



중학생 대상 생성형 AI 기반 프로그래밍 수업의 효과 분석

Analysis of the Effectiveness of Generative AI-based Programming Classes for Middle School Students

심현철[†] · 김경백^{††}
Hyunchul Shim[†] · Kyungbaek Kim^{††}

요약

본 연구는 중학생을 대상으로 생성형 AI 기반 프로그래밍(Python) 수업을 설계·운영하여 인공지능 리터러시와 학업 성취도에 미치는 영향을 검증하는 데 목적이 있다. 이를 위해 광주광역시 소재 중학교 1학년 6개 학급을 실험집단 3개, 통제집단 3개로 구분하고, 문제 기반 학습(Problem-Based Learning)에 기반한 5차시 수업을 적용하였다. 실험집단은 ChatGPT를 활용하여 프롬프트 설계, 코드 작성, 오류 수정 활동을 수행하였고, 통제집단은 동일 단원을 교과서 중심의 전통적 설명·연습 방식으로 학습하였다. 인공지능 리터러시는 6개 하위 영역으로 구성된 24문항(5점 Likert 척도) 설문으로, 학업 성취도는 단위 성취 기준에 근거한 객관식 16문항의 검사로 측정하였다.

사전·사후 검사를 통한 분석 결과, 실험집단은 인공지능 리터러시의 모든 하위 영역에서 통계적으로 유의미한 향상을 보였으며, 특히 프롬프트 작성 능력, 윤리 인식, 비판적 사고력에서 두드러진 성장($t > 2.7, p < .01$)을 나타냈다. 학업 성취도 총점 또한 78.5점에서 85.3점으로 유의하게 상승하였다($t = 2.51, p = .015$). 이러한 결과는 생성형 AI를 통합한 PBL 기반 프로그래밍 수업이 중학생의 개념 이해와 문제 해결 과정에 긍정적으로 작용했음을 시사하지만, 장기적 효과나 고차사고 전반에 대한 인과관계는 본 연구의 범위를 넘어서는 것으로 신중하게 해석할 필요가 있다. 그럼에도 본 연구는 생성형 AI를 중학교 정보과 수업에 체계적으로 통합한 수업 모형과 그 효과를 실험·통제 비교를 통해 제시함으로써, 향후 수업 설계와 교사 연수, AI 윤리 교육 통합을 위한 실천적 기초 자료를 제공한다.

주제어 생성형 인공지능, ChatGPT, 인공지능 리터러시, 프로그래밍 교육, 문제 기반 학습

ABSTRACT

This study aims to design and implement a generative AI-based programming (Python) class for middle school students and to examine its effects on their AI literacy and academic achievement. To this end, six first-grade classes at a middle school in Gwangju were divided into an experimental group (three classes) and a control group (three classes), and a five-session Problem-Based Learning (PBL) program was applied. The experimental group used ChatGPT to carry out activities such as prompt design, code writing, and error correction, whereas the control group learned the same unit through textbook-centered traditional instruction consisting of teacher explanation and practice. AI literacy was measured with a 24-item questionnaire composed of six sub-domains on a 5-point Likert scale, and academic achievement was measured with a 16-item multiple-choice test developed on the basis of the curriculum achievement standards. Analysis of the pre- and post-tests showed that the experimental group demonstrated statistically significant improvement in all sub-domains of AI literacy, with particularly notable gains in prompt-writing ability, ethical awareness, and critical thinking ($t > 2.7, p < .01$). The total academic achievement score also increased significantly from 78.5 to 85.3 points ($t = 2.51, p = .015$). These results suggest that a PBL-based programming class integrated with generative AI has a positive effect on middle school students' conceptual understanding and problem-solving processes, although causal claims about long-term effects or higher-order thinking as a whole go beyond the scope of the present study and should be interpreted with caution. Nevertheless, by experimentally comparing an AI-integrated class with a control group, this study presents a concrete instructional model for systematically integrating generative AI into middle school information classes and provides practical foundational data for future lesson design, teacher professional development, and the integration of AI ethics education.

Keywords Generative Artificial Intelligence; ChatGPT; AI Literacy, Programming Education, Problem-Based Learning

†정회원 전남대학교 교육대학원 전산교육과 교육학 석사
††중신회원 전남대학교 인공지능학부 교수 (교신저자)
논문투고 2025년 08월 07일
심사완료 2025년 11월 28일
게재확정 2025년 12월 16일
발행일자 2026년 03월 22일

1. 서론

현대 사회는 4차 산업혁명의 물결과 함께 디지털 전환이 급격하게 이뤄지고 있는 시대적 배경 속에서 살아가고 있다. 인공지능(Artificial Intelligence, AI)은 이러한 변화의 핵심 기술로 자리 잡았으며, 단순한 정보 처리 도구를 넘어 인간의 사고를 모방하고 새로운 산출물을 생성하는 수준에 이르고 있다. 특히 최근에는 사용자의 자연어 질의에 기반하여 텍스트, 코드, 이미지 등을 생성할 수 있는 생성형 인공지능(Generative AI)이 등장하면서, 일상생활뿐 아니라 교육, 산업, 문화 전반에서 그 영향력이 빠르게 확장되고 있다 [1-3].

ChatGPT와 같은 생성형 AI 도구는 학습자의 질문에 대해 실시간으로 응답하고, 다양한 예시·설명·코드를 제시할 수 있다는 점에서 기존 검색 도구와는 다른 학습 경험을 제공한다. 프로그래밍 교육의 관점에서 보면, 이러한 도구는 코드 자동 생성, 오류 메시지 해석, 디버깅 전략 제안 등에서 학습을 지원할 수 있어 학습자의 인지적 부담을 줄이고, 반복적인 실습을 통해 개념 이해를 심화하는 데 도움을 줄 잠재력을 지닌다. 동시에, 생성형 AI는 근거가 불충분한 응답이나 사실과 다른 정보를 제시할 수 있고, 학습자가 도구에 지나치게 의존하게 될 위험도 존재한다. 따라서 교육 현장에서 생성형 AI를 활용할 때에는 도구의 장점과 한계를 균형 있게 이해하고, 교사의 지도와 학습자의 비판적 사고를 전제로 한 활용 방안 모색이 요구된다[4-7].

중학교 단계는 구체적 조작기에서 형식적 조작기로 이행하는 시기로, 추상적·논리적 사고 능력과 자기주도 학습 역량이 본격적으로 발달하기 시작한다. 동시에, 이 시기의 학생들은 디지털 기기와 온라인 환경에 익숙한 ‘디지털 네이티브’ 세대로서 정보기술을 활용한 학습에 높은 수용성과 흥미를 보이는 경향이 있다. 이러한 특성을 고려할 때, 중학교 정보 교과에서 생성형 AI를 수업 도구로 적절히 통합한다면 학습자의 흥미와 참여를 유발하고, 프로그래밍 개념 이해와 문제 해결 경험을 풍부하게 할 수 있을 것으로 기대된다. 그러나 실제 중학교 정보과 수업에서 생성형 AI를 구조적으로 통합하여 수업을 설계·운영하고, 그 효과를 실험집단과 통제집단을 비교하는 방식으로 검증한 연구는 아직 많지 않다 [8-10].

또한 인공지능 시대를 살아갈 학생들에게는 단순히도구를 ‘잘 쓰는 능력’을 넘어서, 인공지능의 개념과 원리를 이해하고, 사회·윤리적 영향을 성찰하며, AI가 제공하는 정보를 비판적으로 평가·활용하는 인공지능 리터러시(AI literacy)가 필수 역량으로 부각되고 있다. 그럼에도 불구하고 실제 학교 현장에서는 인공지능 리터러시를 체계적으로 길러 줄 수 있는 수업 모형과 평가 도구가 충분히 보급되어 있지 않으며, 생성형 AI를 활용한 수업이 중학생의 인공지능 리터러시와 학업 성취도에 어떤 변화를 가져오는지에 대한 실증적인 자료도 부족한 실정이다[11-16].

이와 같은 맥락에서, 본 연구자는 광주광역시 소재 중학

교 1학년 6개 학급을 대상으로 Python 프로그래밍 단원에 생성형 AI(ChatGPT)를 통합한 문제 기반 학습(Problem-Based Learning, PBL) 수업을 설계·적용하였다. 그리고 수업 전·후 인공지능 리터러시 설문(6개 하위 영역, 24문항)과 학업 성취도 검사(단원 성취 기준에 근거한 객관식 16문항)를 실시하여, 생성형 AI 기반 프로그래밍 수업이 중학생에게 어떤 영향을 미치는지를 통계적으로 검증하고자 하였다. 이를 통해 실제 학교 현장에서 활용 가능한 수업 모형과 근거 자료를 제시하고, 생성형 AI를 교육적으로 통합하는 방향에 대한 논의를 구체화하여 내용들을 전달하고자 한다 [17].

1.1 연구 목적 및 연구 문제

앞서 살펴본 바와 같이, 생성형 AI는 프로그래밍

교육에 새로운 가능성을 제공하지만, 중학교 정보 교과의 실제 수업 맥락에서 그 효과를 검증한 연구는 부족하다. 특히 인공지능 리터러시와 학업 성취도를 동시에 고려하여 생성형 AI 기반 수업의 영향을 살펴본 연구는 찾아보기 어렵다. 이에 본 연구의 궁극적인 목적인 중학생을 대상으로 한 생성형 AI 기반 프로그래밍 수업이 인공지능 리터러시와 학업 성취도에 미치는 영향을 실증적으로 분석하는 것이다. 이를 위해 다음과 같은 연구 문제를 설정하였다[18, 19].

- RQ1). 생성형 AI를 통합한 PBL 기반 프로그래밍 수업은 중학생의 인공지능 리터러시 향상에 효과가 있는가?
- RQ2). 생성형 AI를 통합한 PBL 기반 프로그래밍 수업은 중학생의 프로그래밍 학업 성취도 향상에 효과가 있는가?

이에 따라 검증 가능한 가설을 다음과 같이 도출하였다.

- H1). 생성형 AI 기반 PBL 수업을 적용받은 실험집단은 전통적인 설명·연습 중심 수업을 받은 통제집단에 비해 인공지능 리터러시 점수에서 유의미한 향상을 보일 것이다.
- H2). 생성형 AI 기반 PBL 수업을 적용받은 실험집단은 통제집단에 비해 프로그래밍 학업 성취도 점수에서 유의미한 향상을 보일 것이다.

본 연구는 이러한 가설을 검증함으로써, 생성형 AI를 프로그래밍 수업에 통합했을 때 나타나는 교육적 효과를 수치와 자료에 근거해 제시하고자 한다.

1.2 연구의 필요성 및 기대 효과

본 연구는 다음과 같은 측면에서 필요성과 의의를 가진다.

첫째, 중학교 정보과 수업 맥락에서의 실증 연구라는 점에서 의의가 있다. 많은 선행연구가 대학생이나 성인을 대상으로 생성형 AI에 대한 인식, 수용 의도, 활용 사례를 조사하는 수준에 머물러 있는 반면, 본 연구는 실제 중학교 1학년 6개 학급을 대상으로 단원 전체를 설계하여 수업에 적용하고, 실험집단과 통제집단을 비교하는 양적 연구 설계를 사용하였다. 이를 통해 “교실 수업에서 생성형 AI를 어떻게

통합할 것인가”에 대해 보다 구체적인 시사점을 제공한다.

둘째, 인공지능 리터러시와 학업 성취도를 동시에 분석했다는 점에서 기존 연구와 차별성이 있다. 본 연구는 인공지능 개념 이해, 프롬프트 작성 능력, 윤리 인식, 사회적 영향 이해, 자기 주도성, 비판적 사고력으로 구성된 6개 하위 영역별 리터러시 점수와 단위 성취 기준에 근거해 개발한 학업 성취도 검사 결과를 함께 분석하였다. 이를 통해 생성형 AI 기반 프로그래밍 수업이 학습자의 태도·인식 차원뿐 아니라 실제 성취 수준에도 어떤 영향을 미치는지 종합적으로 파악할 수 있다.

셋째, 수업 설계와 프롬프트 활용 사례를 구체적으로 제시함으로써 현장 교사의 활용 가능성을 높였다. 본 연구는 5차시에 걸친 차시별 학습 목표와 활동, 실험집단에서 ChatGPT를 활용한 프롬프트 예시와 코드 디버깅 절차를 상세히 기록하였다. 이는 향후 교사가 유사한 수업을 설계하거나 변형·적용할 때 참고할 수 있는 실질적인 자료로 활용될 수 있다.

넷째, 생성형 AI 활용에 대한 비판적 관점과 윤리 교육의 중요성을 함께 다루었다는 점에서도 의의가 있다. 연구 결과 해석에서 생성형 AI 통합 수업의 긍정적인 효과뿐 아니라, 도구 의존에 대한 학생들의 우려, 응답의 정확성에 대한 불안 등을 함께 제시함으로써, AI 도입에 따른 교육적 장점과 한계를 균형 있게 제시하고 있다. 이는 향후 AI 윤리 및 정보 윤리 교육 내용을 구체화하는 데 중요한 기초 자료가 될 수 있다.

1.3 연구 윤리 및 IRB 승인

본 연구는 전남대학교 생명윤리위원회(IRB)의 정식 심의를 거쳐 승인되었다(승인 번호 : 1040198-241212-

HR-210-02). 연구 참여 전, 학생과 보호자에게 연구 목적·절차·AI 도구 사용 범위·개인정보 보호·자발적 참여 및 중도 철회 권리를 서면으로 고지하고 동의서를 확보하였다. 수업 중 ChatGPT 활용은 교사 감독하에 학교 규정을 준수하여 이루어졌으며, 평가의 공정성을 위해 AI 산출물은 학습 자료로만 활용하고 성취도 평가는 학교 성취 기준 기반의 별도 검사로 실시하였다. 모든 데이터는 비식별화·익명 처리되었고, 연구자의 윤리적 책임하에 보안 저장·관리되었다. 연구 절차 전반은 학생의 학습권과 안전을 최우선으로 하여 설계·운영되었다.

2. 이론적 배경

2.1 생성형 인공지능과 프로그래밍 교육

최근 등장한 생성형 인공지능은 대규모 언어모델을 기반으로 사용자의 자연어 프롬프트에 반응하여 텍스트, 코드, 이미지 등의 다양한 결과물을 생성하는 기술이다. ChatGPT와 같은 생성형 AI 도구는 질문에 대한 설명, 예시 코드, 오류 수정 방안 등을 실시간으로 제시할 수 있어, 학습

자의 탐구 활동과 자기주도 학습을 지원하는 도구로 주목받고 있다.

국내에서 김태원(2023)은 ChatGPT를 포함한 GPT 계열 언어 모델의 구조와 특성을 분석하고, 자연어 기반 상호 작용과 코드 생성 기능이 교육적으로 가지는 잠재력을 논의하였다. 이명숙(2024)은 생성형 AI를 활용한 소프트웨어 교육 수업 모형을 설계·적용하고, 프로그래밍 학습에서 예시 코드 생성과 디버깅 지원이 학습자의 부담을 줄이고 몰입을 높일 수 있음을 보고하였다. 두 연구 모두 학교 현장에서 생성형 AI 활용 가능성과 사례를 정리하면서, 교사의 수업 설계와 학생의 학습 방식이 변화할 수 있음을 지적하였다.

한편 변순용(2023)은 생성형 AI의 교육적 활용이 새로운 윤리 쟁점을 동반한다는 점을 강조하였다. 이들은 편향된 데이터에 기반한 응답, 사실과 다른 정보 제시, 저작권 및 표절 문제, 학습자의 과의존과 같은 위험 요소를 지적하며, AI 도구를 수업에 도입할 때는 윤리 교육과 비판적 이해가 반드시 병행되어야 한다고 보았다.

한편, 국내 연구동향을 분석한 몇몇 문헌은 생성형 AI 수업 연구의 쏠림과 공백을 함께 지적한다. 예를 들어, 최근 동향 분석 연구들에서는 국내 생성형 AI 교육 연구가 국어·작문, 고등·대학 수준의 수업 사례, 교사·예비 교사의 인식 조사에 집중되어 있으며, 중학교 정보과·프로그래밍을 대상으로 한 실험 연구는 상대적으로 부족하다고 보고한다. 또한 많은 연구가 단일 집단 사전·사후 설계나 수업 사례 보고에 머물러 있어, 명시적인 통제집단을 두고 학업 성취도와 AI 리터러시 변화를 동시에 검증한 연구는 거의 찾아보기 어렵다고 지적한다.

프로그래밍 교육의 관점에서 보면, 생성형 AI는 코드 예시 제공, 오류 메시지 해석, 디버깅 전략 제안, 코드 리팩터링 등에서 학습을 지원할 수 있다. 그러나 학습자가 도구가 제안한 코드를 이해하지 못한 채 그대로 사용하는 경우, 오히려 개념 이해가 약화되거나 문제 해결 과정을 형식적으로만 수행할 가능성도 존재한다. 이와 관련해, 중학교 프로그래밍 수업에서 학습 포기 요인을 분석한 연구들은 추상적 개념에 대한 불안, 오류 수정 과정에서의 좌절감, 정답 중심 평가에 대한 부담 등을 반복적으로 보고하며, 이러한 요인들이 학습 포기나 소극적 참여로 이어질 수 있음을 보여준다. 생성형 AI를 수업에 통합할 때에는 이러한 위험을 줄이기 위해, 도구가 제공하는 피드백을 학습자의 사고 과정과 연결하도록 수업을 설계하고, 교사가 AI 응답의 적절성과 한계를 점검해 주는 구조가 필요하다.

본 연구는 이러한 논의를 바탕으로, 중학교 정보과의 ‘문제 해결과 프로그래밍’ 단원에서 생성형 AI를 문제 기반 학습(Problem-Based Learning, PBL) 구조 속에 통합하였다. 즉, 학생들이 실제적인 문제 상황에서 스스로 프롬프트를 설계하고, ChatGPT의 응답을 비판적으로 수용하며, 이를 기반으로 코드를 작성·수정하도록 수업 절차를 구성하였다. 생성형 AI가 단순한 정답 제공 도구가 아니라, 문제 분석 → 해결 전략 수립 → 코드 구현 → 오류 수정의 전 과정

을 함께 탐색하는 학습 파트너로 기능하도록 설계한 것이 특징이다.

본 연구는 이러한 논의를 바탕으로, 프로그래밍 단원에서 생성형 AI를 문제 기반 학습(Problem-Based Learning, PBL) 구조 속에 통합하였다. 즉, 학생들이 실제적인 문제 상황에서 스스로 프롬프트를 설계하고, ChatGPT의 응답을 비판적으로 수용하며, 이를 기반으로 코드를 작성·수정하도록 수업 절차를 구성하였다. 이를 통해 생성형 AI가 단순한 정답 제공 도구가 아니라, 문제 분석과 해결 전략 수립을 돕는 학습 파트너로 기능하도록 하는 것을 목표로 하였다[20, 21].

Table 1. Affordances and Risks of Generative AI in Programming

Dimension	Affordances in Class	Potential Risks	Instructional Moves
Learning Support	Instant feedback; example code and explanations; iterative practice	Over-reliance; passive copying	Prompt scaffolding; think-aloud/self-explanation
Error Handling	Interpreting errors; suggesting fixes; code restructuring	Hallucination; hidden bias	Verification checklists; peer review
Engagement & Motivation	Personalization; task difficulty tuning	Uncritical acceptance	Critical consumption routines; reflection
Ethics & Copyright	Awareness of source use	Plagiarism; license issues	Critical consumption routines; reflection

Table 1은 프로그래밍 수업에서 생성형 AI가 학습 지원, 오류 처리, 참여·동기, 윤리·저작권 네 차원에서 동시에 장점과 위험을 갖고 있음을 정리한 것이다. 즉각적인 피드백과 예시 코드 제공, 오류 해석과 수정 제안, 개인화된 과제 제시, 출처 인식 제고와 같은 수업 상 이점이 있는 반면, 과의존과 수동적 베끼기, 환각과 편향, 무비판적 수용, 표절 및 라이선스 문제와 같은 위험도 존재한다. 이에 대응하여 표에서는 프롬프트 설계 발판 제공, 사고 과정 말하기와 자기 설명, 검증 체크리스트와 동료 검토, 비판적 수용 절차와 성찰, 출처 표기와 인용 기록 등과 같은 교수 전략을 제안함으로써, 생성형 AI의 장점을 활용하면서도 잠재적 위험을 체계적으로 완화할 수 있는 방향을 제시한다[22-24].

2.2 인공지능 리터러시의 개념과 구성 요소

인공지능 리터러시는 인공지능 기술을 이해하고, 다양한 맥락에서 적절하게 활용하며, 그 사회·윤리적 영향을 성찰하고 비판적으로 판단하는 종합적 역량을 의미한다. AI 리터러시를 구성하는 핵심 역량과 설계 고려 요소를 정리하면서, 단순한 도구 사용 능력을 넘어 개념 이해, 사회적 영

향 인식, 윤리적 판단, 비판적 사고 등을 포함해야 한다고 주장하였다. Ng 등(2021) K-12 학생을 위한 AI 리터러시 프레임워크를 제안하고, 교육과정 설계와 수업 개발에 활용할 수 있는 요소들을 구조화하였다. Wang 등(2022)은 인공지능 리터러시를 인식(Awareness), 사용(Usage), 평가(Evaluation), 윤리(Ethics) 네 영역으로 나누어 측정 도구를 구성하였으며, 중학생을 대상으로 하는 AI 리터러시 척도를 개발하여 타당도와 신뢰도를 검증하였다.

국내에서도 변순용(2023)은 생성형 AI 시대의 AI 리터러시 교육 방향을 제시하면서, 학생들이 AI의 원리와 한계를 이해하고, 사회·윤리적 쟁점을 성찰하며, 도구를 비판적으로 활용할 수 있는 역량을 길러야 한다고 강조하였으며, 생성형 AI 활용 수업에서 학생이 단순히 결과물을 받아들이는 수준을 넘어, 프롬프트를 설계하고 응답의 적절성을 판별하는 능력이 중요하다고 지적하였다.

본 연구에서는 이러한 선행연구를 바탕으로, 중학생의 인공지능 리터러시를 다음의 여섯 가지 하위 영역으로 조작적으로 정의하였다. 이 구성은 Ng 등(2021)의 프레임워크와 척도를 참고하되, 실제 수업에서 관찰 가능한 학생 행동과 응답 양식에 맞추어 재구성한 것이다.

- ① AI 개념 이해(Understanding of AI Concepts) : 인공지능의 원리, 작동 방식, 활용 사례 등에 대한 기초적인 지식
- ② 프롬프트 작성 능력(Prompt Writing Skills) : 원하는 결과를 얻기 위해 AI에게 효과적으로 질문하고 지시문을 구성하는 능력
- ③ AI 윤리 의식(Awareness of AI Ethics) : 개인정보, 저작권, 편향, 사실 여부 등 AI 사용과 관련된 윤리적 쟁점에 대한 인식
- ④ 사회적 영향 이해(Understanding of Social Impact) : 인공지능이 미치는 경제적·사회적 영향에 대한 비판적 성찰
- ⑤ 자기 주도성(Self-directed Learning) : AI 도구를 학습과 탐구 수단으로 활용하여, 스스로 학습 목표를 세우고 실행하는 능력
- ⑥ 비판적 사고력(Critical Thinking) : AI가 제공하는 정보의 정확성과 타당성을 검토하고, 필요할 경우 수정·보완하는 인지적 역량

Table 2. AI Literacy Components and Operational Definitions (This Study)

Component	Operational Definition
Understanding of AI Concepts	Foundational knowledge of AI principles, mechanisms, and application cases
Prompt Design Skills	Ability to craft effective prompts to elicit desired outputs from AI tools
Awareness of AI Ethics	Recognition of privacy, copyright, bias, and factual accuracy issues in AI use
Understanding of Social Impact	Critical reflection on economic/societal effects of AI

Component	Operational Definition
Self-directed Learning	Using AI for exploratory, iterative learning and self-monitoring
Critical Thinking	Evaluating validity, logic, and errors in AI outputs; revising accordingly

본 연구에서는 위와 같은 구성 요소를 기준으로 사전·사후 설문을 실시하여, 생성형 AI 기반 프로그래밍 수업이 각 하위 영역에 어떠한 변화를 가져오는지 분석하였다[25-29].

2.3 생성형 AI 수업 관련 선행연구 분석

생성형 AI와 관련된 교육 연구는 크게 세 가지 흐름으로 정리할 수 있다.

첫째, 생성형 AI 도구의 교육적 가능성과 활용 방안을 탐색한 연구이다. 학교 현장에서 생성형 AI를 활용할 수 있는 수업 아이디어와 사례를 제시하면서, 교사와 학생이 AI를 어떻게 수업 맥락 속에서 사용할 수 있는지에 대해 논의하였다. 김태원(2023)은 ChatGPT 활용 사례와 전망을 통해, 자연어 대화 기반의 AI 도입이 교수자의 역할과 학습자의 학습 전략에 어떤 변화를 가져올 수 있는지 분석하였다. GPT 계열 언어모델의 학습 구조와 교육적 활용 가능성을 검토하면서, 텍스트 생성과 코드 작성 기능이 교육과정의 다양한 영역에 연계될 수 있다고 보았다. 이명숙(2024)은 생성형 AI를 활용한 소프트웨어 교육 수업 모형을 실제 수업에 적용하여, 수업 흐름과 학생 반응을 보고하였다. 이들 연구는 생성형 AI의 활용 가능성과 장점을 제시하는 데 기여하지만, 대부분 사례 제시 또는 수업 모형 제안 수준에 머무르거나, 학습자의 리터러시와 성취 변화를 통계적으로 검증하지는 않았다.

둘째, AI 리터러시 측정 도구와 교육 방향을 제시한 연구이다. Long & Magerko(2020)는 AI 리터러시의 핵심 역량과 설계 고려 사항을 정리하여 교육 프로그램 설계의 출발점을 제시하였다. Ng 등(2021)은 K-12 맥락에서 활용 가능한 AI 리터러시 프레임워크를 제안하였고, Wang 등(2022)은 인식, 사용, 평가, 윤리 네 영역을 중심으로 리터러시 측정 도구를 구성하였다. 중학생을 대상으로 AI 리터러시 척도를 개발하여 신뢰도와 타당도를 검증함으로써, 학교 현장에서 활용 가능한 평가 도구의 기반을 마련하였다. 변순용(2023)은 이러한 논의를 토대로, 국내 교육 환경에서 AI 리터러시 교육이 지향해야 할 방향을 제시하였다. 그러나 이들 연구는 주로 리터러시의 개념과 측정 틀에 초점을 두고 있어, 실제 생성형 AI 수업과 연계된 실증 분석은 상대적으로 부족하다.

셋째, 프로그래밍 및 정보과 교육 맥락에서 수업 모형과 학습자 변인을 다룬 연구이다. 교육부(2020)는 중학교 정보과 교육과정을 분석하여 인공지능 교육 내용 구성 방향을 제시하였고, 이는 본 연구에서 학업 성취도 검사 문항과 수업 내용을 설계하는 근거가 되었다. 고영희(2013)는 중학생의 자기주도학습 능력에 영향을 미치는 요인을 탐색하여, 수업 설계 시 고려해야 할 학습자 특성을 제시하였다. 여상희, 김

구연, 김동식(2023)은 중학교 프로그래밍 수업에서의 학습 포기 요인을 분석함으로써, 프로그래밍 수업에서 학습자의 부담을 줄이고 지속적인 참여를 유도할 필요성을 강조하였다. 김성원, 이은경(2022)은 인공지능 윤리 교육 프로그램을 개발·적용하여, AI 윤리 인식 향상을 위한 교수·학습 전략을 제안하였다. 이러한 연구는 본 연구의 수업 설계, 문항 구성, 분석 관점(자기 주도성, 학습 포기 요인, 윤리 인식 등)을 설정하는 데 중요한 참고가 되었다[30-33].

종합하면, 기존 연구는 생성형 AI의 교육적 가능성과 위험, AI 리터러시의 개념과 측정 도구, 정보과 수업 모형과 학습자 변인을 각각 심도 있게 다루고 있다. 그러나 중학교 정규 정보과 수업에서 생성형 AI를 프로그래밍 단원에 직접 통합하고, 인공지능 리터러시의 하위 영역과 학업 성취도를 동시에 측정하여 실험집단과 통제집단을 비교 분석한 연구는 매우 드물다. 본 연구는 이러한 공백을 보완하기 위해, ChatGPT를 통합한 PBL 기반 Python 수업을 실제로 설계·운영하고, 그 효과를 양적·질적 자료를 통해 분석함으로써 생성형 AI 수업 연구를 한 단계 구체화하고자 한다.

Table 3. Prior Studies on AI Literacy and Generative AI in Education (Synthesis)

Study	Focus & Context	Main Contribution	Limitations vis-à-vis This Study
Ng et al. (2021)	K-12 AI literacy framework	Conceptual pillars for school AI literacy	Theoretical; limited classroom effect testing
Wang et al. (2022)	Measuring AI literacy (awareness, usage, evaluation, ethics)	Structured measurement approach	Middle-school regular-class context underrepresented
Lee & Park (2023)	Middle-school literacy scale	Structured measurement approach	Not an AI-integrated classroom experiment
Long & Magerko (2020)	AI literacy competencies; design considerations	Broad competency model	General scope; limited classroom intervention
Kim (2023)	Generative-AI use cases & caveats in Korea	Balanced view of benefits/risks	Limited experimental verification
Park (2023)	GPT model structure & educational affordances	Technical-pedagogical linkage	Case-level discussion; no control comparison

2.4 윤리·사회적 쟁점과 교육과정 맥락

생성형 AI의 확산은 개인정보 보호, 저작권 및 출처, 편향과 공정성, 책임성과 투명성, 일자리·불평등 등 사회적 영향을 포함한 복합적 논점을 야기한다. 이에 따라 수업에서는 윤리·비판적 수용 교육이 필수 요소로 편성되어야 하며, 사실 검증 절차(출처 대조, 체크리스트)와 역할 분담에 기반

한 책임성 확보가 병행되어야 한다. 또한 2022 개정 정보과 교육과정의 실생활 문제 해결 역량과 정렬하여, ChatGPT는 보조 도구로서 문제 탐색-해결-검증의 전 과정을 지원하되, 평가의 공정성은 교실 내 평가 체제(별도 검사·루브릭)로 보장하는 것이 바람직하다. 관련 내용은 <Table 4>에서 제시하고자 한다.

Table 4. Ethical/Societal Issues and Classroom Implementation in Secondary Education

Issue Area	Typical Risks	Classroom Strategies	Assessment/Documentation
Privacy & Copyright	Data leakage; misuse of copyrighted content	Minimize personal data; teach citation & licensing	Maintain prompt/output logs; enforce attribution
Bias & Accuracy	Stereotypes; hallucination	Cross-checking with multiple sources; bias awareness	Verification checklist; evidence tagging
Responsibility & Transparency	Ambiguous ownership of decisions	Role assignment; peer review; revision protocols	Version control; reflective notes
Societal Impact	Inequality; job displacement; fairness	Debates; case analyses; scenario-based ethics	Rubrics including ethical reasoning

2.5 소결

요약하면, 생성형 AI는 프로그래밍 학습에서 즉각 피드백 - 오류 수정 - 반복 실습의 이점을 제공하나, 환각·편향·과 의존의 위험을 동반하므로 교사 감독과 윤리·비판적 수용 교육이 반드시 병행되어야 한다. 또한 인공지능 리터러시는 원리 이해 - 프롬프트 - 윤리 - 사회적 영향 - 자기 주도 - 비판적 사고의 다차원 구조로 정립되어 있으며, 본 연구는 이를 측정(6개 하위 영역·24개 문항 $\alpha=.89$)과 수업 설계(PBL 순환 내 AI 내재화)에 일관되게 반영하였다. 이러한 이론적 토대는 3장에서 제시할 연구 대상·절차·도구(리터러시 설문과 성취도 검사), 4장 정량·정성 결과 해석의 준거가 된다.

3. 연구 방법

3.1 연구 설계 개요

본 연구는 광주광역시 소재 K 중학교 1학년 6개 학급 학생 전체를 대상으로 하였다. 학교와 학급은 연구자가 근무 중인 학교 중 협조가 가능한 학습을 편의 표집하여 선정하였으며, 학습 단위로 실험집단 3개, 통제집단 3개로 배정하였다. 모든 학급은 동일한 학교, 동일 학년, 동일 단원을 학습하고 있었으며, 담당 교사는 동일한 정보 교사 1인이었다. 이를 통해 학교 환경과 교사 변인을 최소화하고, 수업 처치(생성형 AI 통합 여부)의 효과를 비교할 수 있도록 하였다.

연구 참여 전, 학생과 보호자에게 연구의 목적, 수업 내용, 자료 수집 방법을 충분히 설명하였고, 자발적 참여 의사를 확인한 후 동의서를 회수하였다. 연구는 전남대학교 생명윤리위원회(IRB 승인 번호 : 1040198 -241212-HR-210-02)의 승인을 받은 후 진행되었으며, 학생 개인정보는 익명 처리하여 분석에 사용하였다.

사전 검사에서는 인공지능 리터러시와 학업 성취도 점수를 활용하여 실험·통제 집단의 동질성을 확인하였다. 사전 검사 결과, 두 집단 간 평균값 차이는 통계적으로 유의하지 않아 사전 수준이 유사한 집단 간 비교가 가능함을 전제로 하였다. 사전 동질성 검토에 대한 상세 결과는 4장에서 제시한다.

Table 5. Research Design Overview

Design	Groups	Duration	Key Features
Quasi-experimental (pre/post, comparison)	Experimental (3 intact classes), Control (3 intact classes)	Five sessions (one unit)	Experimental: PBL + ChatGPT; Control: Textbook-based instruction; Pre/post measures

3.2 수업 설계 및 실험 절차

본 연구는 사전-사후 검사를 포함한 실험집단-통제집단 설계를 기본으로 적용하였다. 두 집단 모두 동일한 학습 단원과 성취 기준을 학습하지만, 교수·학습 방법에서 차이를 두어 생성형 AI 활용 여부가 학습 결과(학업성취도 및 인공지능 리터러시)에 미치는 영향을 검증하고자 하였다.

Table 6. Participants and Grouping

School Level	Grade	Classes	Grouping	Teacher
Middle school	1st grade	6	Exp. = 3 classes; Ctrl. = 3 classes	Same teacher across groups

3.3 수업 처치(Intervention)

통제집단과 실험집단의 수업 처치(Intervention)에 관한 내용은 <Table 7>, <Table 8>에 내용을 정리하여 제시하였다.

3.3.1 공통 학습 목표

두 집단 모두 순차·선택·반복 구조에 기반한 알고리즘을 이해하고 코드로 구현하며, 문제 해결 과정을 설계·실행하고, 오류를 수정하여 피드백을 반영해 개선하는 것을 공통 학습 목표로 설정하였다.

3.3.2 실험집단 : PBL + ChatGPT

실험집단은 매 차시 문제 제시 → 프롬프트 설계 → AI 피드백 반영 → 코드 실행/수정 → 결과 공유·성찰의 순환으

로 운영하였다. ChatGPT는 예시 코드 생성, 오류 메시지 해석, 리팩터링 아이디어 제안에 한정해 사용하였고, 평가는 별도의 성취도 검사로 진행하였다.

3.3.3 통제집단 : 전통적 설명 - 학습

통제집단은 교과서 예제 풀이와 개별/짝 활동 중심으로 진행하였고, 오류 수정과 피드백은 교사 주도로 제공하였다. 생성형 AI 도구는 사용하지 않음을 명확히 안내하였다.

Table 7. Instructional Conditions by Group

Component	Experimental (PBL + ChatGPT)	Control (Textbook-based)
Problem framing	Realistic tasks; student-generated prompts	Teacher-presented examples
Tool use	ChatGPT for examples, debugging, refactoring	No AI tools; textbook & IDE only
Feedback	AI + teacher guidance	Teacher guidance only
Collaboration	Pair/whole-class sharing & reflection	Pair/individual practice
Assessment	Separate unit test; open-ended survey	Same unit test; open-ended survey

Table 8. Lesson Structure by Session (Applied to Both Groups; AI Use in Experimental Only)

Session	Topic	Core Activities (Experimental: AI-integrated)
1	Python Output	Prompt "How to print in Python?" → create/test print() variants
2	Variables & Data Types	Ask/compare variable vs. type; analyze sample codes
3	Functions	Prompt "Write a function returning the sum of two numbers" → run/revise
4	Conditionals	Draft conditionals; AI-assisted error checking; guided exercises
5	Loops	Generate loop-based code; iterative execution and refinement

3.4 측정 도구

3.4.1 인공지능 리터러시 설문

선행연구를 바탕으로 6개 하위 영역 × 4문항 = 24 문항의 5점 리커트 척도를 구성하였다. 내적 일관성은 Cronbach's $\alpha = .89$ 로 양호했다. 문항은 중학생 수준에 맞게 어휘 난이도와 맥락을 조정하였다.

Table 9. AI Literacy Questionnaire Blueprint

Component	Items (n)	Sample Focus	Key References
AI Concepts	4	Understanding principles/mechanisms	
Prompt Design	4	Crafting effective prompts	

Component	Items (n)	Sample Focus	Key References
Ethics	4	Privacy, copyright, bias, accuracy	
Social Impact	4	Economic/societal implications	
Self-directed Learning	4	Exploratory, iterative learning	
Critical Thinking	4	Evaluation/revision of AI outputs	
Total	24	5-point Likert	$\alpha = .89$

3.4.2 학업 성취도 검사

'알고리즘과 프로그래밍' 단원의 성취 기준에 부합하도록 객관식 16문항으로 구성하였다. 문항의 내용타당성은 경력 10년 이상 교사 2인 검토를 통해 확보하였고, 문항 출처와 난이도 참조를 위해 국내 중학교 정보 교과서(씨마스, 천재교과서, 비상교과서)를 참고하였다.

Table 10. Programming Achievement Test Composition

Format	Items (n)	Content Coverage	Scoring
Multiple Choice	10	Output, Variables/Types, Functions, Conditionals, Loops	1 point each
Constructed Response	2	Code explanation/correction; algorithm design	Rubric-based
Total	12	Unit-aligned competencies	Max. score reported in results

<Table 10>은 본 연구에서 사용한 프로그래밍 성취도 검사의 구성 요소를 제시한 것이다. 평가 형식은 객관식 16문항으로 출력, 변수와 자료형, 함수, 조건문, 반복문 등 단원 핵심 개념을 고르게 다루었다. 모든 문항 모두 채점 기준표(rubric)에 따라 채점하였다. 전체 문항 수는 16문항이며, 단원 성취기준에 부합하는 역량을 종합적으로 측정하도록 설계되었고, 연구 결과에서는 이 검사의 총점을 중심으로 분석하였다.

3.4.3 개방형 설문

수업 종료 후 3문항의 개방형 설문을 실시하여, 즉각 피드백 유용성·반복 실습 경험·정확성·의존성 등 AI 활용 경험과 인식을 수집하였다.

3.5 자료 수집 절차

사전 검사(인공지능리터러시 검사·학업성취도 검사) → 5차시 수업 처치 → 사후 검사(인공지능리터러시 검사·학업성취도 검사) → 개방형 설문 순으로 진행하였다. 실험/통제는 동일 일정·동일 교사로 운영했으며, 생성형 AI 사용의 범위·규칙은 실험집단 오리엔테이션에서 명확히 고지하였다.

Table 11. Data Collection Timeline and Procedures

Phase	Week/Session	Activities	Artifacts Collected
Pre-test	Week 0	Orientation; consent confirmation; baseline tests	AI literacy (24), achievement (12)
Intervention	Weeks 1-3 (5 sessions)	Lessons per Table 3-4; AI use in Exp. only	Prompt drafts, code versions (Exp.)
Post-test	Week 3	Post-tests; open-ended survey	AI literacy (24), achievement (12), survey responses

3.6 분석 방법

정량 분석은 SPSS 26.0을 사용하여 집단 내 대응 표본 t-검정, 집단 간 독립표본 t-검정을 실시하였다(유의수준 $\alpha = .05$). 척도 신뢰도(α)를 산출하고, 주요 분석 전 기초 통계를 점검하였다. 정성 자료는 준거 기반 내용분석으로 코딩하되, 정확성·유용성·몰입·의존 등의 명시적 기준을 사용하여 범주화하고 서술적 빈도를 보고하였으며, 정성 자료에 대해 추론통계는 수행하지 않았다.

Table 12. Analysis Plan Mapping

Research Question	Measure	Statistic/Test	Interpretation Note
RQ1: AI literacy gains	6-subscale scores (24 items)	Paired t (within), Independent t (between)	Subscale and total scores reported
RQ2: Achievement gains	Unit test (16 items)	Paired t (within), Independent t (between)	MC + CR combined total

3.7 윤리적 고려 및 데이터 무결성

본 연구는 IRB 승인(승인 번호:1040198 -241212-HR-210-02)을 획득하였다. 학생·보호자에게 연구 목적·절차·AI 도구 사용 범위·개인정보 보호·자발적 참여·철회 권리를 서면으로 고지하고 동의를 받았다. 수업 중 ChatGPT 활용은 교사 감독하에 활용이 이루어졌으며, 평가는 별도의 성취도 검사로 실시하여 공정성을 확보하였다. 데이터는 비식별화·익명 처리 후 보안 저장하였고, 명령프롬프트·출력 로그는 수업 성찰 자료로만 사용하였다.

4. 연구 결과

사전 검사에서 인공지능 리터러시와 학업 성취도 점수를 비교한 결과, 실험집단과 통제집단 간 평균의 차이는 통계적으로 유의하지 않았다($\alpha=.05$). 따라서 두 집단은 수업 처치 이전에 유사한 수준의 인공지능 리터러시와 프로그래밍

성취도를 지닌 동질 집단으로 간주하고 이후 분석을 진행하였다.

4.1 인공지능 리터러시 변화

인공지능 리터러시 설문 하위 영역별 사전·사후 점수를 비교한 결과는 <Table 13> 과 같다. 전체 리터러시 점수에서 실험집단은 사전 3.12점에서 사후 4.21점으로 상승하였고, 이 차이는 통계적으로 유의미하였다($t = 3.48, p = .002$). 반면 통제집단은 사전 3.18점에서 사후 3.34점으로 소폭 증가하였으나, 그 차이는 유의미하지 않았다($p > .05$).

하위 영역별로 살펴보면, 실험집단의 프롬프트 작성 능력은 사전 2.95점에서 사후 4.28점($t = 3.12, p = .004$)으로, AI 윤리 인식은 3.10 점에서 4.10 점($t = 2.78, p = .007$)으로, 비판적 사고력은 3.02점에서 4.35점($t = 2.95, p = .005$)으로 각각 유의미하게 향상되었다. 이는 생성형 AI를 활용한 PBL 수업이 학생들이 AI에게 질문을 구성하고, 응답의 적절성을 판단하며, 그 과정에서 윤리적 쟁점을 고려하도록 하는 활동과 밀접하게 연결되어 있음을 시사한다. 다만 본 연구에서 측정된 것은 해당 설문 점수의 변화이므로, 이를 고차사고 전반의 향상으로 일반화하는 데에는 신중한 해석이 필요하다.

Table 13. Changes in AI Literacy by sub-domain(Pre-post)

Sub-domain	Group	Pre-test Mean	Post-test Mean	t	P
Overall AI Literacy	Experimental	3.12	4.21	3.48	.002
	Control	3.18	3.34	1.02	.115
Prompt Writing Ability	Experimental	2.95	4.28	3.12	.004
AI Ethics Awareness	Experimental	3.10	4.10	2.78	.007
Critical Thinking	Experimental	3.02	4.35	2.95	.005

4.2 학업 성취도 변화

학업 성취도 검사의 사전·사후 점수 비교 결과는 <Table 14> 와 같다. 실험집단의 총점은 사전 78.5점에서 사후 85.3점으로 상승하였고, 그 차이는 통계적으로 유의미하였다($t = 2.51, p = .015$). 반면 통제집단은 사전 77.9점에서 사후 79.2점으로 소폭 증가하였으나, 유의미한 차이는 나타나지 않았다($p = .129$).

Table 14. Comparison of Academic Achievement Score

Group	Pre-test Mean	Post-test Mean	t	P
Experimental	78.5	85.3	2.51	.015
Control	77.9	79.2	1.04	.129

이 결과는 ChatGPT를 활용한 PBL 수업이 Python 기

본 문법과 구조를 이해하고, 문제 해결 과정을 코드로 구현하는 데 있어 단기적으로 긍정적인 영향을 미쳤음을 보여준다. 특히 실험집단 학생들은 프롬프트를 활용해 여러 버전의 코드를 생성·비교하고, 오류가 발생했을 때 AI 피드백을 참고하여 코드를 수정하는 과정을 반복적으로 경험하였다. 이러한 활동은 단원 범위 안에서의 개념 이해와 정답 도출 능력 향상에 기여했을 가능성이 있다. 다만 학업 성취도 검사는 해당 단원에 한정된 검사이므로, 결과를 프로그래밍 전반의 능력 향상이나 장기적인 성취로 일반화하는 데에는 제한이 있다[32].

4.3 정성 분석 결과

개방형 응답 146건 분석 결과, 긍정 75.3%, 중립 15.8%, 부정 8.9% 비율이 확인되었다. 학생들은 “즉시 답변이 유익”, “여러 번 시도해 자신감 상승” 등을 장점으로, “정확성 우려”, “과의존 걱정”을 한계로 지적했다. 본 분석은 빈도 중심 서술에 한정하였다. 관련 응답은 <Table 15>, <Table 16>에 제시하였다.

Table 15. Types of Postive Students Responses(Examples)

Type	Example Response
Satisfaction with Immediate Feedback	“It was nice to study alone because it answered my questions right away.”
Encouraged by Repetition	“Even if I make mistakes, I can try again, so I felt more confident.”
Support for Creative Expression	“It was fun to create different versions of code using prompts.”

Table 16. Types of Students Concerns(Examples)

Type	Example Response
Doubts About Accuracy	“I’m not sure if the answer is accurate or not.”
Concern About Over-reliance	“If I keep using AI, I’m worried I won’t be able to do it on my own later.”

4.4 추가 심화 분석

본 연구에서는 리터러시 핵심 3문항(프롬프트 작성·자기 주도성·비판적 사고)에 대해 사전·사후 세부 변화와 대응 표본 t-검정을 추가로 실시하였다. 그 결과는 <Table 17>에 제시하였다.

Table 17. Item-level changes in Key AI Literacy Competencies(Experimental Group)

Item (Competency)	Pre	Post	Δ	t	P
Prompt Writing Ability	2.84	4.21	+1.37	8.72	<.001
Self-directed Learning	3.02	4.21	+1.19	9.13	<.001

4.5 소결

정량 분석에서 AI 리터러시 전 영역과 학업 성취도의 유의한 향상이 확인되었고, 정성 분석은 즉시 피드백·반복 실습의 이점을, 동시에 정확성·과의존에 대한 학생들의 비판적 인식을 보여주었다. 이는 AI 활용의 장점을 살리되, 수업 내 검증 절차·윤리 교육이 병행되어야 함을 시사한다.

5. 결론 및 제언

5.1 연구 요약

본 연구는 광주광역시 소재 중학교 1학년 6개 학급을 대상으로, 생성형 AI(ChatGPT)를 통합한 PBL 기반 Python 프로그래밍 수업이 중학생의 인공지능 리터러시와 학업 성취도에 미치는 영향을 분석하였다. 실험집단은 ChatGPT를 활용하여 프롬프트를 설계하고, 예시 코드를 생성·수정하며, 오류 메시지를 해석하고 해결 방안을 탐색하는 활동을 수행하였고, 통제집단은 교과서와 교사의 설명을 중심으로 한 전통적 설명-연습 수업을 진행하였다.

첫째, 인공지능 리터러시 전체 점수와 하위 영역별 변화 분석 결과, 실험집단은 사전보다 사후에 통계적으로 유의미한 향상을 보였다. 특히 프롬프트 작성 능력, AI 윤리 인식, 비판적 사고력 영역에서의 향상이 두드러졌으며, 통제집단에서는 유의한 변화가 나타나지 않았다. 이는 생성형 AI를 활용한 수업이 학생들로 하여금 AI에게 ‘어떻게 질문할 것인지’를 직접 고민하게 하고, 응답의 타당성을 점검·수정하는 과정을 반복하도록 한 수업 구조와 관련이 있는 것으로 해석된다.

둘째, 학업 성취도 검사에서도 실험집단은 사전 대비 사후 점수가 유의미하게 상승한 반면, 통제집단의 증가는 유의하지 않았다. 두 집단 모두 동일한 단원과 목표를 학습했음을 고려하면, ChatGPT를 활용한 반복 실습과 즉각 피드백, 여러 버전의 코드를 비교·수정하는 활동이 Python 기본 개념 이해와 문제 해결 과정에 긍정적 영향을 미쳤다고 볼 수 있다.

셋째, 개방형 설문을 통한 정성 분석 결과, 학생들은 생성형 AI 수업에 대해 “바로 질문에 답해 준다.”, “실수를 해도 다시 시도할 수 있어 자신감이 생겼다.”, “여러 버전의 코드를 만들어 보는 것이 흥미로웠다”와 같이 즉각 피드백과 반복 실습, 창의적 표현 경험을 긍정적으로 평가하였다. 동시에 “답의 정확성이 항상 확실하지 않다”, “계속 AI에 의존하게 될지 걱정된다.”와 같은 응답도 제시되어, 생성형 AI 활용 수업이 학습 편의성과 함께 정확성·의존성에 대한 비판적 인식 역시 촉발하고 있음을 확인하였다.

이상의 결과를 종합하면, 생성형 AI를 PBL 구조 속에 통합한 프로그래밍 수업은 중학생의 인공지능 리터러시와 단원 수준의 학업 성취도 향상에 일정 부분 기여하는 것으로 나타났다. 다만 이러한 결론은 본 연구의 대상과 기간, 측정 도구 범위 안에서 얻은 결과라는 점을 전제로 해석해야 한

다.

5.2 교육적 시사점

첫째, 생성형 AI는 프로그래밍 수업에서 즉각적인 피드백과 예시 코드 제공, 반복 실습 환경을 제공하는 도구로 활용될 수 있다. 그러나 본 연구 결과가 보여주듯이, 단순히 도구를 제공하는 것만으로는 충분하지 않으며, 문제 제시-프롬프트 설계-AI 응답 검토-코드 재구성-결과 공유로 이어지는 구조화된 학습 절차 속에 통합될 때 리터러시와 성취의 향상으로 연결될 가능성이 크다. 따라서 교사는 생성형 AI를 “정답 제공자”가 아니라 “사고를 촉진하는 질문 파트너”로 활용하는 수업 설계를 고민해야 한다.

둘째, 프롬프트 작성 능력, 윤리 인식, 비판적 사고력 영역에서 실험집단의 향상이 두드러졌다는 결과는, 생성형 AI 활용 수업이 인공지능 리터러시 교육과 자연스럽게 결합될 수 있음을 시사한다. 예를 들어 학생이 작성한 프롬프트와 ChatGPT의 응답을 함께 검토하면서, 개인정보·저작권·편향·사실 여부와 같은 윤리적 쟁점을 함께 논의하는 활동은 리터러시의 여러 하위 요소를 동시에 자극하는 교육적 장면이 될 수 있다. 이는 기존에 별도 차시로 분리되어 진행되던 AI 윤리·정보 윤리 교육을 프로그래밍 수업 맥락 속에서 통합적으로 구현할 수 있는 가능성을 보여준다.

셋째, 본 연구에서 실험집단 학생들은 오류가 발생했을 때 ChatGPT의 설명을 참고하여 스스로 코드를 수정하는 과정을 반복적으로 경험하였다. 이는 교사가 모든 오류를 일일이 해결해 주는 방식에서 벗어나, 학생이 AI의 도움을 받으며 오류 진단-수정 전략 탐색-코드 재작성이라는 순환 과정을 주도적으로 수행하도록 하는 수업 방식이 가능함을 의미한다. 따라서 향후 정보 교사는 생성형 AI와 전통적인 디버깅 지도를 적절히 결합하여, 학생이 “AI에 묻기 전에 스스로 확인해야 할 것”과 “AI 응답을 검증하는 방법”을 명시적으로 지도할 필요가 있다.

넷째, 교사 연수 차원에서 보면, 생성형 AI의 기술적 특성을 이해하는 것뿐 아니라 프롬프트 설계, AI 응답의 교육적 재구성, 윤리·저작권 이슈 지도 방법 등을 포함하는 실습 중심 연수가 요구된다. 교사가 자신 있는 영역에서 직접 프롬프트를 설계하고, 학생 수준에 맞게 응답을 재구성해 보는 경험을 통해, 실제 수업에서의 활용 장벽을 낮출 수 있을 것이다.

5.3 연구의 한계

첫째, 학급 단위 배정으로 인한 비무작위성이 존재한다. 사전 동질성 검토와 동일 교사·동일 단원 통제로 보완했으나, 완전한 무작위 배정이 아니므로 인과 해석은 제한적이다.

둘째, 단기 개입(5차시)으로 장기적 내재화·전이 효과를 판단하기 어렵다.

셋째, 단일 학교·학년 표본으로 일반화 가능성이 제한된다.

넷째, 정성 분석은 서술적 빈도에 그쳤으며, 질적 추론의 깊이를 더하기 위한 삼각 검증(관찰·면담·포트폴리오)이 후속 연구에서 요구된다.

5.4 후속 연구 제안

1) 도구 다양화 비교: Claude, Gemini, Copilot 등 여러 생성형 AI의 과제 유형별 적합성과 학습 결과 차이를 비교한다.

2) 학습자 특성 조절 효과: 자기 주도성·기초 코딩 수준·디지털 활용 역량이 효과에 미치는 상호작용을 분석한다.

3) 장기 중단 설계: 학기·학년 단위의 추적 연구로 리터러시의 내재화/전이를 확인한다.

4) 평가 혁신: 설명·정당화 중심 서술형, 오류 탐지·수정 기반 수행평가, 협업 기록 분석 등을 결합한 믹스드 평가 체계를 개발한다.

5) 윤리·정책 연구: 출처 표기·저작권·개인정보를 포함한 학교 규정 모델과 수업용 체크리스트/로그 표준을 제안·검증한다.

5.5 결론

생성형 AI를 PBL 중심 프로그래밍 수업에 구조적으로 내재화하면, 중학생의 AI 리터러시 전 영역과 학업 성취에서 유의한 향상을 관찰할 수 있었다. 이는 프롬프트 설계-AI 피드백-코드 재구성-성찰로 이어지는 순환이 개별화된 즉각 피드백과 반복 실습을 가능하게 하여 과정 중심 학습을 촉진했기 때문으로 해석된다. 동시에 정확성·편향·과의존의 위험이 확인된 만큼, 윤리·비판적 수용 교육과 검증 절차의 병행이 필수적이다. 본 연구는 중학생 정규 수업 맥락에서 AI-학습자 상호작용의 교육적 가치를 실증적으로 제시했으며, 현장 교사와 정책 입안자에게 설계·평가·윤리 측면의 구체적 실행 지침을 제공한다.

참고문헌

- [1] Kang, D. (2023). The emergence of ChatGPT and responses in Korean language education. *Korean Language and Literature*, 82, 469-496.
- [2] Kang, J.(2023), *Artificial intelligence literacy seen through academic discourse*. Keum Hee-jo(et al.) Generative AI literacy and citizenship in the era of chatGPT(pp.11-34). Gwangju:Indiepub.
- [3] Ko, Y. (2014). *The effects of a problem-solving enhancement program on specialized high school students' self-efficacy, self-directed learning ability, and problem-solving process*[Doctoral dissertation, Wonkwang University]. KCI Central Library.
- [4] Ministry of Education. (2020). *Science-mathematics-informatics-convergence education comprehensive plan ('20-'24)*. Sejong: Ministry of Education. (Retrieved May 26, 2020,

- from Ministry of Education press release.) from <https://www.moe.go.kr/boardCnts/viewRenew.do?boardID=294&boardSeq=80718&lev=0&m=020402&opType=N&s=mo e&statusYN=W>
- [5] Kim, K., & Lee, J. (2016). Analysis of the effectiveness of programming learning based on computational thinking. *Journal of the Korean Association of Computer Education*, 19(1), 27–39, <http://dx.doi.org/10.32431/kace.2016.19.1.004>.
- [6] Kim, S. (2023, July 20). *79% of students have used generative AI; 68% have not received instruction on its use or ethics*. Yonhap News Agency. Retrieved March 27, 2024, from <https://yna.co.kr/view/AKR20230720066600530>
- [7] Kim, S. (2023). Development of code generation prompts for programming education using generative AI. *Journal of the Korean Association of Computer Education*, 26(5), 107–117.
- [8] Kim, Y. (2009). *Media literacy in the digital era*. Seoul: CommunicationBooks.
- [9] Kim, Y. (2023). *The shock of ChatGPT: Generative AI and the future of education*. Bucheon: Freelec.
- [10] Kim, Y. (2024). *AI literacy*. Seoul: Freelec.
- [11] Kim, J. (2024). Regulation of generative artificial intelligence and media literacy. *Journal of International Politics*, 27(1), 1–21.
- [12] Kim, T. (2023). *Will ChatGPT be a driver of innovation?: Case studies and prospects of ChatGPT utilization*. The AI Report 2023-1. Daegu: National Information Society Agency.
- [13] Kim, H. (2024). *A study on the development of AI mathematics teaching and learning materials using ChatGPT*. Master's thesis, Ajou University Graduate School.
- [14] Byun, S. (2023). *Directions for AI literacy education in the era of generative AI. Intelligent Information Ethics Issue Report*, Winter, 4–10. Daegu: National Information Society Agency.
- [15] Seo, Y. (2023, December 30). *Divorced, driven to suicide, and cabinet resignation... What AI did in the Netherlands*. JoongAng Ilbo. <https://www.joongang.co.kr/article/25218663> (Accessed January 9, 2024).
- [16] Yeo, S., Kim, G., & Kim, D. (2023). Exploring Student Learning Opportunities with Programming Tasks in Middle School Computer Science Textbooks. *Journal of Educational Information and Media*, 29(1), 1–30.
- [17] Yoon, S., & Park, S. (2023). Status of Artificial Intelligence Education for Youth in Island Regions : Insights from the Perception of Elementary and Middle School Arter-Schoil Teachers. *GRI Review*, 25(2), 249–270.
- [18] Lee, D., & Jeong, J. (2019). The effects of integrated lessons on middle school mathematics statistics and Python programming on problem-solving ability and interest in the subject. *Journal of the Korea Academia-Industrial Cooperation Society*, 20(4), 336–344.
- [19] Lee, M. (2024). A study on software education instructional models using generative AI – Focusing on ChatGPT. *Journal of Practical Engineering Education*, 16(3), 275–282.
- [20] Lee, S. (2024). *A study on developing teaching and learning materials for “Artificial Intelligence Mathematics” using ChatGPT*. Master's thesis, Ajou University Graduate School, Gyeonggi-do.
- [21] Lee, Y. (2021). *A study on the characteristics of literacy in the AI era: Focusing on AI literacy and relational literacy*. 語文研究 [Eomun Research], 110, 281–302.
- [22] Lee, Y. (2022). A new paradigm and literacy in the digital age: Focusing on digital literacy and AI literacy. *Journal of Liberal Arts Studies, Davinci Future Liberal Arts Research Institute*, 20, 35–60.
- [23] Lee, C., Mo, S., Bae, S., & Lee, S. (2024). *A study on the actual use of generative AI among adolescents and measures to enhance literacy*. Korea Youth Policy Institute Research Report, 24-Basic02. Sejong: Korea Youth Policy Institute. from <https://www.nypi.re.kr/repository/handle/2022.oak/6393>
- [24] Lee, C., Mo, S., & Choi, H. (2023). *How should we foster adolescent digital talent? Korea Youth Policy Institute Research Report*, 23-Basic1. Sejong: Korea Youth Policy Institute. from <https://www.nypi.re.kr/repository/handle/2022.oak/6255>
- [25] Lee, H., & Yoo, J. (2024). Exploring university students' experiences and perceptions of the educational use of generative AI: A case study of University A. *Journal of the Korea Contents Association*, 24(1), 428–437.
- [26] Cha, S., Kim, J., Kim, J., & Lee, W. (2009). A study on changes in perception of the necessity of programming education. *Journal of the Korean Association of Computer Education*, 12(1), 1–13.
- [27] Choi, S. (2022). A study on an AI literacy framework. *Journal of the Korean Association of Computer Education*, 25(5), 73–84.
- [28] Hwang, Y. (2023). *Understanding the digital ecosystem that dominates communication is essential*. Media Literacy, 26, 6–11. Seoul: Korea Press Foundation.
- [29] Kim, S., & Lee, Y. (2022). The artificial intelligence literacy scale for middle school students. *Journal of the Korea society of computer and information*, 27(3), 225–238.
- [30] Chunjae Education (Kim, H.), CMAS (Seo, T.), & Visang Education (Lim, H.). (2022). Middle school informatics textbooks. Seoul: Chunjae Education, CMAS, & Visang Education.
- [31] Lee, S., & Park, G. (2023). *Exploring the impact of ChatGPT literacy on user satisfaction: The mediating role of user motivations*. Cyberpsychology, Behavior, and Social Networking, 26(12), 913–918.
- [32] Ng, D., Leung, J., Chu, S., & Qiao, M. (2021). *AI literacy: Definition, teaching, evaluation and ethical issues*. Proceedings of the Association for Information Science and Technology, 58(1), 504–509.
- [33] Wang, B., Rau, P., & Yuan, T. (2022). *Measuring user competence in using artificial intelligence: Validity and reliability of artificial intelligence literacy scale*. Behaviour & Information Technology.

**심현철**

- 2015년 세한대학교 컴퓨터교육학과(이학사)
- 2025년 전남대학교 전산교육전공 (교육학석사)
- 現 금당중학교 정보 교사

✦ 관심분야 : 인공지능, 정보교육, 교육과정, 교육평가
✉ hyunchul5649@korea.kr

**김경백**

- 現 전남대학교 인공지능학부 교수
- 現 전남대학교 정보화본부장
- 現 한국스마트미디어학회 학회장

✦ 관심분야 : 지능형 분산시스템, 빅데이터 플랫폼,
SDN/NFV
✉ kyungbaekim@jnu.ac.kr