

디자인 비전공 대학생 대상 생성형 AI 기반 지역문제해결디자인 교육의 AI 리터러시 향상 효과

Effects of a Generative AI-Based Community Problem-Solving Design
Education on AI Literacy Enhancement Among Non-Design Major
University Students

주 저 자 : 박현진 (Park, Hyeon Jin) 초당대학교 앵커(ANCHOR)사업단 교수
fredonpark@cdu.ac.kr

<https://doi.org/10.46248/kidrs.2026.3.434>

접수일 2026. 07. 20. / 심사완료일 2026. 09. 09. / 게재확정일 2026. 09. 11. / 게재일 2026. 09. 30.

Abstract

As generative AI becomes increasingly pervasive across society, AI literacy has emerged as a core competency in higher education; however, prior research grounded in cognitive load theory and scaffolding suggests that non-major students may struggle to learn and solve problems with AI in the absence of structured instruction. Building on these theories, this study designed a 15-week generative AI-based community problem-solving design education program and examined pretest-posttest changes in self-reported AI literacy across four dimensions (knowledge, use, critical thinking, and ethics) among 25 non-design major university students. Results showed significant, large increases in the knowledge, use, and critical thinking dimensions, whereas the ethics dimension showed no significant change. However, because this study employed a single-group pretest-posttest design without a control group, these increases cannot be attributed solely to the education program, and reliance on self-report measures is a further limitation. These findings suggest that practice-oriented AI education can raise self-assessed technical and practical competencies but requires explicit instructional design to cultivate ethical awareness.

Keyword

AI 리터러시(AI literacy), 생성형AI(Generative AI), 디자인 교육(Design Education), 인지부하 이론(Cognitive Load Theory), 스캐폴딩(Scaffolding)

요약

생성형 AI가 사회 전반으로 확산되면서 AI 리터러시는 고등교육의 핵심 역량으로 부상하였으나, 구조화된 교육 없이 AI를 다루는 비전공자는 학습과 문제 해결에 어려움을 겪을 수 있다는 인지부하 이론 및 스캐폴딩 관련 선행연구의 지적이 있다. 본 연구는 이 이론들에 근거하여 15주 생성형 AI 기반 지역문제해결디자인 교육 프로그램을 설계하고, 디자인 비전공 대학생 25명을 대상으로 지식·활용·비판·윤리의 네 영역으로 구성된 AI 리터러시 자기평가 점수의 사전·사후 변화를 분석하였다. 그 결과 지식·활용·비판 세 영역은 유의하고 효과크기가 큰 상승을 보였으나, 윤리 영역은 유의한 변화가 없었다. 다만 본 연구는 통제집단이 없는 단일집단 사전·사후 설계로 수행되어 이러한 상승을 교육 프로그램만의 효과로 단정하기는 어려우며, 자기보고식 측정이라는 한계도 지닌다. 이는 실습 중심의 AI 교육이 기술적·실천적 역량에 대한 자기평가를 높일 수 있으나, 윤리 의식 함양에는 별도의 명시적 교육 설계가 필요함을 시사한다.

목차

1. 서론

- 1-1. 연구 배경
- 1-2. 연구의 공백과 목적
- 1-3. 연구 문제(Research Questions) 및 기대효과

2. 이론적 배경

- 2-1. AI 리터러시와 생성형 AI 교육

- 2-2. 인지부하 이론과 교수학습 스캐폴딩
- 2-3. 제약 기반 프롬프팅과 디자인 시스템
- 2-4. 종합 소결

3. 연구 방법

- 3-1. 연구 설계 및 근거
- 3-2. 연구 대상 및 맥락
- 3-3 15주 교육과정
- 3-4 측정 도구

4. 연구 결과

4-1 예비 분석: 신뢰도 및 정규성 검토

4-2 AI 리터러시의 사전-사후 변화

4-3 하위 요인 간 변화 양상의 차이

5. 논의

5-1 세 역량의 향상과 선행연구와의 비교

5-2 향상폭과 향상률의 비대칭

5-3 윤리 영역의 무변화와 선행연구와의 불일치

5-4 종합

6. 결론

참고문헌

1. 서론

1-1. 연구의 배경

생성형 AI가 사회 전반과 전문 영역으로 빠르게 확산되면서, 대학생이 AI를 이해하고 비판적으로 평가하며 협업할 수 있는 능력, 즉 AI 리터러시가 고등교육의 핵심역량으로 자리 잡고있다.¹⁾ Long과 Magerko(2020)는 AI 리터러시를 단순한 도구가 아닌, AI와 효과적으로 상호작용하고 비판적으로 평가하는 일련의 역량으로 정의하였으며,²⁾ 후속 연구들은 이 개념이 지식·활용·평가·윤리와 같은 하위 요소로 구성됨을 정교화하였다³⁾.

이러한 관점은 대학생을 대상으로 한 여러 측정 도구 개발로 이어졌는데, 비전공자를 위한 척도⁴⁾나 인식·활용·평가·윤리의 4차원을 제안한 척도⁵⁾등이 그 예이다.

특히 자연어 프롬프트(prompt)만으로 작동 가능한 결과물을 생성하는 최신 생성형 AI 인터페이스의 등장은, 전통적으로 높은 기술적 진입 장벽이 존재하던 디자인 및 웹 프로토타이핑 영역에서도 디자인 비전공자가 직접 결과물을 구현할 수 있는 가능성을 열었다. 실제로 생성형 AI를 디자인·창작 교육에 도입한 최근 연구들은 AI 활용이 학습자의 창의적 인지와 정의적 태도에 긍정적 영향을 미칠 수 있음을 보고하고 있다.⁶⁾

1) Laupichler, M. C., Aster, A., Schirch, J., & Raupach, T., 'Artificial intelligence literacy in higher and adult education: A scoping literature review', Elsevier, Computers and Education: Artificial Intelligence, 제3권, 2022, p.100101; Long, D., & Magerko, B., 'What is AI literacy? Competencies and design considerations', ACM SIGCHI, Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems, 2020, pp.1-16; Ng, D. T. K., Leung, J. K. L., Chu, S. K. W., & Qiao, M. S., 'Conceptualizing AI literacy: An exploratory review', Elsevier, Computers and Education: Artificial Intelligence, 제2권, 2021, p.100041.

2) Long, D., & Magerko, B., op. cit.

3) Ng, D. T. K., Wu, W., Leung, J. K. L., Chiu, T. K. F., & Chu, S. K. W., 'Design and validation of the AI literacy questionnaire: The affective, behavioural, cognitive and ethical approach', British Educational Research Association, British Journal of Educational Technology, 제55권 제3호, 2024, pp.1082-1104.

4) Laupichler, M. C., Aster, A., Haverkamp, N., & Raupach, T., "'Scale for the assessment of non-experts' AI literacy"의 개발: 탐색적 요인분석', Elsevier, Computers in Human Behavior Reports, 제12권, 2023, p.100338.

5) Wang, B., Rau, P.-L. P., & Yuan, T., 'Measuring user competence in using artificial intelligence: Validity and reliability of artificial intelligence literacy scale', Taylor & Francis, Behaviour & Information Technology, 제42권 제9호, 2023, pp.1324-1337.

6) Hwang, Y., & Wu, Y., 'The influence of generative artificial intelligence on creative cognition of design students: A chain mediation model of self-efficacy and anxiety', Frontiers Media SA, Frontiers in Psychology, 제15권, 2024, p.1455015; Peng, C., Zhang, Z., Zhang, J., Liu, Y., Jia, R., & Xu, X., 'AI literacy and creative self-beliefs among Chinese college students: Examining the underlying mechanisms and boundary conditions', Frontiers Media SA, Frontiers in Psychology, 제17권, 2026,

이는 비전공자를 대상으로 한 AI 리터러시 교육의 잠재력을 시사한다. 그러나 이러한 기술적 가능성이 곧바로 비전공자의 성공적인 학습 경험과 역량 향상으로 이어지는 것은 아니다.

1-2. 연구의 공백과 목적

생성형 AI가 진입 장벽을 낮춘다는 기대와 달리, 비전공자가 실제로 효과적인 프롬프트를 설계하는 일은 결코 단순하지 않다. Zamfirescu-Pereira 외는 AI 비전문가들이 대형언어모델(LLM)을 활용해 원하는 결과를 얻으려 할 때 체계적이지 못하고 기회주의적인 방식으로 프롬프트를 시도하며 반복적으로 실패한다는 점을 실증적으로 보여주었다.⁷⁾ 즉 프롬프팅 환경이 곧바로 비전공자의 성공적 결과물로 이어지지는 않는다.

이러한 현상은 교육심리학의 인지 이론으로 설명될 수 있다. Sweller의 인지부하 이론(cognitive load theory)에 따르면, 인간의 작업기억은 용량이 제한되어 있어, 과제 자체의 복잡성에서 비롯되는 본질적 인지부하와 부적절한 학습 환경에서 비롯되는 외재적 인지부하가 과도하면 학습과 문제 해결이 저해된다.⁸⁾ 더 나아가 Kirschner, Sweller, Clark는 충분한 선행 지식이 없는 초심자에게 최소한의 안내만 제공하는 자유도 높은 교수 방식이 오히려 비효율적이며, 명시적 안내와 구조화된 지원이 학습에 더 효과적임을 광범위한 증거를 통해 논증하였다.⁹⁾ 디자인과 웹 구조에 대한 인지적 스키마(schema)가 형성되지 않은 비전공자에게 아무런 제약 없는 생성형 AI 환경을 제공하는 것은, 바로 ‘초심자에게 최소한의 안내만 주는’ 상황에 해당한다.

p.1719970.

7) Zamfirescu-Pereira, J. D., Wong, R. Y., Hartmann, B., & Yang, Q., 'Why Johnny can't prompt: How non-AI experts try (and fail) to design LLM prompts', ACM SIGCHI, Proceedings of the 2023 CHI Conference on Human Factors in Computing Systems, 2023, Article 437, pp.1-21.

8) Sweller, J., 'Cognitive load during problem solving: Effects on learning', Cognitive Science Society, Cognitive Science, 제12권 제2호, 1988, pp.257-285.

9) Kirschner, P. A., Sweller, J., & Clark, R. E., 'Why minimal guidance during instruction does not work: An analysis of the failure of constructivist, discovery, problem-based, experiential, and inquiry-based teaching', American Psychological Association Division 15, Educational Psychologist, 제41권 제2호, 2006, pp.75-86.

학습과학 연구는 이처럼 복잡한 과제를 다룰 때 학습자의 사고 과정을 구조화하는 스캐폴딩(scaffolding)이 필요함을 일관되게 강조해 왔으며¹⁰⁾, 그 효과는 대규모 메타분석으로도 뒷받침된다.¹¹⁾ 또한 컴퓨팅 사고 연구는 문제를 하위 요소로 분해하고 논리적으로 구조화하는 사고 도구가 복잡한 과제 수행의 토대가 됨을 보여주었다.¹²⁾ 한편 창의성 연구에서는 적절히 부과된 제약이 오히려 학습자의 인지적 방향을 줄이고 창의적 산출을 촉진할 수 있음이 보고되었다.¹³⁾

이상의 논의를 종합하면 다음과 같은 연구의 공백이 드러난다. 기존 연구는 AI 리터러시의 개념과 측정을 정교화하고, 생성형 AI가 교육에 갖는 긍정적 잠재력을 보여주었으며, 비전공자가 자유로운 프롬프팅에서 겪는 실패를 진단했지만, 인지부하 이론에 근거하여 비전공자의 인지부하를 체계적으로 통제함으로써 실제 문제를 해결하는 결과물 도출로 이끄는 구체적인 교수학습 설계를 제안하고, 그 효과를 다차원적 AI 리터러시 지표로 실증한 연구는 부재하다. 본 연구는 이 공백을 메우는 것을 목적으로 한다.

이를 위해 본 연구는 인지부하 이론(Sweller, 1988; Sweller et al., 2019)과 스캐폴딩 원리(Reiser, 2004)

10) Reiser, B. J., 'Scaffolding complex learning: The mechanisms of structuring and problematizing student work', International Society of the Learning Sciences, Journal of the Learning Sciences, 제13권 제3호, 2004, pp.273-304.

11) Belland, B. R., Walker, A. E., Kim, N. J., & Lefler, M., 'Synthesizing results from empirical research on computer-based scaffolding in STEM education: A meta-analysis', American Educational Research Association, Review of Educational Research, 제87권 제2호, 2017, pp.309-344; Kim, N. J., Belland, B. R., & Walker, A. E., 'Effectiveness of computer-based scaffolding in the context of problem-based learning for STEM education: Bayesian meta-analysis', Springer, Educational Psychology Review, 제30권 제2호, 2018, pp.397-429.

12) Grover, S., & Pea, R., 'Computational thinking in K-12: A review of the state of the field', American Educational Research Association, Educational Researcher, 제42권 제1호, 2013, pp.38-43; Wing, J. M., 'Computational thinking', Association for Computing Machinery(ACM), Communications of the ACM, 제49권 제3호, 2006, pp.33-35.

13) Tromp, C., & Baer, J., 'Creativity from constraints: Theory and applications to education', Elsevier, Thinking Skills and Creativity, 제46권, 2022, p.101184.

에 기반하여, 학습자가 곧바로 프롬프팅에 돌입하기 전에 문제를 하위 기능으로 분해·구조화하는 컴퓨팅 사고 기반 활동(Grover & Pea, 2013; Wing, 2006)과, 표준화된 디자인 시스템 규칙을 프롬프트의 제약 조건으로 부과하는 활동(Tromp & Baer, 2022)을 결합한 15주 교육 프로그램을 제안한다. 그리고 이 프로그램이 디자인·컴퓨터공학 비전공 대학생의 AI 리터러시—지식 및 원리 이해, 활용 및 실무 기능, 비판적 사고 및 탐지, 윤리 및 사회적 책임의 네 개 하위 영역—에 미치는 효과를 사전·사후 설계로 검증한다.

1-3 연구 문제(Research Questions) 및 기대효과

‘1.1’ 및 ‘1.2’에서 논의한 결과를 근거로 다음과 같은 연구 문제를 설정하였다.

RQ: 생성형 AI 기반 지역문제해결디자인 교육 프로그램에 참여한 후, 비전공자 대학생의 AI 리터러시(지식 및 원리 이해, 활용 및 실무 기능, 비판적 사고 및 탐지, 윤리 및 사회적 책임)는 통계적으로 유의하게 향상되는가?

본 연구의 기대효과는 다음과 같다. 첫째, 이론적 기여 측면이다. 본 연구는 인지부하 이론(Kirschner et al., 2006; Sweller, 1988)과 스캐폴딩 원리(Reiser, 2004)를 생성형 AI 기반 교육이라는 새로운 맥락으로 확장하여, 문제의 구조화와 디자인 시스템 제약이 비전공자의 인지부하를 통제하는 교수학습 기제를 제시한다. 특히 그 효과를 단일 지표가 아니라 지식·활용·비판·윤리의 다차원적 AI 리터러시 지표로 검증함으로써, 본 교육과정 내 리터러시의 어떤 하위 영역에 어떻게 작용하는지를 규명한다.

둘째, 실용적 기여 측면이다. 본 연구는 디자인·개발 지식이 없는 비전공자도 생성형 AI를 활용해 실제 문제를 해결하는 결과물을 안정적으로 도출할 수 있도록 돕는 구체적 교육 모델과 15주 교육과정을 제공함으로써, 고등교육에서 AI 리터러시 교과를 설계하려는 교수자에게 활용 가능한 사례를 제시한다.

2. 이론적 배경

2-1 AI 리터러시와 생성형 AI 교육

2-1-1 AI 리터러시의 개념화와 측정

AI 리터러시의 개념적 토대는 Long과 Magerko가

마련하였다. 이들은 AI 리터러시를 AI를 비판적으로 평가하고, AI와 소통·협업하며, AI를 도구로 활용하는 역량의 집합으로 정의하고, 이를 구성하는 핵심 역량과 설계 고려사항을 제시하였다.¹⁴⁾ Ng 외는 이 정의를 탐색적으로 검토하여 ‘알고 이해하기, 활용하기, 평가하기, 윤리’라는 다차원 구조로 구분하였다.¹⁵⁾ 이후 측정 도구 개발 연구들이 이 다차원성을 공유하면서도 차원의 수와 구성에서 서로 다른 선택을 했다는 것이다.

Wang 외는 디지털 리터러시 이론에 기반하여 인식·활용·평가·윤리의 4개 구성으로 12문항 척도(AILS)를 개발하였고,¹⁶⁾ Laupichler 외는 형식 교육을 받지 않은 비전공자를 겨냥하여 기술적 이해·비판적 평가·실천적 활용의 3요인 척도(SNAIL)를 탐색적 요인분석으로 도출하였다.¹⁷⁾ Ng 외는 한 걸음 더 나아가 정의적(affective)·행동적(behavioural)·인지적(cognitive)·윤리적(ethical) 차원을 포괄하는 설문(ABCE 접근)을 설계·검증함으로써, 기존 척도들이 상대적으로 소홀히 했던 정의적·행동적 차원을 명시적으로 포함시켰다.¹⁸⁾ 이처럼 척도들이 차원 구성에서 차이를 보인다는 사실은, AI 리터러시가 단일하게 합의된 구성개념이 아니라 측

14) Long, D., & Magerko, B., 'What is AI literacy? Competencies and design considerations', ACM SIGCHI, Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems, 2020, pp.1-16.

15) Ng, D. T. K., Leung, J. K. L., Chu, S. K. W., & Qiao, M. S., 'Conceptualizing AI literacy: An exploratory review', Elsevier, Computers and Education: Artificial Intelligence, 제2권, 2021, p.100041.

16) Wang, B., Rau, P.-L. P., & Yuan, T., 'Measuring user competence in using artificial intelligence: Validity and reliability of artificial intelligence literacy scale', Taylor & Francis, Behaviour & Information Technology, 제42권 제9호, 2023, pp.1324-1337.

17) Laupichler, M. C., Aster, A., Haverkamp, N., & Raupach, T., 'Development of the "Scale for the assessment of non-experts' AI literacy" - An exploratory factor analysis', Elsevier, Computers in Human Behavior Reports, 제12권, 2023, p.100338.

18) Ng, D. T. K., Wu, W., Leung, J. K. L., Chiu, T. K. F., & Chu, S. K. W., 'Design and validation of the AI literacy questionnaire: The affective, behavioural, cognitive and ethical approach', British Educational Research Association, British Journal of Educational Technology, 제55권 제3호, 2024, pp.1082-1104.

정 목적·대상·맥락에 따라 달리 조직화되는 개념임을 시사한다.

또한 고등·성인 교육 맥락의 문헌을 종합한 검토에서도 확인되듯, 이들 연구의 다수는 학습자의 현재 리터러시 '수준'을 측정하는 데 초점을 두며,¹⁹⁾ 그 역량이 특정 교수학습 개입을 통해 '어떻게 변화하는가'를 실증하는 개입 연구는 상대적으로 드물다. 특히 SNAIL은 AI 리터러시 교육의 효과성 평가를 위해 설계되었음에도, 이를 실제 15주 규모의 대학 정규 교과에 적용해 사전·사후 변화를 검증한 사례는 충분하지 않다.

2-1-2 디자인 교육에서의 생성형 AI

생성형 AI가 디자인·창작 교육에 도입되면서, 연구의 초점은 AI가 학습자의 창의적 인지와 정의적 태도에 미치는 영향으로 확장되었다. 그러나 이들 연구의 결론은 일률적이지 않으며, 활용 방식에 따라 긍정적 효과와 잠재적 위험이 엇갈린다. 긍정적 측면에서, Hwang과 Wu는 디자인 전공 학생을 대상으로 생성형 AI 활용이 자기효능감과 불안에 매개로 창의적 인지에 영향을 미치는 연쇄 매개 모형을 제시하였고,²⁰⁾ Peng 외는 AI 리터러시가 문제 해결 능력을 매개로 창의적 자기신념을 예측함을 보고하였다.²¹⁾

반면, 구조화되지 않은 생성형 AI 활용의 위험을 지적하는 연구도 있다. Vartiainen과 Tedre는 텍스트·이미지 생성 모델을 활용한 창작 교육에서, 학습자가 AI 결과물을 비판적으로 평가하지 못한 채 그대로 수용하는 경향과 결과의 예측 불가능성에서 비롯되는 긴장을 관찰하고, AI 기반 프로젝트를 신중하게 구조화해야 할 필요성을 강조하였다.²²⁾ 이러한 상반된 결과는, 생성

형 AI의 교육적 효과가 도구 자체보다 그것을 둘러싼 교수학습 설계에 의해 좌우됨을 보여준다. 특히 디자인 비전공자는 시각 언어와 결과물 구조에 대한 스키마가 부족하므로, 아무런 구조 없이 AI를 제공받을 경우 그 잠재적 위험에 더 크게 노출될 수 있다.

2.2 인지부하 이론과 교수학습 스케폴딩

2-2-1 인지부하의 구조

비전공자의 학습 과정을 설명하는 강력한 이론적 렌즈는 인지부하 이론이다. Sweller(1988)는 인간의 작업 기억이 용량과 지속 시간에서 제한된다는 인지 아키텍처에 근거하여, 학습이 이 한계 안에서 이루어진다고 이론화하였다.²³⁾ 인지부하는 세 유형으로 구분된다. 본질적 인지부하, 외재적 인지부하, 본유적 인지부하로 구분된다.²⁴⁾ 이론의 핵심 처방은 외재적 부하를 최소화하여 가용 인지 자원을 본유적 처리로 전환하는 것이다.

2-2-2 안내 논쟁과 스케폴딩

스케폴딩은 학생이 스스로 문제를 해결할 수 있도록 교수자나 앞선 또래가 단계별로 도움을 주는 학습 전략이다. 학습자의 수준에 맞춰 꼭 필요한 만큼만 지원을 제공하는 것이다.

이 관점에서 가장 첨예한 논쟁은 초심자에게 어느 정도의 안내를 제공할 것인가이다. Kirschner, Sweller, Clark는 구성주의·발견학습·문제기반학습 등 최소 안내 교수법이 선행 지식이 부족한 초심자에게 비효율적이며, 명시적·구조화된 안내가 더 효과적이라고 광범위한 증거로 논증하였다.²⁵⁾ 이 주장은 활발한 반론을 불러 일으켰으나, 후속 연구들은 학습자의 사전 지식 수준이 적정 안내 수준을 좌우한다는 절충적 결론으로 수렴해 왔다. 즉 초심자일수록 더 구조화된 지원이 필요하다는 것이다. 이러한 지원의 구체적 기제를 Reiser는 학습자가 복잡한 과제의 핵심 요소를 따라가도록 과제를 구조화하는 것과, 학습자가 중요한 내용에 주목하도록 과제를 문제화하는 것으로 정식화하였다.²⁶⁾

19) Laupichler, M. C., Aster, A., Schirch, J., & Raupach, T., 'Artificial intelligence literacy in higher and adult education: A scoping literature review', Elsevier, Computers and Education: Artificial Intelligence, 제3권, 2022, p.100101.

20) Hwang, Y., & Wu, Y., 'The influence of generative artificial intelligence on creative cognition of design students: A chain mediation model of self-efficacy and anxiety', Frontiers Media SA, Frontiers in Psychology, 제15권, 2024, p.1455015.

21) Peng, C., Zhang, Z., Zhang, J., Liu, Y., Jia, R., & Xu, X., 'AI literacy and creative self-beliefs among Chinese college students: Examining the underlying mechanisms and boundary conditions', Frontiers Media SA, Frontiers in Psychology, 제17권, 2026, p.1719970.

22) Vartiainen, H., & Tedre, M., 'Using artificial

intelligence in craft education: Crafting with text-to-image generative models', Taylor & Francis, Digital Creativity, 제34권 제1호, 2023, pp.1-21.

23) Sweller, J., op. cit.

24) Sweller, J., van Merriënboer, J. J. G., & Paas, F., 'Cognitive architecture and instructional design: 20 years later', Springer, Educational Psychology Review, 제31권 제2호, 2019, pp.261-292.

25) Kirschner, P. A., Sweller, J., & Clark, R. E., op. cit.

2-2-3 스캐폴딩의 실증적 근거

스캐폴딩의 효과는 대규모 메타분석으로 뒷받침된다. Belland 외는 STEM 교육에서 컴퓨터 기반 스캐폴딩을 다룬 다수의 실험 연구를 종합하여, 인지적 학습 성과에 대해 일관되게 정적인 효과가 나타나며 이 효과가 맥락·연령·평가 수준 전반에 견고함을 보고하였다.²⁷⁾ 후속 연구 역시 문제기반학습 맥락에서 컴퓨터 기반 스캐폴딩의 유의한 효과를 확인하였다. 종합하면, 인지부하 이론과 스캐폴딩 연구는 자유도 높은 복잡 과제일수록 통제된 교육이 학습 성과를 높인다는 일관된 결론을 제시한다.²⁸⁾

2-3 제약 기반 프롬프팅과 디자인 시스템

2-3-1 비전공자의 프롬프팅 어려움

생성형 AI를 활용한 문제 해결은 본질적으로 자유도가 높은 복잡 과제다. Zamfirescu-Pereira 외는 AI 비전문가들이 LLM과 상호작용할 때 체계적 전략 없이 기회주의적으로 프롬프트를 시도하며, 자신의 의도를 결과물로 전환하는 데 반복적으로 실패함을 실증적으로 보여주었다.²⁹⁾ 이는 프롬프팅이 단순한 자연어 입력이 아니라 구조화된 역량을 요구하는 활동이며, 2.2에서 논의한 ‘구조화된 안내의 부재’ 문제가 생성형 AI 맥락에서 그대로 재현됨을 보여준다.

2-3-2 패턴과 제약을 통한 프롬프트 구조화

프롬프트 엔지니어링 연구는 이 문제에 대응하여 프롬프트를 재사용 가능한 형태로 구조화하려 시도해 왔다. White 외는 소프트웨어 공학의 디자인 패턴 개념을 차용하여, 출력 형식 지정·맥락 제어·역할 부여 등 반복되는 문제에 대한 해결책을 ‘프롬프트 패턴(prompt patterns)’으로 정리하였다.³⁰⁾ 이는 프롬프트

가 일종의 프로그래밍이며 명시적 규칙과 제약을 통해 결과물을 통제할 수 있음을 보여준다.

2-3-3 창의성의 자원으로서의 제약

과제의 제약을 부과하는 것이 결과물의 품질뿐 아니라 창의성 자체를 촉진할 수 있다는 점은 창의성 연구에서 흔히 지지되는 사실이다. 적절히 부과된 제약은 학습자의 인지적 방향을 줄이고 해결 공간을 구조화함으로써 창의성을 촉진한다.³¹⁾ 디자인 시스템은 바로 이러한 제약 장치로 기능할 수 있다. 컴포넌트·색상·타이포그래피·레이아웃에 관한 표준 규칙을 프롬프트 제약으로 부과하면 생성형 AI 결과물의 시각적 무작위성을 통제할 수 있으며, 이는 인지부하 이론의 관점에서 외재적 부하를 경감하는 동시에 창의성 연구의 관점에서 생산적 제약으로 작용하는 이중 효과를 가진다.

2-4 종합 소결

세 갈래의 문헌은 하나의 지점으로 수렴한다. AI 리터러시 연구는 개념과 측정을 정교화했으나 비전공자의 역량 획득 과정에 대한 처방적 설계가 부족하고, 생성형 AI 교육 연구는 그 효과가 교수학습 구조에 좌우됨을 보여주었다(Hwang & Wu, 2024; Vartiainen & Tedre, 2023). 인지부하·스캐폴딩 연구는 자유도 높은 과제에서 구조화된 지원이 학습을 최적화함을 메타분석 수준에서 입증했으며(Belland et al., 2017; Kirschner et al., 2006; Reiser, 2004), 프롬프트·제약 연구는 명시적 제약이 결과물 품질과 창의성을 동시에 높일 수 있음을 시사한다(Tromp & Baer, 2022; White et al., 2023). 본 연구는 이 세 흐름을 통합하여, 비전공자가 생성형 AI로 실제 문제를 해결하는 과정에서 문제의 구조화(가능 분해)와 디자인 시스템 제약이라는 두 층위의 지원을 제공하고, 이 프로그램 참여 전후 지식·활용·비판·윤리의 네 개 AI 리터러시 영역에서 자기 보고식 점수의 변화를 검증한다.

3. 연구 방법

3-1 연구 설계 및 근거

본 연구는 단일집단 사전·사후 설계를 채택하였다. 이는 동일 학습자 집단을 대상으로 교육 전후의 변화를 비교하는 설계로, 통제집단을 두지 않는 전실험설계

26) Reiser, B. J., op. cit.

27) Belland, B. R., Walker, A. E., Kim, N. J., & Lefler, M., op. cit.

28) Kim, N. J., Belland, B. R., & Walker, A. E., op. cit.

29) Zamfirescu-Pereira, J. D., Wong, R. Y., Hartmann, B., & Yang, Q., op. cit.

30) White, J., Fu, Q., Hays, S., Sandborn, M., Olea, C., Gilbert, H., Elnashar, A., Spencer-Smith, J., & Schmidt, D. C., 'A prompt pattern catalog to enhance prompt engineering with ChatGPT', arXiv(Cornell University), arXiv preprint arXiv:2302.11382, 2023.

31) Tromp, C., & Baer, J., op. cit.

에 해당한다. 이 설계는 성숙효과나 역사효과 등 내적 타당도 위협을 통제하지 못한다는 근본적 한계를 지니지만, 본 연구가 통제집단을 설정하지 않은 것은 교육 현장의 윤리적 제약에 따른 의도적 선택이다. 본 연구는 정규 학점 강좌의 일부로 운영되었으며, 동일 강좌에서 일부 학습자에게만 제공하고 나머지에 이를 보류하는 것은 교육적 형평성에 위배된다. 따라서 본 연구는 이러한 설계상의 한계를 인지한 상태에서, 수강생 전원에게 동일한 교육을 제공하고, 각 학습자가 자신의 사전 상태를 비교 기준으로 삼는 사전-사후 비교를 수행하였다.

3-2 연구 대상 및 맥락

본 연구의 참여자는 C대학교에 재학 중인 디자인 비전공 대학생 25명이다. C대학교의 학과 편제 기준으로 인문/사회/이학/공학으로 구분하였다. 이들은 단일 분반으로 운영된 정규 학점 강좌의 수강생이며, 데이터 수집은 2026학년도 1학기에 이루어졌다. 참여자의 일반적 특성은 표 1과 같다. 연구 참여에 동의하지 않는 5명은 분석에서 제외하였다.

[표 1] 연구 참여자의 일반적 특성 (N = 25)

구분	범주	빈도 (n)	비율 (%)
성별	남성	13	52.0
	여성	12	48.0
전공 계열	이학계열	10	40.0
	인문계열	5	20.0
	사회계열	5	20.0
	공학계열	5	20.0
학년	1학년	17	68.0
	2학년	4	16.0
	3학년	2	8.0
	4학년 이상	2	8.0
수강 전 생성형 AI 사용 빈도	주 1~2회 이하	7	28.0
	주 3~4회 이하	14	56.0
	주 5~6회 이하	4	16.0
	전혀 없음 / 주 7일 이상	0	0.0
지역사회 문제 사전 관심도	평균(SD), 7점 척도	3.84 (1.43)	—

모든 참여자에게는 데이터 수집에 앞서 연구의 목적과 절차, 데이터의 활용 범위, 참여의 자율성, 익명성 보장에 관한 사항을 안내하고 서면 사전동의를 확보하였다. 개인 식별 정보(성명·학번 등)는 수집하지 않았으며, 사전-사후 응답의 연결을 위해 참여자 본인만 알 수 있는 익명 매칭 코드(부모 성씨 1자 + 본인 생일 4자리)를 사용하였다. 수집된 자료는 암호화하여 보관하였고, 법정 보관 기간(연구 종료 후 3년) 경과 후 폐기한다. 연구

참여 여부는 성적 평가나 학업 수행에 어떠한 영향도 미치지 않으며, 참여자는 언제든지 중단할 수 있고 이 경우 해당 데이터는 분석에서 제외됨을 명시하였다.

3-3 15주 교육과정

본 교육과정은 생성형 AI(Gemini)를 활용한 ‘지역문 제해결디자인’ 교육으로, 15주에 걸쳐 이론 단계(1-7주차)와 실습 단계(9-15주차)로 운영되었다. 이론 단계에서는 AI 리터러시의 기초 개념, 프롬프트 엔지니어링 원리, 컴퓨팅 사고, 디자인 시스템 규칙을 다루었으며, 실습 단계에서는 교수자의 1:1 피드백을 중심으로 학습자가 지역사회 문제를 해결하는 프로토타입을 구현하였다(표 2참조).

[표 2] 15주 교육과정 구성

주차	단계	핵심 주제 및 활동
1	이론	오리엔테이션, AI 리터러시 개관
2	이론	생성형 AI와 프롬프팅의 이해
3	이론	컴퓨팅 사고와 문제 분해
4-5	이론	기능 구조화 및 명세 작성 실습
6-7	이론	디자인 시스템 규칙과 프롬프트 제약 변환
8	중간고사	생성형 AI 개념 확인을 위한 지필고사 실시
9-10	실습	프로토타입 구현 착수, 1:1 피드백
11-12	실습	디자인 일관성 적용 인터랙션 구현, 1:1 피드백
13-14	실습	프로토타입 정교화·오류 수정, 1:1 피드백
15	기말고사	최종 발표 및 마무리

3-4 측정 도구

3-4-1 AI 리터러시 척도

학습자의 AI 리터러시는 선행연구에서 개발·타당화된 척도를 본 교육 맥락에 맞게 재구성한 자기보고식(self-report) 설문으로 측정하였다.

본 척도는 비전공자를 위한 AI 리터러시 척도(SNALI; Laupichler et al., 2023)의 기술적 이해·비판적 평가·실천적 활용 차원과, 인식·활용·평가·윤리의 4 차원을 제안한 AI 리터러시 척도(Wang et al., 2023), 그리고 정의적·행동적·인지적·윤리적 접근의 AI 리터러시 문항(Ng et al., 2024)을 이론적 근거로 삼아 네 개 하위 요인으로 구성하였다.

각 문항은 7점 리커트 척도(1 = 매우 그렇지 않다 ~ 7 = 매우 그렇다)로 응답하며, 총 35개 문항으로 구성되었다. 하위 요인별 구성은 표 3과 같다.

본 연구에서의 척도 신뢰도는 각 하위 요인별 Cronbach's α 로 검토하였으며, 그 값은 결과(4.1절)에 보고하였다.

[표 3] AI 리터러시 척도의 하위 요인 구성

순번	척도명	문항	근거
1	지식·원리 이해	AI와 기존 소프트웨어의 차이를 설명할 수 있는가?	SNAIL(기술적 이해); Ng et al.(인지적)
2		머신러닝이 데이터를 통해 학습하는 원리를 이해하는가?	
3		거대언어모델(LLM)이 무엇인지 설명할 수 있는가?	
4		AI 성능이 학습 데이터의 양과 질에 의존함을 인지하는가?	
5		AI 시스템의 주요 구성 요소(입력·처리·출력)를 식별할 수 있는가?	
6		인공지능 기술의 역사와 발전 과정을 이해하는가?	
7		'딥러닝' 개념을 쉬운 용어로 설명할 수 있는가?	
8		AI 의사결정에서 알고리즘의 역할을 이해하는가?	
9		'범용 인공지능(AGI)'과 '특화 인공지능'을 구분할 수 있는가?	
10		과업에 가장 적합한 AI 도구를 선택할 수 있는가?	
11	활용·실무 가능	원하는 결과를 위해 효과적인 프롬프트를 설계할 수 있는가?	SNAIL(실천적 활용); Wang et al.(활용)
12		지역 문제 해결 아이디어를 얻는 데 AI를 활용할 수 있는가?	
13		'바이브코딩'으로 간단한 프로그램을 구현할 수 있는가?	
14		초기 결과가 미흡할 때 프롬프트를 수정·반복 설계할 수 있는가?	
15		지역 이슈 데이터를 분석·요약하는 데 AI를 활용할 수 있는가?	
16		AI 생성 결과물을 보고서·프로토타입에 통합할 수 있는가?	
17		다양한 생성형 AI 인터페이스 사용에 자신감이 있는가?	
18		자신의 요구사항을 AI에 효과적으로 전달할 수 있는가?	
19		지역 문제 해결책을 프로토타입으로 제작하는 데 AI를 활용할 수 있는가?	
20		AI 정보의 오류·환각 현상을 식별할 수 있는가?	SNAIL(비판적 평가); Wang et al.(평가)
21	비판적 사고·탐지	AI 정보의 신뢰성 확인을 위해 다른 출처와 대조하는가?	
22		AI 결과물의 숨겨진 논리와 근거를 평가할 수 있는가?	
23		AI 모델의 잠재적 편향성(성별·인종 등)을 인지할 수 있는가?	
24		AI 생성 코드에 버그·보안 취약점이 있을 수 있음을 인지하는가?	
25		AI 해결책이 실제 지역 문제에 적용 가능한지 평가할 수 있는가?	
26		인간 작성 콘텐츠와 AI 생성 콘텐츠를 구분하는 안목이 있는가?	
27		인공지능 기술의 기술적 한계를 비판적으로 사고할 수 있는가?	
28	윤리·사회적 책임	AI 활용 시 개인정보 보호·데이터 보안 원칙을 준수하는가?	Wang et al.(윤리); Ng et al.(윤리적)
29		AI 생성 콘텐츠 사용 시 저작권 침해 가능성에 주의하는가?	
30		AI 도움을 받았을 때 이를 출처로 명확히 밝히는가?	
31		AI가 지역 공동체의 권익·존엄성에 미칠 영향을 고려하는가?	
32		거대 AI 모델이 환경에 미치는 영향을 인지하는가?	
33		AI가 공동 이익·사회 정의 방향으로 쓰여야 한다고 믿는가?	
34		사람에 관한 결정에 AI를 쓰는 것의 윤리적 함의를 이해하는가?	
35		인공지능을 책임감 있고 투명하게 활용하려는 태도를 갖추는가?	

3-4-2 인구통계 및 배경 변수

인구통계학적 특성으로 성별, 전공 계열(인문·사회·이학·공학), 학년을 조사하였다. 또한 교육 이전의 개인차를 통제하기 위한 배경 변수로 수강 전 생성형 AI 사용 빈도(5점 범위)와 지역사회 문제에 대한 사전 관심도(7점 척도)를 함께 측정하였다.

3-5 자료 분석 절차

수집된 사전·사후 설문 데이터는 Python(pandas, NumPy, SciPy라이브러리)을 사용하여 분석하였다. 먼저 측정 도구의 내적 일관성 신뢰도를 확인하기 위해 네 개 하위 요인별로 Cronbach's α 를 산출하였다. 다음으로 각 하위 요인의 사전·사후 점수 차이값에 대해 Shapiro-Wilk 검정을 실시하여 정규성 가정의 충족 여부를 확인하였다. 정규성 가정을 충족한 요인에는 대응

표본 t-검정을, 충족하지 못한 요인에는 비모수 검정인 윌콕슨 부호순위검정을 적용하였다. 효과크기는 대응표본 t-검정을 적용한 요인에 대해 Cohen's d 를, 윌콕슨 부호순위검정을 적용한 요인에 대해 $r(= Z/\sqrt{N})$ 을 산출하였다. 네 개 하위 요인에 대해 반복적으로 검정을 수행함에 따라 발생할 수 있는 제1종 오류의 증가를 통제하기 위해 Holm-Bonferroni 방법으로 유의확률을 보정하였다.

4. 연구 결과

4-1 예비 분석: 신뢰도 및 정규성 검토

본 분석에 앞서 측정 도구의 내적 일관성 신뢰도를 검토하기 위해 Cronbach's α 값을 확인하였다. 그 결과 네 개 하위 요인의 신뢰도는 사전 .854~.931, 사후

.815~.910으로 모두 양호한 수준(.80 이상)을 보였으며, 전체 35개 문항의 신뢰도 또한 사전 .964, 사후 .946으로 매우 높게 나타나 척도가 신뢰할 만한 것으로 확인되었다(표 4).

[표 4] AI 리터러시 척도의 신뢰도

하위 요인	문항 수	Cronbach's α (사전)	Cronbach's α (사후)
지식·원리 이해	10	.913	.885
활용·실무 기능	9	.918	.910
비판적 사고·탐지	8	.931	.897
윤리·사회적 책임	8	.854	.815
전체	35	.964	.946

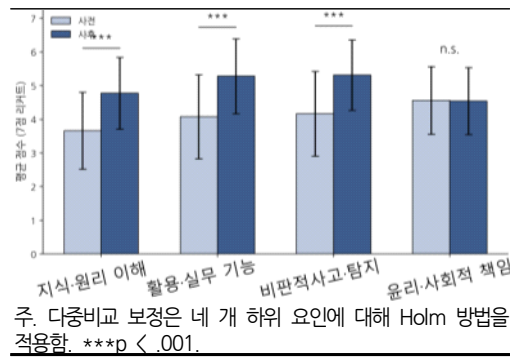
이어 사전·사후 차이값의 정규성을 Shapiro-Wilk 검정으로 확인하였다. 그 결과 지식·원리 이해($p = .009$), 활용·실무 기능($p = .014$), 비판적 사고·탐지($p = .004$) 세 요인은 정규성 가정을 충족하지 못하여 비모수 검정인 윌콕슨 부호순위검정(Wilcoxon signed-rank test)을 적용하였으며, 윤리·사회적 책임 요인($p = .165$)은 정규성을 충족하여 대응표본 t-검정(paired t-test)을 적용하였다.

4-2 AI 리터러시의 사전·사후 변화

교육 프로그램 참여 전후의 AI 리터러시 변화를 분석한 결과는 표 5와 같다. 네 개 하위 요인 중 세 개 요인에서 통계적으로 유의한 향상이 나타났다.

[표 5] AI 리터러시의 사전·사후 비교 (N = 25)

하위 요인	사전 M (SD)	사후 M (SD)	향상	통계량	p	효과 크기	p (Holm 보정)
지식·원리 이해	3.66 (1.14)	4.78 (1.06)	+1.11	0.00	<.001 ₁	$r = .875$	<.001**
활용·실무 기능	4.08 (1.25)	5.28 (1.11)	+1.20	0.00	<.001 ₁	$r = .858$	<.001**
비판적 사고·탐지	4.16 (1.26)	5.32 (1.05)	+1.15	0.00	<.001 ₁	$r = .860$	<.001**
윤리·사회적 책임	4.57 (1.00)	4.54 (0.99)	-0.02	-0.69	.495	$d_z = .139$	.495 (n.s.)
전체 AI 리터러시	4.09 (1.03)	4.97 (0.87)	+0.88	0.00	<.001 ₁	$r = .875$	—



구체적으로, 지식·원리 이해는 사전 3.66에서 사후 4.78로 1.11점 향상되었으며 이는 통계적으로 유의하였다($p < .001$, $r = .875$). 활용·실무 기능은 4.08에서 5.28로 1.20점 향상되어 네 요인 중 가장 큰 상승폭을 보였고($p < .001$, $r = .858$), 비판적 사고·탐지 역시 4.16에서 5.32로 1.15점 유의하게 향상되었다($p < .001$, $r = .860$). 이 세 요인의 효과크기는 모두 $r \approx .86$ 이상으로 매우 큰 효과에 해당하며, Holm 방법으로 다중비교를 보정한 후에도 유의성이 유지되었다.

반면 윤리·사회적 책임요인은 사전 4.57에서 사후 4.54로 실질적 변화가 없었으며 통계적으로 유의하지 않았다($p = .495$). 전체 AI 리터러시 평균은 4.09에서 4.97로 0.88점 유의하게 향상되었다($p < .001$, $r = .875$).

4-3 하위 요인 간 변화 양상의 차이

네 하위 요인의 변화 양상은 뚜렷하게 구분되었다. 지식·활용·비판의 세 인지적·기능적 역량은 모두 1점 이상 큰 폭으로 향상된 반면, 윤리·사회적 책임은 변화가 거의 없었다. 이러한 차이는 두 가지 측면에서 해석의 단서를 제공한다.

첫째, 사전 점수의 분포를 보면 윤리·사회적 책임 요인은 사전 평균이 4.57로 네 요인 중 가장 높았던 반면, 지식·원리 이해는 3.66으로 가장 낮았다. 즉 교육 이전에 학습자들은 AI 윤리에 대해서는 이미 상대적으로 높은 인식을 지니고 있었던 반면, AI의 기술적 지식과 원리에 대해서는 이해가 가장 부족한 상태였다. 이는 향상 폭이 가장 컸던 영역이 사전 수준이 가장 낮았던 영역과 대응함을 시사한다.

둘째, 본 교육 프로그램이 생성형 AI를 활용한 문제 해결의 실천적 수행에 중점을 두었다는 점을 고려하면,

프롬프트 설계 결과물 구현과 직접 관련된 활용·실무 기능이 가장 크게 향상된 것은 교육과정의 성격과 일관된 결과로 볼 수 있다.

5. 논의

5-1 세 역량의 향상과 선행연구와의 비교

분석 결과 AI 리터러시의 인지적·실천적·비판적 세 영역이 모두 유의하게 향상되었다. 향상폭은 활용·실무 기능이 +1.20으로 가장 컸고, 비판적 사고·탐지(+1.15), 지식·원리 이해(+1.11) 순이었으며, 세 영역 모두 효과크기 $r \approx .86$ 이상의 큰 효과를 보였다($p < .001$). 이 향상은 특정 학습자에 국한되지 않고 집단 전반에서 관찰되었는데, 지식 영역은 참여자 25명 전원이, 활용과 비판 영역은 각각 24명이 향상되었으며 전체 향상의 개인차는 작았다($SD = 0.31$).

이 결과는 문제해결 기반 AI 교육의 효과를 보고한 선행연구와 일치한다. 특히 문제해결 중심의 AI 리터러시 수업에서 네 개 하위 차원 모두 큰 효과크기($r = .76 \sim .88$)로 향상되었다는 보고와 본 연구의 효과크기 수준($r \approx .86$)은 유사한 범위에 있다. 또한 본 결과는 비전공자가 구조적 지원 없이 자유롭게 프롬프팅할 경우 반복적으로 실패한다는 Zamfirescu-Pereira 외(2023)의 관찰과 대비를 이룬다. 본 연구에서 초심자에게 최소한의 안내보다 구조화된 지원이 효과적이라는 Kirschner, Sweller, Clark(2006)의 논지 및 컴퓨터 기반 스캐폴딩의 학습 효과를 보고한 메타분석(Belland et al., 2017; Kim et al., 2018)과 방향을 같이한다.

세 영역이 서로 다른 성격의 역량임에도 함께 향상된 점은, 본 교육이 특정 역량에 국한되지 않고 문제해결 과정 전반에 작용하였을 가능성을 시사한다. 학습자가 문제를 하위 기능으로 분해하고(지식), 이를 프롬프트로 구현하며(활용), 결과물을 점검하는(비판) 과정이 하나의 수행 절차 안에서 반복되었기 때문으로 해석할 수 있다. 다만 본 연구의 설계로는 세 역량 간 향상의 인과적 선후관계를 규명할 수 없으므로, 이 해석은 잠정적이다.

5-2 향상폭과 향상률의 비대칭

향상의 양상을 세분하면 절대적 향상폭과 상대적 향상률이 서로 다른 순위를 보인다. 절대 향상폭은 활용·실무 기능(+1.20)이 가장 컸으나, 사전 대비 향상률은

지식·원리 이해가 30.3%로 가장 높았다. 이는 지식 영역의 사전 점수($M = 3.66$)가 네 영역 중 가장 낮았던 반면 활용 영역의 사전 점수($M = 4.08$)는 상대적으로 높았던 데 기인한다. 즉 학습자들은 교육 이전에 AI를 활용하는 능력에 비해 그 작동 원리에 대한 이해가 더 취약한 상태에서 출발하였고, 가장 취약했던 영역에서 가장 큰 비율의 향상이 나타났다.

이 결과는 초보 학습자에게 추상적 AI 개념을 직접 전달하기보다 실제 문제해결을 통한 접근이 효과적이라는 선행연구의 지적과 부합한다. 본 연구에서 지식 영역이 별도의 개념 강의를 아니라 문제의 기능적 분해라는 수행 활동을 통해 가장 큰 비율로 향상되었다는 점은, 실천적 수행이 개념적 이해를 견인할 수 있음을 보여준다. 또한 비판적 사고·탐지 영역이 활용과 거의 동등하게 향상된 점(+1.15)은 주목할 만하다. 비판적 역량은 일반적으로 별도의 명시적 훈련을 요한다고 논의되어 왔으나, 본 연구에서는 AI 결과물을 반복 검토·수정하는 실천 과정 자체가 이 역량의 향상과 연관된 것으로 보인다.

5-3 윤리 영역의 무변화와 선행연구와의 불일치

본 연구에서 가장 두드러진 결과는 윤리·사회적 책임 영역의 정체이다. 이 영역은 사전 4.57에서 사후 4.54로 통계적으로 유의한 변화가 없었으며($p = .495$), 참여자 25명 중 6명만이 향상을 보여 다른 세 영역(24~25명 향상)과 뚜렷이 구분되었다.

이 결과는 대학생 대상 AI 리터러시 프로그램에서 윤리 인식이 유의하게 향상되었다고 보고한 Kong 외의 결과와 불일치한다.³²⁾ 이러한 불일치는 교육과정 설계의 차이로 설명될 수 있다. Kong 외의 프로그램은 프로젝트 기반 학습 과정에 AI 윤리 쟁점에 대한 논의를 명시적으로 포함하였던 반면,³³⁾ 본 연구는 문제해결을 위한 실천적 수행에 집중하였고 윤리적 쟁점을 성찰하도록 하는 구성 요소를 포함하지 않았다. 즉 두 연구의 상반된 결과는, AI 윤리 인식의 변화가 교육 기

32) Kong, S.-C., Cheung, W. M.-Y., & Zhang, G., 'Evaluating an artificial intelligence literacy programme for developing university students' conceptual understanding, literacy, empowerment and ethical awareness', International Forum of Educational Technology & Society, Educational Technology & Society, 제26권 제1호, 2023, pp.16-30.

33) Kong, S.-C., Cheung, W. M.-Y., & Zhang, G., op. cit.

간이나 대상의 문제라기보다 윤리를 명시적으로 다루는 교육 설계의 유무에 달려 있음을 시사한다. 이는 AI 윤리 교육에서 명시적 커리큘럼이 있을 때 윤리 인식이 변화하며 실습의 부산물로는 향상되기 어렵다는 선행 논의와도 일관된다.

윤리 영역은 사전 점수가 네 영역 중 가장 높아, 교육 이전에 학습자들이 이미 상당한 수준의 윤리 인식을 지니고 있었다. 사전 수준이 높은 척도에서 추가적 향상을 관찰하기 어렵다는 점은 AI 리터러시 측정에서 반복적으로 지적되어 왔다. 다만 본 연구에서 윤리 영역의 사전 점수(4.57)가 7점 척도의 상한에 근접했다고 보기는 어렵다. 따라서 윤리 영역의 정체는 높은 사전 수준과 윤리를 자극하지 않은 교육 설계가 함께 작용한 결과로 해석하는 것이 타당하다.

5-4 종합

이상의 결과를 종합하면, 본 교육은 디자인 비전공자의 기술적·실천적·비판적 AI 리터러시 자기평가 점수를 큰 폭으로 향상시켰으나 윤리적 역량에는 유의한 변화가 나타나지 않았다. 이는 실습 중심의 AI 활용 교육이 도구적·비판적 역량에 대한 자기평가를 높이는 데는 유리할 수 있으나, 윤리 의식의 변화는 별도의 명시적 설계를 요구할 수 있음을 시사한다. 기술 역량과 윤리 의식이 동일한 교육에 대해 상이하게 반응한다는 본 연구의 결과는, AI 리터러시를 단일 총점이 아니라 하위 영역별로 구분하여 측정하고 접근해야 한다는 다차원적 관점(Ng et al., 2021, 2024)을 뒷받침한다.

6. 결론

본 연구는 생성형 AI 기반 지역문제해결디자인 교육 프로그램 참여 전후 디자인 비전공 대학생의 AI 리터러시 자기평가 점수 변화를 분석하였다.

문제의 기능적 분해와 디자인 시스템 제약이라는 두 층위의 스캐폴딩을 핵심 원리로 하는 15주 교육을 비전공자 25명에게 적용하고, 지식·원리 이해, 활용·실무 기능, 비판적 사고·탐지, 윤리·사회적 책임의 네 영역으로 구성된 AI 리터러시의 변화를 분석하였다.

분석 결과, AI 리터러시의 네 영역 중 세 영역에서 통계적으로 유의하고 효과크기가 큰 향상이 나타났다. 활용·실무 기능(+1.20), 비판적 사고·탐지(+1.15), 지식·원리 이해(+1.11)가 모두 큰 효과크기($r \approx .86$)로

향상되었으며, 이 향상은 소수의 학습자가 아니라 집단 전반에 걸쳐 고르게 관찰되었다.

반면 윤리·사회적 책임 영역은 유의한 변화를 보이지 않았다($p = .495$). 특히 지식 영역은 사전 점수가 가장 낮았으나 향상률(30.3%)이 가장 높아, 실천적 수행을 통한 접근이 개념적 이해를 견인하였음을 보여주었다.

본 연구의 시사점은 다음과 같다.

이론적으로, 본 연구는 구조화된 스캐폴딩이 자유도는 높은 생성형 AI 과제에서도 학습자의 자기평가 점수가 향상되었음을 보여줌으로써, 인지부하 이론과 스캐폴딩 논의를 인간-AI 협업 교육 맥락으로 확장하는 시사점을 제공하였다.

또한 대학생 대상 AI 리터러시 교육에서 윤리 인식의 향상을 보고한 선행연구(Kong et al., 2023)와 달리 본 연구에서는 윤리 영역이 변화하지 않았는데, 이러한 차이는 교육과정에 윤리적 성찰을 다루는 명시적 설계가 포함되었는지 여부에 기인하는 것으로 해석된다.

이는 AI 리터러시의 하위 영역이 동일한 교육과정에 대해 상이하게 반응함을 보여주며, AI 리터러시를 단일 총점이 아니라 하위 영역별로 구분하여 측정하고 접근해야 한다는 다차원적 관점을 경험적으로 뒷받침한다.

실무적으로, 본 연구는 디자인·개발 지식이 없는 비전공자도 적절한 스캐폴딩이 제공되면 생성형 AI로 작동 가능한 결과물을 도출할 수 있음을 보여주었으며, 동시에 AI 윤리 의식은 실습 중심 교육의 결과물로 향상되지 않으므로 명시적 윤리 교육 요소를 별도로 설계해야 함을 시사하였다.

본 연구는 다음의 한계를 지닌다.

첫째, 통제집단이 없는 단일집단 사전-사후 설계를 채택하여, 성숙 효과나 역사 효과 등 내적 타당도 위협을 완전히 배제하지 못한다. 이는 정규 학점 강좌에서 모든 수강생에게 동일한 교육 기회를 제공해야 한다는 교육적 형평성을 고려한 선택이었으나, 관찰된 향상을 전적으로 교육과정에 귀속하기 위해서는 비교집단을 포함한 설계가 필요하다. 특히 15주간 생성형 AI를 지속적으로 사용하는 경험 자체가 특정 교육과정과 무관하게 자기평가 점수를 높였을 가능성을 배제할 수 없다.

둘째, 표본이 단일 분반 25명으로 제한되어 결과의 일반화에 신중을 요한다.

셋째, AI 리터러시를 자기보고식 설문으로 측정하였

으며, 설문의 익명성 보장을 위해 결과물과 개인별 응답을 연결하지 않아 학습자의 인식과 실제 수행 간의 관계를 직접 검토하지 못하였다. 따라서 관찰된 점수 상승이 실제 AI 리터러시 역량의 향상이 아니라 AI 활용에 대한 자신감(자기효능감)의 상승을 반영했을 가능성을 배제할 수 없다.

넷째, 질적 자료의 분석이 단일 연구자에 의해 수행되어 복수 평가자에 의한 교차 검증에는 이르지 못하였다.

이러한 한계를 바탕으로 향후 연구를 다음과 같이 제언한다.

첫째, 비교집단을 포함한 설계를 통해 교육 효과의 인과성을 보다 엄밀하게 검증할 필요가 있다. 또한 NASA-TLX나 Paas의 정신적 노력 척도 등 측정이 추가적으로 요구된다.

둘째, 표본과 대상 기관을 확대하여 결과의 일반화 가능성을 높여야 한다.

셋째, 사전 동의 절차 내에서 결과물과 설문 응답을 개인 단위로 연결할 수 있는 설계를 마련함으로써, 인식의 변화와 실제 수행의 변화가 어떻게 대응하는지를 직접 검토할 필요가 있다.

넷째, 본 연구에서 변화가 관찰되지 않은 윤리·사회적 책임 영역에 대해, 명시적 윤리 교육 요소를 과정에 포함한 설계의 효과를 별도로 검증할 필요가 있다.

이를 통해 비전공자를 유능하면서도 책임 있는 AI 활용자로 기르는 교육 모델이 정교화되기를 기대한다.

참고문헌

1. Belland, B. R., Walker, A. E., Kim, N. J., & Lefler, M., 'Synthesizing results from empirical research on computer-based scaffolding in STEM education: A meta-analysis', SAGE Publications, Review of Educational Research, 2017
2. Grover, S., & Pea, R., 'Computational thinking in K-12: A review of the state of

the field', SAGE Publications, Educational Researcher, 2013

3. Hwang, Y., & Wu, Y., 'The influence of generative artificial intelligence on creative cognition of design students: A chain mediation model of self-efficacy and anxiety', Frontiers Media SA, Frontiers in Psychology, 2024
4. Kim, N. J., Belland, B. R., & Walker, A. E., 'Effectiveness of computer-based scaffolding in the context of problem-based learning for STEM education: Bayesian meta-analysis', Springer, Educational Psychology Review, 2018
5. Kirschner, P. A., Sweller, J., & Clark, R. E., 'Why minimal guidance during instruction does not work: An analysis of the failure of constructivist, discovery, problem-based, experiential, and inquiry-based teaching', Routledge(Taylor & Francis), Educational Psychologist, 2006
6. Kong, S.-C., Cheung, W. M.-Y., & Zhang, G., 'Evaluating an artificial intelligence literacy programme for developing university students' conceptual understanding, literacy, empowerment and ethical awareness', International Forum of Educational Technology & Society, Educational Technology & Society, 2023
7. Laupichler, M. C., Aster, A., Haverkamp, N., & Raupach, T., 'Development of the "Scale for the assessment of non-experts' AI literacy" - An exploratory factor analysis', Elsevier, Computers in Human Behavior Reports, 2023
8. Laupichler, M. C., Aster, A., Schirch, J., & Raupach, T., 'Artificial intelligence literacy in higher and adult education: A scoping literature review', Elsevier, Computers and Education: Artificial Intelligence, 2022
9. Long, D., & Magerko, B., 'What is AI literacy? Competencies and design considerations', Association for Computing

- Machinery(ACM), Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems, 2020
10. Ng, D. T. K., Leung, J. K. L., Chu, S. K. W., & Qiao, M. S., 'Conceptualizing AI literacy: An exploratory review', Elsevier, Computers and Education: Artificial Intelligence, 2021
 11. Ng, D. T. K., Wu, W., Leung, J. K. L., Chiu, T. K. F., & Chu, S. K. W., 'Design and validation of the AI literacy questionnaire: The affective, behavioural, cognitive and ethical approach', Wiley, British Journal of Educational Technology, 2024
 12. Peng, C., Zhang, Z., Zhang, J., Liu, Y., Jia, R., & Xu, X., 'AI literacy and creative self-beliefs among Chinese college students: Examining the underlying mechanisms and boundary conditions', Frontiers Media SA, Frontiers in Psychology, 2026
 13. Reiser, B. J., 'Scaffolding complex learning: The mechanisms of structuring and problematizing student work', Routledge(Taylor & Francis), Journal of the Learning Sciences, 2004
 14. Sweller, J., 'Cognitive load during problem solving: Effects on learning', Wiley, Cognitive Science, 1988
 15. Sweller, J., van Merriënboer, J. J. G., & Paas, F., 'Cognitive architecture and instructional design: 20 years later', Springer, Educational Psychology Review, 2019
 16. Tromp, C., & Baer, J., 'Creativity from constraints: Theory and applications to education', Elsevier, Thinking Skills and Creativity, 2022
 17. Vartiainen, H., & Tedre, M., 'Using artificial intelligence in craft education: Crafting with text-to-image generative models', Routledge(Taylor & Francis), Digital Creativity, 2023
 18. Wang, B., Rau, P.-L. P., & Yuan, T., 'Measuring user competence in using artificial intelligence: Validity and reliability of artificial intelligence literacy scale', Taylor & Francis, Behaviour & Information Technology, 2023
 19. White, J., Fu, Q., Hays, S., Sandborn, M., Olea, C., Gilbert, H., Elnashar, A., Spencer-Smith, J., & Schmidt, D. C., 'A prompt pattern catalog to enhance prompt engineering with ChatGPT', arXiv(Cornell University), arXiv preprint, 2023
 20. Wing, J. M., 'Computational thinking', Association for Computing Machinery(ACM), Communications of the ACM, 2006
 21. Zamfirescu-Pereira, J. D., Wong, R. Y., Hartmann, B., & Yang, Q., 'Why Johnny can't prompt: How non-AI experts try (and fail) to design LLM prompts', Association for Computing Machinery(ACM), Proceedings of the 2023 CHI Conference on Human Factors in Computing Systems, 2023