드백은 빠른 응답 속도로 실현된다. 이를 통해 학습자의 몰입과 성취도가 높아진다.

### 3.2. 인지부하 이론 (CLT) 및 개인화 원리 (Personalization Principle)

Sweller (1988)의 인지부하 이론은 인간의 작업 기억 용량이 제한적이며, 학습 효율을 높이기 위해서는 학습 내용 자체의 난이도(Intrinsic Load)를 제외한 불필요한 인지적 노력(Extraneous Cognitive Load)을 최소화해야 함을 강조한다 (Sweller, 1988). Mayer (2021)는 이러한 인지부하 이론을 멀티미디어 환경에 적용하여, 학습 자료의 언어적 표현 방식을 최적화하는 '개인화 원리(Personalization Principle)'를 제안하였다 (Mayer, 2021). 이 원리는 학습 자료를 대화체 스타일(Conversational Style)로 제시하는 것이 격식체(Formal Style)보다 학습자의 인지 부하를 줄이고 이해도를 높인다고 주장한다 (Mayer, 2021). 대화체는 학습자가 아바타를 단순한 정보 제공자가 아닌 사회적 파트너로 인식하도록 유도하여, 정보 처리에 필요한 정신적 노력을 경감시킨다 (Mayer, 2021). Ukenova et al. (2025)는 아바타 기반 학습 시스템의 개선 방안으로 언어 구조 정렬(Linguistic Structure Alignment)과 감정 기반 표현(Sentiment-Driven Expressions)의 중요성을 논했으며 (Ukenova et al., 2025), 이는 언어 스타일의 조작이 학습 효율에 직접적인 영향을 미칠 수 있음을

시사한다. 대화체는 학습자와 아바타 사이의 친밀도를 높이고 외재적 인지 부하를 줄임으로써 학습 효율을 개선한다.

### 3.3. AI 아바타, 사회적 실재감 및 학습 효과

아바타는 학습 환경에서 단순한 시각적 표현을 넘어, 학습자들에게 사회적 실재감(Social Presence)을 부여하고 상호작용을 촉진하는 핵심 인터페이스이다. Xia et al. (2024)은 아바타 사용이 학습자의 주의력과 긍정적 감정을 향상시키고, 사회적 불안을 완화함으로써 온라인 학습의 질을 개선한다는 점을 확인하였다 (Xia et al., 2024). Pang et al. (2023)은 아바타 기반 플랫폼이 화학 실험실 세션과 같은 협력적 환경에서 사회적 상호작용을 촉진하는 가능성을 탐색하였다 (Pang et al., 2023). Kodani et al. (2025)은 안드로이드 아바타의 교육적 효과를 탐구하며, 학습자가 아바타의 의인성(Anthropomorphism)과 능력(Competence)을 긍정적으로 인식할 때 주관적인 교육 효과가 향상되지만, 불편함(Discomfort)은 부정적인 영향을 미친다고 보고하였다 (Kodani et al., 2025). 이는 아바타의 설계가 학습자의 심리적 반응에 얼마나 중요한지를 시사한다. Herbert 와 Dożycka (2024)의 연구 역시 인공 교육 에이전트와 시각적 아바타의 사용이 학습 성과, 인식 및 만족도에 미치는 영향을 다루며, 아바타가 갖는 교수학적 역할에 주목하였다 (Herbert & Dożycka, 2024). 메타버스 환경에서 아바타의 역할을 논한 Islam 과 Wang (2025)의 연구는 AI 기술과 아바타 커스터마이징 옵션이 학습에 미치는 영향을 검토하며 (Islam & Wang, 2025), Sinlapanuntakul 과 Zachry (2025)는 비디오 매개 협력 상호작용에서 아바타 표현의 효과를 다루었다 (Sinlapanuntakul & Zachry, 2025). 이러한 연구들은 아바타의 구현 방식과 상호작용의 질이 사회적 실재감 형성에 중요하며, 이는 학습자의 전반적인 행동 및 기술 수용에 영향을 미치는 Briz-Ponce et al. (2017)의 주장과 맥을 같이 한다 (Briz-Ponce et al., 2017).

### 3.4. 시간적 효율성과 언어적 친밀성의 상호작용 메커니즘

이 연구의 온라인 강의 시스템에 적용한 응답 속도와 언어적 표현은 사회적 실재감과 학습 성과에 영향을 준다. 응답 속도는 학습자와 아바타의 상호작용이 얼마나 빠르고 원활하게 이루어지는가에 대한 지표이고, 언어 스타일은 학습자와 아바타의 친근감을 형성하는 요소로 작용할 것이다. 질문에 대한 아바타의 답변이 지연될 수록 학습자는 이를 온라인 강의 시스템이 느리게 반응하거나 제대로 동작하지 않는다는 의미로 받아들일 가능성이 높고, 이로 인하여 학습자는 화면 새로 고침을 클릭하거나 답변을 준비하는 중에 질문을 다시 입력하는 등 다시 시도하는 행동을 보일 것이다. 이러한 행동은 학습 내용과 직접 관련 없는 생각이나 판단을 추가로 하게 만들어서 불필요하게 학습자의 인지 부담을 높게 되고,

학습의 흐름을 끊을 수 있다.

Kodani 등(2025)은 이러한 지연이 학습자에게 불편함을 유발하고, 결과적으로 아바타에 대한 부정적 평가로 이어진다고 보고하였다. 또, 기술적 결함으로 인한 불만은 대화체가 갖는 긍정적 효과를 감쇄시키고 사회적 실재감과 학습 성과를 저해하게 될 것이다.

이에 본 연구에서는 응답 속도와 언어 스타일이 함께 작동할 때 학습자의 경험이 어떻게 달라지는지를 보다 구체적으로 살펴보고자 한다. 다만 두 요인이 항상 일관되게 작용하는 것은 아닐 수 있기에, 실제 데이터에서는 다소 복합적인 양상이 나타날 가능성도 염두에 둔다. 그럼에도 불구하고 전반적으로는 두 요인이 적절히 조화를 이루는 조건에서 학습 효과가 가장 두드러지게 나타날 것이라 가정한다.

#### 4. 연구 가설 설정 (Research Hypotheses)

이 연구는 먼저 실시간 질의응답 제공 여부를 기반으로 두 개의 집단으로 나누며, 실시간 질의응답을 제공하는 집단은 2 X 2 요인 설계를 통해 4 개의 집단으로, 총 5 개의 집단을 대상으로 다음의 가설을 검증하고자 한다.

##### 4.1. 상호작용의 기본 효과 및 매개 가설

**H1: 실시간 질의응답을 활용하는 집단은 활용하지 않는 집단보다 학습 성과가 유의미하게 높을 것이다.**

Flow 이론에 따르면 학습자는 활동 수행 과정에서 즉각적이고 명확한 피드백을 제공받을 때 몰입 상태를 유지하며, 이는 심층적 정보 처리와 학습 성과 향상으로 이어진다(Csikszentmihalyi, 1991; Sherry, 2004). 온라인 학습에서의 실시간 질의응답 기능은 학습자가 강의 중 잘 이해하지 못한 내용을 질문하여 해결할 수 있다는 점에서 학습의 흐름이 끊기지 않게 막아주고, 학습자가 온라인 강의에 집중할 수 있도록 기능을 한다. 또한 인지부하 이론에 따르면, 미해결 질문이나 이해 실패로 인해 발생하는 외재적 인지부하(Extraneous Cognitive Load)는 학습 효율을 저해하는 주요 요인이다(Sweller, 1988; Mayer, 2021). 실시간 질의응답을 제공함으로써 이런 불필요한 인지부하를 낮춰줌으로써, 학습자가 오로지 강의를 듣는 것에 집중할 수 있도록 해준다. 따라서 실시간 질의응답을 활용하는 집단은 그렇지 않은 통제집단에 비해 더 높은 학습 성과를 보일 것이라는 가설을 설정하였다.

**H2: 실시간 질의응답 기능은 몰입과 사회적 실재감에 의해 학습 성과에 정적(+)으로 매개되고, 인지부하에 의해 부적(-)으로 매개될 것이다.**

Flow 이론은 즉각적인 피드백과 과제 집중이 몰입 상태를 유발하며, 이러한 몰입 경험이 학습 성과를 증진시키는 핵심 메커니즘임을 제시한다(Csikszentmihalyi, 1991). 학습자가 강의 중 갖는 의문점에 대해 실시간 질의응답을 활용함으로써 답변을 구할 수 있고, 이를 통해 학습에 오로지 전념할 수 있도록 만들어 학습 성과에 긍정적인 효과를 미칠 것으로 기대한다.

한편, 사회적 실재감 이론에 따르면 학습 환경에서 상호작용 파트너의 사회적 존재감이 높게 지각될수록 학습자는 더 높은 참여도와 정서적 유대를 형성하게 되며, 이는 학습 만족도와 성취로 이어진다(Biocca et al., 2003). 실시간으로 질문에 대해 답변하는 AI 아바타를 통하여 학습자는 온라인 강의 시스템을 단순한 정보 전달 도구가 아닌 학습자와 상호작용하는 개체로 인식하도록 만들어 줌으로써 사회적 실재감을 증가시킬 가능성이 있다.

반대로 인지부하 이론에 따르면, 질문 미해결 상태나 반응 지연으로 인해 발생하는 외재적 인지부하는 학습 자원의 비효율적 소모를 초래하여 학습 성과를 저해한다(Sweller, 1988; Mayer, 2021). 실시간 질의응답 활용은 이런 불필요한 인지부하를 감소시킬 것이고, 이 감소 효과는 학습 성과에 대한 부적(-) 매개 경로로 동작할 것으로 기대된다.

따라서 이 연구는 실시간 질의응답 사용 여부가 학습 성과에 미치는 영향을 몰입과 사회적 실재감의 정적 매개 효과, 인지부하의 부적 매개 효과를 통해 설명하는 매개 모형을 가설로 설정하였다.

## 4.2. 응답 속도 및 언어스타일의 주 효과 가설

**H3: 응답 지연 시간이 짧을수록 학습자의 몰입도와 사회적 실재감이 높아져서 학습 성과가 향상될 것이다.**

Flow 근거 이론에 따르면 학습자는 과제 수행 과정에서 즉각적이고 명확한 피드백을 제공받을 때 몰입 상태를 유지하게 되며, 이는 과제 집중도와 심층적 정보 처리 수준을 높이는 핵심 요인으로 작용한다(Csikszentmihalyi, 1991).

온라인 강의 환경 위에서 즉각적인 피드백을 구현하기 위한 방법이 바로 시스템의 응답 속도를 줄이는 것이며, 응답 시간이 짧을수록 학습자는 실시간으로 상호작용하고 있다고 인식하게 될 것이다. 이러한 반응의 신속함은 학습자가 아바타를 적절한 상호작용 개체로 인식하도록 만들게 되고, 온라인 강의 시스템의 신뢰성과 능력에 대한 긍정적 평가를 강화할 것이다.

**H4: 대화체는 격식체보다 사회적 실재감을 높이고, 인지부하를 낮추어 학습 성과를 향상시킬 것이다.**

Mayer(2021)의 개인화 원리를 살펴보면, 대화체를 사용하여 강의를 진행할 경우 학습자의 주의 집중과 의미 구성 과정을 촉진하는 방향으로 작용하게 된다. 대화체는 인간과 상호작용을 연상시키는

사회적 단서를 제공함으로써 학습자의 사회적 스키마를 활성화하고, 시스템을 단순한 정보 전달 도구가 아닌 나와 상호작용을 하는 개체로 인식하게 만든다(Mayer, 2021). 이런 인식의 변화로 인하여 학습자와 아바타의 사회적 실재감을 증가시키는 요인으로 작용할 가능성이 크다.

또한 인지부하 이론에 따르면, 정보 제시 방식이 학습자의 처리 용량을 불필요하게 소모할 경우 외재적 인지부하가 증가하여 학습 효율이 저하된다(Sweller, 1988). 격식체 표현은 내용 이해보다는 문서 또는 강의 자료를 읽는 것으로 인식하는 것과 같이 말의 형식 또는 표현 방식에 더 신경을 쓰게 만들 수 있기 때문에 학습 중 주위가 흐트러질 가능성이 있다. 반면에 대화체 표현의 경우 이와 반대로 친숙하고 이해하기 쉬운 언어를 사용하기 때문에, 강의 내용을 듣고 이해하는데 필요한 부담을 줄여줄 것이다(Mayer, 2021). 이에 따라 대화체 언어 사용이 격식체에 비해 학습자에게 더 많은 사회적 실재감을 제공하고 인지적 부담을 낮춤으로써 이러한 요인이 온라인 강의에 함께 작용할 경우 학습 성과를 높일 것이라는 가설을 설정하였다.

### 4.3. 상호작용 효과 가설

#### **H5: 응답 속도(빠름 vs 느림)와 언어스타일(대화체 vs 격식체) 사이에는 유의미한 상호작용 효과가 존재할 것이다.**

Mayer(2021)는 설명할 내용을 딱딱하게 전달하기보다는, 마치 상대방과 대화하듯이 말하는 경우에 학습자가 시스템을 단순하게 정보를 제공하는 도구가 아니라, 나와 함께 소통하는 대상으로 인식하게 된다고 설명하였다. 이 관점에서 보면, 학습자가 질문을 했을 때 대화체로 빠르게 답변을 제공받는 경우 기존의 일반적인 온라인 강의보다 강의에 더 집중하기 쉬울 것이다. 예를 들어, 강의를 듣다가 이해하지 않는 부분이 생겼을 때 바로 질문을 했고, 아바타가 “좋은 질문입니다. 다시 설명해 볼게요”라고 하면서 답변한다면 학습자는 강의를 혼자 듣고 있다는 느낌보다 누군가와 함께 공부하고 있다고 여길 것이다. 이러한 환경에서는 학습 과정에서 생기는 부담도 줄어들 가능성이 있다. 질문에 대한 답을 기다리느라 강의를 멈추거나, 인터넷 검색을 따로 하지 않아도 되기 때문이다. 실제로 온라인 강의를 들으면서 답변이 늦어지면 시스템에 문제가 생긴 것은 아닌지 생각하게 되고 화면을 새로 고침하거나 강의에 대한 집중을 잃어버리고 컴퓨터가 정상인지 확인하게 된다. 반대로 질문에 바로바로 답이 오는 경우는 강의 흐름이 끊기지 않아 자연스럽게 집중을 유지할 수 있다. 또한 질문에 빠르게 반응하는 시스템은 만족도와 신뢰도를 함께 높일 수 있다. 여기에 추가로 대화체 언어 스타일을 함께 온라인 강의에 적용할 경우 학습자는 아바타를 딱딱한 설명 도구가 아니라 함께 소통하는 존재로 느끼게 돼서 사회적 실재감이 높아질 수 있다. 이렇듯 이 두 요소가 복합적으로 작동하는 경우에 있어 학습자는 친근하면서도 믿을 수 있는 존재로 인식하게 되면서 이는 결과적으로 학습 성과를 높이는데 긍정적인 영향을 줄 수 있을 것이라

예상한다.

## 5. 연구 방법론 (Proposed Research Methodology and Design)

### 5.1. 연구 설계 및 참여자 모집

이 연구는 총 52 명(남성 18 명, 여성 32 명, 성별 무응답 2 명) 으로 진행되었다. 실험자 연령대는 20 대에서 40 대에 이르는 대학생, 대학원생 및 일반인 52 명을 대상으로 통제된 실험 환경에서 수행되었다. 실험에서 사용한 온라인 강의의 주제인 사회조사 연구방법론에 대한 사전 학습 경험이 있는 참여자는 19 명(36.5%), 없는 참여자는 33 명(63.5%)이었다. 연구의 전 과정은 연구 윤리 규정을 엄격히 준수하여 진행되었다. 실험 설계는 두 개의 독립 변인을 조작한 2 X 2 요인 설계를 적용하였으며, 통제집단을 포함하여 총 5 개의 실험 집단을 설정하여 진행하였다.

#### 5.1.1 연구 설계: 5 개 집단 비교 설계

총 다섯 개의 실험 집단을 설정하여, Q&A 시스템을 일체 제공하지 않는 통제집단을 기준으로 네 가지 실험집단(Q&A 제공)을 비교 분석하였다.

**Table 1. Experimental group classification**

Group Number	Group Type	Q&A Provision	Response Speed	Language Style
Control Group	Q&A Not Provided	Not Provided	N/A	N/A
Experimental Group 1	Q&A Provided (Fast/Conversational)	Provided	Fast (within 3.0 seconds on average)	Conversational
Experimental Group 2	Q&A Provided (Fast/Formal)	Provided	Fast (within 3.0 seconds on average)	Formal
Experimental Group 3	Q&A Provided (Slow/Conversational)	Provided	Slow (6.0 seconds or	Conversational

			longer on average)	
Experimental Group 4	Q&A Provided (Slow/Formal)	Provided	Slow (6.0 seconds or longer on average)	Formal

### 5.1.2 참여자 모집 및 할당

통계적 검정력을 확보하기 위해 총 52 명의 참여자를 모집하였으며, 이들을 5 개의 실험 집단에 무작위 배정(random assignment)하였다. 수집된 데이터는 통제집단과 4 개의 실험집단 간의 학습 성과 및 심리적 변인의 차이를 규명하기 위해 분석되었으며, 가설 검증을 위해 적절한 통계적 분석 방법을 적용하였다.

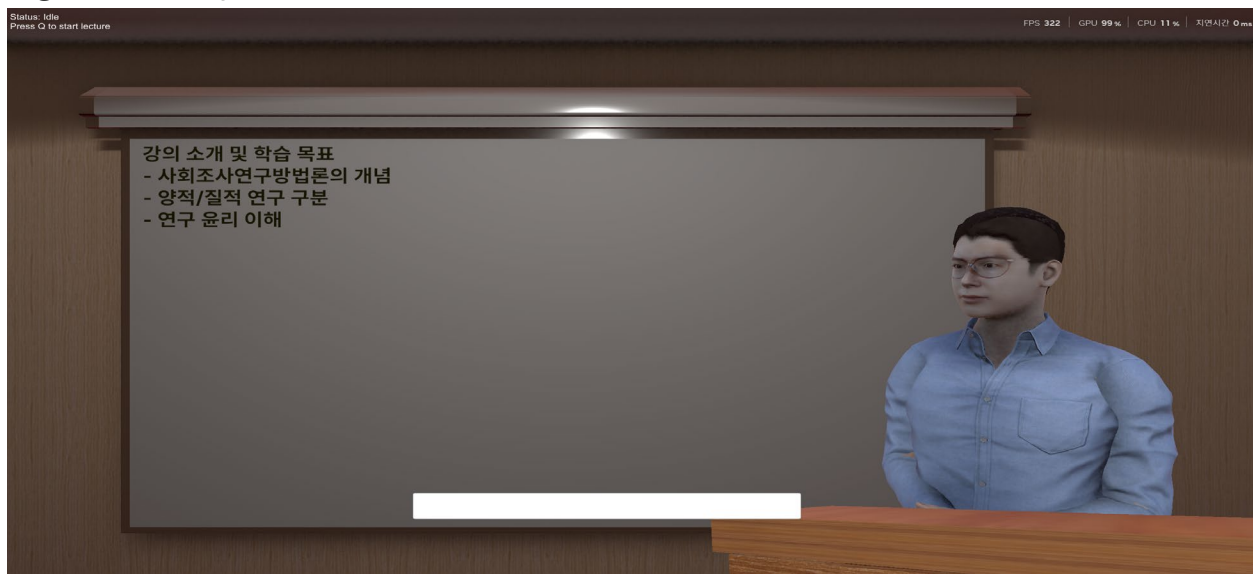
### 5.1.3 실험 절차 및 조작

실험을 위한 온라인 동영상 강의는 유니티(Unity) 6000.2 와 인공지능 기술을 기반으로 자체 개발하였으며, 강의의 주제는 사회조사연구방법론을 주제로 약 7 분 분량의 강의로 구성하였다. 학습자는 아바타와 음성을 통해 질문을 하거나, 키보드를 이용하여 질문을 입력하는 방식으로 질의응답을 수행할 수 있도록 구현하였다. Figure1 을 보면 하단에 하얀색 박스를 볼 수 있는데, 이 부분이 텍스트를 통해 질의할 수 있다. 실험은 시스템 응답 속도와 언어 스타일을 독립 변인으로 하여 조작하였다. 응답 속도는 빠름(Fast) 조건의 경우 평균 3.0 초 이내로 유지하였으며, 느림(Slow) 조건은 평균 6.0 초 이상으로 인위적으로 지연시켰다. 언어스타일은 친근감을 주는 대화체(해요체) 조건과 격식을 갖춘 격식체(합쇼체) 조건으로 명확히 구분하여 적용하였다. 아바타는 Reallusion 의 CC5(Character creator5)를 이용하여 생성하였으며, 시스템의 인공지능 모듈은 Google TTS(Speech-to-Text)와 ElevenLabs 의 TTS(Text-to-Speech)를 활용하였으며, 추론 모델로는 Gemini-flash-latest LLM 을 탑재하였다. 특히, 고성능 추론 모델의 경우 높은 정확도를 보이나, 실시간 상호작용에 필수적인 응답 속도(Latency) 확보를 위해 추론 시간이 빠른 Flash 모델을 채택하였다.

본 실험에 앞서 수행된 10 명의 조작 점검을 위한 테스트 결과를 바탕으로 실험 환경을 개선하였다.

먼저 잘못된 음성 인식으로 인해 엉뚱한 답변을 하는 경우를 대비하여 채팅을 통한 질의가 가능하도록 인터페이스를 개선하였으며, 학습 내용의 이해를 돕기 위해 화면 내 강의의 내용을 간단하게 기술하였다(Figure 1). 또한 응답 속도 조작의 타당성은 시스템 로그를 통해 객관적으로 확인하였다. 빠른 조건의 실제 평균 응답 시간은 약 3 초 이내, 느린 조건은 6 초 이상으로 측정되어 두 조건 간 실험 의도가 정확히 구현되었음을 확인하였다. 언어 스타일 조작은 대화체(해요체)와 격식체(합쇼체)의 프롬프트 설계 및 사전 검토를 통해 확인하였다. 다만 본 실험의 참여자에게 고지는 하였으나, 실제로 참여자가 시각에 기반한 주관적 조작 점검 설문을 별도로 실시하지는 않았다.

**Figure 1. Unity Experimental Environment**



## 5.2. 독립 변인의 조작 및 구현

### 5.2.1. 실시간 질의응답 제공 여부에 기반한 실험

이 연구의 실험 조건은 학습 중 상호작용이 가능한지 여부에 따라 통제집단과 실험집단으로 나뉜다. 먼저 통제집단이 속한 참여자들은 실시간 질의응답 기능이 없는 환경에서 강의를 듣도록 하였다. 즉, 일반적인 온라인 강의처럼 영상을 시청만 하는 것이다. 실험을 시작하기 전에 저자는 이 강의를 듣다가 궁금한 사항이 생길 경우 강의를 중지하고, 인터넷 검색 또는 스마트기기를 이용해서 검색하라고 미리 안내하였고 참여자들도 이를 이해하고 실험에 참여하였다. 반면 실험집단은 강의를 들으면서 질문을 할 수 있고, 질문을 하면 AI 아바타가 이를 인식해 답변을 해 주는 환경에서 학습을 진행하였다. 예를 들어 강의 중간에 질적 연구에 대해 자세하게

설명해주세요 라고 질문할 경우 아바타가 음성으로 설명하는 방식이다. 이 실험 집단은 다시 네 가지 조건으로 나뉘었는데, 이는 아바타가 응답 속도가 빠르지 느리지, 그리고 말투가 대화체인지 격식체인지에 따라 달라진다. 실험 집단의 경우 실제 실험에 들어가기 전에 음성으로 질의를 하는 연습을 통해 질의하는 과정에 익숙해진 상태로 만들고 실험을 진행하였다.

### 5.2.2. 응답 속도 차이에 따른 실험

이 연구에서 독립 변인으로 설정한 시스템 응답 속도는, 학습자가 질문을 했을 때 AI 아바타가 이를 인식한 순간부터 실제로 답변을 보여주기기까지 걸리는 시간을 의미한다. 쉽게 말해, 질문을 던진 뒤 얼마나 빨리 대답해주는지에 관한 것이다. 본 실험에서는 이 응답 속도를 빠름과 느림의 두 가지 조건으로 나누어 조작하였다.

먼저 빠름 조건은 질문을 한 뒤 평균적으로 3 초 이내 답변이 나오도록 설정하였다. 이를 위해 Google Gemini-flash-latest LLM 을 사용하여, 응답에 대한 답변이 바로 생성되도록 하였으며, 문장이 생성되는 시간도 줄이기 위해, 파이썬을 이용하여 문장을 끊어서 짧게 TTS 로 변환시켜 재생시키는 방법으로 마치 즉시 답변하는 것처럼 인식하도록 구현하였다. 이는 질문을 하고 잠깐 기다리면 바로 답이 나오는 정도의 속도로, 일상적으로 메신저에서 상대방이 바로 답장해 주는 상황과 비슷하다. 예를 들어 강의를 듣다가 잘 모르는 개념에 대해 질문했을 때, 화면을 바라보고 있는 사이에 아바타가 바로 설명을 시작하게 된다. 이런 정도의 응답 속도에서는 학습자가 기다리고 있다는 느낌보다는 자연스럽게 대화가 이어진다는 인상을 받기 쉽다.

반대로 느림 조건은 평균적으로 6 초 이상이 지나서 답변이 나오도록 조작하였다. 6 초는 숫자로 보면 짧아 보일 수 있지만 실제로 질문을 던지고 아무 반응 없이 화면을 보고 있으면 생각보다 길게 느껴지는 시간이다. 간단하게 예를 들어, 아무 것도 하지 않은채 1 부터 6 까지 숫자를 세본다면 이 시간이 얼마나 긴지 느낄 수 있을 것이다. 이렇듯 질문을 한 후 몇 초 동안 아무런 변화가 없으면, 학습자는 제대로 질문한 것이 맞나 또는 내 노트북에 문제가 생겼나 와 같은 생각을 하게 될 수도 있다. 이 과정에서 강의 내용에 대한 집중은 자연스럽게 끊길 수 있다.

본 연구에서는 이러한 두 조건을 통해 응답 속도가 학습자의 몰입, 인식, 학습 경험에 어떤 차이를 만들어 내는지를 비교하고자 하였다.

### 5.2.3. 언어스타일 차이에 따른 실험

이 연구에서 언어스타일은 아바타가 강의할 때와 질문에 대한 답을 할 때 어떤 말투와 표현 방식을 사용하는지를 의미한다. 본 연구에서는 Mayer 의 개인화 원리를 참고하여, 아바타의 언어 스타일을

대화체와 격식체 두 가지로 나누어서 설정하였다.

먼저 대화체 조건은 학습자와 아바타 사이의 거리감을 줄이고, 보다 친근한 분위기를 만들기 위해 사용되었다. 이 조건에서는 ~요, ~에요 와 같은 해요체를 사용하였다. 예를 들어 학습자가 삼각 측량법에 대해 질문을 할 경우 하나의 대상이나 현상을 여러 가지 방법이나 자료로 동시에 분석해서 연구 결과의 신뢰성을 높이기 위한 방법이에요 라고 답변하도록 만들었다.

반면 격식체 조건은 정보를 차분하고 공식적으로 전달하기 위해 집중하였다. 이 조건에서는 문장 끝을 ~입니다. ~이다와 같은 형태로 구성하여 전반적으로 딱딱하고 사무적인 느낌을 유지하였다. 예를 들어, 같은 질문에 대해 하나의 대상이나 현상을 여러 가지 방법이나 자료로 동시에 분석해서 연구 결과의 신뢰성을 높이기 위한 방법입니다 라는 방식으로 답변을 제공하도록 설계하였다. 이는 교과서나 공식 강의 자료를 읽는 것과 비슷한 인상을 준다.

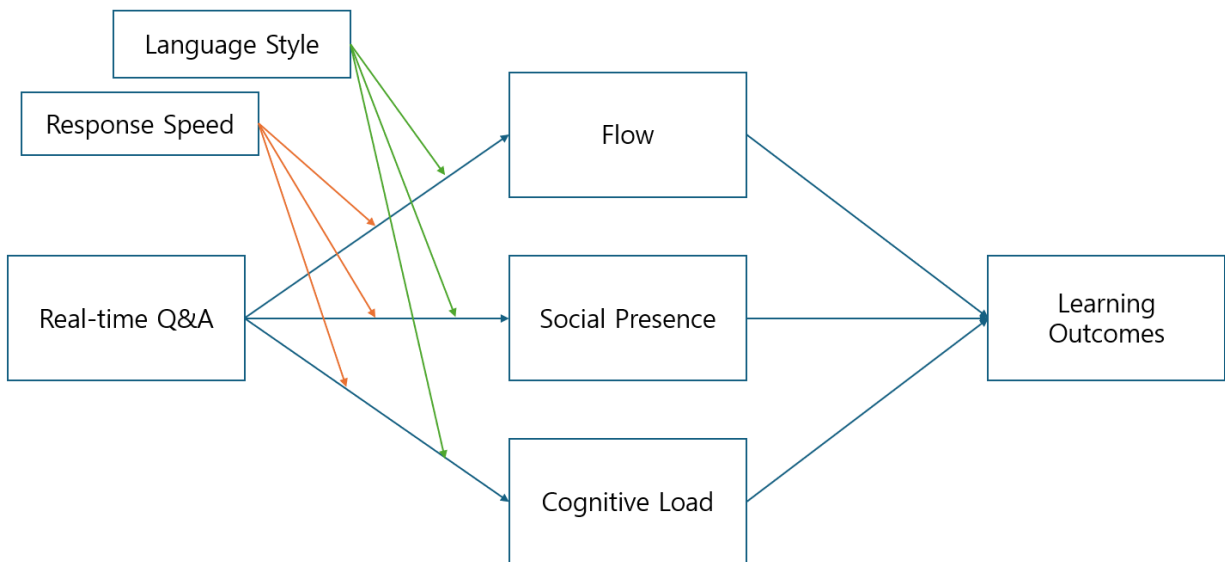
이처럼 본 연구에서는 말투와 표현 방식의 차이만으로도 학습자가 느끼는 친숙함이나 상호작용의 인식이 어떻게 달라지는지 관찰하고자 하였다.

### 5.3. 종속 및 매개 변인의 측정 (Measurement)

이 연구에서는 실험이 끝난 직후 참여자들에게 설문조사를 실시하였으며, 학습 중 느낀 경험에 대해 응답하도록 하였다. 몰입도, 사회적 실재감, 인지부하와 같은 심리적 요인들은 모두 자기보고식 설문 문항으로 측정하였다. 각 문항은 5 점 리커트 척도로 구성되었으며, 참여자들은 전혀 그렇지 않다(1 점) 부터 매우 그렇다(5 점)까지의 범위 내에서 자신의 느낌에 가장 가까운 값을 선택하였다. 예를 들어 강의에 얼마나 집중했다고 느꼈는가? 또는 아바타가 실제로 나와 소통하고 있다고 느꼈는가와 같은 질문에 점수로 답하는 방식이다.

객관적인 학습 성과는 설문이 아니라 간단한 지식 테스트를 통해 측정하였다. 실험 전과 실험 후 비슷한 난이도로 설정된 문제를 풀게 하고, 두 점수의 차이를 계산하여 학습 성과의 변화량을 측정하였다. 이를 통해 단순한 만족도 뿐만 아니라 실제로 학습 내용이 얼마나 이해되었는지를 살펴보고자 하였다. 또한 설문 점수만으로는 알기 어려운 부분을 보완하기 위해서 하나의 주관식 문항을 포함하여, 참여자들이 실험을 위해 구축한 강의를 체험하면서 느낀점을 자유로이 기술할 수 있도록 하였다.

Figure2 은 이 실험의 연구 모형을 보여준다. 실시간 질의응답 여부를 독립 변인으로 하고, 응답 속도와 언어 스타일이 각각 몰입, 사회적 실재감, 인지부하에 영향을 주고, 이러한 요인들이 다시 학습 성과로 이어지는 구조를 확인할 수 있다. 이 그림에 나타난 각 심리적 변인들은 설문과 주관식 응답을 통해 함께 측정함으로써, 실험 조건에 따라 학습자가 실제로 어떤 경험을 했는가를 보다 자세히 분석하고자 하였다.

**Figure 2. Research model**

#### 5.4. 데이터 분석 방법

이 연구는 실시간 질의응답 기능을 제공하는 여부와 실시간 질의응답 기능을 제공하는 집단의 경우 응답 속도, 언어 스타일의  $2 \times 2$  요인 설계를 적용하여, 한 개의 통제 집단과 4 개의 실험 집단의 실험 결과를 종합적으로 검증하고자 하였다. 전처리는 엑셀을 사용하여 수행하였고, 이후 모든 통계 분석은 IBM SPSS Statistics 31.0.1.0 을 사용하였다.

##### 5.4.1. Q&A 제공 효과 검증 (H1)

가설 1(H1)인 ‘실시간 Q&A 시스템 제공 유무에 따른 학습 성과의 차이’를 검증하기 위해 집단을 재구성하였다. 전체 참여자는 Q&A 시스템을 제공받지 않은 통제집단과, 시스템을 활용한 네 개의 실험집단을 포함하는 Q&A 제공 집단으로 범주화 하였다.

두 집단 간 학습 성과의 평균 차이를 통계적으로 검증하기 위해 독립표본 t-검정(Independent Samples t-test)을 실시하였으며, 추가적으로 집단 간 변동성을 보다 포괄적으로 확인하기 위해 일원분산분석(one-way ANOVA)을 병행하였다. 이를 통해 실시간 Q&A 제공 여부가 학습 성과에 미치는 주효과의 존재 여부를 다각적으로 검증하고자 하였다.

### 5.4.2. 요인 및 상호작용 효과 검증 (H3, H4, H5)

시스템의 특성이 학습 성과와 여러 심리적 요인에 어떤 영향을 미치는지 자세히 살펴보기 위해, 본 분석에서는 통제집단을 제외하고 네 개의 실험집단만을 대상으로 분석을 진행하였다. 이 네 집단은 응답 속도가 빠르는지 느린지, 그리고 아바타의 언어스타일이 대화체인지 격식체인지에 따라 나뉜 집단들이다. 구체적으로 응답 속도(빠름/느림)와 언어 스타일(대화체/격식체)에 따른 주 효과(H3, H4)와 두 변인 간의 상호작용 효과(H5)를 검증하고자 2 × 2 이원분산분석(Two-way ANOVA)을 실시하였다.

### 5.4.3. 매개 효과 검증 (H2)

가설 2(H2)에서는 독립변인(Q&A 제공 유무, 응답 속도, 언어 스타일)이 심리적 변인(몰입, 사회적 실재감, 인지부하)을 매개하여 학습 성과에 미치는 간접 효과를 검증하였다. 매개 효과가 유의한지를 확인하기 위해 부트스트래핑 기법을 사용하였다. 이 부트스트래핑은 수집한 데이터가 꼭 정규분포를 따른다는 가정을 하지 않아도 분석을 할 수 있다는 점에서, 이 실험의 경우와 같이 표본 수가 많지 않은 경우에도 상대적으로 안정적인 결과를 얻을 수 있는 장점이 있다. 특히 간접 효과처럼 한 번 더 거쳐서 나타나는 효과를 확인하는데 적합한 분석 방법이라고 판단되어 선정하였다.

분석에는 95% 신뢰구간을 기준으로 결과를 해석하였다. 결과 분석 방법에 대해 기술하자면 부트스트래핑으로 계산된 신뢰구간 안에 0 이 있는가를 확인하여 매개 효과의 유의 여부를 판단할 수 있다. 측정된 신뢰구간에 0 이 포함되지 않는 경우 매개 효과가 통계적으로 의미가 있는 것으로 해석할 수 있다.

## 6. 실험 결과

본 장에서는 AI 아바타 기반 온라인 강의에서 상호작용 방식이 학습 성과와 사용자 경험에 미치는 영향을 정량·정성 자료로 보고한다. 정량 자료는 (1) 지식 테스트(사전·사후)로 측정한 학습 성과(ScoreDiff)와 (2) 설문으로 측정한 몰입(FlowAvr), 사회적 실재감(SocialPresenceAvr), 인지부하(CognitiveLoadAvr), 만족(SatisfactionAvr), 주관적 성취감(GradeFeelAvr), 전반적 UX(UXAvr)로 구성된다. 정성 자료는 자유서술 응답을 의미 단위로 분절하여 범주화한 결과로 제시한다. 본 장의 구성은 다음과 같다. 7.1 에서는 실시간 Q&A 제공 여부가 학습 성과에 미치는 효과(H1)를 검증한다. 7.2 에서는 Q&A 제공 조건 내에서 응답 속도(빠름/느림)와 언어 스타일(대화체/격식체)이 학습 성과와 정의적 경험에 미치는 주효과 및 상호작용 효과(H3-H5)를 이원분산분석으로 검증한다. 7.3 에서는 매개/조절된 매개 분석(H2)을 결과 중심으로 정리한다. 7.4 에서는 조건별 설문 평균 패턴을

표 중심으로 상세 보고한다. 7.5 에서는 정성 의견을 범주별로 확장하여 정량 패턴을 보강한다. 7.6 에서는 정량·정성 자료를 통합하여 결과의 의미를 요약한다.

## 6.1 H1 - 실시간 질의응답 제공 효과 검증

### 6.1.1 데이터 구조 및 변수 정의

이 연구의 전체 참여자는 N=52 명이며, 조건은 통제집단 1 개와 실험집단 4 개로 구성된다. 통제집단은 실시간 Q&A 를 제공하지 않는 일반 강의(온라인강의 시청) 조건이며, 실험집단은 Q&A 제공 조건에서 응답 속도(빠름/느림)와 언어 스타일(대화체/격식체)를 결합한  $2 \times 2$  조건이다. 즉, 실험집단은 B(대화체-빠름), C(격식체-빠름), D(대화체-느림), E(격식체-느림)으로 구분된다. 집단별 표본수는 통제집단 11 명, 실험집단 각 10~11 명으로 비교적 균형에 가깝다.

학습 성과는 사전 지식 테스트 점수(PreTest)와 사후 지식 테스트 점수(PostTest)로 측정했으며, ScoreDiff 는 PostTest-PreTest 로 산출하였다. ScoreDiff 는 “실험 전후 성취 변화”를 나타내는 핵심 종속변수이다. 정의적 경험은 설문 평균 지표(몰입[Flow], 사회적 실재감[Social Presence], 인지 부하[CognitiveLoad], 강의 만족도[Satisfaction], 주관적 학습 성취감[GradeFeel], UIUX 만족도[UX] 등)로 요약하였다.

### 6.1.2 기술 통계량

집단별 평균을 보면, 통제집단은 PreTest 평균 12.45, PostTest 평균 13.00 으로 0.55 점 상승하였다. 실험집단에서는 모든 조건에서 평균 상승이 관찰되며 상승폭은 0.64~1.50 범위이다. 이를 단순 평균 비교 관점에서 보면, Q&A 제공 조건이 통제조건보다 향상폭이 큰 방향성을 보이지만, 이것이 통계적으로 유의한지 여부는 별도의 추론통계로 확인해야 한다. Table 2, Table 3 은 5 개 집단의 핵심 기술통계를 한 표에 모은 것이다.

Table 2, Table 3 은 결과 해석에서 중요한 두 가지 분화를 보여준다. 첫째, 학습 성과(ScoreDiff)는 E 조건이 평균 1.50 으로 가장 큰 상승을 보이며, B 조건이 1.10 으로 그 뒤를 잇는다. 둘째, 정의적 경험(몰입, 실재감, 인지부하, 강의 만족도, 주관적 학습 성취감)은 B 조건이 전반적으로 가장 우수한 프로파일을 보인다. 특히 B 조건의 경우 몰입 점수가 4.08 로 가장 높게 나타났고, 사회적 실재감 역시 3.68 로 네 가지 조건 중 가장 높은 값을 보여주었다. 반면에 인지부하는 1.97 로 가장 낮은, 전반적으로

학습자가 느끼는 경험의 질이 가장 좋은 조건이라고 볼 수 있다. 강의 만족도(4.20)와 주관적인 학습 성취감(4.26) 또한 B 집단이 가장 높게 나타났다. 하지만 성취 점수의 변화량이 가장 큰 조건과 몰입, 사회적 실재감, 인지부하 등의 지표들이 가장 높았던 조건이 항상 일치하진 않았다. 즉 학습 성과가 크게 향상되는 조건이 학습자가 가장 좋은 경험을 했다고 느낀 조건과 같지 않을 수 있다는 결과를 보여주었다. 이러한 차이는 7장에서 온라인 강의 상호작용 설계에 대한 시사점으로 심도있게 논의하고자 한다.

### 6.1.3 독립표본 t-검정

H1은 Q&A 제공 여부가 학습 성과에 영향을 미친다는 가설이다. 이를 검증하기 위해 통제집단(A)과 실험집단군(B-E)을 통합하여 ScoreDiff 평균 차이를 비교하였다. 집단통계량은 다음과 같다. 통제집단은 N=11, 평균 0.5455, 표준편차 5.1257이며, 실험집단군은 N=41, 평균 0.9756, 표준편차 2.1271이다. 두 집단의 표준편차가 크게 다르므로 등분산성(분산 동일) 가정이 성립하는지 점검해야 한다.

**Table 2. Descriptive Statistics<sup>1</sup> for Each Group**

Experimental Groups	N	PreTest	PostTest	ScoreDiff
Control Group (A: No Q&A)	11	12.45	13.00	0.55
Group B (Conversational Tone – Fast Response)	10	13.30	14.40	1.10
Group C (Formal Tone – Fast Response)	10	13.20	13.90	0.70
Group D (Conversational Tone – Slow Response)	11	14.55	15.18	0.64
Group E (Formal Tone – Slow Response)	10	14.10	15.60	1.50

Levene 검정에서  $p=.001$  로 등분산 가정이 충족되지 않았다. 따라서 등분산을 가정하지 않는 t-검정 결과를 기준으로 해석하였다.  $t=-.272$ ,  $df=10.940$ ,  $p=.791$ (양측)로 집단 평균 차이는 통계적으로 유의하지 않았다.

**Table 3. Descriptive Statistics<sup>2</sup> for Each Group**

Experimental Groups	Flow	Social Presence	Cognitive Load	Satisfaction	Grade Feel	UX
Control Group (A: No Q&A)	3.29	3.04	2.73	3.82	3.18	3.78
Group B (Conversational Tone – Fast Response)	4.08	3.68	1.97	4.20	4.26	3.88
Group C (Formal Tone – Fast Response)	3.86	2.98	2.53	4.00	3.72	3.38
Group D (Conversational Tone – Slow Response)	3.91	2.96	2.58	4.00	3.67	3.60
Group E (Formal Tone – Slow Response)	3.68	3.44	2.40	3.90	3.78	3.54

이 결과는 이 연구 표본에서 Q&A 제공이 ScoreDiff 를 유의하게 증가시킨다고 결론내리기 어렵다는 것을 의미한다. 다만 평균 수준에서는 실험집단군이 통제집단보다 높으며( $0.9756 > 0.5455$ ), 동시에 통제집단의 변동성이 매우 크게 나타났다(SD 5.1257). 반대로 실험집단군의 변동성은 상대적으로 작다(SD 2.1271). 즉, 본 자료는 “평균 상승 효과의 유의성”은 확보되지 않았지만, Q&A 제공이 학습 성취의 변동성을 낮추는 방향으로 작용했을 가능성을 시사한다. 이 관찰은 7 장에서 ‘안정성’ 관점의 교육 설계 논의로 연결할 수 있다(Table 4, Table 5).

**Table 4. Group Statistics and Independent Samples t-Test for the Effect of Q&A Availability**

	Statistics				
	QA_Group	N	MEAN	SD	SE
Score	Control group	11	.5455	5.12569	1.54545
Diff	Experimental group	41	.9756	2.12706	.33219

**Table 5. Independent Samples t-Test for SDiff**

Test		Value
Levene	Levene's F	11.562
	Levene's Sig	.001
t-test	t	-.272
	Df	10.940
	One-tailed p	.395
	Two-tailed p	.791
	Mean Difference	-.430
	Std.Error Differernce	1.581
95% CI	95% CI Lower	-3.912
	95% CI Uppder	3.051

\* *Note.* Equal variances not assumed based on Levene's test ( $F = 11.562$ ,  $p = .001$ )

#### 6.1.4 사전 학습 경험(PreStudyExp.)에 따른 변화량 분해

Table 6 은 사전 학습 경험(사회조사 연구 방법론 과목 사전 학습 경험) 유무에 따른 ScoreDiff 평균의 분해 결과를 보여준다. 분석 결과, 집단 조건에 따라 상이한 평균 패턴이 관찰되었다. 이 연구에서 PreStudyExp.= 0 은 사전 학습 경험이 없는 집단을, PreStudyExp.=1 은 사전 학습 경험이 있는 집단을 지칭한다.

사전 학습 경험이 없는 집단(PreStudyExp.= 0)에서 ScoreDiff의 집단별 평균은 통제집단 0.50(n=8), B 조건(대화체-빠름) 1.17(n=6), C 조건(격식체-빠름) 0.14(n=7), D 조건(대화체-느림) 0.14(n=7), E 조건(격식체-느림) 1.00(n=5)으로 나타났다. 이 분해에서는 B 조건이 가장 큰 평균 상승폭을 보였으며, E 조건이 그 다음으로 높은 값을 보였다. 반면, C 및 D 조건에서는 평균 상승폭이 상대적으로 매우 제한적으로 나타났다.

사전 학습 경험이 있는 집단(PreStudyExp.=1)에서는 ScoreDiff의 집단별 평균이 통제집단 0.67(n=3), B 조건 1.00(n=4), C 조건 2.00(n=3), D 조건 1.50(n=4), E 조건 2.00(n=5)으로 관찰되었다. 이 분해에서는 C 조건(격식체-빠름)과 E 조건(격식체-느림)이 가장 높은 평균값(각 2.00)을 보였으며, D 조건(대화체-느림)이 1.50으로 그 뒤를 이었다.

이러한 패턴은 학습자의 사전 지식 수준에 따라 상대적으로 “유리하게 작용하는 상호작용 페르소나(응답 속도-언어스타일 조합)”가 달라질 가능성을 시사한다. 즉, 사전 경험이 없는 학습자에게는 대화체-빠름 조건(B)이 보다 효과적으로 작용하는 반면, 사전 경험이 있는 학습자에게는 격식체 조건(C, E)이 상대적으로 더 큰 성취 상승과 연결되는 경향이 관찰되었다.

다만 각 표본 수가 최소 3명에서 최대 7명 수준으로 매우 작아, 이 결과를 통계적으로 확정적인 결론으로 제시하기에는 한계가 있다. 따라서 본 절에서는 이를 탐색적 수준의 “관찰된 경향(observed tendency)”으로만 기술하며, 8.4 한계 및 8.5 후속 연구 제안 절에서 학습자 수준별 맞춤형 상호작용 설계의 필요성을 추가적으로 논의한다.

**Table 6. Descriptive Statistics for ScoreDiff by Experimental Group and Prior Learning Experience**

ExpGroup	PreStudyExp	MEAN	SD	N
Control Group	No Prior Learning Experience	.5000	5.58058	8
	With Prior Learning Experience	.6667	4.72582	3
	Total	.5455	5.12569	11
Conversational Tone - Fast	No Prior Learning Experience	1.1667	3.31160	6

Response	With Prior Learning Experience	1.0000	2.16025	4
	Total	1.1000	2.76687	10
Formal Tone – Fast Response	No Prior Learning Experience	.1429	2.60951	7
	With Prior Learning Experience	2.0000	1.73205	3
	Total	.7000	2.45176	10
Conversational Tone – Slow Response	No Prior Learning Experience	.1429	1.95180	7
	With Prior Learning Experience	1.5000	2.08167	4
	Total	.6364	2.01359	11
Formal Tone – Slow Response	No Prior Learning Experience	1.0000	1.41421	5
	With Prior Learning Experience	2.0000	.70711	5
	Total	1.5000	1.17851	10
Total	No Prior Learning Experience	.5455	3.30805	33
	With Prior Learning Experience	1.4737	2.16970	19
	Total	.8846	2.95492	52

### 6.1.5 이상치(outlier) 민감도 점검

ScoreDiff에서 큰 변화를 보인 값(-8, 6, 7, 11)을 극단값(outliers)으로 간주하고 해당 사례를 제외한 후, 집단별 사전·사후 점수의 평균을 재산출하였다(Table 7). 이는 등분산성 위배 가능성과 평균 왜곡 위험을 점검하기 위한 강건성 점검 절차의 일환이다.

이상치를 제거한 뒤 집단별 사전 테스트 평균을 살펴보면 A 집단은 13.75였고, B 집단은 13.89, C 집단은 13.20, D 집단은 14.55, E 집단은 14.10으로 나타났다. 사후테스트 평균은 A 집단 13.38, B 집단 14.33, C 집단 13.90, D 집단 15.18, E 집단 15.60으로 측정되었다. 사전·사후 점수 차이를 기준으로 살펴보면, A 집단은 -0.37로 점수가 소폭 감소하였으며, 반면에 B-E 집단은 0.44에서 1.50 사이의 점수 증가를 보여주었다. 즉, 이상치를 제외한 경우에도 실시간 질의응답을 사용한 집단이 전반적으로 더 긍정적인 학업 성취도를 보여주었다고 할 수 있겠다. 특히 E 집단의 평균 변화량은 1.50으로 가장 컸고, 가장 클 것이라 예상했던 B 집단 역시 감소 없이 양의 변화량을 유지하였다.

다만, 이상치 제거 후 집단별 표본수가 8~11명 수준으로 제한되어 있어(통제집단 8명, B 9명, C 10명, D 11명, E 10명), 본 결과를 통계적으로 확정적인 효과로 해석하기에는 한계가 있다. 따라서 본 분석은 주 효과 검증을 대체하기보다는, 결과의 방향성과 안정성을 확인하기 위한 강건성 점검(robustness check) 차원의 참고 결과로 해석하는 것이 타당하다.

**Table 7. Group-wise Pre-test and Post-test Means and Differences**

Group	Pre-test Mean	Post-test Mean	Difference (Post - pre)
Control Group (A: No Q&A)	13.75	13.38	-0.37
Group B (Conversational Tone - Fast Response)	13.89	14.33	0.44
Group C (Formal Tone - Fast Response)	13.20	13.90	0.70
Group D (Conversational Tone - Slow Response)	14.55	15.18	0.63
Group E (Formal Tone - Slow Response)	14.10	15.60	1.50

## 6.2 응답 속도 및 언어 스타일 효과 검증(H3-H5)

본 절에서는 통제집단을 제외한 Q&A 제공 조건(N=41)에 대해, 응답 속도(ResponseSpeed: 빠름/느림)와 언어 스타일(LanguageStyle: 대화체/격식체)의 주효과 및 상호작용 효과를 검증한다. 분석은  $2 \times 2$  이원분산분석(two-way ANOVA)로 수행하였다. 각 종속변수에 대해 (1) 평균과 표준편차 (2) 등분산성 점검(Levene) (3) 분산분석 결과(F, p, 부분 에타 제곱) (4) 결과 해석(유의성 및 효과크기 중심)을 순서대로 기술한다.

### 6.2.1 학습 성과(ScoreDiff)에 대한 이원분산분석

$2 \times 2$  이원 분산 분석을 진행하기 전에, 집단 간의 사전 점수의 동질성 여부를 알아보기 위해 일원분산분석을 실시하였다. 등분산성 점검(Levene) 결과 집단 간 분산의 동질성 가정이 유의하지 않아( $F(4, 47) = 7.54, p < .001$ ), 등분산 가정을 전체하지 않는 Welch's ANOVA 를 적용하였다. 분석 결과 집단별 PreTest 평균은 통제집단이 12.5 로 가장 낮았고, 대화체, 느린 응답의 집단이 14.5 로 가장 높았다. 집단 간 차이는  $p = .477$  로 유의하지 않았다. 이를 통해 실험 처치 이전 집단 간 사전 지식 수준이 동일함을 확인하였다( $F(4, 22.5) = 0.906, p = .477$ ).

ScoreDiff(사전-사후 점수 변화량)에 대해 응답 속도(빠름/느림)와 언어 스타일(대화체/격식체)의 주효과 및 상호작용 효과를 검증하기 위해  $2 \times 2$  이원분산분석을 실시하였다(Table 9). 분석에 앞서 등분산성 가정을 점검한 결과, Levene 검정은  $p = .362$  로 유의하지 않아 등분산 가정은 충족되는 것으로 확인되었다(Table 10).

표본수는 빠름-대화체 10 명, 빠름-격식체 10 명, 느림-대화체 11 명, 느림-격식체 10 명이었으며(Table 8), 평균은 각각 1.10(B), 0.70(C), 0.64(D), 1.50(E)로 나타났다. 표준편차는  $B = 2.77, C = 2.45, D = 2.01, E = 1.18$  로 확인되었다.

이원분산분석 결과, 응답 속도의 주효과는  $F(1, 37) = 0.061, p = .807$ , 부분 에타 제곱  $\eta^2 = .002$  로 유의하지 않았고, 언어 스타일의 주효과 역시  $F(1, 37) = 0.116, p = .736, \eta^2 = .003$  으로 유의하지 않았다. 또한 응답 속도와 언어 스타일의 상호작용 효과도  $F(1, 37) = 0.859, p = .360, \eta^2 = .023$  으로 통계적 유의수준에 도달하지 못하였다(Table 9). 효과크기 관점에서 주효과( $\eta^2 = .002 \sim .003$ )는 매우 작은 수준이며, 상호작용 효과( $\eta^2 = .023$ ) 역시 절대값 기준으로 크지 않은 수준에 해당한다.

다만 기술통계에서의 평균 패턴은 교차 형태에 가깝게 나타났다. 대화체 조건에서는 빠른 응답(B)이

느린 응답(D)보다 높았고( $1.10 > 0.64$ ), 격식체 조건에서는 느린 응답(E)이 빠른 응답(C)보다 높았다( $1.50 > 0.70$ ). 이러한 교차 패턴은 속도 또는 말투의 독립적 효과라기보다, 특정 언어 스타일이 요구하는 상호작용 리듬과 응답 속도의 정합성이 성취 변화와 연결될 가능성을 시사한다.

이 가능성은 사회적 실재감에서 관찰된 상호작용 경향( $p = .051$ )과 함께, 7장에서 AI 아바타 페르소나 설계의 맥락적 일치성(언어 단서-시간 단서 정합성)을 논의하는 근거로 확장하여 해석할 수 있다.

**Table 8. Descriptive Statistics (Means, Standard Deviations) and Cell Sample Sizes for ScoreDiff in the  $2 \times 2$  Design (Dependent Variable: ScoreDiff)**

Language Style	Response Speed: Fast Mean(SD), N	Response Speed: Slow Mean(SD), N
Conversational Tone	1.10 (2.77), n=10	0.64 (2.01), n=11
Formal Tone	0.70 (2.45), n=10	1.50 (1.18), n=10

**Table 9. Summary of the  $2 \times 2$  Two-Way ANOVA Results for ScoreDiff**

Effect	F(1,37)	p	Partial Eta Squared ( $\eta^2$ )
Response Speed (Fast vs. Slow)	0.061	0.807	0.002
Language Style (Conversational vs. Formal)	0.116	0.736	0.003
Response Speed $\times$ Language Style	0.859	0.360	0.023

**Table 10. Results of Levene's Test for Homogeneity of Variance**

Test	p
Levene's Test for Homogeneity of Variances	0.362

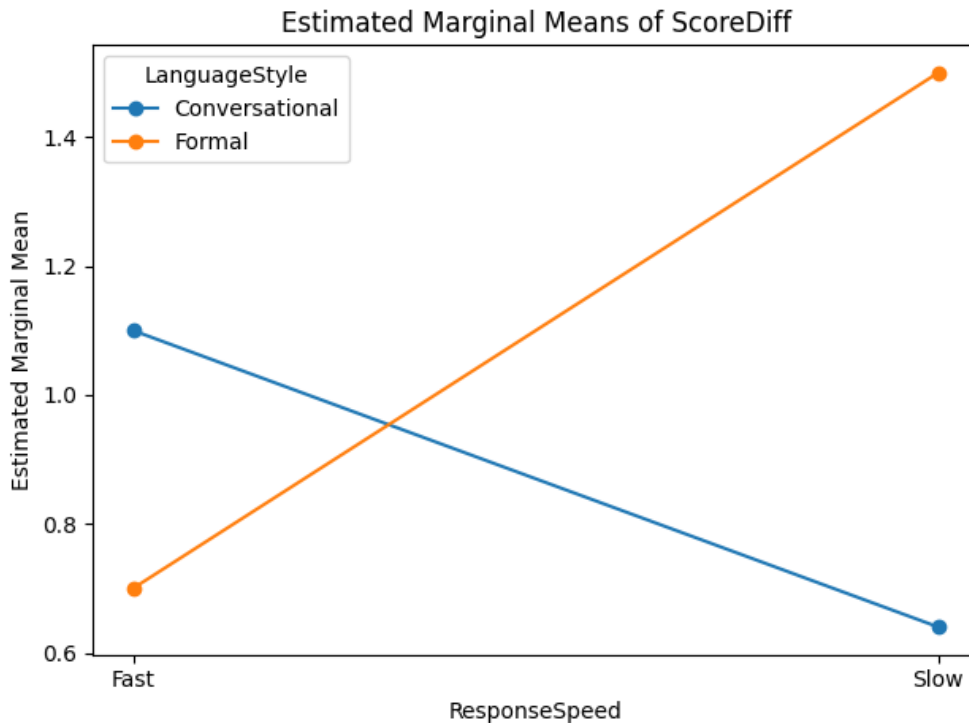
**Figure 3. Estimated Marginal Means of Score Difference(PostTest – PreTest)**

Figure 3은 응답 속도와 언어 스타일이 학습 성과(ScoreDiff)에 미치는 상호작용 양상을 시각화한 것이다. 이원분산분석 결과, 응답 속도의 주효과( $F(1, 37) = 0.061, p = .807$ )와 언어 스타일의 주효과( $F(1, 37) = 0.116, p = .736$ ), 그리고 두 변인의 상호작용 효과( $F(1, 37) = 0.859, p = .360$ ) 모두 통계적으로 유의미하지 않았다(Table 9).

추정 평균을 기준으로 집단별 결과를 살펴보면, 응답 속도와 언어 스타일에 따라 서로 다른 패턴이 나타났음을 알 수 있다. 먼저 대화체 조건에서는 응답이 빠른 경우가 느린 경우에 비해 평균 성취 변화량이 더 크게 나타났다(B 집단 1.10, D 집단 0.64). 이와 반대로 격식체 조건에서는 응답이 느린 경우가 빠른 경우보다 더 높은 성취 변화량을 보여주었다(E 집단 1.50, C 집단 0.70). 즉 응답 속도와 언어 스타일의 효과가 한 방향으로만 나타난 것이 아니라, 두 조건에 어떻게 조합되는지에 따라 결과가 달라지는 모습을 확인할 수 있었다.

이러한 패턴은 학습 성과가 응답 속도나 언어 스타일 중 하나의 영향만으로 결정된다기 보다는, 두 요소가 함께 어떤 방향으로 작용하느냐에 따라 달라질 수 있음을 보여주는 것이라고 볼 수 있다. 통계적으로 상호작용 효과가 유의한 수준에는 도달하지 못했지만( $p = .360$ ), 효과 크기를 보면 상호작용 효과( $\eta^2 = .023$ )가 각 주 효과 크기( $\eta^2 = .002 \sim .003$ )보다 상대적으로 크게 나타났음을 알 수 있다. 이는 표본 수가 더 많았다면 상호작용 효과가 보다 분명하게 나타났을 가능성도 함께 시사한다.

이 결과를 설계 관점에서 이야기한다면, AI 아바타 튜터 시스템에 하나의 정답처럼 적용할 수 있는 응답 방식이 존재하지 않을 수도 있다. 예를 들어 학습자의 몰입이나 빠른 반응이 중요한 상황에서는 대화체에 빠른 응답이 더 적합할 수 있고, 개념을 차분하게 설명하거나 전문적인 내용을 전달해야 하는 경우에는 격식체에 다소 느린 응답이 오히려 긍정적으로 작용할 가능성도 있다. Petty와 Cacioppo(1986)의 정교화 가능성 모델(Elaboration Likelihood Model, ELM)의 관점으로 바라보면, 격식체는 학습자로 하여금 강의 내용을 보다 진지하고 분석적으로 수용하도록 유도하는 사회적 프레임으로 작용했을 가능성이 있다. ELM에 따르면 학습자가 메시지를 중심 경로로 처리할 경우 한층 더 높은 수준의 인지적 정교화가 이루어지며, 이는 장기적 태도와 학습 내면화로 이어진다(Petty & Cacioppo, 1986). 대화체가 친숙하고 편안한 상호작용 관계를 구성하는 반면에, 격식체는 학습자에게 강의를 진중하게 받아들여야 한다는 심리적인 기재로 작용함으로써, 학습 내용에 대해 분석적인 처리를 촉진했을 가능성이 있다. 둘째, Sweller(1988)의 인지부하 이론의 관점에서 들여다보면, 느린 응답이 제공되는 대기 시간은 단순한 공백이 아니라 학습자가 직전에 들은 내용을 자발적으로 되짚고 재구조화하는 시간으로 사용했을 가능성이 있다. 대화체·빠른·응답 조건에서 즉각적인 피드백이 학습 흐름을 유지하는데 유리하게 작용한 반면에 (Csikszentmihalyi, 1991), 격식체·느린 응답 조건에서는 응답을 기다리는 동안 학습자가 해당 질문에 대해 한 번 더 생각하고 다양한 연관 사고를 전개하는 과정이 이루어졌을 가능성이 있다. 즉, 격식체가 만들어내는 진중한 수용 태도와 느린 응답이 유발하는 자발적 사고 확장이 시너지를 이루면서, 결과적으로 학습 성과의 향상으로 이어진 것으로 해석할 수 있다. 다만 본 연구는 학습자의 반추 행동이나 인지 처리 깊이를 직접 측정하지 않았기에, 후속 연구에서 이 가설을 실증적으로 검증할 필요가 있다.

이러한 해석은 사회적 실재감에서 관찰된 상호 작용 경향( $p = .051$ )과도 비슷한 흐름을 보이면서, 언어 표현과 응답 속도가 서로 잘 맞을 때 학습성과에도 영향을 줄 수 있다는 점을 보여주는 근거로 볼 수 있다. 이 부분에 대해서는 7장에서 좀 더 자세하게 논의할 예정이다.

## 6.2.2 몰입(FlowAvr)에 대한 이원분산분석

몰입(FlowAvr)에 대해 응답 속도(빠름/느림)와 언어 스타일(대화체/격식체)의 주효과 및 상호작용 효과를 검증하기 위해  $2 \times 2$  이원분산분석을 실시하였다(Table 12). 분석에 앞서 등분산성 가정을 점검한 결과, Levene 검정은  $p = .421$ 로 유의하지 않아 등분산 가정은 충족되는 것으로 확인되었다(Table 13).

표본수는 빠름-대화체 10명, 빠름-격식체 10명, 느림-대화체 11명, 느림-격식체 10명이었으며, 평균은 각각 4.08(B), 3.86(C), 3.91(D), 3.68(E)로 나타났다. 표준편차는  $B = 0.83$ ,  $C = 0.48$ ,  $D = 0.70$ ,

E = 0.73 으로 확인되었다(Table 11).

이원분산분석 결과, 응답 속도의 주효과는  $F(1, 37) = 0.645, p = .427$ , 부분 에타 제곱  $\eta^2 = .017$  로 유의하지 않았고, 언어 스타일의 주효과 역시  $F(1, 37) = 1.057, p = .311, \eta^2 = .028$  으로 유의하지 않았다. 또한 응답 속도와 언어 스타일의 상호작용 효과는  $F(1, 37) = 0.000, p = .984, \eta^2 = .000$  으로 거의 0 에 가까운 값이 산출되어 통계적으로 유의하지 않았다(Table 12). 효과크기 관점에서 볼 때, 주효과의  $\eta^2$  값(.017~.028)은 작은 수준에 해당하며, 상호작용 효과는 사실상 관찰되지 않았다. 이는 본 표본 범위에서는 말투-속도 조합이 몰입에 시너지 효과를 낸다고 보기 어렵다는 것을 의미한다.

Table 11 을 보면, 대화체 조건이 격식체 조건보다 평균 몰입 점수가 더 높게 나타나는 경향을 보였다. 응답이 빠른 조건에서 대화체의 몰입 점수가 격식체보다 높았으며, 응답이 느린 조건에서도 대화체가 격식체보다 높은 값을 보여주었다. 즉, 응답 속도와 관계없이 말투가 대화체일 때 학습자가 더 몰입하는 경향이 반복해서 나타났다고 해석할 수 있다. 다만 이러한 차이가 통계적으로 유의한 수준까지 도달하지 못한 점으로 미루어 볼 때, 표본 수가 충분하지 않았고 개인차로 인한 점수의 변동이 영향을 주었을 수도 있다 하지만 다수의 인원을 기준으로 집단 평균을 낸 것이므로, 개인차로 인한 점수의 변동은 집단별로 동일하다고 보고 제외하였다. 따라서 몰입과 관련해서 “말투 효과가 존재할 수 있으나 연구 자료만으로는 확정하기엔 어려움이 있다”는 제한적인 결론을 내는 것이 적절하다고 판단하였다.

또한 몰입은 상호작용 즉시성 외에도 콘텐츠 난이도, 학습자의 사전 흥미, 시청 환경, 음성 합성 품질, 아바타의 비언어적 표현(표정·제스처·립싱크) 등 다양한 요인의 영향을 받을 수 있다. 실제로 정성 의견에서 음성 발음 및 억양의 부자연성, 아바타 동작의 경직성 등이 몰입을 저해하는 요소로 지적되었다는 점은, 말투·속도 조작의 잠재 효과가 이러한 잡음 요인에 의해 약화되었을 가능성을 시사한다. 이와 같은 해석은 8.4 절의 연구 한계 및 8.5 절의 후속 연구 제안에서 몰입 결정 요인의 다변량적 통제가 필요하다는 논의로 확장할 수 있다.

**Table 11. 2 × 2 Cell Descriptive Statistics for Flow**

Language Style	Response Speed: Fast Mean(SD), N	Response Speed: Slow Mean(SD), N
Conversational Tone	4.08 (0.83), n = 10	3.91 (0.70), n = 11
Formal Tone	3.86 (0.48), n = 10	3.68 (0.73), n = 10

**Table 12. Summary of the 2 × 2 Two-Way ANOVA Results for Flow**

Effect	F(1,37)	p	Partial Eta
--------	---------	---	-------------

			Squared ( $\eta^2$ )
Response Speed (Fast vs. Slow)	0.645	0.427	0.017
Language Style (Conversational vs. Formal)	1.057	0.311	0.028
Response Speed × Language Style	0.000	0.984	0.000

**Table 13. Levene's Test for Flow**

Test	p
Levene's Test for Homogeneity of Variances	0.421

**Figure 4. Estimated Marginal Means of Flow**

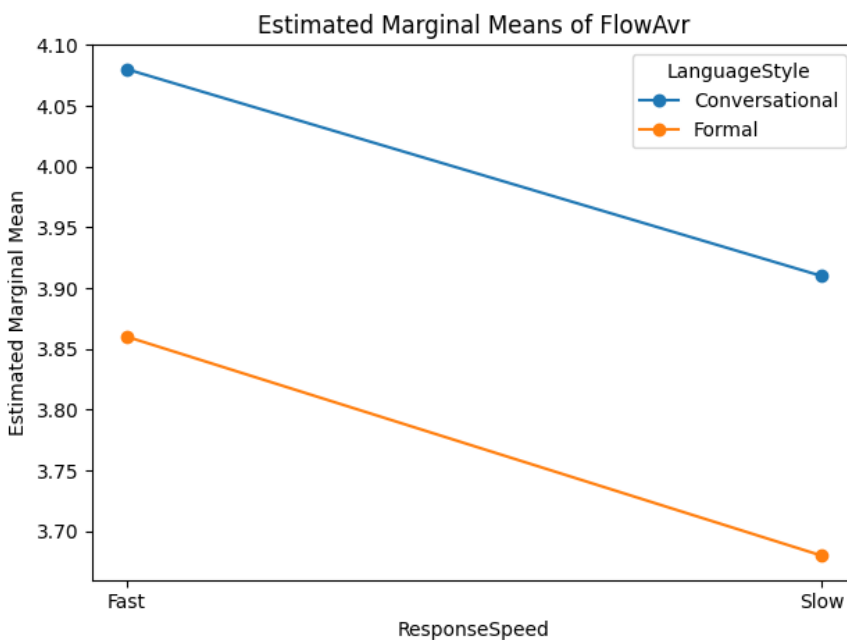


Figure 4는 응답 속도(Response Speed)와 언어 스타일(Language Style)에 따른 학습자의 몰입도(Flow)의 추정 주변 평균을 시각화한 것이다. 그래프를 보면, 두 언어 스타일 조건 모두에서 응답

속도가 빠를 때 몰입도가 더 높고, 느려질수록 소폭 감소하는 유사한 하강 패턴이 관찰된다. 대화체 조건에서는 빠른 응답에서 평균 몰입도가 4.08 로 가장 높았으며, 느린 응답에서는 3.91 로 감소하였다. 격식체 조건 역시 빠른 응답에서 3.86, 느린 응답에서 3.68 로 나타나, 응답 속도에 따른 변화 방향이 대화체 조건과 동일하게 유지되었다.

특히 두 조건의 선이 거의 평행한 형태를 이루고 있다는 점은, 응답 속도 효과가 언어 스타일에 따라 크게 달라지지 않음을 시각적으로 보여준다. 이는 두 변인의 조합이 몰입을 추가적으로 증폭하거나 저해하는 상호작용 효과가 거의 없을 가능성을 시사하며, 실제 이원분산분석에서 상호작용 효과가 거의 0 에 가까운 수준으로 나타난 결과( $F(1, 37) = 0.000, p = .984$ )와도 일관된다.

다만 평균 수준만을 놓고 볼 때, 대화체 조건은 모든 응답 속도 구간에서 격식체 조건보다 일관되게 높은 몰입도를 유지하고 있다는 점이 주목된다. 빠른 응답 조건에서는 대화체(4.08)가 격식체(3.86)보다 높았고, 느린 응답 조건에서도 대화체(3.91)가 격식체(3.68)보다 높게 나타났다. 이러한 패턴은 비록 통계적으로 유의하지는 않았으나, 친근한 언어 스타일이 학습자의 주관적 몰입 경험을 전반적으로 상향 조정하는 경향성을 가질 가능성을 시사한다.

종합하면, Figure 4 는 응답 속도와 언어 스타일의 조합이 몰입에 구조적인 상호작용 효과를 형성하지는 않았음을 보여주는 동시에, 대화체 기반 상호작용이 전반적으로 더 높은 몰입 수준과 연관될 수 있다는 약한 방향성을 시각적으로 제시한다. 이는 몰입이 단일 상호작용 단서보다는 콘텐츠 난이도, 음성·아바타 품질, 학습 맥락 등 다양한 요인의 영향을 복합적으로 받는 지표일 가능성을 뒷받침하며, 본 실험 환경에서는 이러한 잡음 요인들이 언어 스타일 및 응답 속도의 잠재 효과를 희석시켰을 가능성도 함께 시사한다.

### 6.2.3 사회적 실재감(Social Presence)에 대한 이원분산분석

사회적 실재감(Social Presence; SocAvr)에 대해 응답 속도(빠름/느림)와 언어 스타일(대화체/격식체)의 주효과 및 상호작용 효과를 검증하기 위해  $2 \times 2$  이원분산분석을 실시하였다. 분석에 앞서 등분산성 가정을 점검한 결과, Levene 검정은  $p = .781$  로 유의하지 않아 등분산 가정은 충족되는 것으로 확인되었다(Table 16). 표본수는 빠름-대화체 10 명, 빠름-격식체 10 명, 느림-대화체 11 명, 느림-격식체 10 명이었으며, 평균은 각각 3.68(B), 2.98(C), 2.96(D), 3.44(E)로 나타났다. 표준편차는  $B = 0.82, C = 0.88, D = 1.07, E = 0.93$  이었다(Table 14).

이원분산분석 결과, 응답 속도의 주효과는  $F(1, 37) = 0.193, p = .663$ , 부분 에타 제곱  $\eta^2 = .005$  로 유의하지 않았고, 언어 스타일의 주효과 역시  $F(1, 37) = 0.147, p = .704, \eta^2 = .004$  로 유의하지 않았다. 반면, 응답 속도와 언어 스타일의 상호작용 효과는  $F(1, 37) = 4.060, p = .051$ ,

$\eta^2 = .099$ 로 나타나 5% 유의수준에서는 경계선에 위치하였으나, 10% 유의수준에서는 유의한 상호작용으로 해석할 수 있다(Table 15). 다만 사회과학분야의 관례적 기준인 5% 유의수준에도달하지 못하였으므로, 통계적으로 유의한 상호작용이 있었다고 선불리 단정하기는 어렵다. 다만 효과크기 관점에서 볼 때, 상호작용 효과의  $\eta^2$  값(.099)은 두 주효과( $\eta^2 = .004\sim.005$ )에 비해 현저히 크며, 표본 수가 충분히 확보된 후속 연구에서 유의성이 확인될 가능성을 배제하기 어렵다. 따라서 이하의 논의는 확정적 주장이 아닌, 후속 연구에서 검증이 필요한 탐색적 가능성의제안으로 제한하며 해석한다. 또한 이 상호작용 효과 크기는 실험 결과를 분석한 정의적 지표들 중 가장 큰 효과크기에 해당한다. 이는 사회적 실재감이 응답 속도나 언어 스타일의 단독 효과보다, 두 요인의 결합 방식에 더 민감할 가능성을 시사한다.

상호작용의 방향은 교차 형태(crossover pattern)로 나타났다. 대화체 조건에서는 빠른 응답(B)이 느린 응답(D)보다 사회적 실재감이 높았고( $3.68 > 2.96$ ), 반대로 격식체 조건에서는 느린 응답(E)이 빠른 응답(C)보다 더 높았다( $3.44 > 2.98$ ). 즉, 응답 속도가 빠르다고 항상 실재감이 높아지는 것은 아니며, 말투가 설정하는 기대와 실제 응답 속도의 정합성이 사회적 실재감 경험을 좌우할 수 있음을 시사한다. 이 패턴은 결과 장에서 과도한 이론화를 피하면서 다음과 같이 기술할 수 있다. “대화체 조건은 대화 상호작용의 즉시성을 기대하게 만들며, 이 기대가 충족되는 빠른 응답에서는 실재감이 높아지지만, 지연이 발생하면 실재감이 크게 저하되는 경향이 관찰되었다. 반대로 격식체 조건에서는 지연이 반드시 실재감 저하로 연결되지 않았으며, 느린 조건에서 오히려 실재감이 더 높게 나타났다.”

이러한 결과는 사회적 실재감이 단순한 반응 속도나 언어 톤의 문제가 아니라, 상호작용 맥락 전반에 대한 학습자의 해석 과정과 연관될 가능성을 보여준다. 즉, 대화체는 인간 대화 규범에 가까운 즉시성 기대를 강하게 유발하는 반면, 격식체는 상대적으로 숙고된 반응을 허용하는 인지적 프레임을 형성할 수 있다. 다만 이 연구는 이러한 메커니즘을 직접 조작하거나 측정하지 않았으므로, 인과적 설명은 7장에서 ‘가능성’ 수준으로 제한하여 논의하는 것이 타당하다. 이러한 해석은 후속 연구에서 대기 상태 표시, 답변 준비 신호, 언어 톤의 세분화 등을 추가 조작 요인으로 포함할 필요성을 제기하는 근거로 확장될 수 있다.

**Table 14. Cell Descriptive Statistics for Social Presence**

Language Style	Response Speed: Fast Mean(SD), N	Response Speed: Slow Mean(SD), N
Conversational Tone	3.68 (0.82), n = 10	2.96 (1.07), n = 11
Formal Tone	2.98 (0.88), n = 10	3.44 (0.93), n = 10

**Table 15. Two-Way ANOVA Results for Social Presence**

Effect	F(1,37)	p	Partial Eta Squared ( $\eta^2$ )
Response Speed (Fast vs. Slow)	0.193	0.663	0.005
Language Style (Conversational vs. Formal)	0.147	0.704	0.004
Response Speed × Language Style	4.060	0.051	0.099

**Table 16. Levene’s Test for Social Presence**

Test	p
Levene’s Test for Homogeneity of Variances	0.781

**Figure 5. Estimated Marginal Means of Social Presence**

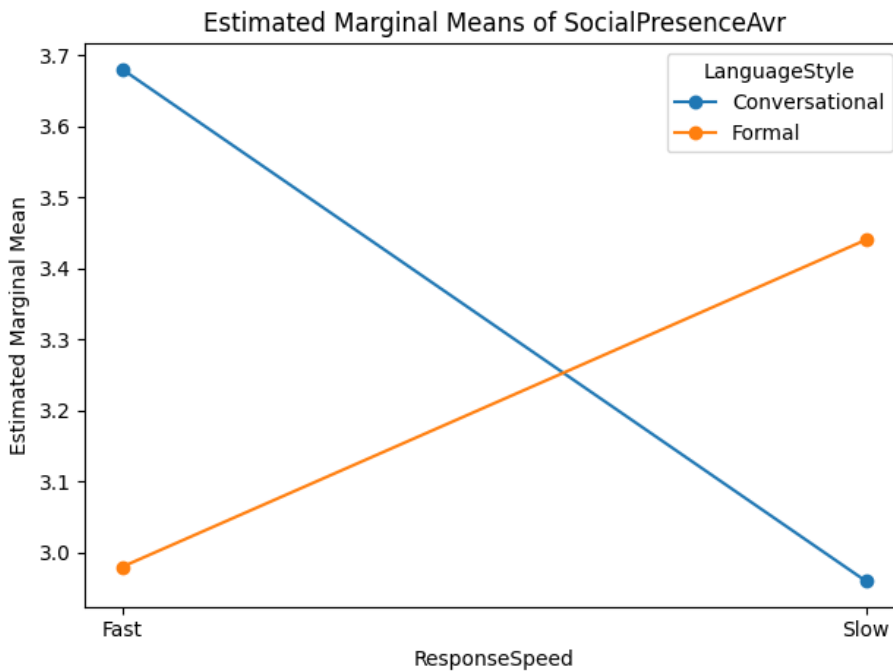


Figure 5 는 응답 속도(ResponseSpeed)와 언어 스타일(LanguageStyle)에 따른 학습자의 사회적 실재감(SocialPresenceAvr)의 추정 주변 평균을 나타낸 것이다. 그래프를 보면 두 변인 간에 뚜렷한 교차 패턴(crossover pattern)이 관찰된다. 대화체(Conversational) 조건에서는 빠른 응답에서 사회적 실재감이 가장 높게 나타났으며( $M=3.68$ ), 응답이 느려질 경우 실재감이 크게 감소하였다( $M=2.96$ ). 반면, 격식체(Formal) 조건에서는 빠른 응답보다 느린 응답에서 오히려 사회적 실재감이 더 높게 나타났다( $M=3.44 > 2.98$ ). 즉, 사회적 실재감은 “응답이 빠를수록 항상 높아진다”는 단순한 선형 관계를 따르지 않고, 언어 스타일이 설정하는 상호작용 맥락에 따라 응답 속도의 효과 방향이 반전되는 양상을 보였다.

이러한 결과는 초기 가설에서 예측한 “빠른 응답이 사회적 실재감을 일관되게 증진시킬 것”이라는 가정과는 부분적으로 어긋나는, 예상 외의 결과로 해석될 수 있다. 그러나 이를 언어 스타일과 시간적 단서의 맥락적 정합성 관점에서 재해석할 경우, 오히려 일관된 행동 패턴으로 이해될 여지가 있다. 대화체 조건은 학습자에게 일상적 대화 맥락과 유사한 상호작용 기대(expectation of immediacy)를 형성하며, 이 기대는 빠른 응답과 결합될 때 충족된다. 이 경우 학습자는 AI 아바타를 실제 대화 상대에 가까운 사회적 존재로 지각하게 되어 사회적 실재감이 극대화되는 것으로 해석할 수 있다. 반대로 대화체 조건에서 응답이 느려질 경우, 이러한 즉시성 기대가 좌절되면서 실재감이 급격히 저하되는 패턴이 나타난다.

한편 격식체 조건에서는 상반된 양상이 관찰되었다. 격식 있는 말투는 전문성, 신중함, 정보의 정확성을 강조하는 사회적 단서를 내포하고 있으며, 이 맥락에서는 다소 느린 응답 속도가 오히려 “깊이 사고한 후 답변하는 과정”으로 해석될 가능성이 있다. 실제로 격식체 조건에서 느린 응답이 빠른 응답보다 더 높은 사회적 실재감을 형성한 결과는, 학습자가 AI 아바타의 지연 응답을 기술적 결함이 아니라 신중한 인지적 처리 과정으로 재해석했을 가능성을 시사한다. 이는 격식체 페르소나가 갖는 ‘전문가적 이미지’가 응답 지연을 사회적으로 용인 가능한 단서로 전환시키는 완충 효과(buffering effect)를 가졌음을 의미한다.

종합하면, 본 그림이 보여주는 핵심 메시지는 사회적 실재감이 응답 속도나 언어 스타일의 독립적 효과에 의해 결정되기보다는, 두 요인이 결합되어 형성하는 통합된 상호작용 페르소나에 의해 좌우된다는 점이다. 대화체 페르소나는 빠른 응답과 결합될 때, 격식체 페르소나는 상대적으로 느린 응답과 결합될 때 각각 최대의 사회적 실재감을 형성하였다. 이러한 결과는 학습자가 AI 아바타의 언어적 태도와 반응 리듬을 분리된 요소로 인식하기보다, 하나의 일관된 사회적 행동 양식으로 통합하여 해석하고 있음을 시사한다.

따라서 이 연구의 사회적 실재감 결과는, AI 아바타 설계 시 단순히 응답 속도를 극대화하는 접근이 항상 최적의 전략이 아님을 보여준다. 오히려 설정된 언어 스타일에 부합하는 응답 리듬을 함께 설계하는 것이 학습자의 사회적 존재감 인식을 극대화하는 보다 정교한 설계 원칙이 될 수 있다. 이는 이후 7 장에서 논의되는 ‘언어 단서-시간 단서 정합성(Social Cue Alignment)’ 개념을 경험적 차원에서 지지하는 핵심

근거로 기능한다.

#### 6.2.4 인지부하(CognitiveLoadAvr)에 대한 이원분산분석

**Table 17. 2 × 2 Cell Descriptive Statistics for Cognitive Load**

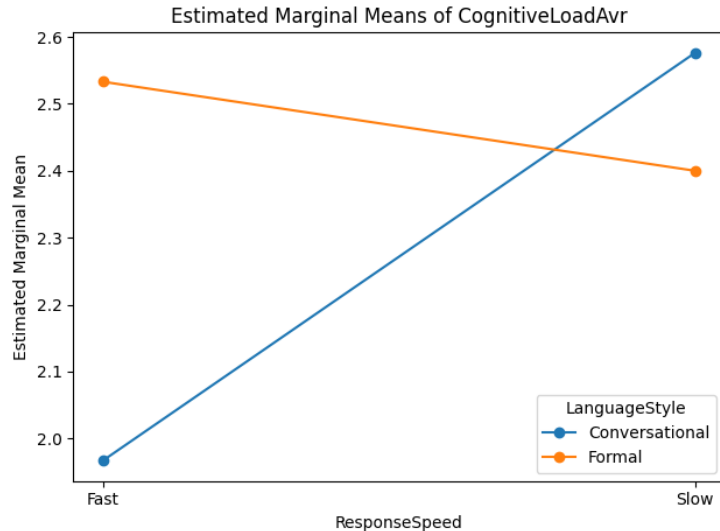
Language Style	Response Speed: Fast	Response Speed: Slow
Conversational Tone	1.97 (0.60), n = 10	2.58 (0.68), n = 11
Formal Tone	2.53 (0.91), n = 10	2.40 (0.72), n = 10

**Table 18. Two-Way ANOVA Results for Cognitive Load**

Effect	F(1,37)	p	Partial Eta Squared ( $\eta^2$ )
Response Speed (Fast vs. Slow)	1.077	0.306	0.028
Language Style (Conversational vs. Formal)	0.727	0.399	0.019
Response Speed × Language Style	2.622	0.114	0.066

**Table 19. Levene's Test for Cognitive Load**

Test	p
Levene's Test for Homogeneity of Variances	0.601

**Figure 6. Estimated Marginal Means of Cognitive Load**

인지부하(Cognitive Load; CognitiveLoadAvr)를 종속변수로 하여 응답 속도(Response Speed: 빠름/느림)와 언어 스타일(Language Style: 대화체/격식체)의 효과를 검증하기 위해  $2 \times 2$  이원분산분석을 실시하였다. 분석에 앞서 등분산성 가정을 점검한 결과, Levene 검정은  $p = .601$  로 유의하지 않아 등분산 가정은 충족되는 것으로 확인되었다(Table 19). 표본수는 빠름-대화체 10 명, 빠름-격식체 10 명, 느림-대화체 11 명, 느림-격식체 10 명이었으며, 평균은 각각  $B = 1.97$ ,  $C = 2.53$ ,  $D = 2.58$ ,  $E = 2.40$  으로 나타났다. 표준편차는  $B = 0.60$ ,  $C = 0.91$ ,  $D = 0.68$ ,  $E = 0.72$  였다(Table 17).

Table 18 의 이원분산분석 결과를 보면, 응답 속도의 주 효과는 통계적으로 유의하지 않았다( $F(1, 37) = 1.077$ ,  $p = .306$ ,  $\eta^2 = .028$ ). 언어 스타일 역시 주효과가 유의한 수준에 도달하지 못했으며( $F(1, 37) = 0.727$ ,  $p = .399$ ,  $\eta^2 = .019$ ), 두 변인의 상호작용 효과도 마찬가지로 통계적으로 유의하지 않았다( $F(1, 37) = 2.622$ ,  $p = .114$ ,  $\eta^2 = .066$ ). 다만 효과 크기를 기준으로 보면, 상호작용 효과의 값이 두 주효과보다 상대적으로 크게 나타났다는 점은 눈여겨볼 만하다. 이는 인지부하가 하나의 요인보다는 응답 속도와 언어 스타일이 함께 어떻게 조합되는지에 따라 달라질 가능성을 내포하고 있다.

Figure 6 을 통해서 인지부하의 추정 평균을 보면, 사회적 실재감에서 나타난 결과와 비슷한 교차 형태가 확인된다. 대화체 조건에서 응답이 빠를 때, 인지부하( $B = 1.97$ )가 가장 낮았지만, 응답이 느려지면 인지부하( $D = 2.58$ )가 크게 증가하는 모습을 보였다. 반면 격식체 조건에서 빠른 응답인 경우와, 느린 응답인 경우 인지부하( $C = 2.53$ ,  $E = 2.40$ ) 차이가 비교적 작아서 응답 속도 변화에 따른 인지 부하의 변화는 크지 않다고 보는 것이 타당하다. 즉, 대화체에서 응답이 늦어질수록 부담이 확연히 커짐을 볼 수 있는 반면에, 격식체의 경우 같은 지연이 상대적으로 덜 부담스럽게 느껴졌다고 볼 수 있다.

이러한 결과는 응답 지연이 단순히 기다리는 시간이 늘어나는 문제만은 아니라는 점을 보여준다. 특히 대화체 말투는 학습자로 하여금 곧 답이 올 것이라라는 기대를 갖게 만드는 경향이 있는데, 이 기대가

충족되지 않으면 사용자는 시스템이 제대로 작동하고 있는지에 대해 고민하게 된다. 예를 들어 응답이 한 동안 없는 경우, 오류가 난 것인지 또는 질문이 전달되지 않은 것인지 등을 추측하게 되며, 이 과정 자체가 추가적인 부담으로 작용할 가능성이 있다. 이런 이유로 대화체 - 느낌 조건에서 인지부하가 가장 높게 나타난 것으로 해석할 수 있다.

반대로 격식체 조건에서는 같은 수준의 응답 지연이 반드시 부정적으로 인식되지 않았다. 격식체 말투는 상대적으로 신중하고 공식적인 인상을 줄 수 있기 때문에 응답이 늦어질 경우에도 이를 시스템 문제라기보다는 생각하고 답변하는 과정으로 받아들여질 가능성이 있다고 본다. 그 결과 응답이 느려져도 인지부하가 크게 증가하지 않는 패턴이 나타난 것으로 판단된다.

이러한 해석은 6.5 절에서 다룬 정성적 의견과도 연결된다. 실제로 여러 참여자들은 응답이 늦을 때 멍통처럼 느껴졌다거나 대기 상태를 알려주는 표시가 필요하다는 의견을 제시하였다. 정리하면, 인지부하에 대한 본 연구의 결과는 응답 속도나 언어 스타일이 각각 따로 작동하기보다는, 두 요소가 서로 얼마나 잘 맞는지가 더 중요할 수 있음을 보여준다고 하겠다. 특히 대화체 언어 스타일을 사용하는 AI 아바타의 경우, 빠른 응답이 함께 제공되지 않으면 오히려 학습자의 부담을 증가시키는 결과로 이어질 수 있다. 이는 AI 아바타 기반 교육 시스템을 설계할 때, 친근한 말투를 사용할수록 응답 지연을 줄이거나, 로딩 표시나 답변 생성 중이라는 상태 안내 기능을 함께 제공하는 것이 필요하다는 점을 시사한다.

### 6.3 매개 및 조절된 매개 효과 검증(H2)

H2는 상호작용 특성(Response Speed  $\times$  Language Style)이 몰입(FlowAvr), 사회적 실재감(SocialPresenceAvr), 인지부하(CognitiveLoadAvr)를 통해 학습 성과(ScoreDiff)에 간접적으로 영향을 미칠 수 있다는 매개 가설을 포함한다. 이를 검증하기 위해 PROCESS Macro(Model 8)를 사용한 조절된 매개 분석을 실시하였으며, 그 결과를 Table 20와 Table 21에 제시하였다.

첫째, 몰입(FlowAvr)을 매개로 하는 경로에서 조건부 간접효과의 95% 부트스트랩 신뢰구간은 모든 조건에서 0을 포함하여 통계적으로 유의하지 않았다. 또한 조절된 매개 지수 역시 유의하지 않았다. 이는 7.2.2 절에서 몰입의 상호작용 효과가 거의 0에 가까운 수준( $F(1,37)=0.000$ ,  $p=.984$ )으로 나타난 결과와 정합적이며, 상호작용 특성이 몰입을 경유해 성취 변화로 이어진다고 보기 어렵다는 점을 시사한다.

둘째, 사회적 실재감(SocialPresenceAvr)을 매개로 하는 경로에서도 조건부 간접효과는 모든 조건에서 유의하지 않았다. 비록 7.2.3 절의 이원분산분석에서 실재감 단계에서는 상호작용 효과가 경계선 수준( $F(1,37)=4.060$ ,  $p=.051$ )으로 관찰되었으나, 이 효과가 학습 성과(ScoreDiff)까지 연쇄적으로 매개된다는 통계적 증거는 확보되지 않았다. 즉, 이 연구 자료는 “실재감의 상호작용 패턴” 자체는

보여주지만, 이것이 “학습 성취 변화”로 구조적으로 매개된다고 확정하기에는 충분하지 않다.

셋째, 인지부하(CognitiveLoadAvr)를 매개로 하는 경로에서도 조건부 간접효과와 조절된 매개 지수는 모두 유의하지 않았다. 이는 7.2.4 절에서 관찰된 인지부하의 교차 패턴이 학습 성과로까지 통계적으로 매개되지는 않았음을 의미한다.

종합하면, 이 연구에서는 응답 속도와 언어 스타일의 상호작용 특성이 몰입, 사회적 실재감, 인지부하를 통해 학습 성과에 간접적으로 영향을 미친다는 H2의 매개 가설은 지지되지 않았다. 따라서 본 결과는 매개 메커니즘을 통계적으로 확정하기보다는, 정량적 패턴과 정성 의견을 결합하여 가능한 작동 경로를 7장의 고찰에서 탐색적으로 논의하는 것이 타당하다. 이는 이 연구가 파일럿 연구라는 점을 감안할 때, 향후 표본 크기를 확장한 후속 연구에서 매개 구조를 재검증할 필요성을 시사한다.

#### 6.4 설문 정량 결과의 상세 보고

본 절은 7.2의 가설 검정과 구분하여, “조건별 평균 프로파일”을 더 촘촘히 제시한다. 심사 관점에서 이 절의 역할은, 유의하지 않은 결과가 존재하더라도 데이터가 보여주는 실질적 패턴을 투명하게 보고하는 데 있다.

**Table 20. Moderated Mediation of Response Speed (PROCESS Model 8)**

Predictor → Mediator → Outcome	Conditional Indirect Effect (Conversational)	95% BootCI (LL, UL)	Conditional Indirect Effect (Formal)	95% BootCI (LL, UL)	Index of Moderated Mediation	95% BootCI (LL, UL)
RxSpeed → FlowAvr → SDiff	-0.1413	[- 1.0089, 0.5023]	-0.1488	[- 0.7680, 0.3644]	-0.0075	[- 0.8827, 0.9547]
RxSpeed → SocPresence →	-0.2468	[- 1.0308, 0.2773]	0.1585	[- 0.2325, 0.7080]	0.4053	[- 0.4207, 1.4953]

SDiff						
RxSpeed → CogLoad → SDiff	-0.0836	[-0.6376, 0.3985]	0.0183	[-0.2903, 0.3641]	0.1019	[-0.5616, 0.8566]

**Table 21. Moderated Mediation of Language Style (PROCESS Model 8)**

Predictor → Mediator → Outcome	Conditional Indirect Effect (Fast Response)	95% BootCI (LL, UL)	Conditional Indirect Effect (Slow Response)	95% BootCI (LL, UL)	Index of Moderated Mediation	95% BootCI (LL, UL)
LanStyle → FlowAvr → SDiff	-0.1819	[-1.0597, 0.3840]	-0.1894	[-0.8800, 0.3948]	-0.0075	[-0.8651, 0.9446]
LanStyle → SocAvr → SDiff	-0.2412	[-0.8392, 0.3072]	0.1641	[-0.1999, 0.9170]	0.4053	[-0.4153, 1.5528]
LanStyle → CogLoad → SDiff	-0.0778	[-0.5818, 0.3803]	0.0241	[-0.2550, 0.3746]	0.1019	[-0.5563, 0.8137]

**6.4.1 Q&A 제공 여부 기준의 평균 비교(통제 vs 실험 통합)**

Table 3 을 Q&A 제공 여부로 재구성하면, 통제집단은 몰입 3.29, 사회적 실재감 3.04, 인지부하 2.73 이다. 실험집단 중에서는 B 조건이 몰입 4.08, 실재감 3.68, 인지부하 1.97 로 가장 긍정적이다. 즉, 평균 수준에서 Q&A 제공은 몰입과 만족을 높이고 인지부하를 낮추는 방향의 프로파일을 보인다. 다만 이 결론은 “평균 패턴”이며, 유의성 여부는 각 변인별 분산분석 결과에 따라 별도로 판단해야 한다.

#### 6.4.2 말투 기준 평균 비교(대화체 vs 격식체)

대화체(B, D)의 평균은 몰입이 상대적으로 높고(B 4.08, D 3.91), 격식체(C, E)는 상대적으로 낮다(C 3.86, E 3.68). 사회적 실재감은 대화체가 빠름에서는 매우 높지만(B 3.68), 느림에서는 낮아져(D 2.96) 평균이 단순하지 않다. 인지부하는 대화체가 빠름에서는 최저(B 1.97)지만 느림에서는 상승(D 2.58)한다. 이 결과는 말투만을 독립 요인으로 보는 것보다, “말투가 속도 기대를 어떻게 설정하는지”를 함께 고려해야 함을 강화한다.

#### 6.4.3 속도 기준 평균 비교(빠름 vs 느림)

빠름 조건(B, C)은 사회적 실재감이 말투에 의해 크게 갈린다(B 3.68 vs C 2.98). 느림 조건(D, E)에서도 말투에 따라 실재감이 크게 갈린다(D 2.96 vs E 3.44). 즉, 속도도 단독으로는 설명력이 제한적이며, 말투와 결합된 상태에서 경험을 형성한다는 패턴이 반복된다.

#### 6.4.4 학습 성취와 사용자 경험의 분리: “최고 성취” vs “최고 경험”

학습 성취 점수 변화량을 보면 E 조건이 평균 1.50 으로 가장 높게 나타났다. 반면 몰입, 사회적 실재감, 인지부하, 강의 만족도, 주관적 성취감과 같은 사용자 경험 측면에서는 B 집단이 전반적으로 가장 좋은 평가를 받았다. 즉, 시험 점수가 가장 많이 오른 집단과 학습자가 가장 편하고 좋게 느낀 집단이 서로 같지는 않았다는 점을 확인할 수 있다.

이 결과는 학습자에게 가장 즐겁고 부담이 적은 학습 경험이 반드시 가장 큰 점수 향상으로 이어지지 않을 수 있음을 보여준다. 하지만 이를 연구의 한계나 단점으로 보기보다는, AI 튜터 설계 자체가 여러 목표를 동시에 고려해야 하는 문제로 해석할 수 있다. 다시 말해, 학습 성취를 우선할 것인지, 학습 경험의 질을 중시할 것인지에 따라 아바타의 말투나 반응 방식과 같은 페르소나를 다르게 설정할 필요가 있음을 시사한다.

## 6.5 설문 정성 결과 분석

본 절에서는 자유서술 응답을 범주별로 정리하되, 정량 결과에서 관찰된 패턴(특히 지연-대화체에서의 실재감 하락/부담 증가)을 설명하는 근거로 사용한다. 정성 자료는 통계적 유의성을 대체하지 않으며, “왜 이런 패턴이 나타났을 수 있는지”를 해석 가능하게 만드는 보강 자료로 위치시킨다.

### 6.5.1 응답 지연과 상태 가시성(대기 표시)의 중요성

정성 의견을 살펴보면, 가장 많이 반복된 요구는 질문 후 **대기 상태가 보이지 않는 점**에 대한 불편함이었다. 응답이 바로 나오지 않을 경우, 참여자들은 이를 아바타가 생각하고 있는 과정으로 받아들이기보다는 시스템 오류나 먹통 상태로 오해하는 경우가 많았다. 실제로 설문에는 “질문 후 아바타가 바로 대답하지 않으면 오류라고 생각했음. 생각 중 상태가 직관적으로 표시되면 좋겠음”이나 “답변 대기 시간에 어떤 표시가 있으면 좋겠음(먹통 되었나 생각됨)”과 같은 의견이 반복적으로 나타났다.

이러한 반응은 6.2.3 절에서 확인된 사회적 실재감의 상호작용 패턴, 특히 대화체-느림 집단에서 실재감이 낮아진 결과와도 연결된다. 대화체 말투는 사람과 대화하는 상황과 비슷한 기대를 만들기 때문에, 사용자는 자연스럽게 즉각적인 반응을 기대하게 된다. 이 상태에서 응답이 지연되면, 사용자는 상호작용이 끊겼다고 느끼거나 시스템이 제대로 작동하지 않는다고 판단할 가능성이 커진다. 그 결과 아바타와의 상호작용에 대한 실재감이 낮아질 수 있다.

반면 응답이 늦더라도 “답변 생성 중”과 같은 대기 상태가 명확하게 표시된다면, 사용자는 상호작용이 계속 진행되고 있다고 인식할 수 있다. 이는 응답 지연으로 인한 실재감 저하를 어느 정도 완화하는 역할을 할 가능성이 있다. 다만 이러한 해석은 정성 의견에 기반한 가능성 수준의 논의이며, 본 연구에서는 대기 상태 표시를 직접 조작하지 않았기 때문에, 7장에서 설계 제안 차원으로 논의하는 것이 적절하다고 판단된다.

### 6.5.2 음성 품질(발음·억양)과 오류 인식

정성 의견 중에는 아바타의 발음이 정확하지 않을 경우 이를 시스템 오류로 오해했다는 내용도 포함되어 있었다. 실제로 한 참여자는 “아바타의 발음이 부정확한 부분이 있어 오류가 난 줄 알았다”고 답하였다. 이처럼 음성이 자연스럽게 않거나 발음이 또렷하지 않을 경우, 사용자는 단순한 음성 품질 문제를 시스템 이상으로 받아들일 수 있다.

이러한 반응은 음성합성 품질이나 말하는 속도, 억양과 같은 요소가 사용자 경험에 중요한 영향을 미친다는 점을 보여준다. 특히 답변이 조금 늦어지는 상황에서 음성까지 어색하게 들리면, 사용자는 두

문제를 함께 인식하면서 “시스템이 제대로 작동하지 않는다”는 판단을 더 쉽게 하게 된다. 이 과정에서 아바타에 대한 신뢰가 떨어질 수 있고, 동시에 상황을 해석하려는 추가적인 생각이 필요해져 인지적 부담도 커질 수 있다. 따라서 음성의 자연스러움과 발음의 명확성은 사회적 실재감과 인지부하 모두에 영향을 줄 수 있는 중요한 요소로 볼 수 있다.

### 6.5.3 비언어 표현(립싱크, 표정, 눈·손 제스처)의 실재감 기여

정성 의견을 살펴보면, 아바타의 움직임이나 외형이 자연스럽지 않아서 실재감이 떨어진다는 의견도 여러 번 나타났다. 예를 들어 “아바타의 모습이 부자연스럽다”거나, “눈 움직임이나 손짓 같은 동작이 추가되면 더 실제 사람처럼 느껴질 것 같다”는 응답이 있었다. 이는 학습자가 아바타를 볼 때 말투나 응답 속도뿐만 아니라, 시각적으로 보이는 요소도 함께 중요하게 인식하고 있음을 보여준다.

이러한 의견은 사회적 실재감이 단순히 언어 스타일이나 응답 속도만으로 결정되지 않는다는 점을 분명히 한다. 즉, 말투와 응답 속도의 조합에서 의미 있는 패턴이 나타나더라도, 아바타의 표정이나 시선, 손짓과 같은 비언어적 표현이 충분히 자연스럽지 않다면 전체적인 실재감은 제한될 수 있다. 따라서 후속 연구에서는 이러한 비언어적 표현의 품질을 일정하게 통제하거나, 별도의 실험 조건으로 설정하여 언어 스타일과 응답 속도의 효과와 구분해 살펴볼 필요가 있다. 이러한 점은 향후 AI 아바타 기반 학습 시스템을 설계하는 데 중요한 참고 근거로 활용될 수 있다.

### 6.5.4 자막/텍스트 지원 요구: 이해 안정성과 접근성

정성 의견 중에는 아바타의 답변을 자막으로 함께 제공해 달라는 요구도 분명하게 나타났다. 실제로 “질문에 대한 대답이 자막으로 나오면 좋겠다”거나, “아바타 음성이 자막으로 같이 나오면 이해하기 쉬울 것 같다”는 의견이 있었다. 이는 단순히 편의를 위한 기능이라기보다, 학습 과정에서 느끼는 불편함을 줄이기 위한 필요로 볼 수 있다.

자막이 제공될 경우, 소리만으로 이해하기 어려운 부분을 다시 확인할 수 있어 도움이 된다. 특히 발음이나 억양이 자연스럽지 않을 때, 자막은 내용을 정확하게 이해할 수 있는 보조 수단이 될 수 있다. 또한 답변이 바로 나오지 않는 상황에서도 자막이나 텍스트 표시가 함께 제공되면, 학습자는 “곧 설명이 나올 것”이라는 신호를 받아 보다 안정적으로 기다릴 수 있다. 이런 점에서 자막 지원은 학습자의 부담을 줄이고, 아바타와의 상호작용을 보다 자연스럽게 만드는 데 기여할 수 있다.

따라서 자막 제공은 단순한 편의 기능을 넘어, 실재감과 인지부하에 영향을 줄 수 있는 설계 요소로 볼

수 있다. 특히 본 연구에서 확인된 대화체-지연 집단에서의 인지부하 증가 현상을 완화할 수 있는 하나의 방법으로, 7장에서 설계 제안 차원에서 논의할 수 있을 것으로 판단된다.

### 6.5.5 정성적 사용자 의견 요약

결과 장에서 정성 의견을 제시할 때에는, 원인을 단정하기보다는 관찰된 내용을 그대로 전달하는 수준에서 서술하는 것이 적절하다. 따라서 다음과 같은 표현 방식이 타당하다고 볼 수 있다. 예를 들어, “응답이 지연될 경우 이를 시스템 오류로 오해할 수 있다는 응답이 일부 존재하였다”와 같이 응답 내용의 존재를 보고하는 방식이다. 또한 “응답을 기다리는 동안 대기 상태 표시나 자막 지원이 필요하다는 요구가 제시되었다”는 식으로 사용자 의견을 그대로 정리할 수 있다.

마찬가지로, “아바타의 눈 움직임이나 손짓, 표정과 같은 비언어적 표현이 강화될 경우 실재감에 도움이 될 수 있다는 의견이 관찰되었다”와 같은 표현도 사용할 수 있다. 이러한 서술 방식은 실제로 수집된 응답을 정리하는 수준에 머물며, 정성 자료를 근거로 인과 관계를 과도하게 해석하지 않는다는 점에서 결과 장에 적합하다.

### 6.6 정성적 사용자 의견 요약 및 정량·정성 통합 해석

이 절에서는 설문지의 자유응답 문항을 통해 수집한 사용자 의견을 정리하고, 이를 정량 분석 결과와 함께 해석하고자 한다. 여기서 제시하는 내용은 인과 관계를 단정하기보다는, 실험 과정에서 학습자들이 실제로 어떤 경험을 했는지와 그에 따른 설계상의 시사점을 정리하는 데 목적이 있다.

정성 의견을 분석한 결과, 몇 가지 공통된 반응이 반복적으로 나타났다. 먼저 응답이 바로 나오지 않을 경우, 이를 아바타가 생각하고 있는 상태로 보기보다는 시스템 오류나 먹통으로 느꼈다는 의견이 많았다. 일부 참여자들은 “응답이 늦어지면 멈춘 것처럼 느껴졌다”거나, “질문이 제대로 전달된 건지 알기 어려웠다”고 응답하였다.

또한 응답을 기다리는 동안 현재 상태를 알 수 있는 표시가 필요하다는 요구도 나타났다. 예를 들어 “답변 생성 중이라는 안내가 있었으면 좋겠다”, “로딩 애니메이션이나 진행 바가 있으면 덜 불안할 것 같다”는 의견이 있었으며, 음성과 함께 자막을 제공해 달라는 요구도 함께 언급되었다.

이와 함께 아바타의 움직임이나 표정과 같은 비언어적 표현이 더 자연스러워졌으면 좋겠다는 의견도 확인되었다. 눈 깜박임, 고개 움직임, 손짓, 표정 변화 등이 추가된다면 아바타가 더 실제 사람처럼 느껴질 것이라는 응답이 많았다.

이러한 정성 결과는 7 장의 정량 분석에서 나타난 패턴을 이해하는 데 도움을 준다. 정량 분석에서는 사회적 실재감에서 말투와 응답 속도가 함께 작용하는 경향( $p = .051$ )이 관찰되었다. 대화체 조건에서는 응답이 빠를 때 실재감이 높았지만(B 집단,  $M = 3.68$ ), 응답이 늦어지면 실재감이 크게 낮아지는 모습(D 집단,  $M = 2.96$ )이 나타났다. 반면 격식체 조건에서는 응답이 느린 경우(E 집단,  $M = 3.44$ )에 오히려 실재감이 더 높게 나타났다. 인지부하 역시 대화체-자연 집단에서 가장 크게 증가하여, 사회적 실재감과 비슷한 패턴을 보였다.

정성 의견을 함께 보면 이러한 결과를 보다 쉽게 이해할 수 있다. 응답이 지연될 때 학습자들은 이를 ‘생각 중’이라기보다는 오류나 시스템 불안정으로 해석하는 경우가 많았고, 대기 상태 표시가 없을수록 이러한 인식이 더 강해졌다. 또한 발음이 부정확하거나 아바타의 움직임이 어색하고, 자막이 없는 경우에는 시스템에 대한 신뢰감이 떨어질 수 있다는 의견도 반복적으로 나타났다. 이는 대화체 말투가 만들어내는 ‘곧바로 반응할 것이라는 기대’와 실제 응답 속도, 그리고 화면에서 제공되는 신호가 서로 맞지 않을 때 사회적 실재감이 낮아지고 인지적 부담이 커질 수 있음을 보여준다.

정리하면, 동일한 수준의 응답 지연이라 하더라도 언어 스타일이 만들어내는 기대와 대기 상태 표시, 자막, 비언어적 표현과 같은 UX 요소가 어떻게 결합되는지에 따라 학습자가 느끼는 경험의 질은 달라질 수 있다. 이는 응답 속도나 말투 하나만의 문제가 아니라, 상호작용 전반의 조합과 맥락이 사회적 실재감과 인지부하에 중요한 역할을 한다는 점을 시사한다.

## 7. 결론 및 고찰

이번 절에서는 6 장에서 제시한 정량 및 정성 분석 결과를 종합하여, 연구 질문과 가설에 대한 결론을 정리하고 이론적·설계적 함의를 논의하고자 한다. 이 연구는 설정한 가설이 5% 유의수준에서 모두 기각되었다는 점을 전제로, 통계적 유의성의 유무를 과하게 포장하지 않으면서, 설문지의 기술통계 데이터 패턴, 설정한 변인이 어느 정도 영향력을 가지고 있는지 알 수 있는 효과의 크기, 실험자의 피드백인 정성 자료의 정합성을 결합하여 AI 아바타 기반 온라인 강의 상호작용 설계에 대한 제한적이지만 실질적인 시사점을 알아보기로 한다.

### 7.1 실시간 질의응답 제공 효과의 안정성 측면에서의 재해석

이 연구에서 실시간 질의응답 기능의 제공 여부는 학습자의 학습 성과에 대하여 통계적으로 유의미한 평균 차이를 보여주진 않았다. 하지만 기술통계 결과를 보면, 실시간 질의응답을 사용한 집단의 평균 학습

성취 변화량(0.9756)은 통제집단(5455)보다 높게 나타났다. 특히 통제집단의 표준편차( $SD = 5.1257$ )가 매우 크게 나타난 반면, 실험집단의 표준편차( $SD = 2.1271$ )는 상대적으로 작음을 알 수 있다. 이 연구는 이 지점을 통해 실시간 질의응답 기능이 학습자의 평균 성취 점수를 끌어올리는 기능 외에도 학습 성취의 변동성을 낮추는 방향으로도 작용했을 것이라고 보고 있다.

이를 보다 구체적으로 확인하기 위해 통제집단과 실시간 질의응답 제공 집단의 사전·사후 퀴즈 점수 분포를 비교해 보았다. 통제집단의 경우 사전 점수 분포가 비교적 넓게 퍼져 있었고, 사후 점수 역시 고득점과 저득점이 혼재된 형태를 유지한 반면, 실시간 질의응답이 제공된 집단들에서는 사전 퀴즈 점수 분포가 집단별로 들쭉날쭉한 양상을 보였음에도 불구하고, 사후 퀴즈 점수는 대부분 만점 부근으로 수렴하는 형태를 나타냈다. 이 연구는 이 분포 변화를 통해 실시간 Q&A 제공이 학습자 간 성취 격차를 완화하는 방향으로 작용했을 가능성을 확인할 수 있다.

이와 같은 경향은 단기기억에 따른 우연의 일치로 실험군에 속한 실험자들의 기억력이 상대적으로 좋아서 발생한 효과일 가능성도 배제할 수 없지만, 이 연구는 실시간 질의응답 과정에서 학습자가 자신이 잘 모르는 부분을 실시간 질의응답 기능을 활용하여 의문점을 해소하는 과정에서 핵심 개념이 보다 명확히 머리속에 기억되었을 가능성 또한 함께 고려할 필요가 있다고 본다. 즉, 실시간 질의응답 기능의 제공은 모든 학습자의 성취를 동일한 방향으로 끌어올리면서, 학습 실패 또는 급격한 성취 저하의 가능성을 완화하고, 최소 성취 수준을 끌어올리는 안정화 장치로 기능했을 수 있다. 이 연구는 이 점에서 실시간 질의응답의 효과를 단순한 평균 상승 효과가 아니라, 학습 성취 분포를 줄이는 분산 감소 효과로도 해석할 수 있다고 판단하였다.

실험 종료 후 실험자와 진행한 인터뷰에서 얻은 정성 의견 중 “질의응답이 가능하여 직업 수업에 참여하는 것 같아 집중이 잘 되었다”, “모르는 부분을 즉시 확인할 수 있어 몰입이 잘 되었다”는 의견이 있었다. 이는 정량적인 결과에서 나타난 분산 축소 패턴과 부합한다고 볼 수 있다. 따라서 AI 아바타 기반 온라인 강의 설계에서 실시간 질의응답 기능을 제공하는 것은 학습자의 평균 성취 점수의 상승이기도 하면서, 성과의 신뢰도와 일관성을 확보하는 구조적 장치로도 볼 수 있다고 사료된다. 현재 다수 온라인 플랫폼 등에서 쉽게 접할 수 있는 온라인 동영상 강의는 이를 이용하는 학습자가 이해되지 않는 지점을 즉시 해소하기 어렵다는 한계를 지니고 있다. 특히 학습 중 발생한 의문이 누적된다면 이후 강의 내용의 이해가 잘 될 수 없다. 또한 학습자가 의문을 해소하기 위해 강의를 중지시키고 스마트기기나 인터넷 등을 사용하여 검색 과정은 학습 흐름을 단절시키는 요인으로 작용할 수밖에 없다. 이러한 관점에서 실시간 질의응답의 가치는 ‘즉시 답변’ 그 자체만이 아니라, 학습자가 질문을 제기할 수 있는 기회를 강의 안에 명시적으로 포함시키는 데서 한층 업그레이드될 수 있다. 예를 들어, 강의 진행 중 일정한 간격으로 “질문 있나요?”와 같은 음성을 제공하며, 학습자의 질문을 기다린다거나, 오프라인 학습의 쉬는 시간과 같이 별도의 질문 시간을 제공함으로써 학습자는 질문을 ‘방해 요소’가 아니라 ‘정규 절차’로 인식하게 되고,

이는 “질문을 언제 해야하지”라는 심리적 장벽을 낮추어 보다 적극적인 상호작용을 유도할 수 있다. 종합해보면 실시간 질의응답 기능의 효과를 얻기 위해서는 단순하게 기능을 탑재하는 수준을 넘어서, 교수자가 강의를 하는 중간 중간 쉬는 시간을 통해 질문을 받는 것과 같이 강의 내에 질문 가능 구간을 설계하여 학습자가 편하게 질문할 수 있도록 시간적 틈을 부여하는 것이 중요하다고 하겠다.

## 7.2 학습 성과와 경험 품질의 분리 가능성

이 실험의 결과를 통계분석한 결과 학습 성과(사후 테스트 - 사전 테스트)와 정의적 경험 지표(몰입, 사회적 실재감, 인지부하, 만족, 체감 성취감)가 같은 조건에서 동시에 최적화되지 않을 가능성을 보여주었다. 구체적으로 기술하면, 학업 성취 변화가 가장 크게 나타난 조건은 격식체와 느린 응답을 사용하는 조건(E)이었으나, 몰입, 사회적 실재감, 만족, 체감 성취감 등 사용자 경험 품질 지표는 대화체와 빠른 응답을 사용한 조건(B)에서 전반적으로 높게 나타났음을 알 수 있다. 이는 학업 성취 변화량이 가장 큰 조건과 사용자가 가장 좋게 느낀 조건이 일치하지 않을 수 있음을 의미하는 것이다.

이 점은 AI 아바타 온라인 강의 설계에서 유일한 최적 전략이 존재하지 않을 수 있음을 시사하는 것이다. 학습자의 학습 성과를 최대화하는 방법과 학습자의 사용자 경험 품질을 최대화하는 전략은 서로 다른 방향성을 가질 수 있고, 이에 따라 온라인 강의 시스템 설계 목표를 성취 점수인 단일 지표로 한정하는 것보다, 학습자의 학업 성취와 학습 경험의 질을 동시에 고려하는 측면의 최적화 관점이 필요함을 이론적으로 도출할 수 있다.

이는 향후 AI 교육 시스템이 “가장 점수가 잘 오르는 인터페이스”와 “가장 사용자가 머무르고 싶어하는 인터페이스” 사이에서 균형점을 설계적으로 탐색해야 함을 의미한다.

## 7.3 말투-속도 조합의 맥락 의존적 최적성

사회적 실재감과 인지부하 분석에서 공통적으로 관찰된 중요한 패턴은, 응답 속도나 언어 스타일의 단독적인 효과보다 두 요인의 결합된 방식이 사용자 경험을 좌우할 가능성이 크다는데 있다. 사회적 실재감에서는 언어 스타일과 응답 속도 상호작용 경향성( $p = .051$ )이 관찰되었고, 대화체와 빠른 응답을 사용한 집단에서는 아바타를 실재 존재하는 것으로 느끼는 사회적 실재감이 높았으나, 대화체와 느린 응답을 사용하는 집단에서는 사회적 실재감이 낮았다. 반대로 격식체를 사용하는 온라인 강의 조건에서는 응답이 느린 경우에도 사회적 실재감이 상대적으로 비슷하게 관측되었다.

인지부하에서도 유사한 교차 구조를 확인할 수 있었다. 대화체를 사용한 집단에서는 응답이 느려질

경우 인지부하가 급격히 증가하였지만, 격식체를 사용하는 집단에서는 동일한 지연이 상대적으로 완만한 부담 증가로 해석되는 경향을 보여주었다.

이러한 결과는 “빠른 응답이 항상 최선”이라는 단순 규칙이 성립하지 않음을 보여주며, 오히려 언어 스타일이 설정하는 기대 프레임에 맞춰 응답 지연 시간을 조절하는 설계가 더 중요한 요인일 수 있다는 점을 부각시켜준다.

최근 LLM 을 학습에 사용하는 것이 익숙해지면서 추론 과정에 필요한 시간이 필요한 부분이 당연시되면서, 응답 지연에 대해 이미 적응한 사용자도 상당수 존재하였다. 이처럼 기존 선행 연구의 내용처럼 모든 응답 지연이 부정적으로 인식되는 것은 아니다. 여기서 중요한 것은 응답 지연의 절대적 길이가 아니라, 그 지연이 현재 상호작용 페르소나와 의미적으로 부합하느냐인 것이다..

따라서 기술통계 기반 설계 시사점은 다음과 같이 요약될 수 있다. 대화체를 사용할 경우에는 가능한 한 짧은 응답 지연을 고려하는 것이 바람직하며, 격식체를 사용할 경우에는 다소 긴 응답도 학습 성과 및 실재감을 반드시 저해하지 않을 수 있다는 점이다. 이 결과는 단일한 고정 응답 속도 전략이 아닌, 말투에 따른 가변 지연 전략이 필요함을 의미한다.

#### 7.4 몰입(Flow)에서의 대화체 효과의 일관성과 이론적 정합성

설문한 몰입 점수를 기반으로 집단별로 분석한 결과 응답 속도와 언어 스타일의 상호작용 효과가 없다고 관찰되었다. 통계 분석 결과 거의 0 에 가까운 수준으로 나타났으며( $p = .984$ ), 통계적으로 유의한 주효과 역시 관찰할 수 없었다. 그러나 기술통계에 기반하여 모든 응답 속도 조건에서 대화체가 격식체에 비해서 실험자는 평균적으로 강의에 몰입을 하였다고 기술하였다는 점에서 몰입이라는 지표는 응답 속도보다는 언어 스타일에 더 민감하다는 가능성을 시사한다고 볼 수 있다. 즉 AI 아바타를 이용한 온라인 강의 설계시 학습자의 몰입을 위해 대화체를 사용하는 것이 학습자의 정서를 안정적으로 상향 조정하는 효과를 가질 수 있다.

이러한 저자의 분석 결과는 Mayer(2021)의 개인화 원리와 부합한다. 이 연구자는 비형식적이고 대화적인 언어가 학습자의 외재적 인지부하를 감소시키고, 사회적 유대감과 학습 동기를 강화함으로써 학습 경험의 질을 향상시킬 수 있음을 주장하였다. 따라서 본 연구의 몰입 결과는 이러한 이론적 주장과 방향성 측면에서 일치하는 경향을 보여준다. 따라서 설계적 관점의 측면에서, 몰입을 핵심 목표로 하는 강의는 대화체 언어를 기본값으로 채택하는 것이 합리적이다.

#### 7.5 학습자 수준 기반 페르소나 차별화 가능성 및 한계

사전 학습 경험 여부를 고려한 분석에서 학습자 수준에 따라 유리한 상호작용 조건이 달라질 가능성이 있다는 점이 관찰되었다. 사회조사 연구 방법론의 사전 학습 경험이 없는 학습자 집단에서는 대화체 및 빠른 응답(B) 집단이 가장 높은 평균 상승폭을 보였으며, 사전 경험이 있는 학습자 집단에서는 격식체 및 빠른 응답(C)과 격식체 및 느린 응답(E) 집단에서 평균이 가장 높았다는 점에서 볼 때 학습자 수준에 따라 유리한 상호작용 페르소나가 달라질 가능성이 있다는 점을 시사한다고 본다. 또한 해당 과목 또는 주제에 대해서 처음으로 학습하는 학습자의 경우 친근하고 즉각적인 피드백을 제공하는 대화체와 빠른 응답 시스템을 활용하는 것이 효과적일 수 있으며, 복습을 하는 학습자에게는 격식체와 느린 응답이 적용된 강의를 제공하는 것이 효과가 있을 수 있다.

다만 각 표본 수가 매우 작아(3~7 명 수준) 본 결과는 통계적 확정 결론으로 제시하기 어렵다. 따라서 이 연구에서는 이를 “관찰된 경향”으로만 기술하며, 후속 연구에서는 표본 확대와 수준별 무작위 배정을 통해 이 가설을 체계적으로 검증할 필요가 있다.

## 7.6 UX 설계의 보완점

실험자와의 인터뷰를 통해 얻은 메시지를 확인한 결과 아바타의 눈, 손짓, 표정 등의 비언어 표현을 추가하는 것이 몰입에 도움이 될 것이라는 의견, 강의 내용을 자막으로 제공해달라는 의견이 있었으며, 응답 지연 시 “먹통처럼 느껴졌다”는 의견과 “대기 상태 표시가 필요하다”는 요구 있음을 확인하였다. 이것은 특히 인지부하 상승 패턴을 해석하는데 중요하다. 질문에 대한 아바타의 응답 지연 그 자체보다, 현재 시스템이 정상 작동 중이라는 신호가 없는 상태가 학습자의 인지적 불확실성을 크게 만들 가능성이 크다. 예를 들어, 인터넷 서치 중 화면이 일정 시간 내에 바뀌지 않거나, LLM 에 질문을 했는데 답변이 생성되지 않는 경우 인터넷 이상을 예상하고 페이지 새로 고침을 하는 것과 동일하다고 볼 수 있다.

이 연구는 위 문제를 완화하기 위하여 보조적 UX 로써, 아바타가 생각하는 표정을 보여줌으로써 로딩중임을 알 수 있도록 하고, 아바타의 말을 메시지, 자막 등으로 함께 표시하는 UX 신호는 선택적으로 적용하도록 검토할 필요성이 있다고 보았다. 이 점은 단순히 기술적 보완으로써가 아닌, 인지부하 및 실재감을 조절하는 심리적 장치라고 볼 수 있기 때문이다. 특히 자막은 청각의 기능을 보조할 뿐만 아니라, 발음·역양 오해를 감소시키고, 응답 지연 시 “정보가 곧 제공된다”는 신호 제공이라는 이중적 기능을 수행할 수 있다, 이 연구에서는 제외되었지만 후속 연구에서 별도의 조작 요인으로 추가하여 실험적으로 검증할 가치가 있다고 판단하였다.

## 7.7 결론

이 연구는 AI 아바타 기반 온라인 강의 환경에서 응답 속도, 언어 스타일을 조절 변인으로 하는 실시간 질의응답 제공 여부의 독립 변인이 학습자의 학습 성과와 학습경험에 미치는 영향을 실험을 통해 검증하였다. 실험 전에 선행 연구를 기반으로 설정한 모든 가설이 통계적으로 유의하진 않았지만, 실시간 질의응답을 사용한 집단의 평균이 통제 집단에 비해 높았다는 점과 대화체로 강의를 들은 집단이 격식체로 강의를 들은 집단에 비해 몰입 점수가 높았다는 점은 의의가 있다고 본다. 추가적으로 집단별 평균 변화량과 효과 크기 그리고 학습자와의 인터뷰에서 나타난 응답을 분석한 결과 다음 여섯 가지로 정리할 수 있다.

첫째, 실시간 질의응답을 활용한 집단은 통제집단에 비해 평균적으로 실험 전후 점수 변화폭은 크지 않았지만, 점수의 변화량 지표인 분산은 더 작게 나타났음을 볼 수 있었고, 이는 학습 성과의 변동성을 낮추는 방향으로 작용했을 가능성이 있음을 내포한다.

둘째, 실험 전후 점수 변화폭이 가장 컸던 조건과 학습자가 느낀 몰입, 사회적 실재감 점수가 높았던 조건이 항상 일치하지 않을 수 있다는 점을 다시 한 번 확인할 수 있었다. 이는 학습자의 학습 성과와 경험이 동일한 조건에서 동시에 극대화되지 않을 수 있음을 시사한다고 볼 수 있으며, 하나의 지표로만 근거해서 상호작용 조건을 설계하기 보다는 서로 다른 지표들 사이의 균형을 고려할 필요가 있다. 또한 본 연구는 예상한 결과와 다른 방향의 결과가 충분히 나올 수 있음을 보여주어, 보다 유연한 설계 관점이 필요하다는 점도 시사한다

셋째, 아바타의 빠른 응답 속도가 학습자에게 항상 유리한 방향으로 작동하지 않는다는 점이다. 아바타의 언어스타일 중 대화체 집단일 경우 학습자에게 긍정적 방향성을 보인 반면에, 격식체 사용 집단에서는 효과가 제한적으로 나타나 응답 속도의 효과와 언어 스타일과의 조합에 따라 학습자의 성취도, 몰입, 사회적 실재감 등이 달라질 수 있다는 점을 확인하였다.

넷째, 온라인 강의에 몰입 정도는 모든 응답 속도 조건에서 대화체 언어를 사용한 강의를 학습한 집단이 격식체 언어를 사용한 강의를 학습한 집단보다 높은 평균값을 보여, 언어 스타일이 정의적 경험에 독립적인 영향을 미칠 가능성이 관찰되었다.

다섯번째, 온라인 강의의 주제였던 사회조사 연구 방법론을 학습한 적이 없는 집단에서는 대화체 및 응답 속도 빠른 집단에서 실험 전후 점수 상승폭이 상대적으로 컸으며, 이전에 학습한 적이 있는 집단에서는 격식체 집단에서 더 큰 상승폭을 보여주었다, 이는 학습자 수준에 따라 필요한 상호작용 페르소나가 다를 수 있음을 시사한다.

여섯번째, 실험 종료 후 짧은 인터뷰를 통해 강의 내용의 자막 표시, 응답 대기 상태 아이콘의 표시, 아바타의 시선 이동과 같은 비언어적 신호에 대한 필요함을 알 수 있었다. 이를 통해 AI 아바타 온라인 강의 설계시 학습자의 인지부하를 완화하고 사회적 실재감을 보완하는 UX 설계 요소의 보조적 수단을 추가한다면 학습자가 좀 더 나은 성취도를 얻을 수 있는 가능성을 보여준다.

다만, 연구의 한계로서 총 52 명의 참여자를 모집하여 5 개 실험 집단에 무작위 배정하였다. 집단당 표본 수는 10~11 명으로, 이는 통계적 검정력 확보를 위한 권고 표본 수엔 미치지 못한다는 점이다. 이 점에서 본 연구는 처음부터 확증적 검증보다는 변인 조합의 방향성과 효과 크기를 탐색하는데 목적을 둔 탐색적 파일럿 연구로 설계되었음을 명시한다. 실제로 본 연구에서 관찰된 효과 크기와 평균 패턴의 구조적 일관성은 표본이 충분히 커질 경우 통계적 유의성이 확보될 가능성을 내포한다고 볼 수 있다. 따라서 본 연구의 통계 결과는 가설의 기각 또는 채택으로 결론짓기보다, 후속 대규모 실험에서의 표본 수 산정, 변인 조합 최적화, 실험 설계 세분화를 위한 기초 근거로 활용하는 것이 적절해 보인다.

이 연구는 응답 속도와 언어 스타일이라는 미시적 설계 변인이 학습 성과와 경험 품질에 서로 다른 방향으로 작용할 수 있음을 실험적으로 보여주었으며, 향후 연구에서 표본 수 확장, 변인 조합 세분화, 매개·조절효과 검증을 포함한 실험 설계를 위한 출발점으로써, 학습자와 상호작용할 수 있는 AI 아바타를 이용한 온라인 강의 기반 구축에 도움이 될 수 있기를 기대한다.

## 8. 참고문헌

- Biocca, F., Harms, C., & Burgoon, J. K. (2003). Toward a More Robust Theory and Measure of Social Presence: Review and Suggested Criteria. *Presence: Teleoperators and Virtual Environments*, 12(5), 456–480. <https://doi.org/10.1162/105474603322761270>
- Briz-Ponce, L., Pereira, A., Carvalho, L., Juanes-Méndez, J. A., & García-Peñalvo, F. J. (2017). Learning with mobile technologies – Students’ behavior. *Computers in Human Behavior*, 72, 612–620. <https://doi.org/10.1016/j.chb.2016.05.027>
- Csikszentmihalyi, M. (1991). *Flow: The psychology of optimal experience* (1. HarperPerennial ed). HarperPerennial.
- Georgiou, Y., & Kyza, E. A. (2018). Relations between student motivation, immersion and learning outcomes in location-based augmented reality settings. *Computers in Human Behavior*, 89, 173–181. <https://doi.org/10.1016/j.chb.2018.08.011>
- Herbert, C., & Dołżycka, J. D. (2024). Teaching online with an artificial pedagogical agent as a

teacher and visual avatars for self-other representation of the learners. Effects on the learning performance and the perception and satisfaction of the learners with online learning: Previous and new findings. *Frontiers in Education*, 9, 1416033.

<https://doi.org/10.3389/feduc.2024.1416033>

Islam, M. Z., & Wang, G. (2025). Avatars in the educational metaverse. *Visual Computing for Industry, Biomedicine, and Art*, 8(1), 15. <https://doi.org/10.1186/s42492-025-00196-9>

Kodani, N., Uchida, T., Kameo, N., Sakai, K., Funayama, T., Minato, T., Kikuchi, A., & Ishiguro, H. (2025). Android avatar improves educational effects by embodied anthropomorphization. *Frontiers in Robotics and AI*, 11, 1469626.

<https://doi.org/10.3389/frobt.2024.1469626>

Mayer, R. E. (2021). *Multimedia learning* (Third edition). Cambridge University Press.

<https://doi.org/10.1017/9781316941355>

Nah, F. F.-H. (2004). A study on tolerable waiting time: How long are Web users willing to wait? *Behaviour & Information Technology*, 23(3), 153–163.

<https://doi.org/10.1080/01449290410001669914>

Pang, H., Tang, S., Han, J. Y., & Fung, F. M. (2023). Exploring the Use of an Avatar-Based Online Platform to Facilitate Social Interaction in Laboratory Sessions. *Journal of Chemical Education*, 100(10), 3832–3840. <https://doi.org/10.1021/acs.jchemed.3c00093>

*Response Time Limits: Article by Jakob Nielsen*. (n.d.). Nielsen Norman Group. Retrieved November 4, 2025, from <https://www.nngroup.com/articles/response-times-3-important-limits/>

Petty, R. E., & Cacioppo, J. T. (1986). The Elaboration Likelihood Model of Persuasion. In *Advances in Experimental Social Psychology* (Vol. 19, pp. 123–205). Elsevier.

[https://doi.org/10.1016/S0065-2601\(08\)60214-2](https://doi.org/10.1016/S0065-2601(08)60214-2)

Sherry, J. L. (2004). Flow and Media Enjoyment. *Communication Theory*, 14(4), 328–347.

<https://doi.org/10.1111/j.1468-2885.2004.tb00318.x>

Sinlapanuntakul, P., & Zachry, M. (2025). Perception in Pixels: Effects of Avatar Representation in Video-Mediated Collaborative Interactions. *Proceedings of the 4th Annual Symposium on Human-Computer Interaction for Work*, 1–16.

<https://doi.org/10.1145/3729176.3729183>

Sweller, J. (1988). Cognitive Load During Problem Solving: Effects on Learning. *Cognitive Science*, 12(2), 257–285. [https://doi.org/10.1207/s15516709cog1202\\_4](https://doi.org/10.1207/s15516709cog1202_4)

Ukenova, A., Bekmanova, G., Zaki, N., Kikimbayev, M., & Altaibek, M. (2025). Assessment and Improvement of Avatar-Based Learning System: From Linguistic Structure Alignment to Sentiment-Driven Expressions. *Sensors*, 25(6), 1921.

<https://doi.org/10.3390/s25061921>

Xia, T., Tan, M., & Guo, J. (2024). Avatars improve the quality of online learning: The role of attention and positive emotions. *Education and Information Technologies*, 29(18), 24203–24222.

<https://doi.org/10.1007/s10639-024-12790-0>