

생성형 AI에 대한 대학생 학습자의 개념적 이해와 활용 기대

대학 AI 리터러시 교육에의 함의

University Students' Conceptualizations of Generative AI and Their Expectations for Its Use
: Implications for AI Literacy Education in Higher Education

변경가*

국문요약 본 연구는 대학생 학습자가 생성형 AI를 어떻게 개념적으로 이해하고 있으며, 이러한 이해가 활용 기대와 어떠한 관계를 보이는지를 탐색하였다. 이를 위해 대학생 38명을 대상으로 ChatGPT의 개념, 작동 원리, 향후 활용 계획에 대한 서면 응답을 수집하고 정신 모형의 관점에서 분석하였다. 연구 결과, 참여자들은 전반적으로 생성형 AI의 작동 원리를 정교하게 이해하기보다는 단순하고 직관적인 정신 모형에 의존하는 경향을 보였다. 대표적인 오개념으로는 생성형 AI를 검색엔진처럼 외부 정보를 탐색·정리하는 시스템으로 이해하는 검색엔진 모형과, 내부 정보를 조회·조합하는 시스템으로 이해하는 데이터베이스 모형이 나타났다. 또한 이러한 정신 모형은 활용 기대와도 연결되어, 생성형 AI를 정보 탐색과 자료 정리 중심으로 활용하려는 경향이 두드러졌다. 반면 비교적 정교한 이해를 보인 일부 참여자는 학습, 글쓰기, 아이디어 생성, 문제 해결 등 보다 폭넓은 활용 가능성을 제시하였다. 이에 따라 향후 대학 교육에서는 생성형 AI의 작동 원리와 한계를 이해하도록 돕는 개념적 AI 리터러시 교육과, 이를 학습의 전 과정에서 비판적으로 활용하도록 돕는 실천적 AI 리터러시 교육이 이루어질 필요가 있다.

핵심어 생성형 AI, AI 리터러시, 정신 모형, 대학 교육, 매체 교육, 매체 인식

- 차례**
1. 서론
 2. 이론적 배경
 3. 연구 방법
 4. 연구 결과
 5. 결론 및 시사점

한 학습 활동을 지원하는 도구로 활용되고 있으며, 대학생들은 학습 과정에서 이를 점점 더 적극적으로 활용하고 있다. 생성형 AI는 대학 교육에 영향을 미치는 새로운 매체 환경으로 자리 잡고 있다.

이에 따라 생성형 AI가 교수·학습에 미치는 영향과 교육적 활용 가능성을 탐색하는 과정에서 생성형 AI에 대한 학습자의 인식을 탐색하는 연구들이 활발하게 이루어지고 있다. 인간-컴퓨터 상호작용(Human-Computer Interaction, HCI) 연구에서는 이용자가 기술을 어떻게 이해하고 해석하는지가 실제 기술 활용 방식에 중요한 영향을 미친다는 관점에서, 주로 기술수용모형(Technology Acceptance Model, TAM)을 바탕으로 AI에 대한 태도, 신뢰, 기대 등의 요인이 AI의 활용 및 수용에 미치는 인과 관계를

1. 서론

최근 생성형 인공지능(Generative AI)은 대학 교육 환경 전반에 빠르게 확산되고 있다. ChatGPT와 같은 챗봇 기반 생성형 AI는 자료 탐색, 정보 요약, 텍스트 생성 등 다양

* 서원대학교 국어교육과 조교수

분석하는 연구가 지속적으로 수행되어 왔다.¹ 또한 교육 연구에서도 생성형 AI 활용에 대한 긍정적 인식과 부정적 인식, 생성형 AI가 학습에 제공할 수 있는 교육적 가능성과 한계, 생성형 AI 활용에 대한 기대와 우려 등을 중심으로 학습자가 생성형 AI에 대해 어떤 태도를 형성하고 있는지를 분석한 연구가 이루어져 왔다.² 이러한 연구는 생성형 AI가 교육 환경에 도입되는 과정에서 학습자의 생성형 AI 수용 양상을 이해하고 예측하는 데 중요한 기초 자료를 제공해 왔다.

그러나 대학생들이 실제 교수·학습 맥락에서 생성형 AI를 어떠한 방식으로 활용하는지를 이해하기 위해서는 이러한 태도 중심의 인식 조사만으로는 충분하지 않다. 이러한 접근은 생성형 AI에 대한 전반적인 태도나 수용 양상을 설명하는 데 유용하지만, 이용자가 기술 그 자체를 어떻게 이해하고 해석하는지를 설명하기에는 한계가 있기 때문이다. 생성형 AI는 기술적 복잡성이 높고 작동 원리가 이용자에게 완전히 투명하게 제시되지 않는 특성을 지니므로, 이용자는 제한된 경험과 정보에 기반하여 해당 기술에 대한 나름의 설명 체계를 형성하게 된다. 선행 연구에 따르면 학습자는 새로운 기술이나 개념을 접하기 이전

에도 이미 일정한 이해 구조를 형성하는 경우가 많다. 예컨대 아동과 청소년은 인터넷, 코드, 유비쿼터스 컴퓨팅과 같은 추상적인 디지털 기술에 대해 학교 교육 이전부터 다양한 개념을 형성하는 것으로 보고된다.³ 이러한 경향은 AI 기술에 대해서도 유사하게 나타난다.⁴ 특히 대학생의 경우 생성형 AI에 대한 경험과 이해 수준이 이질적으로 분포되어 있어, 해당 기술을 이해하는 방식 또한 다양하게 나타날 가능성이 높다. 이에 따라 생성형 AI를 활용하는 방식 역시 이러한 이해의 차이에 따라 달라질 수 있다. 따라서 대학생 학습자의 생성형 AI 활용을 보다 정교하게 이해하기 위해서는, 이들이 해당 기술의 작동 방식과 특성을 어떻게 이해하고 있는지를 함께 살펴볼 필요가 있다.

나아가 학습자가 특정 기술에 대해 초기 단계에서 형성한 오개념이 이후 학습 과정에서도 지속되거나 강화될 수 있다는 점에서,⁵ 이는 AI 리터러시 연구에서도 중요한 문제로 간주된다.⁶ 실제로 다수의 AI 리터러시 연구에서는 학습자가 인공지능에 대해 형성한 개념이 이후 기술 활용과 학습에 중요한 영향을 미칠 수 있음을 지적하며, 학습자가 보다 적절한 이해를 형성하도록 지원하는 것을 주요 교육적 목표로 제시하고 있다.⁷

1 김근영·김승인, 「AI 영상 편집 도구 사용자 경험 분석: 지각된 유희성을 반영한 확장 TAM 모델을 중심으로」, 『산업진흥연구』 10(3), 산업진흥원, 2025, 99~109쪽; 손석·등기령·왕옥뢰·Liu Ziyang, 「젊은 사용자들이 정신 건강을 위해 AI 챗봇을 지속적으로 사용하는 방법 이해하기: TAM과 PSI 관점」, 『한국컴퓨터정보학회논문지』 30(11), 한국컴퓨터정보학회, 2025, 329~339쪽.

2 김보현, 「대학생의 학술 발표 과제를 위한 생성형 AI 활용 양상과 인식 연구」, 『교양학연구』 33, 다빈치미래교양연구소, 2025, 189~225쪽; 박종향·신나민, 「인공지능기술과 인공지능교사에 대한 인식 분석: 초·중·고등학생의 관점에서」, 『한국교원교육연구』 34(2), 한국교원교육학회, 2017, 169~192쪽; 변경가, 「글쓰기 맥락에서 대학생 필자의 생성형 AI 수용 인식 연구: 인공지능 기기 수용(AIDUA) 모형에 기반한 수용 구조 탐색과 타당도 검증」, 『청람어문교육』 106, 청람어문교육학회, 2025, 183~220쪽; 송윤희, 「ChatGPT의 학습 및 수업 활용에 대한 대학생의 인식 분석: K 신학대학교 사례를 중심으로」, 『기독교교육정보』, 한국기독교교육정보학회, 2024, 109~136쪽; 오의경, 「대학생의 AI 활용 특성과 인식 분석」, 『한국문헌정보학회지』 59(1), 한국문헌정보학회, 2025, 671~692쪽; 이한샘·유지원, 「대학생의 생성형 AI의 교육적 사용경험 및 인식 탐색: A대학 사례를 중심으로」, 『한국콘텐츠학회논문지』 24(1), 한국콘텐츠학회, 2024, 428~437쪽.

3 Mertala, P., Fagerlund, J., & Calderon, O., "Finnish 5th and 6th grade students' pre-instructional conceptions of artificial intelligence," *Computers and Education: Artificial Intelligence* 3, 2022, 100095.

4 Kreinsen, M. & Schulz, S., "Students' conceptions of artificial intelligence," *Proceedings of the 16th Workshop in Primary and Secondary Computing Education*, 2021, pp.1~2; Mertala, P., Fagerlund, J., & Calderon, O., "Finnish 5th and 6th grade students' pre-instructional conceptions of artificial intelligence," *Computers and Education: Artificial Intelligence* 3, 2022, 100095.

5 Biber, C., Tuna, A. & Korkmaz, S., "The mistakes and the misconceptions of the eighth grade students on the subject of angles," *European Journal of Science and Mathematics Education* 1(2), 2013, pp.50~59.

6 Kreinsen, M. & Schulz, S., "Students' conceptions of artificial intelligence," *Proceedings of the 16th Workshop in Primary and Secondary Computing Education*, 2021, pp.1~2.

7 Long, D. & Magerko, B., "What is AI literacy? Competencies and design considerations," *Proceedings of the 2020 CHI Conference*, 2020, pp.1~16; Su, J., Zhong, Y., & Ng, D. T. K., "A meta-review of literature on educational approaches for teaching AI at the K-12

이러한 논의는 최근 교육 분야에서 강조되고 있는 AI 리터러시 논의와도 밀접하게 관련된다. 최근 유럽 연합(EU) 집행위원회, OECD 등이 공동으로 제시한 「Empowering Learners for the Age of AI」 보고서⁸는 AI 리터러시를 미래 교육의 핵심 역량으로 제시하며, 이를 'AI와의 상호작용(Engaging with AI)', 'AI와의 창작(Creating with AI)', 'AI 활용 관리(Managing AI)', 'AI 설계 이해(Designing AI)'의 네 가지 영역으로 구조화하였다. 이 프레임워크는 각 영역에서 요구되는 역량을 지식, 기능, 태도의 측면에서 제시하는데, 특히 지식 영역에서 AI가 데이터를 처리하는 방식, 인간 사고와의 차이, AI 시스템에서 편향이 발생하는 원리 등 AI의 기술적 기반과 작동 원리에 대한 개념적 이해를 핵심 요소로 강조하였다.

이와 유사한 관점은 미국 노동부(Department Of Labor) 산하 교육훈련청(Employment and Training Administration)이 발표한 「AI Literacy Framework」⁹에서도 확인된다. 이 프레임워크는 청소년, 학생, 교직원, 구직자, 재직자 등 다양한 집단을 대상으로 설계된 범용 역량 체계로서,¹⁰ AI 리터러시를 생성형 AI를 중심으로 AI 기술을 책임감 있게 활용하고 평가하기 위한 기초 역량으로 정의하고, 이를 'AI 원리 이해', 'AI 활용 탐색', 'AI 효과적 지시', 'AI 결과 평가', '책임 있는 AI 사용'의 내용 영역으로 제시하였다. 특히 AI 원리 이해 영역에서는 AI가 데이터의 통계적 패턴을 기반으로 결과를 생성하는 방식, 학습과 추론의 과정, 환각과 같은 기술적 한계 등 AI의 작동 원리에 대한 기초적인 이해를 중요한 학습 내용으로 포함하고 있다.

이처럼 OECD와 미국 노동부에서 체계화한 AI 리터러

시 프레임워크는 적용 대상과 맥락에 차이가 있지만, 공통적으로 AI의 원리와 작동 방식에 대한 이해를 AI 리터러시의 핵심 요소로 강조한다는 점에서 주목할 필요가 있다. 이는 학습자가 AI 기술을 어떻게 개념화하여 이해하고 있는지가 AI 리터러시 형성의 중요한 기반이 됨을 시사하며, 학습자의 생성형 AI에 대한 개념적 이해 양상을 구체적으로 탐색할 필요성을 제기한다.

이에 본 연구는 대학생이 생성형 AI를 어떻게 개념적으로 이해하고 있는지를 분석하고, 이러한 이해가 생성형 AI 활용에 대한 기대와 어떠한 양상을 보이는지를 탐색하고자 한다. 여기서 개념적 이해란 학습자가 특정 기술이 무엇이며 어떻게 작동한다고 생각하는지에 대해 형성하는 이해 구조를 의미한다. 이를 위해 생성형 AI에 대한 개념적 이해를 정신 모형의 관점에서 유형화하고, 대학 교수·학습 맥락에서 생성형 AI의 역할과 활용 가능성에 대한 기대를 분석한 뒤, 양자 간의 관계를 질적으로 탐색하고자 한다. 이를 통해 생성형 AI 활용에 대한 대학생 학습자의 인식 구조를 보다 정교하게 이해하고, 대학 교육에서의 생성형 AI 활용 논의를 위한 기초적 시사점을 제공하고자 한다.

본 연구에서 설정한 연구문제는 다음과 같다.

첫째, 대학생은 생성형 AI를 어떻게 개념적으로 이해하고 있는가? 이는 어떻게 유형화되는가?

둘째, 대학생은 생성형 AI의 역할과 활용 가능성을 어떻게 기대하고 있는가? 이는 어떻게 유형화되는가?

셋째, 생성형 AI에 대한 개념적 이해 유형과 활용 기대 유형 사이에는 어떠한 관계가 나타나는가?

2. 이론적 배경

2.1. 생성형 AI에 대한 개념적 이해와 정신 모형

본 연구는 생성형 AI에 대한 학습자의 이해를 생성형

levels," *Computers and Education: Artificial Intelligence* 3, 2022, 100065.

8 OECD, *Empowering Learners for the Age of AI: An AI Literacy Framework for Primary and Secondary Education*, OECD, 2025.

9 U.S. Department of Labor, *The U.S. Department of Labor's Artificial Intelligence Literacy Framework*, 2026.

10 한국과학창의재단, 「미국 노동부(DOL) AI 리터러시 프레임워크 발표 및 전국 확산 추진」, 정책 브리프, 2026.

AI의 작동 원리와 활용을 어떻게 개념화하는지를 중심으로 살펴보고자 한다. 이러한 개념적 이해를 설명하기 위해 인지과학, 심리학, 컴퓨터 과학 등 다양한 학문 분야에서 인간의 이해 구조를 설명하는 여러 인지적 개념들이 논의되어 왔다. 예컨대 내적 표상(internal representations)이나 내적 모형(internal models)은 인간이 외부 세계를 이해하기 위해 형성하는 인지적 표상 구조를 설명하는 데 활용되어 왔으며, 스키마(schema) 개념 역시 이러한 이해 과정을 설명하는 주요 이론적 틀로 제시되어 왔다. 스키마 이론에 따르면 새로운 지식은 기존 지식 구조를 기반으로 형성되며, 학습자는 기존 스키마와 새로운 경험 사이의 동화와 조절 과정을 통해 자신의 이해 구조를 재구성한다. 이외에도 학습자가 형성한 개념을 설명하기 위해 대체 개념(alternative conception), 직관적 이론, 신념, 아이디어 등의 용어가 사용되기도 한다.¹¹

이처럼 인간의 이해를 설명하기 위한 다양한 개념들이 제시되어 왔으나, 본 연구에서는 복잡한 기술 시스템에 대한 이용자의 이해를 분석하는 데 적합한 이론적 틀로서 정신 모형(mental model) 개념에 주목하고자 한다. 정신 모형은 Johnson-Laird(1980, 1983)에 의해 체계화된 개념으로,¹² 인간이 주변 세계를 이해하고 설명하며 예측하기 위해 구성하는 인지적 표상을 의미한다. 정신 모형은 현대의 사회 기술적 시스템(socio-technical systems)과 같이 복잡하고 비가시적인 기술을 인간이 어떻게 이해하는지를 설명하는 데 유용한 개념으로, 기술 시스템에 대한 이용자의 이해를 분석하는 주요 이론적 틀로 자리 잡으며 HCI 연구를 비롯한 다양한 분야에서 폭넓게 활용되어 왔다.¹³ 또한 미디어 및

AI 리터러시 연구에서도 이용자의 정신 모형을 분석하는 접근의 중요성이 점차 강조되고 있으며, Tedre et al.(2021)은 AI 교육 연구의 주요한 향후 과제로 정신 모형 연구를 제시한 바 있다.¹⁴

정신 모형은 인간이 외부 세계의 대상이나 시스템이 어떻게 작동하는지를 이해하기 위해 형성하는 비교적 안정적인 지식 구조로 정의된다.¹⁵ 즉, 이는 이용자가 특정 기술이 무엇이며 어떻게 작동하는지에 대해 구성한 설명 체계로 이해할 수 있다. 인간은 동일한 대상에 대해서도 개인의 경험과 지식에 따라 서로 다른 방식으로 이해하고 설명하는데, 정신 모형은 이러한 인지적 차이를 설명하는 기저 구조로 작용한다. 이용자는 기술의 실제 작동 원리를 완전히 알지 못하더라도 자신의 경험과 지식을 바탕으로 기술에 대한 설명 구조를 형성하며, 이러한 정신 모형은 기술 사용 과정에서의 기대와 판단을 조직하는 역할을 한다.

그러나 정신 모형은 실제 시스템의 구조를 그대로 반영하기보다는 부분적이고 단순화된 형태로 구성되는 경향이 있다. Norman(1983)은 정신 모형이 현실을 축약한 불완전한 표상임을 지적하였으며,¹⁶ 시스템 제어와 관련된 연구에서도 전문가조차 실제 시스템과 완전히 일치하지 않는 정신 모형을 형성하는 것으로 나타났다.¹⁷ 이는 전

of shared mental models in human-AI teams: A theoretical review," *Theoretical Issues in Ergonomics Science* 24(2), 2023, pp.129~175.

14 Marx, E., Leonhardt, T., & Bergner, N., "Secondary school students' mental models and attitudes regarding artificial intelligence: A scoping review," *Computers and Education: Artificial Intelligence* 5, 2023, 100169에서 재인용.

15 Johnson-Laird, P. N., *Mental models: Towards a cognitive science of language, inference, and consciousness*, Harvard University Press, 1983; Rouse, W. B. & Morris, N. M., "On looking into the black box: Prospects and limits in the search for mental models," *Psychological Bulletin* 100(3), 1986, pp.349~363.

16 Norman, D. A., "Some observations on mental models," *Mental Models*, 1983.

17 Rouse, W. B. & Morris, N. M., "On looking into the black box: Prospects and limits in the search for mental models," *Psychological Bulletin* 100(3), 1986, pp.349~363.

11 Marx, E., Leonhardt, T., & Bergner, N., "Secondary school students' mental models and attitudes regarding artificial intelligence: A scoping review," *Computers and Education: Artificial Intelligence* 5, 2023, 100169.

12 Johnson-Laird, P. N., "Mental models in cognitive science," *Cognitive Science* 4(1), 1980, pp.71~115; Johnson-Laird, P. N., *Mental models: Towards a cognitive science of language, inference, and consciousness*, Harvard University Press, 1983.

13 Andrews, R. W., Lilly, J. M., Srivastava, D., & Feigh, K. M., "The role

문성을 갖춘 경우에도 시스템의 실제 작동 원리를 완전하게 반영한 이해를 형성하기 어렵다는 점을 시사하며, 상대적으로 경험과 지식이 제한된 일반 이용자나 초보 학습자의 경우에는 실제 작동 원리와 더욱 큰 차이를 보이는 이해 구조를 형성할 가능성이 있음을 보여준다.

이러한 관점에서 본 연구는 대학생 학습자가 생성형 AI와 상호작용하는 과정에서 해당 기술을 어떻게 개념화하는지에 주목한다. 생성형 AI는 확률적 언어 생성 메커니즘을 기반으로 작동하는 기술로서 그 작동 과정이 직접적으로 드러나지 않는 특성을 지닌다. 생성형 AI는 대규모 언어 모델에 기반하여 입력된 맥락을 바탕으로 다음에 올 단어의 확률을 예측하고 이를 순차적으로 생성하는 방식으로 작동하는 기술이다.¹⁸ 이는 외부 정보를 검색하거나 데이터베이스를 조회하는 방식과는 구별되는 생성 메커니즘으로, 학습된 통계적 패턴을 바탕으로 새로운 텍스트를 산출한다는 점에서 특징적이다.

이러한 생성형 AI는 웹 텍스트, 도서, 위키피디아 등 다양한 코퍼스를 기반으로 학습되며, 토큰화와 임베딩, 트랜스포머 구조를 활용한 학습, 그리고 인간 피드백을 반영한 미세 조정 과정을 거쳐 구축된다.¹⁹ 그러나 이러한 학습 과정과는 별개로, 실제 사용 단계에서 생성형 AI는 주어진 입력에 대해 확률적 예측을 반복하는 방식으로 응답을 생성하며, 그 작동 과정이 이용자에게 직접적으로 드러나지 않는 특성을 지닌다.

따라서 학습자는 생성형 AI의 실제 작동 원리를 완전히 이해하기보다는 제한된 경험과 정보를 바탕으로 나름의 설명 체계를 형성하게 된다. 특히 생성형 AI와 같이 작동 원리가 불투명한 기술의 경우, 학습자가 기존 경험과 지식을 바탕으로 의미를 구성하는 과정에서 서로 다른 개념적 이해를 형성할 가능성이 크며, 이 과정에서 생성형 AI에

대한 특정한 이해 방식이나 오개념이 형성되어 실제 활용 방식에 영향을 미칠 수 있다. 이러한 점에서 학습자가 생성형 AI를 어떻게 개념적으로 이해하고 있는지를 탐색하는 것은 생성형 AI 활용에 대한 기대와 사용 양상을 이해하는 데 중요한 단서를 제공할 뿐 아니라, 대학 교육 맥락에서 요구되는 AI 리터러시를 정교하게 논의하기 위한 기초를 제공한다.

2.2. 생성형 AI에 대한 학습자의 정신 모형 연구 : 방법론과 주요 동향

정신 모형은 인간이 기술이나 시스템을 어떻게 이해하는지를 설명하는 중요한 이론적 틀로 활용되어 왔다. 그러나 정신 모형은 직접적으로 관찰 가능한 대상이 아닌 인지적 구조이기 때문에 이를 직접 측정하거나 확인하는 데에는 본질적인 한계가 따른다. 이와 관련하여 Norman(1987)은 실제 시스템, 시스템에 대한 설계자의 모형, 개인이 형성한 정신 모형, 개인의 정신 모형을 바탕으로 연구자가 구성한 연구자의 모형이라는 네 가지 수준을 구분하였다. 이에 따르면 연구자가 관찰하거나 분석할 수 있는 것은 개인의 실제 정신 모형 자체가 아니라, 개인의 발화나 행동을 바탕으로 재구성한 ‘모형에 대한 모형’에 해당한다. 이러한 점에서 정신 모형 연구는 개인의 이해 구조를 직접적으로 포착하기보다는 다양한 자료를 통해 이를 간접적으로 추론하는 접근을 취한다.

이와 같은 특성에 따라 정신 모형을 탐색하기 위한 방법으로는 관찰, 면담, 설문 조사, 개념 기술 방법 등이 활용되어 왔다.²⁰ 관찰 방법은 작업 수행 과정에서 나타나는 행동을 바탕으로 사고 과정을 추론하는 방식으로, 실제 작업 환경이나 시뮬레이션 상황에서 적용된다. 면담은 참여자의 설명을 통해 정신 모형을 탐색하는 방법으로, 구조화된 면담부터 반구조화, 비구조화 면담에 이르기까지 다양한

18 Brown, T., Mann, B., & Ryder, N., "Language models are few-shot learners," *Advances in Neural Information Processing Systems* 33, 2020, pp.1877~1901.

19 김성우, 『인공지능은 나의 읽기-쓰기를 어떻게 바꿀까』, 유투, 2024.

20 이하에서 제시하는 각 방법에 대한 설명과 관련 선행 연구는 Andrews et al.(2023)의 정리를 바탕으로 재구성한 것이다.

방식으로 이루어진다. 설문 조사는 특정 개념에 대한 이해를 정량적으로 파악하기 위해 활용되며, Likert 척도나 지식 평가 문항이 주로 사용된다. 또한 개념 지도, 카드 분류, 군집 분석과 같은 개념 기술 방법은 개념 간 관계를 구조적으로 드러내는 방식으로 정신 모형의 조직 구조를 분석하는 데 활용된다.

이러한 방법론을 바탕으로 AI에 대한 학습자의 정신 모형을 탐색하려는 연구가 이루어져 왔다.²¹ 초기 연구에서는 AI 기술이 일상화되기 이전 단계에서 사람들이 AI를 무엇으로 인식하고 어떻게 개념화하는지를 탐색하는데 초점이 맞추어졌다. Schank(1991)은 성인들이 AI를 ‘마법의 총알’, ‘추론 엔진’, ‘기계를 넘어서는 능력’, ‘학습하는 기계’ 등 다양한 방식으로 이해하고 있음을 제시하였다. 이후 AI 기술이 보편화되면서 일반 이용자의 이해를 경험적으로 분석하는 연구들이 수행되었으며,²² Alizadeh, Stevens, & Esau(2021)는 이용자들이 AI를 로봇이나 음성 비서와 같은 구체적 장치로 이해하거나 ‘자동화’, ‘주체성’, ‘개인화 서비스’와 같은 의미를 부여하는 경향이 있음을 보고하였다.²³

교육 연구에서도 학습자의 AI에 대한 정신 모형을 탐색하는 연구가 이루어져 왔다. Mertala, Fagerlund, & Calderon(2022)은 초등학생들이 AI를 인간과 유사한 인지 능력을 지닌 존재로 의인화하거나 센서를 통해 환경을 인식하는 기술로 이해하는 경향이 있음을 밝혔다.²⁴ 또한

Kreinsen & Schulz(2021)는 학습자들이 AI를 ‘프로그래밍된 데이터 저장 및 검색 시스템’으로 이해하는 경우가 많다고 보고하였으며²⁵, 이는 컴퓨터를 전지적 데이터베이스로 인식하는 기존 연구²⁶와 유사한 양상을 보인다고 설명하였다. 최근에는 생성형 AI 기술에 대한 이용자의 이해를 탐색하는 연구도 이루어지고 있다. Singh et al.(2023)은 대학생을 대상으로 생성형 AI 기반 글쓰기 도구의 작동 원리를 설명하도록 한 결과, 학습자들이 이를 ‘인간 언어 기반 시스템’, ‘규칙 기반 시스템’, ‘무작위적이고 창의적인 시스템’ 등으로 다양하게 개념화하고 있음을 확인하였다.²⁷ Anderson(2023) 역시 생성형 AI에 대한 이해를 은유적 프레임으로 분석하면서 이를 ‘글쓰기용 계산기’, ‘통계 소프트웨어’, ‘협력적 파트너’와 같은 방식으로 설명할 수 있음을 제시하였다.²⁸

이상의 연구들은 학습자가 AI를 이해하는 방식이 개인의 경험과 맥락에 따라 다양하게 구성되며, 실제 기술적 작동 원리와는 차이를 보일 수 있음을 공통적으로 시사한다. 특히 학습자들은 AI를 정보 검색 시스템, 자동화 도구, 또는 인간과 유사한 존재로 이해하는 경향을 보이며, 이러한 이해는 실제 활용 방식과 상호작용 양상에 영향을 미친다. 그러나 기존 연구는 주로 일반적인 AI 개념이나 특정 기술 사례를 중심으로 이루어져 왔으며, 대학 교육 맥락에서 대학생 학습자가 챗봇 기반 생성형 AI를 어떻게 이해하고 있는지를 정신 모형의 관점에서 분석한 연구는 아직 제한적인 실정이다. 따라서 생성형 AI의 작동 원리와 특성에

21 AI에 대한 이용자의 이해를 탐색한 연구들은 정신 모형 접근 외에도 민속 이론(naive theory), 개념적 은유(conceptual metaphor) 등 다양한 인지적 틀을 함께 활용하는 경향이 있다. 이에 본 연구에서는 이 가운데 정신 모형과 관련된 논의를 포함하는 연구를 중심으로 인용하였음을 밝힌다.

22 Schank, R. C., "Where's the AI?," *AI Magazine* 12(4), 1991, pp.38~38.

23 Alizadeh, F., Stevens, G., & Esau, M., "I don't know, is AI also used in airbags? An empirical study of folk concepts and people's expectations of current and future artificial intelligence," *I-com* 20(1), 2021, pp.3~17.

24 Mertala, P., Fagerlund, J., & Calderon, O., "Finnish 5th and 6th grade students' pre-instructional conceptions of artificial intelligence," *Computers and Education: Artificial Intelligence* 3, 2022, 100095.

25 Kreinsen, M. & Schulz, S., "Students' conceptions of artificial intelligence," *Proceedings of the 16th Workshop in Primary and Secondary Computing Education*, 2021, pp.1~2.

26 Rucker, M. T. & Pinkwart, N., "Review and discussion of children's conceptions of computers," *Journal of Science Education and Technology* 25(2), 2016, pp.274~283.

27 Singh, N., Bernal, G., Savchenko, D., & Glassman, E. L., "Where to hide a stolen elephant: Leaps in creative writing with multimodal machine intelligence," *ACM Transactions on Computer-Human Interaction* 30(5), 2023, pp.1~57.

28 Anderson, S. S., "Places to stand: Multiple metaphors for framing ChatGPT's corpus," *Computers and Composition* 68, 2023, 102778.

대한 학습자의 개념적 이해를 분석하는 것은 AI 리터러시 교육의 관점에서 필요성을 지닌다.

3. 연구 방법

3.1. 연구 참여자

본 연구는 충북 소재 A대학교에서 개설된 교양 강의를 수강하는 대학생 38명을 연구 참여자로 선정하였다. 참여자는 교양 의사소통 교과목을 수강하는 1학년 학생들로, 이공 계열 전공 학생으로 구성된 1개 분반과 인문 계열 전공 학생으로 구성된 1개 분반, 총 2개 분반에서 모집되었다. 참여자 전원은 생성형 AI(ChatGPT)를 활용해 본 경험이 있었으며, 일상에서의 생성형 AI 활용 빈도는 '자주 활용한다'에서 '종종 활용한다' 사이에 분포하였다. 이러한 점에서 본 연구의 참여자는 생성형 AI에 대한 일정 수준의 사용 경험을 갖춘 집단으로, 생성형 AI에 대한 개념적 이해와 활용 기대를 탐색하기에 적합하게 표집되었다.

3.2. 자료 수집

본 연구는 생성형 AI에 대한 대학생의 개념적 이해와 활용 기대를 탐색하기 위해 개방형 질문지를 활용한 서면 조사를 실시하였다. 조사에서는 생성형 AI의 대표적 사례로 ChatGPT를 제시하고, 이에 대한 참여자의 인식을 중심으로 자료를 수집하였다. 이는 ChatGPT가 현재 가장 널리 활용되는 생성형 AI 서비스이며, 연구 참여자 전원이 해당 도구를 사용해 본 경험이 있다는 점을 고려한 것이다.

먼저 생성형 AI에 대한 개념적 이해를 확인하기 위해 Mertala et al.(2022)과 Singh et al.(2023)의 연구를 참고하여 다음의 두 가지 질문을 제시하였다. 첫째, 'ChatGPT가 무엇인지 친구에게 설명한다면 어떻게 설명할 것인가'에 대해 자유 서술형으로 응답하도록 하고, 둘째, 'ChatGPT가

어떻게 작동한다고 생각하는가'에 대해 글과 그림을 활용하여 설명하도록 하였다. 이러한 문항 구성은 참여자가 생성형 AI에 대해 형성하고 있는 이해를 언어적·시각적으로 표현하도록 함으로써, 해당 기술에 대한 학습자의 정신 모형을 다각적으로 포착하기 위한 것이다. 다음으로 대학교육 맥락에서의 활용 기대를 확인하기 위해 '향후 대학 생활에서 ChatGPT를 어떠한 상황에서, 어떠한 방식으로 활용할 것인가'에 대해 서술하도록 하였다. 이를 통해 생성형 AI에 대한 개념적 이해와 실제적 활용 기대 간의 관계를 함께 분석하고자 하였다.

3.3. 자료 분석

자료 분석에는 연역적 추론과 귀납적 추론을 병행하는 귀추적 접근을 활용하였다. 질적 연구에서 일반적으로 활용되는 지속적 비교 분석법을 기반으로 자료를 반복적으로 검토하면서 범주를 도출하는 동시에, II장에서 검토한 선행 연구를 참고하여 분석 틀을 보완하고 해석의 타당성을 확보하였다. 특히 대학 글쓰기 교육 맥락에서 생성형 AI에 대한 학습자의 정신 모형을 분석한 Singh et al.(2023)의 연구를 주요 준거로 삼아, 생성형 AI에 대한 학습자의 정설명 방식과 이해 구조를 보다 체계적으로 분류하고자 하였다.

구체적인 분석 절차는 다음과 같다. 먼저 참여자의 서술 응답과 그림 설명 자료를 Excel에 텍스트로 전사 및 정리한 후, 생성형 AI에 대한 설명에서 나타나는 의미 단위를 중심으로 1차 코딩을 실시하였다. 이후 의미 단위 간의 유사성과 차이를 비교하여 범주를 구성하고, 각 범주의 특징을 분석하였다. 마지막으로 도출된 범주를 토대로 학습자가 생성형 AI를 어떻게 개념화하고 있으며, 이를 바탕으로 어떠한 활용 계획을 형성하고 있는지를 유형화하였다.

4. 연구 결과

4.1. 생성형 AI에 대한 대학생 학습자의 개념적 이해

4.1.1. 생성형 AI 정신 모형의 유형과 분포

본 연구는 대학생 학습자가 생성형 AI(ChatGPT)의 작동 원리를 어떻게 이해하고 있는지를 분석하기 위해, 참여자의 서술 응답을 바탕으로 정신 모형을 유형화하였다. 이때 Singh et al.(2023)의 틀을 참고하여, 설명의 초점(기능 중심 vs. 작동 원리 중심)과 설명의 기술적 깊이(단순 vs. 정교)의 두 차원을 적용하였다.²⁹ 이에 따라 참여자의 생성형 AI에 대한 이해 양상은 ‘기능-단순’, ‘기능-정교’, ‘작동-단순’, ‘작동-정교’의 네 가지 유형으로 분류되었다. 그 분포는 <표 1>과 같다.

<표 1> 생성형 AI(ChatGPT)에 대한 정신 모형의 분포

구분	기능 중심 모형	작동 원리 모형	합계
단순한 설명	15	20	35
정교한 설명	1	2	3
합계	16	22	38

분석 결과, 먼저 설명의 초점 차원에서 ‘작동 원리’를 중심으로 설명한 응답이 22명(57.9%)으로 ‘기능 중심’ 설명의 16명(42.1%)보다 다소 높은 비중을 차지하였다. 이는 상당수 참여자가 생성형 AI의 내부 작동 원리에 대해 나름의 설명을 시도하고 있음을 보여준다.

한편 이러한 결과는 질문지에서 작동 원리 설명을 명시적으로 요구하였다는 점을 고려할 때, 절반에 가까운 여전히 적지 않은 학습자가 기능 중심 설명에 머물러 있다는 점에서 주목할 필요가 있다. 기능 중심 모형은 생성형 AI

가 수행하는 결과나 역할(예: 요약, 설명, 답변 생성)에 초점을 두는 설명으로, 기술의 작동 메커니즘에 대한 이해가 결여된 상태를 반영한다. 이는 학습자들이 생성형 AI를 ‘무엇을 하는 도구인가’의 차원에서 이해하고 있으며, ‘어떻게 작동하는가’에 대한 인식은 상대적으로 제한적임을 시사한다.

다음으로 설명의 기술적 깊이 차원에서 보면, 전체 참여자 중 35명(92.1%)이 단순한 설명 수준에 해당하였으며, 정교한 설명은 3명(7.9%)에 불과하였다. 이는 학습자들이 생성형 AI의 작동 원리에 대해 일정한 설명을 시도하고 있음에도 불구하고, 그 설명이 개념적으로 충분히 구조화되거나 기술적으로 정확한 수준에는 도달하지 못하고 있음을 보여준다.

이를 종합하면, 가장 높은 비중을 차지한 유형은 ‘작동-단순’으로 20명(52.6%)이었다. 이는 참여자들이 생성형 AI의 작동 원리를 설명하려는 시도는 하고 있으나, 그 이해가 단편적이며 불완전한 수준에 머물러 있음을 의미한다. 특히 이 유형은 작동 원리에 대한 설명을 포함하고 있음에도 불구하고 오개념을 내포하거나 기존 기술 경험에 기반한 직관적 해석에 의존하는 경향을 보인다는 점에서, 대학생 학습자의 생성형 AI 이해가 개념적으로 충분히 정교화되지 않았음을 드러낸다.

이에 따라 이어지는 내용에서는 ‘작동-단순’ 유형을 중심으로, 참여자들이 생성형 AI의 작동 원리를 어떠한 방식으로 이해하고 있으며, 그 과정에서 어떠한 오개념이 나타나는지를 구체적으로 분석하고자 한다.

4.1.2. 생성형 AI 정신 모형에서 나타난 오개념 유형

앞서 살펴본 바와 같이, 다수의 참여자가 ‘작동-단순’ 유형에 해당하여 생성형 AI의 작동 원리를 설명하려는 시도를 보였으나, 그 설명에는 실제 기술적 메커니즘과 차이가 있는 다양한 오개념이 포함되어 있었다. 이에 본 연구에서는 생성형 AI에 대한 학습자의 이해 양상에서 나타난

²⁹ 기능 중심은 생성형 AI가 수행하는 역할이나 결과에 초점을 두어 설명하는 방식이며, 작동 원리 중심은 이러한 결과가 어떠한 기술적 메커니즘을 통해 생성되는지를 설명하는 방식이다. 단순 설명은 개념이 부분적이거나 모호한 수준을, 정교 설명은 작동 과정을 비교적 정확하고 구조화된 개념으로 설명하는 수준을 의미한다.

대표적인 오개념 유형을 도출하였으며, 그 결과는 <표 2>와 같다.

<표 2> 생성형 AI(ChatGPT)에 대한 오개념 유형

구분	빈도	주요 답변 사례
검색엔진모형	12	질문을 받으면 인터넷 검색을 통해 내용을 취득한 후 간추리는 작업을 거쳐 나에게 답을 한다(S11) ChatGPT가 질문 받은 것을 토대로 인터넷에서 구글링을 하여 정보를 정리 후 출력한다(S18) AI가 인터넷에서 자료를 찾아 분석하고 답한다(S23) 네이버, 구글 등을 찾아 답변할 것 같다(S38)
데이터베이스모형	6	입력돼 있는 데이터베이스에서 정보를 찾아 출력한다(S6) ChatGPT 내부에서 질문과 관련된 자료가 한 번에 나오고 그것들을 정리하여 답을 알려준다(S9) 방대한 데이터를 가지고 있어 키워드와 일치시키고 답을 한다(S33)
기타	2	단 몇 초 사이에 나의 말에 답변해준다. 매우 많은 데이터베이스를 기반으로 답변해주는데 초기와는 다르게 데이터는 계속해서 쌓이는 것이라고 생각해서 실시간으로 데이터는 추가되기도 할 것이다(S13) 질문을 하면 여러 명의 GPT들이 집단 지성으로 토론을 열며 그 가운데 가장 좋은 GPT가 대표를 맡아 토론 결과를 알려준다(S22)

분석 결과, 가장 두드러지게 나타난 유형은 ‘검색엔진 모형’(12명)이었다. 본 연구에서는 참여자들이 생성형 AI의 작동 원리를 인터넷 검색엔진과 유사한 절차와 기능으로 이해하고 있다는 점에서 이를 ‘검색엔진 모형’으로 명명하였다. 해당 참여자들은 ChatGPT가 질문을 입력받으면 인터넷 검색을 통해 관련 정보를 탐색하고 이를 정리 및 요약하여 제시한다고 설명하며, 생성형 AI의 작동 과정을 ‘외부 정보 탐색 → 수집 → 요약’의 절차로 이해하는 경향을 보였다. 그러나 생성형 AI의 기반이 되는 대규모 언어 모델은 외부 데이터베이스를 실시간으로 검색하는 방식이 아니라, 사전 학습을 통해 형성된 확률적 언어 모델을 바탕으로 문맥에 따라 다음 토큰을 예측함으로써 텍스트를 생성한다. 이러한 점에서 검색엔진 모형 유형은 생성형 AI가 외부 정보를 실시간으로 탐색한다고 보는 외부 정보 의존에 대한 오해와, 텍스트 생성 과정을 기존 정보를 정리하는 과정으로 환원하는 생성 과정에 대한 오해를 포함하고 있다. 이는 생성형 AI의 작동을 정보 검색 및 처리 과정으로 한정함으로써, 핵심 원리인 확률적 생성을 충분

히 이해하지 못한 결과로 해석할 수 있다.

다음으로 빈번하게 나타난 유형은 ‘데이터베이스 모형’으로(6명), 참여자들이 생성형 AI의 작동 원리를 내부에 저장된 정보의 조회와 조합 과정으로 이해하고 있다는 점에서 ‘데이터베이스 모형’으로 명명하였다. 해당 참여자들은 ChatGPT가 내부에 축적된 데이터를 기반으로 질문과 관련된 정보를 찾아 이를 출력한다고 설명하며, 생성형 AI의 작동을 ‘정보 저장 → 조회 → 조합’의 절차로 이해하는 모습을 보였다. 그러나 이러한 이해는 생성형 AI의 핵심 메커니즘인 문맥 기반 확률적 토큰 예측에 의한 텍스트 생성을 정확하게 반영하지 못한 것이다. 생성형 AI는 내부에 명시적으로 저장된 정보를 조회하여 응답을 생성하는 시스템이 아니라, 분산된 파라미터에 통계적 패턴의 형태로 지식을 인코딩하고, 이를 바탕으로 문맥에 따라 다음 토큰을 확률적으로 예측함으로써 텍스트를 생성한다. 이와 같은 확률적 생성 방식으로 인해 동일한 입력에 대해서도 다양한 결과가 산출될 수 있다. 따라서 데이터베이스 모형 유형은 생성형 AI의 작동을 명시적으로 저장된 정보를 조회하고 선택하는 과정으로 파악하여, 분산된 파라미터와 토큰 예측에 기반한 생성 메커니즘을 충분히 이해하지 못하였다고 볼 수 있다.

이 외에도 일부 응답에서 생성형 AI의 작동 원리 및 특성을 잘못 이해 방식이 확인되었다. 예컨대 생성형 AI가 이용자와의 상호작용을 통해 지속적으로 학습한다고 보거나(S13), 여러 AI가 협력하여 답변을 도출한다고 이해하는 경우(S22)이다. 이러한 설명은 생성형 AI의 작동 원리를 인간의 학습이나 사회적 의사결정 과정에 비유하여 해석한 것으로, 기술에 대한 의인화가 반영된 사례로 볼 수 있다.

이상의 결과는 대학생 학습자들이 생성형 AI의 작동 원리에 대해 상당한 수준의 오개념을 형성하고 있음을 보여준다. 특히 참여자들은 기존에 익숙한 디지털 기술 경험, 즉 검색엔진이나 데이터베이스 시스템에 대한 이해를 바

탕으로 생성형 AI를 해석하는 경향을 보였다. 이러한 경향은 인공지능을 ‘프로그래밍된 데이터 저장 및 검색 시스템’으로 이해하거나, 컴퓨터를 ‘전지전능한 데이터베이스’로 인식하는 기존 연구 결과³⁰와도 맥을 같이한다. 종합하면, 참여자들은 생성형 AI의 작동 원리에 대해 일정 수준의 설명을 시도하였으나, 그 이해는 개념적으로 충분히 정교화되지 못한 채 직관적이고 유추적인 정신 모형에 크게 의존하고 있었으며, 이 과정에서 다양한 오개념이 형성되고 있었다.

4.2. 생성형 AI에 대한 대학생 학습자의 개념적 이해와 활용 기대 간 관계

대학생 학습자의 생성형 AI에 대한 개념적 이해와 활용 기대 간의 관계를 확인하기 위해 참여자의 응답에서 나타난 정신 모형 유형과 향후 대학 교육 맥락에서의 활용 계획을 반복적으로 읽으며 교차 분석을 실시하였다. 이를 통해 두 변인 간의 조합 패턴을 도출하고, 그중 빈도가 높은 상위 패턴을 중심으로 분석하였다. 특히 모든 조합을 나열하기보다는, 참여자 수 기준으로 빈도가 높은 상위 조합 패턴과 기타 특징적인 패턴을 선별하여 제시하였으며, 그 결과는 <표 3>과 같다.

<표 3> 생성형 AI(ChatGPT)에 대한 정신 모형과 활용 기대 간 관계

구분	빈도	주요 답변 사례
작동-단순-정보탐색	11	필요한 정보 또는 사전 지식이 부족한 정보에 대해 조사를 해야 할 때 빠른 시간 내에 기초자료를 찾기 위해 사용할 것이다(S4) 보고서를 작성할 때 자료를 한눈에 보기 쉽게 정리가 필요할 때 사용할 것 같다(S12) 보고서 정리나 논문 자료 검색에 사용할 것 같다(S15)

작동-단순-글쓰기 보조 및 피드백	6	글을 쓸 때 구성을 잡기 어려울 때 도움을 받을 것 같다(S13) 내가 적은 글을 피드백 받고 싶을 때 사용할 것이다(S23)
작동-정교-학습 전반에 적극 활용	2	과제나 보고서를 시작할 때부터 진행하고 마무리하는 과정까지 전반적으로 참고용으로 활용할 것 같다. 자료를 찾거나 방향을 잡을 때뿐 아니라, 진행 중 막힐 때도 계속 도움을 받을 수 있을 것 같다(S2) 글쓰기나 창작뿐 아니라 아이디어 구상, 내용 전개, 수정 과정까지 활용할 것 같다. 아이디어가 떠오르지 않거나 글이 잘 이어지지 않을 때 지피티를 이용해 흐름을 자연스럽게 만들고 전체 과정을 보완할 수 있을 것 같다. (S31)
작동-단순-활용 계획 없음	1	SI도 사람과 대화하여 데이터를 습득하여 우리에게 답변하고 우리도 SI를 통해 필요한 정보를 습득 ... 활용 계획이 없다. 왜냐하면 점점 나의 능력이 퇴화될 것 같다. 그리고 SI에게 너무 의지하게 될 것 같다(S34)

분석 결과, ‘작동 - 단순 - 정보 탐색’이 11명, ‘작동 - 단순 - 글쓰기 보조 및 피드백’이 6명으로 가장 지배적인 패턴으로 나타났다. 참여자들은 생성형 AI를 주로 대학 학습 맥락에서 요구되는 과제 수행과 글쓰기 활동을 지원하는 도구로 활용하고자 하는 경향을 보였다. 특히 정보 탐색과 자료 정리를 중심으로 한 활용이 두드러지게 나타났는데, 이렇듯 활용 기대가 정보 탐색 중심으로 집중되어 있다는 점은 생성형 AI에 대한 참여자의 정신 모형이 검색엔진 모형이나 데이터베이스 모형에 기반하고 있기 때문으로 해석할 수 있다. 즉, 생성형 AI를 정보 탐색 및 정리 도구로 이해하는 인식이 그대로 활용 기대로 이어지고 있으며, 이는 참여자들이 ChatGPT를 새로운 생성 기술로서가 아니라 기존 정보 탐색 도구의 확장된 형태로 인식하고 활용하고 있을 것임을 시사한다.

한편 이러한 지배적인 패턴 외에도 몇 가지 특징적인 사례가 확인되었다. 첫째, 생성형 AI에 대해 비교적 정교한 이해를 보인 참여자(S2, S31)는 학습, 글쓰기, 아이디어 생성, 문제 해결 등 다양한 상황에서의 활용 가능성을 폭넓게 제시하였다. 이는 생성형 AI에 대한 이해 수준이 높을수록 활용 범위 역시 확장되는 경향이 있음을 시사한다. 선행 연구에서는 생성형 AI에 대한 정교한 설명이 실제 활용 수준과 유의미한 관련을 보이지 않는다고 보고된 바 있

30 Kreinsen, M. & Schulz, S., "Students' conceptions of artificial intelligence," *Proceedings of the 16th Workshop in Primary and Secondary Computing Education*, 2021, pp.1~2; Rücker, M. T. & Pinkwart, N., "Review and discussion of children's conceptions of computers," *Journal of Science Education and Technology* 25(2), 2016, pp.274~283.

으나(Singh et al., 2023), 본 연구에서는 정교한 이해를 보인 참여자들이 보다 다양한 학습 및 창작 활동에 활용하고자 하는 경향이 확인되었다. 둘째, 생성형 AI에 대한 활용 계획이 없는 사례도 확인되었다(S34). 해당 참여자는 생성형 AI가 이용자와의 상호작용을 통해 지속적으로 학습한다고 이해하는 오개념을 보였으며, 이러한 이해는 생성형 AI 기술에 대한 불확실한 인식을 형성하여 활용을 유보하거나 회피하는 태도로 이어진 것으로 볼 수 있다. 이는 생성형 AI에 대한 이해 방식이 활용 기대뿐 아니라 활용 회피와도 연결될 수 있음을 시사한다.

이러한 결과를 종합하면, 참여자들은 생성형 AI의 작동 원리에 대해 주로 검색엔진 모형이나 데이터베이스 모형과 같은 기존 기술 기반의 정신 모형을 통해 이해하고 있었으며, 이러한 이해는 활용 기대에서도 정보 탐색과 자료 정리 중심의 기능적 활용으로 이어지는 경향을 보였다. 일부 글쓰기 보조 및 피드백 활용이 나타나기는 했으나 그 범위는 제한적이었다. 이는 대학생 학습자가 생성형 AI의 생성 메커니즘과 한계를 충분히 이해하지 못한 상태에서, 제한된 활용 기대를 형성하고 있음을 보여준다.

5. 결론 및 시사점

본 연구는 대학생 학습자가 생성형 AI를 어떻게 개념적으로 이해하고 있는지, 그리고 이러한 이해가 대학 교육 맥락에서의 생성형 AI 활용 기대와 어떠한 관계를 보이는지를 탐색하고자 하였다. 이를 위해 생성형 AI의 대표적 사례인 ChatGPT를 중심으로 대학생 38명의 서면 응답을 수집하고, 생성형 AI의 작동 원리에 대한 설명과 향후 활용 계획을 정신 모형의 관점에서 분석하여, 대학생 학습자가 생성형 AI를 어떠한 방식으로 이해하고 있으며 그 이해가 실제 활용 기대와 어떻게 연결되는지를 확인하였다.

주요 연구 결과는 다음과 같다. 첫째, 대학생 학습자는

전반적으로 생성형 AI의 작동 원리에 대한 정교한 개념화에 이르지 못한 채, 단순하고 직관적인 정신 모형에 머물러 있는 경향을 보였다. 참여자 가운데 절반 이상은 생성형 AI의 작동 원리를 설명하려는 시도를 보였으나, 그 설명은 부분적이고 모호한 수준에 머물렀으며, 기술적 메커니즘을 정확하게 반영하지 못하였다. 특히 가장 높은 비중을 차지한 ‘작동-단순’ 유형은 생성형 AI의 작동 원리에 대한 설명을 포함하고 있음에도 불구하고, 오개념이나 기존 기술 경험에 기반한 유추적 해석을 포함하고 있었다. 이는 대학생 학습자가 생성형 AI를 사용해 본 경험이 있음에도 불구하고, 그 작동 원리와 특성에 대해서는 충분히 구조화된 이해를 형성하고 있지 못함을 보여준다.

둘째, 생성형 AI에 대한 대표적인 오개념 유형으로 검색엔진 모형과 데이터베이스 모형이 나타났다. 참여자들은 생성형 AI를 외부 정보를 탐색·수집·정리하는 검색엔진이나, 내부에 저장된 정보를 조회·조합하여 출력하는 데이터베이스와 유사한 시스템으로 이해하고 있었다. 이는 참여자들이 질문 입력과 답변 출력과 같은 외형적 상호작용에는 주목하였으나, 그 사이에서 이루어지는 확률적 토큰 예측과 텍스트 생성 과정과 같은 핵심 메커니즘에는 주목하지 못하였음을 의미한다. 즉, 대학생 학습자가 생성형 AI를 확률적 언어 생성 시스템으로 이해하기보다는 기존에 익숙한 검색엔진이나 데이터베이스 시스템의 확장된 형태로 해석하는 경향이 확인되었다.

셋째, 생성형 AI에 대한 정신 모형은 생성형 AI 활용 기대와도 밀접하게 연결되어 나타났다. 생성형 AI를 검색엔진이나 데이터베이스로 이해하는 경우, 활용 기대 역시 정보 탐색과 자료 정리 중심의 기능적 활용에 집중되는 경향을 보였다. 반면 생성형 AI의 작동 원리에 대해 비교적 정교한 이해를 보인 일부 참여자는 학습, 글쓰기, 아이디어 생성, 문제 해결 등 보다 폭넓은 활용 가능성을 제시하였다. 이는 생성형 AI에 대한 개념적 이해 수준이 활용 기대의 범위와 질과도 관련될 수 있음을 시사한다.

Anthropic에서 개발한 생성형 인공지능 Claude를 기반으로 대학생의 학업 관련 대화 약 100만 건을 분석한 연구에 따르면³¹, 대학생의 생성형 AI 활용은 즉각적인 문제 해결, 즉각적인 결과물 창작, 협력적인 문제 해결, 협력적인 결과물 창작의 네 가지 유형으로 구분되며, 이 중 약 47%가 단일 질문-결과 중심의 제한적인 상호작용에 머무르는 것으로 나타났다. 이는 본 연구에서 확인된 정보 탐색 중심의 활용 경향과 상응하는 결과이다. 나아가 해당 양상은 생성형 AI를 고도화된 방식으로 활용하거나 결과를 비판적으로 평가하는 단계에 이르지 못한 채, 검색엔진과 유사한 방식으로 활용하는 ‘일반 사용자’³²의 특성에도 부합한다. 이러한 결과는 대학생 학습자의 생성형 AI 활용이 여전히 제한된 수준에 머물러 있음을 보여주며, 체계적인 AI 리터러시 교육의 필요성을 뒷받침한다.

이러한 점에서 대학 교육에서의 AI 리터러시 교육은 다음과 같은 방향에서 이루어질 필요가 있다. 첫째, 생성형 AI를 검색엔진이나 데이터베이스와 구별되는 기술로 이해하도록 돕는 개념적 AI 리터러시 교육이 요구된다. 이는 생성형 AI의 확률적 생성 메커니즘, 한계, 오류 가능성 등에 대한 이해를 포함한다. 특히 학습자의 AI 이해가 비공식적 경험에 기반하여 형성될 경우 과학적으로 정확한 개념 형성이 어려울 수 있다는 점을 고려할 때,³³ 대학 교육 차원에서 체계적인 개념 형성 지원이 이루어질 필요가 있다. 둘째, 생성형 AI를 단순 질의응답 도구가 아니라 학습의 전 과정에서 비판적으로 협력할 수 있는 자원으로 활용하도록 돕는 실천적 AI 리터러시 교육이 요구된다. 특히 학습자가 개념 설명 요청, 과제 구조 설계, 글쓰기 수정, 문제 해결 등 다양한 학습 단계에서 생성형 AI를 성찰적으로

활용할 수 있도록 지원할 필요가 있다. 이를 통해 생성형 AI를 사고를 확장하고 조정하는 인지적 도구로 활용할 수 있도록 해야 한다.

한편 본 연구는 몇 가지 한계를 지닌다. 본 연구는 특정 대학의 소수 학습자를 대상으로 수행된 사례 연구라는 점에서 결과를 일반화하는 데 제한이 있으며, 학습자의 서술 응답을 중심으로 분석을 수행하여 실제 생성형 AI 활용 과정에서 나타나는 행동이나 상호작용을 직접적으로 관찰하지는 못하였다. 따라서 향후 연구에서는 다양한 학습자 집단을 대상으로 연구를 확장하고, 실제 활용 맥락에서 나타나는 상호작용 과정과 인식 변화를 함께 분석할 필요가 있다. 그럼에도 불구하고 본 연구는 대학생 학습자의 생성형 AI 이해가 기존 디지털 기술 경험에 기반하여 형성되고 있으며, 이러한 이해가 활용 기대의 범위와 질을 규정하고 있음을 밝힘으로써, 대학 교육에서의 AI 리터러시 교육 방향을 모색하는 데 기초 자료를 제공한다는 점에서 의의를 지닌다.

31 Anthropic, 「Anthropic Education Report: How University Students Use Claude」, Anthropic, 2025.

32 양소은·최석영·김병준, 『생성형 AI 활용 격차와 AI 리터러시』, 한국언론진흥재단, 2025.

33 Mertala, P., Fagerlund, J., & Calderon, O., "Finnish 5th and 6th grade students' pre-instructional conceptions of artificial intelligence," *Computers and Education: Artificial Intelligence* 3, 2022, 100095.

참고문헌

1. 단행본 및 논문

- 김근영·김승인, 「AI 영상 편집 도구 사용자 경험 분석: 지각된 유희성을 반영한 확장 TAM 모델을 중심으로」, 『산업진흥연구』 10(3), 산업진흥원, 2025.
- 김보현, 「대학생의 학술 발표 과제를 위한 생성형 AI 활용 양상과 인식 연구」, 『교양학 연구』 33, 다빈치미래교양연구소, 2025.
- 김성우, 『인공지능은 나의 읽기-쓰기를 어떻게 바꿀까』, 유유, 2024.
- 박종향·신나민, 「인공지능기술과 인공지능교사에 대한 인식 분석: 초·중·고등학생의 관점에서」, 『한국교육연구』 34(2), 한국교육연구학회, 2017.
- 변경기, 「글쓰기 맥락에서 대학생 필자의 생성형 AI 수용 인식 연구: 인공지능 기기 수용(AIDUA) 모형에 기반한 수용 구조 탐색과 타당도 검증」, 『청람어문교육』 106, 청람어문교육학회, 2025.
- 손석·등기령·왕옥뢰·Liu Ziyang, 「젊은 사용자들이 정신건강을 위해 AI 챗봇을 지속적으로 사용하는 방법 이해하기: TAM과 PSI 관점」, 『한국컴퓨터정보학회 논문지』 30(11), 한국컴퓨터정보학회, 2025.
- 송운희, 「ChatGPT의 학습 및 수업 활용에 대한 대학생의 인식 분석: K 신학대학교 사례를 중심으로」, 『기독교교육정보』, 한국기독교교육정보학회, 2024.
- 양소은·최석영·김병준(2025), 「생성형 AI 활용 격차와 AI 리터러시」, 한국언론진흥재단.
- 오의경, 「대학생의 AI 활용 특성과 인식 분석」, 『한국문헌정보학회지』 59(1), 한국문헌정보학회, 2025.
- 이한샘·유지원, 「대학생의 생성형 AI의 교육적 사용경험 및 인식 탐색: A대학 사례를 중심으로」, 『한국콘텐츠학회논문지』 24(1), 한국콘텐츠학회, 2024.
- 한국과학창의재단, 「미국 노동부(DOL) AI 리터러시 프레임워크 발표 및 전국 확산 추진」, 정책브리프, 2026.
- Alizadeh, F., Stevens, G., & Esau, M., "I don't know, is AI also used in airbags? An empirical study of folk concepts and people's expectations of current and future artificial intelligence," *I-com* 20(1), 2021.
- Anderson, S. S., "Places to stand: Multiple metaphors for framing ChatGPT's corpus," *Computers and Composition* 68, 2023.
- Andrews, R. W., Lilly, J. M., Srivastava, D., & Feigh, K. M., "The role of shared mental models in human-AI teams: A theoretical review," *Theoretical Issues in Ergonomics Science* 24(2), 2023.
- Anthropic, 「Anthropic Education Report: How University Students Use Claude」, Anthropic, 2025.
- Biber, C., Tuna, A. & Korkmaz, S., "The mistakes and the misconceptions of the eighth grade students on the subject of angles," *European Journal of Science and Mathematics Education* 1(2), 2013.
- Brown, T., Mann, B., & Ryder, N., "Language models are few-shot learners," *Advances in Neural Information Processing Systems* 33, 2020.
- Johnson-Laird, P. N., "Mental models in cognitive science," *Cognitive Science* 4(1), 1980.
- Johnson-Laird, P. N., *Mental models: Towards a cognitive science of language, inference, and consciousness*, Harvard University Press, 1983.
- Kreinsen, M. & Schulz, S., "Students' conceptions of artificial intelligence," *Proceedings of the 16th Workshop in Primary and Secondary Computing Education*, 2021.
- Long, D. & Magerko, B., "What is AI literacy? Competencies and design considerations," *Proceedings of the 2020 CHI Conference*, 2020.
- Marx, E., Leonhardt, T., & Bergner, N., "Secondary school students' mental models and attitudes regarding artificial intelligence: A scoping review," *Computers and Education: Artificial Intelligence* 5, 2023.
- Mertala, P., Fagerlund, J., & Calderon, O., "Finnish 5th and 6th grade students' pre-instructional conceptions of artificial intelligence," *Computers and Education: Artificial Intelligence* 3, 2022.
- Norman, D. A., "Some observations on mental models," *Mental Models*, 1983.
- _____, "Some observations on mental models," *Human-Computer Interaction: A Multidisciplinary Approach*, 1987.
- OECD, *Empowering Learners for the Age of AI: An AI Literacy Framework for Primary and Secondary Education*, OECD, 2025.
- Rouse, W. B. & Morris, N. M., "On looking into the black box: Prospects and limits in the search for mental models," *Psychological Bulletin* 100(3), 1986.
- Rücker, M. T. & Pinkwart, N., "Review and discussion of children's conceptions of computers," *Journal of Science Education and Technology* 25(2), 2016.
- Schank, R. C., "Where's the AI?," *AI Magazine* 12(4), 1991.
- Singh, N., Bernal, G., Savchenko, D., & Glassman, E. L., "Where to hide a stolen elephant: Leaps in creative writing with multimodal machine intelligence," *ACM Transactions on Computer-Human Interaction* 30(5), 2023.
- Su, J., Zhong, Y., & Ng, D. T. K., "A meta-review of literature on educational approaches for teaching AI at the K-12 levels," *Computers and Education: Artificial Intelligence* 3, 2022.
- U.S. Department of Labor, *The U.S. Department of Labor's Artificial Intelligence Literacy Framework*, 2026.

Abstract**University Students' Conceptualizations of Generative AI and Their Expectations for Its Use**

Implications for AI Literacy Education in Higher Education

Byun Kyeong-ga | Seowon University

This study investigates how university students conceptually understand generative AI and how such understanding shapes their expectations for its use. Written responses from 38 students on ChatGPT, its operating mechanisms, and future use were analyzed through the framework of mental models. The findings show that most participants relied on simple and intuitive rather than well-developed understandings of generative AI. Two misconceptions were especially prominent: the search engine model, which treats generative AI as a tool for retrieving and organizing external information, and the database model, which assumes that it recombines internally stored data. These mental models were closely related to students' expectations, with many viewing generative AI primarily as an instrument for information search and organization. By contrast, students with more refined understandings envisioned broader applications, including learning, writing, idea generation, and problem-solving. The study highlights the need for higher education to foster both conceptual and practical AI literacy.

Keywords Generative AI, AI literacy, mental models, higher education, media education, media perception
