



생성형 AI 기반 과학 수업에서 ChatGPT의 스캐폴딩 가능성 탐색: 중학생의 문제 풀이 학습을 중심으로*

Exploring the Scaffolding Potential of ChatGPT in Generative AI-Based Science Instruction: A Focus on Middle School Students' Problem-Solving Learning

김미나† · 김성백††

MiNa Kim[†] · SeongBaeg Kim^{††}

요약

최근 인공지능(Artificial Intelligence, AI)의 발전은 교육 현장에 새로운 교수·학습 가능성을 제시하고 있으며, ChatGPT와 같은 생성형 AI는 학습자와 상호작용하며 학습 과정을 단계적으로 지원하는 스캐폴딩 도구로 활용될 수 있다. 과학 수업의 문제 풀이 활동은 개념 이해와 오개념 교정에 중요하지만, 실제 교실에서는 질문에 대한 부담감이나 오류 인식의 어려움으로 인해 학습자의 참여가 제한되는 경우가 많다. 이에 본 연구는 중학교 과학 수업의 문제 풀이 학습 맥락에서 ChatGPT 기반 프롬프트 스캐폴딩이 학습자의 개념 이해와 문제 해결에 미치는 영향을 탐색하고자 하였다. 중학교 과학 '여러 가지 힘' 단원을 중심으로 사람 스캐폴딩 집단과 ChatGPT 프롬프트 스캐폴딩 집단을 대상으로 형성평가 결과와 학습자 인식 조사를 분석한 결과, 두 집단 모두 사전 대비 사후 형성평가에서 정답률이 유의미하게 향상되었으며, 일부 문항에서는 집단 간 향상 정도의 차이가 통계적으로 유의하지 않은 것으로 나타났다. 이는 ChatGPT 기반 프롬프트 스캐폴딩이 중학교 과학 수업에서 인간 기반 스캐폴딩을 보완하는 학습 지원 도구로 활용될 수 있으며, 질문 중심의 대화형 상호작용을 통해 학습자의 개념 점검과 재구성을 지원하는 생성형 AI 스캐폴딩의 교육적 가능성을 실제 교실 수업 맥락에서 확인했다는 점에서 의의가 있다.

주제어 인공지능, 생성형 AI, 스캐폴딩, 프롬프트 스캐폴딩, 과학 문제 풀이 학습

ABSTRACT

Recent advances in artificial intelligence (AI) have expanded new possibilities for teaching and learning in education. Generative AI tools such as ChatGPT can function as scaffolding tools by interacting with learners and providing step-by-step learning support. In science education, problem-solving activities are crucial for conceptual understanding and the correction of misconceptions; however, learners' participation is often limited by reluctance to ask questions or difficulty in recognizing errors. Therefore, this study examined the effects of ChatGPT-based prompt scaffolding on learners' conceptual understanding and problem-solving in middle school science classes. Focusing on the unit Forces, formative assessment results and learner perception surveys were analyzed for a human scaffolding group and a ChatGPT prompt scaffolding group. The results showed that both groups demonstrated significant improvements in correct response rates from pre- to post-assessment, and for some items, the differences in improvement between the groups were not statistically significant. These findings indicate that ChatGPT-based prompt scaffolding can complement human scaffolding in middle school science instruction and highlight the educational potential of generative AI scaffolding in supporting learners' concept monitoring and reconstruction through question-centered dialogue in authentic classroom contexts.

Keywords Artificial Intelligence, Generative AI, Scaffolding, Prompt Scaffolding, science problem-solving learning

†정회원 제주대학교 대학원 컴퓨터교육과 박사수료

††중신회원 제주대학교 컴퓨터교육과 교수(교신저자)

논문투고 2025년 11월 14일

심사완료 2026년 02월 03일

게재확정 2026년 03월 08일

발행일자 2026년 04월 30일

* 본 논문은 2025년 컴퓨터교육학회 하계학술대회에서 "생성형 AI 기반 과학 수업이 중학생의 문제 풀이 학습에 미치는 영향"의 제목으로 발표된 논문을 확장한 것임.

1. 서론

4차 산업혁명 시대를 맞아 첨단 테크놀로지의 발전은 교수·학습의 방식과 교육의 본질적 목표에 근본적인 변화를 일으키고 있다. 테크놀로지는 학습자가 지식을 구성하고 문제를 해결하며 타인과 상호작용하는 과정 전반에 영향을 미치며, 그 중심에는 인공지능(Artificial Intelligence, AI)의 급속한 발전이 있다. 특히 인간의 인지 구조를 모방한 AI 기술은 기존의 획일화된 학교 교육을 넘어 개인화된 학습 경험을 제공할 수 있다는 점에서 주목받고 있다. 특히 2022년 OpenAI에서 개발한 ChatGPT는 이러한 변화를 상징하는 대표적인 생성형 AI로, 교육 분야에서도 활용 가능성이 활발히 논의되고 있다[1].

딥러닝 기반의 생성형 AI 기술은 사회 전반의 다양한 영역에 영향을 미치며, 교육 영역에서도 AI 소양 교육과 학습자 역량 강화를 위한 연구가 활발하게 진행되고 있다. 학교 현장에서는 생성형 AI를 교수·학습 과정에 적용하기 위한 실천적 시도가 확산되고 있고[2, 3] ChatGPT와 같은 생성형 AI는 가상 교사 또는 멘토의 역할을 수행하여 학습자가 필요로 하는 정보를 신속하게 획득하고 학습 방법에 대한 조언을 실시간으로 받을 수 있게 한다[4].

전통적인 교실 상황에서 학생이 교사에게 질문할 기회는 제한적이며, 질문의 내용 또한 소수의 적극적인 학습자에게 집중되는 경향이 있다. 많은 학생들이 수업 중 질문이 다른 학습자의 흐름을 방해할 수 있다고 생각하거나, 잘못된 질문으로 인한 부정적 평가를 우려하기 때문이다. 반면, 생성형 AI를 활용하면 다른 사람의 시선을 의식하지 않고 자유롭게 질문할 수 있으며, 질문이 미숙한 학습자에게는 AI가 적절한 단서를 제시하여 사고를 확장하도록 도울 수 있다. 이러한 상호작용은 학습자의 흥미와 자신감을 향상시키는 긍정적 요인으로 작용한다[5].

그러나 생성형 AI의 급속한 확산은 긍정적 기대와 함께 교육의 본질에 대한 우려도 불러일으키고 있다. AI가 인간의 학습과 사고를 대체할 수 있는가에 대한 논의는 여전히 지속 중이며, ChatGPT를 교육 현장에 실제로 적용한 실증 연구는 매우 제한적이다. 특히 공교육 맥락에서 ChatGPT 프롬프트 스캐폴딩(prompt scaffolding)의 효과를 검증하고, 학습자가 문제 해결 과정에서 이를 어떻게 활용하는지를 구체적으로 분석한 연구는 부족한 실정이다[6, 7].

이에 본 연구는 중학교 과학 교과의 「여러 가지 힘」 단원에서 생성형 AI를 스캐폴딩 도구로 적용하고, ChatGPT 프롬프트 스캐폴딩이 중학생의 문제 풀이 학습에 미치는 영향을 실증적으로 분석하였다. 본 연구의 목적은 사람 스캐폴딩과 생성형 AI 스캐폴딩의 상대적 우수성을 비교·판단하는 데 있지 않다. 대신, 기존 스캐폴딩 이론에서 제시한 인지적·질문 중심 지원 전략이 생성형 AI 기반 학습 환경에서도 실제 수업 맥락 속에서 구현될 수 있는지를 탐색적으로 검증하는 데 연구의 초점을 둔다. 본 연구를 통해 향후 AI 기반 스캐폴딩의 교육적 시사점을 제시하고 효과적인 학습 환경

구축에 기여하고자 한다.

2. 이론적 배경 및 선행연구

2.1 생성형 인공지능의 개념과 교육적 의의

인공지능(AI)의 발전은 교육 현장 전반에 급격한 변화를 불러왔다[6, 7]. 특히 인간이 만든 결과물과 구별하기 어려울 정도로 정교한 산출물을 생성할 수 있는 생성형 인공지능(Generative AI)이 그 중심에 있다[6]. 생성형 AI는 대규모 데이터의 패턴을 학습하여 텍스트, 이미지, 음성, 영상 등 다양한 디지털 콘텐츠를 새롭게 생산할 수 있는 기술로, 단순한 데이터 재현을 넘어 사용자의 요구나 입력(프롬프트)에 기반한 창의적 결과물을 생성한다[8, 9].

이러한 기술은 학습자의 요구에 맞추어 개인화된 학습 경험을 제공하고, 교수·학습 방법과 평가 전반에 혁신을 가져올 잠재력을 지닌다[10]. 교육적 측면에서 생성형 AI는 학습자의 인지 수준에 맞춘 맞춤형 피드백, 문제 풀이 지원, 자료 추천, 번역 및 보고서 초안 작성, 코드 생성 등 기존 테크놀로지와는 차별화된 방식으로 학습을 지원한다. 또한 인간과 자연스러운 언어 기반 상호작용을 통해 학습자의 흥미를 자극하고 자기주도적 탐구를 촉진한다[10].

이처럼 생성형 AI는 교육의 목표, 내용, 방법, 평가 방식 등 여러 영역에 근본적인 영향을 미칠 것으로 전망된다. 한편, AI가 생산하는 콘텐츠는 할루시네이션(hallucination) 문제나 편향된 데이터로 인한 신뢰성 저하 등 교육적 위험 요소도 내포하고 있다. 과도한 AI 의존이 학습자의 비판적 사고력 저하나 동기 약화로 이어질 수 있다는 연구도 제기되었다. 따라서 교육 현장에서의 생성형 AI 활용은 효율성뿐 아니라 정보의 정확성, 학습자의 자율성과 사고력 유지라는 측면에서도 균형 잡힌 접근이 요구된다.

2.2 생성형 AI(ChatGPT)와 교육에서의 활용 연구

ChatGPT와 같은 생성형 AI는 학습자와 대화형으로 상호작용하며 실시간으로 학습을 지원할 수 있다는 점에서 교육 분야에 새로운 가능성을 제시한다. 이러한 AI는 단순한 정보 제공을 넘어 학습자의 수준과 맥락에 맞는 응답을 생성하여 개념 이해, 피드백, 상담, 평가 등 다양한 교수·학습 단계에서 활용될 수 있다[11].

조현국(2023)은 과학 교육에서 텍스트 기반 생성형 AI가 학습자의 질의응답 과정에서 개념 이해를 돕고, 학습자 맞춤형 피드백 제공에 기여할 수 있음을 제시하였다[12]. 김주인과 유훈(2023)은 물리교육 맥락에서 ChatGPT를 활용하는 여섯 가지 영역—강의 계획 수립, 교육 자료 준비, 물리 문제 풀이, 응답의 신뢰성 검증, 평가 문항 개발, 서술형 답안 채점—을 탐색하여 ChatGPT가 교육의 보조 수단으로 충분히 활용 가능함을 확인하였다[12].

이러한 연구들은 생성형 AI가 과학 교육에서 교사의 설명을 보조하거나 학습자의 탐구 과정에서 개념적 이해를

강화할 수 있는 가능성을 보여준다. 그러나 대부분의 연구가 개별 사례 분석에 그치거나 ChatGPT의 ‘활용 가능성’을 중심으로 논의되어, 실제 학습자의 문제 해결 과정에서 AI가 어떠한 인지적 스캐폴딩 역할을 수행하는지는 충분히 검증되지 않았다. 본 연구는 이러한 한계를 보완하고자 ChatGPT 프롬프트를 스캐폴딩 도구로 활용하여 중학생의 과학 문제 풀이 학습에서 그 효과를 실증적으로 탐색하였다.

2.3 과학 교육에서의 생성형 AI 활용과 시사점

과학 교육은 개념적 이해와 탐구 능력 함양을 동시에 요구하는 학문 영역으로, 학습 과정에서 질문과 피드백이 핵심적 역할을 한다. 생성형 AI는 이러한 상호작용을 모사하여 학습자의 개념 정교화와 오개념 수정에 도움을 줄 수 있다. 최근 ChatGPT는 실험 설계, 개념 설명, 오류 진단, 과학적 의사소통 훈련 등 다양한 과학 학습 활동에서 지원 도구로 활용되고 있다[11, 12].

특히 문제 풀이 학습에서 ChatGPT는 학생의 사고 과정을 단계적으로 유도하는 질문을 제시하고, 학습자의 반응을 기반으로 추가 설명이나 피드백을 제공함으로써 ‘대화형 스캐폴딩’ 기능을 수행한다. ‘대화형 스캐폴딩’이란 비고츠키의 ZPD, Rogoff의 3층 분석틀을 바탕으로, 전문가(교사)가 학습자의 현재 수준에 맞게 도움을 조정하면서도 그 수준을 ‘늘려 주는’ 대화적 지원을 스캐폴딩이다. 즉, 질문, 재진술, 설명 요구, 피드백 등 상호작용적 발화 행위를 통해 학습자의 사고를 확장시키고, 점차적으로 책임을 학습자에게 이전하는 임시적 언어적 지원을 dialogic scaffolding이라 정의한다. 다시 말해 “학습자가 혼자서는 수행하기 어려운 이해·추론 활동을 할 수 있도록, 교사나 또래가 대화적 상호작용을 통해 민감하게 조정된 언어적 도움을 제공하고, 시간이 지나며 점차 그 도움을 줄여 가는 과정으로서의 스캐폴딩”이라고 볼 수 있다[13]. 이는 학습자가 교사나 또래의 도움 없이도 자신의 오개념을 인식하고 수정할 수 있는 환경을 제공한다. 그러나 ChatGPT의 응답 품질은 프롬프트의 구체성과 명확성에 따라 달라지므로, 교사의 프롬프트 설계 및 지도 전략이 필수적이다. 이러한 맥락에서 ‘프롬프트 스캐폴딩’ 개념이 주목받고 있다.

2.4 프롬프트 스캐폴딩(Prompt Scaffolding)의 개념과 특징

프롬프트(prompt)는 사용자가 인공지능 언어모델에 제공하는 명령어나 질문을 의미하며, 프롬프트 엔지니어링(prompt engineering)은 이를 효과적으로 설계·활용하는 과정을 말한다[6]. 프롬프트의 품질은 AI의 응답 정확도와 학습 지원 효과를 결정하는 핵심 요소로, 명확하고 구체적인 프롬프트는 유의한 학습 피드백을 유도하지만, 부적절한 프롬프트는 비논리적이거나 무의미한 결과를 초래할 수 있다.

장진호(2024)는 프롬프트 스캐폴딩을 “교사가 학습자의

요구를 반영하여 ChatGPT와의 상호작용 과정에서 학습자가 정확하고 목적에 부합하는 응답을 얻을 수 있도록 지원하는 교수 전략”으로 정의하였다[6, 14]. 이는 프롬프트 엔지니어링의 개념을 교육적 맥락에 적용한 것으로, 교사는 학습자에게 적절한 프롬프트 작성법을 안내하고 구체적 예시를 제시함으로써 학습자의 질문을 명료화하고 ChatGPT와의 상호작용을 원활하게 한다.

연구 결과, 프롬프트 스캐폴딩은 학습자의 AI 의사소통 역량(인공지능이 이해하기 쉬운 언어로 사고를 표현하고 활용하는 능력)을 향상시키는 것으로 나타났다. 또한 학습자는 교사의 프롬프트 지원을 통해 ChatGPT와의 대화를 보다 명확하게 인식하고, 인공지능 활용에 대한 긍정적 태도를 형성하였다. 이러한 맥락에서 프롬프트 스캐폴딩은 단순한 도구적 지원을 넘어 학습자의 사고력과 자기주도성을 강화하는 새로운 형태의 학습 지원 전략으로 주목받고 있다.

본 연구에서 활용한 ChatGPT 프롬프트 스캐폴딩은 단순한 정보 제공이나 정답 제시에 머무르지 않고, 학습자의 응답에 따라 질문의 수준과 방향이 조정되는 대화형·적응형 스캐폴딩을 핵심으로 한다. 이는 기존 스캐폴딩 이론에서 강조한 인지적 지원 요소를 생성형 AI의 상호작용 특성에 맞게 재구성한 것이다. 구체적으로 본 연구의 ChatGPT 프롬프트 스캐폴딩은 다음과 같은 네 가지 전략으로 구분된다.

첫째, 인지 점검형 스캐폴딩(cognitive checking scaffolding)은 문제 해결 과정에서 학습자의 현재 사고 상태와 개념 이해 수준을 점검하기 위한 질문을 제시하는 전략이다. 이는 학습자가 자신의 사고 과정을 인식하고, 문제 해결의 출발점을 명확히 하도록 돕는다.

둘째, 오개념 인식 유도형 스캐폴딩(misconception awareness scaffolding)은 학습자의 응답에 포함된 오류를 직접적으로 지적하기보다는, 조건을 변형한 질문이나 반례 제시를 통해 학습자가 스스로 오개념을 인식하도록 유도하는 전략이다.

셋째, 단계적 추론 유도형 스캐폴딩(stepwise reasoning scaffolding)은 문제 해결 과정을 여러 단계의 질문으로 분절하여, 학습자가 논리적 추론 과정을 순차적으로 구성하도록 지원하는 전략이다. 이 과정에서 ChatGPT는 학습자의 응답 내용에 따라 후속 질문을 조정함으로써 적응적 피드백을 제공한다.

넷째, 개념 재구성 촉진형 스캐폴딩(conceptual reconstruction scaffolding)은 문제 해결 이후 학습자가 개념을 자신의 언어로 재정의하거나 설명하도록 요구함으로써, 개념을 정교화하고 장기 기억으로 전이하도록 돕는 전략이다.

이와 같은 프롬프트 스캐폴딩은 교사나 또래의 즉각적 설명 중심 스캐폴딩과 달리, 학습자의 반응을 기반으로 질문과 피드백이 조정된다는 점에서 생성형 AI 환경에 특화된 스캐폴딩 전략으로 볼 수 있다.

3. 연구 방법

3.1 연구 대상

본 연구는 O중학교 1학년 학생 총 250명(남학생 141명, 여학생 109명)을 대상으로 실시하였다. 연구 참여 학생들은 1학년 과학 교과와 '여러 가지 힘' 단원 중 '중력'과 '부력' 개념을 중심으로 수업에 참여하였다. 연구 집단은 학급 단위로 사람 스캐폴딩 집단과 ChatGPT 프롬프트 스캐폴딩 집단으로 구성하였다. 집단 간 학업 성취도의 차이가 연구 결과에 영향을 미치지 않도록, 이전 학기 과학 성취도를 기준으로 학급의 전반적인 성취 수준을 분석한 후 두 집단의 성취 수준이 유사하도록 실험집단과 통제집단을 구성하였다. 이를 통해 두 집단 간 초기 학업 수준의 동질성을 확보하고자 하였다. 수업은 스캐폴딩 유형(사람 스캐폴딩과 AI 스캐폴딩)에 따라 구분하여 운영되었으며, 두 집단 모두 동일한 학습 목표와 학습 내용을 바탕으로 문제 풀이 활동을 수행하였다. 또한 생성형 AI가 과학 문제 풀이 과정에서 스캐폴딩으로 기능할 수 있는지를 검증하기 위해, 수업 전 과정을 성실히 이수하고 사전·사후 설문 에 모두 응답한 학생의 자료를 연구 대상으로 선정하였다.

3.2 연구 절차

본 연구의 절차는 다음과 같다.

첫째, 중학교 1학년 과학 교과와 「여러 가지 힘」 단원 중 '중력'과 '부력' 개념에 대한 학습자의 이해 수준을 측정하기 위해 동일한 형식의 형성평가를 각 개념별로 2회분 준비하였다.

둘째, 문제 풀이 과정에서 학습자가 제기할 수 있는 다양한 질문을 예측하고, 이를 생성형 인공지능(ChatGPT)과의 상호작용 형태로 구현할 수 있도록 프롬프트 문장을 설계하였다. 프롬프트는 중학교 1학년 수준의 언어와 사고 수준에 맞추어 개발하였다.

셋째, 과학 수업을 통해 '중력'과 '부력'의 기본 개념을 학습한 후, 학생들의 사전 개념 이해 정도를 확인하기 위한 1차 형성평가를 실시하였다.

넷째, 1차 형성평가 후 학생들은 스스로 답안을 채점하고, 오답 문항이나 개념이 불명확한 부분을 확인하였다. 학생들은 해당 문항을 다시 해결하기 위해 과학 개념을 재학습하였으며, 이때 스캐폴딩 유형에 따라 사람 스캐폴딩(교사 또는 또래의 도움)과 생성형 AI 스캐폴딩(ChatGPT 프롬프트 활용)으로 나뉘어 지원을 받았다. 사람 스캐폴딩은 수업 중 교실 현장에서 자연스럽게 이루어지는 상호작용을 의미하며, 사전에 구조화된 절차나 고정된 발화 예시로 제공되는 개입이 아니라 학습자의 필요에 따라 유동적으로 제공되는 지원이다. 본 연구에서는 모둠별 문제 풀이 활동 과정에서 학습자가 형성평가 문항을 틀린 이유에 대해 본인보다 유능한 또래 또는 교사에게 질문하고, 정답을 직접 제공받기보다는 과학 개념과 원리에 대한 질의응답을 통해 자신의 사고를 점검·보완하는 과정으로 스캐폴딩이 이루어졌다. 각 집단의 학생들은 제공된 스캐폴딩을 바탕으로 오개념의 원인을 탐색하고, 과학 개념을 재구성하며 문제 해결 과정을

정교화하였다. 문제 풀이 과정에서 생성형 AI 스캐폴딩을 제공받은 학생들은 사전에 설계된 단일 프롬프트를 일괄적으로 적용받은 것이 아니라, 학습자의 응답에 따라 질문의 유형과 난이도가 조정되는 프롬프트 스캐폴딩을 경험하였다. ChatGPT는 학습자의 초기 응답을 기반으로 인지 점검, 오개념 탐색, 단계적 추론 유도, 개념 재구성을 위한 질문을 순차적으로 제시함으로써 문제 해결 과정을 지원하였다.

이러한 방식은 기존 연구에서 제시한 인지적·질문 중심 스캐폴딩 전략을 생성형 AI의 대화적 특성을 활용하여 구현한 것으로, 학습자가 교사의 직접 개입 없이도 자신의 사고를 점검하고 개념을 재구성할 수 있는 학습 환경을 제공한다. Figure 1은 학생들에게 제공된 ChatGPT 프롬프트 스캐폴딩의 예시를 나타낸다.

활동① 형성평가를 풀고 힌트 받기 <프롬프트 예시>	
1번 문항	(예시1)지구 중력의 방향은 어디인지 설명해주세요 (예시2)지구 중력 방향에 대해 설명해주세요 (예시3)지구 중력 방향을 중학교 1학년 수준으로 설명해주세요
2번 문항	(예시1)질량이 10kg인 물체의 물체는 지구와 달에서 각각 몇N인지 계산하는 문제에 대한 힌트를 주세요. 답은 알려주지 마세요. 중학교 1학년 수준으로 설명해주세요. (단, 중력가속도는 9.8) (예시-1)달의 중력가속도 값을 1.6이라고 한 경우, 추가 질문 왜 달에서 중력가속도 값은 1.6인가요?
3번 문항	(예시1)질량이 10kg인 물체의 물체의 달에서의 질량과 무게를 각각 계산하는 문제에 대한 힌트를 주세요. 답은 알려주지 마세요. 중학교 1학년 수준으로 설명해주세요. (단, 중력가속도는 9.8)
4번 문항	(예시1)지구의 중력이 1/2로 줄어든다면 10kg의 물체의 무게를 구하는 문제에 대한 힌트를 주세요. 답은 알려주지 마세요. 중학교 1학년 수준으로 설명해주세요. (단, 중력가속도는 9.8)

Figure 1. Example of ChatGPT Prompt Scaffolding (ChatGPT prompts presented to students as scaffolding during science problem-solving activities)

다섯째, 학생들의 학습 효과를 확인하기 위해 '중력'과 '부력' 개념에 대한 2차 형성평가를 실시하였다.

여섯째, 사람 스캐폴딩 집단과 생성형 AI 스캐폴딩 집단의 1차·2차 형성평가 결과를 비교·분석하여 개념 이해도와 문제 해결력의 향상 정도를 검증하였다.

본 연구는 프로그램의 효과성을 검증하기 위해 Table 1 과 같이 이질집단 사전·사후 검사 설계(non-equivalent pretest-posttest design)를 적용하였다.

Table 1. Research Design

Pre-test	Scaffolding	Post-test
O1	X1	O2
O3	X2	O4
O ₁ , O ₂ :	Pre-test on scientific concepts of force (gravity, buoyancy)	
X ₁ :	Problem-solving activity with human scaffolding (teacher or peer support)	
X ₂ :	Problem-solving activity with AI-based prompt scaffolding using ChatGPT	
O ₃ , O ₄ :	Post-test on scientific concepts of force (gravity, buoyancy)	

3.3 연구 도구

본 연구에서 사용한 도구는 다음과 같다. 본 연구에서는 생성형 AI(ChatGPT)를 활용한 프롬프트 스캐폴딩이 과학 문제 풀이 학습에 미치는 영향을 검증하기 위해 ‘중력’과 ‘부력’ 개념에 대한 형성평가 문항을 활용하였다. 형성평가는 총괄평가를 위한 측정 도구가 아니라, 학습자의 개념 이해 상태를 진단하고 이후 제공되는 스캐폴딩의 방향을 설정하기 위한 형성적·진단적 평가 도구로 활용되었다. 이에 따라 문항의 난이도나 변별도를 통계적으로 세분화하기보다는, 중학교 과학 교육과정 성취기준에 근거하여 기본 개념 이해 여부를 확인할 수 있는 중간 수준의 난이도로 문항을 설계하고, 학습자의 오답 유형과 개념 이해 여부를 파악하는 데 분석의 초점을 두었다. ‘중력’ 개념 형성평가는 2022 개정 교육과정의 성취기준과 다수의 교과서를 분석하여 연구자가 직접 문항을 개발한 뒤, 과학교육 전문가 2인의 검토를 통해 내용 타당도를 확보하였다. 중력의 방향, 질량과 무게의 구분, 중력의 크기 등 중학교 수준에서 요구되는 핵심 개념 이해 여부를 진단하기 위한 형성평가 문항으로 구성되었다. 또한 서로 구분되는 개별 개념 요소에 대한 이해를 측정하도록 구성되었다. 각 문항은 하나의 핵심 개념 요소에 초점을 두어, 학습자의 개념별 이해 변화가 문항 반응을 통해 드러나도록 설계되었다. Figure 2는 중력 형성평가의 예시 문항을 나타낸다.

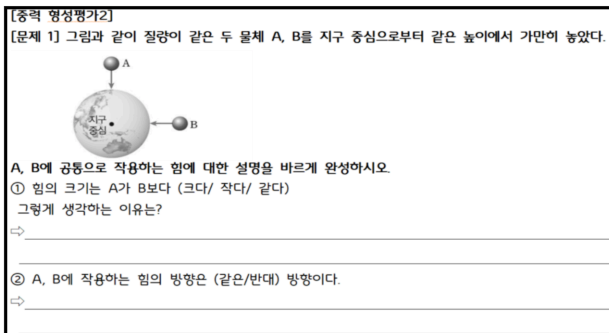


Figure 2. Example of Formative Assessment on Gravity (Question on the Magnitude and Direction of Gravitational Force Acting on Two Objects with Equal Mas)

‘부력’ 개념 형성평가는 김이영, 김지나(2012)의 「중학생들의 부력 개념 분석」에서 제시한 검사지 중 일부 문항을 본 연구의 수업 맥락에 맞게 수정·보완하여 사용하였다. 해당 문항은 선행 연구에서 중학생의 부력 개념 이해를 진단하기 위해 개발·검증된 도구로, 본 연구에서는 문항의 타당도와 신뢰도를 새롭게 산출하기보다는 선행 연구에서 확보된 검증 결과를 근거로 활용하였다. 문항 내용은 물과 공기 중에서 물체의 무게 비교, 물체를 물에 넣는 방향 변화에 따른 부력 변화, 무게가 같고 모양이 다른 물체의 부력 크기 비교, 물에 잠긴 물체의 부피에 따른 부력 크기 비교 등으로 각 문항은 부력 개념을 구성하는 서로 다른 개념 요소(매질, 부피, 잠긴 정도)에 대한 이해를 개별적으로 측정하도록 설계되었다. Figure 3은 부력 형성평가의 예시 문항을 나타낸다.

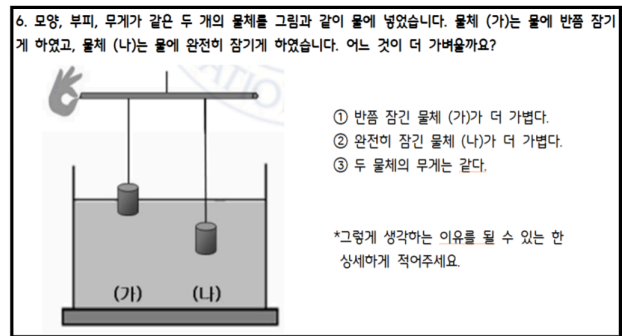


Figure 3. Example of Formative Assessment on Buoyancy (Comparing the Buoyant Force of Two Objects with Different Submerged Volumes)

본 연구에서는 형성평가 문항의 난이도나 변별도를 새롭게 분석하기보다, 학습자의 오답 유형과 개념 이해 여부를 파악하는 데 초점을 두었으며, 이를 통해 생성형 AI 기반 스캐폴딩 제공 과정에서의 학습 지원 효과를 분석하고자 하였다.

1차 및 2차 형성평가는 동형 문항으로 구성되었으나, 본 연구의 분석 초점은 총점 또는 평균 정답률의 변화가 아니라 개별 문항 반응을 통해 드러나는 개념 요소별 이해 변화를 탐색하는 데 두었다. 이는 생성형 AI 기반 스캐폴딩이 학습자의 문제 풀이 과정에서 특정 과학 개념 이해에 어떠한 영향을 미치는지를 미시적으로 분석하고자 한 연구 목적에 따른 선택이다. Table 2는 중력, 부력 형성평가에서 문항별로 묻고 있는 과학 개념 내용을 나타낸다.

Table 2. Item-by-Item Scientific Concept Learning Content

		Scientific Concept Learning Content
Gravity Formative Assessment	①	Direction of Gravity
	②	Mass and Weight
	③	Magnitude of Gravity
	④	Comparison of the Magnitude of Gravity on the Earth and the Moon
Buoyancy Formative Assessment	①	Comparison of an Object's Weight in Air and Water
	②	Comparison of Buoyant Force According to the Direction of Submersion in Water
	③	Comparison of Buoyant Force in Water According to the Area of the Bottom Surface
	④	Comparison of Buoyant Force in Water According to the Depth of Submersion
	⑤	Comparison of Buoyant Force in Water According to Object Size (External Dimensions)
	⑥	Comparison of Buoyant Force in Water According to the Submerged Volume
	⑦	Comparison of Buoyant Force in Water According to Object Volume

효과성 분석은 동형의 형성평가를 스캐폴딩 제공 전후로 두 차례 실시하여, 각 문항의 정답률 향상 정도를 비교, 분석하는 방식으로 진행하였다. 1차 형성평가는 중력과 부력에 대한 개념 이해 수준을, 2차 형성평가는 스캐폴딩 적용 후의

중력과 부력에 대한 개념 이해 수준을 측정하였다. 이를 통해 사람 스캐폴딩과 AI 스캐폴딩 간의 학습 효과 차이를 통계적으로 검증하였다.

본 연구에서는 교사 스캐폴딩과 또래 스캐폴딩을 분리된 변인으로 설정하지 않고, 실제 수업에서 학습자가 경험하는 인간 기반 스캐폴딩 전반을 포괄하는 개념으로 ‘사람 스캐폴딩’을 설정하였다. 이는 중학교 정규 과학 수업 맥락에서 교사의 개입을 완전히 배제하는 것이 현실적으로 불가능하며, 모둠 활동 중심의 문제 풀이 수업에서는 또래 간 상호작용 역시 의도적으로 차단하기 어렵다는 수업 환경의 특성을 반영한 설계적 판단이다. 이러한 설정은 생성형 AI 기반 스캐폴딩이 인간의 도움을 완전히 대체하는지를 검증하기보다, 현실적인 수업 상황 속에서 인간 기반 스캐폴딩을 보완하는 학습 지원 도구로 기능할 수 있는지를 탐색하고자 한 본 연구의 목적에 부합한다.

4. 연구 결과

본 연구에서는 생성형 AI 기반 스캐폴딩의 효과성을 분석하기 위해 형성평가 결과에 대한 통계적 검증을 실시하였다. 1차 및 2차 형성평가 결과의 차이를 비교하기 위해 대응 표본 t-검정을 사용하였으며, 통계적 유의수준은 0.05로 설정하였다.

‘중력’과 ‘부력’ 개념을 중심으로 형성평가를 사전·사후로 실시하였으며, 사람 스캐폴딩과 AI 스캐폴딩 집단의 성취 변화와 개념 이해 향상 정도를 비교, 분석하였다. 분석 결과는 Table 3과 Table 4에 제시하였다. 첫째, 중력 문제 풀이 학습에서의 연구 결과이다. 교사나 자신보다 숙련된 또래에게 스캐폴딩을 제공받은 집단과 ChatGPT 프롬프트 스캐폴딩을 제공받은 집단 모두 1차 형성평가보다 2차 형성평가에서 정답률이 전반적으로 향상되었다.

Table 3. Accuracy Rates by Scaffolding Type in the Gravity Concept Test

Type	Formative Assessment on Gravity					
	1st Formative Assessment Accuracy (①)	2nd Formative Assessment Accuracy (②)	Accuracy Improvement (②-①)	t-value	p (<.05)	
Human Scaffolding	①	68.5%	75.3%	6.8%	-1.244	0.2149
	②	39.8%	54.4%	14.6%	-1.994	0.0475
	③	35.2%	53.5%	18.3%	-2.55	0.0115
	④	36.1%	49.5%	13.4%	-2.15	0.0356
ChatGPT Prompt Scaffolding	①	74.2%	90.0%	15.8%	-3.392	0.00081
	②	46.0%	68.2%	22.2%	-3.68	0.00029
	③	50.8%	58.2%	7.4%	-1.882	0.061
	④	46.0%	55.40%	9.43%	-2.449	0.015

중력 형성평가의 각 문항별 정답 향상률을 비교한 결과,

사람 스캐폴딩이 2문항에서 더 높은 향상률을 보였으며, ChatGPT 프롬프트 스캐폴딩이 2문항에서 더 높은 향상률을 보였다. 즉, 두 스캐폴딩 방식 모두 학습 효과를 나타냈으며, 생성형 AI를 활용한 프롬프트 스캐폴딩이 중력 개념 학습에서도 사람 스캐폴딩과 유사한 수준의 학습 지원 도구로 기능할 가능성이 있음을 확인할 수 있었다.

통계 분석 결과, 사람 스캐폴딩 집단의 1번 문항에서는 1차 형성평가와 2차 형성평가 간의 평균 차이가 통계적으로 유의하지 않았다. 이는 학습자들이 해당 개념을 이미 어느 정도 이해하고 있었기 때문에, 추가적인 스캐폴딩이 유의미한 변화를 이끌지 못했을 가능성을 보여준다.

그러나 2번부터 4번 문항에서는 스캐폴딩 제공 이후 정답률이 통계적으로 유의하게 향상되었다. 특히 3번 문항의 경우 가장 큰 향상 폭을 보여, 교사 또는 유능한 또래의 개입이 학습자의 오개념 교정과 문제 이해도를 높이는 데 실질적인 도움을 주었음을 알 수 있다.

한편, ChatGPT 프롬프트 스캐폴딩을 제공받은 집단에서도 대부분의 문항에서 정답률이 향상되었으며, 특히 중력의 방향과 크기, 힘의 관계를 묻는 문항에서 유의미한 향상을 보였다. 이는 ChatGPT가 학생의 응답에 따라 즉각적인 피드백과 보충 설명을 제공함으로써, 학습자의 사고 과정을 구체화하고 개념을 재구성하도록 도왔기 때문으로 해석된다. 분석 결과, 사람 스캐폴딩 집단과 ChatGPT 프롬프트 스캐폴딩 집단 모두 사전 대비 사후 형성평가에서 정답률이 향상되었으며, 일부 문항에서는 두 집단 간 향상 폭의 차이가 통계적으로 유의하지 않은 것으로 나타났다. 이는 기존 스캐폴딩 이론에서 제시한 질문 중심 인지적 지원 전략이 생성형 AI 환경에서도 유사하게 구현될 수 있음을 시사한다. 특히 ChatGPT 스캐폴딩은 질문-응답 기반 학습 상황에서 오개념 교정과 개념 정교화에 효과적인 지원 도구로 작용할 가능성을 보여주었다.

둘째, 부력 문제 풀이 학습에서의 연구 결과에 대한 설명이다. 사람 스캐폴딩과 ChatGPT 프롬프트 스캐폴딩을 제공받은 모든 학생들은 1차 형성평가보다 2차 형성평가에서 정답률이 전반적으로 향상되었다. 사람 스캐폴딩 집단의 경우, 7개 문항 중 5개 문항에서 통계적으로 유의한 향상이 나타났다. 특히 2번, 3번, 4번 문항에서 통계적으로 유의한 차이가 확인되었다($p < .001$). 이는 교사 또는 또래의 구체적인 피드백과 설명이 학습자의 오개념을 교정하고, 부력의 원리 및 작용 방향에 대한 개념적 이해를 강화한 결과로 해석된다.

Table 4. Accuracy Rates by Scaffolding Type in the Buoyancy Concept Test

Type	Formative Assessment on Buoyancy					
	1st Formative Assessment Accuracy (①)	2nd Formative Assessment Accuracy (②)	Accuracy Improvement (②-①)	t-value	p (<.05)	
Human Scaffolding	①	94.51%	96.94%	2.43%	-0469	0.639
	②	63.74%	94.90%	31.16%	-5.726	4.06×10 ⁻⁸
	③	56.04%	89.80%	33.75%	-5.942	1.36×10 ⁻⁸
	④	38.46%	86.73%	48.27%	-7.703	7.68×10 ⁻¹³
	⑤	67.03%	91.84%	24.80%	-4.157	4.92×10 ⁻⁵
	⑥	85.71%	88.78%	3.06%	-0.385	0.701
	⑦	41.76%	80.61%	38.85%	-5.155	6.44×10 ⁻⁷
ChatGPT Prompt Scaffolding	①	93.2%	94.9%	1.8%	-0.549	0.583
	②	69.2%	89.8%	20.6%	-3.855	0.00015
	③	51.3%	83.1%	31.8%	-5.631	5.15×10 ⁻⁸
	④	35.0%	81.4%	46.3%	-8.303	8.43×10 ⁻¹⁵
	⑤	74.4%	86.4%	12.1%	-3.33	0.0010
	⑥	81.2%	89.0%	7.8%	-1.861	0.064
	⑦	47.0%	79.7%	32.7%	-5.189	4.62×10 ⁻⁷

반면 1번과 6번 문항은 통계적으로 유의하지 않았는데, 이는 두 문항의 정답률이 이미 85~95% 수준으로 높아 천장 효과(ceiling effect)가 발생했기 때문으로 보인다.

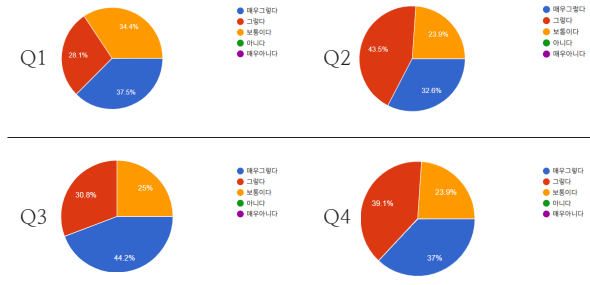
ChatGPT 프롬프트 스캐폴딩 집단 역시 대부분의 문항에서 유의한 향상을 보였다. 특히 3번, 4번, 5번 문항에서 매우 높은 통계적 유의성이 나타났다. 이는 ChatGPT가 학습자의 질문에 즉각적으로 반응하고, 부력의 방향, 크기 등을 단계적으로 설명함으로써 학습자의 사고를 구체화하고 개념적 혼란을 줄이는 데 효과적으로 작용했음을 시사한다. 다만 1번과 6번 문항에서는 유의한 차이가 확인되지 않았다. 이는 학습자가 이미 해당 개념을 충분히 이해하고 있었거나 문항 난이도가 낮았기 때문으로 해석되며, 스캐폴딩 효과를 측정하는 데 있어 측정 도구의 민감도가 낮은 영역이 있었음을 시사한다. 즉, 정답률이 이미 높거나, 문항 난이도가 낮아 ChatGPT 스캐폴딩의 효과가 통계적으로 드러나지 않은 것으로 판단된다.

양 집단의 향상률을 비교한 결과, 사람 스캐폴딩이 더 높은 상승폭을 보인 문항은 6개, ChatGPT 프롬프트 스캐폴딩이 더 높은 상승폭을 보인 문항은 1개였다. 그러나 그 차이는 평균적으로 2~3% 수준에 불과해 통계적으로 유의한 격차는 나타나지 않았다. 따라서 ChatGPT 스캐폴딩이 사람 스캐폴딩에 비해 열등하지 않으며, 실제 수업 맥락에서 학습자의 자기주도적 탐구를 지원하는 보완적 스캐폴딩 도구로 활용 가능성이 있음을 확인할 수 있다. 결과적으로, 사람 스캐폴딩 집단과 ChatGPT 프롬프트 스캐폴딩 집단 모두 사전 대비 사후 형성평가에서 정답률이 향상되었으며, 일부 문항에서는 두 집단 간 향상 폭의 차이가 통계적으로 유의하지 않은 것으로 나타났다. 더 나아가 1차 형성평가에서 과학 개념에 대한 이해가 부족하여 문제를 틀린 학생 중

ChatGPT 프롬프트 스캐폴딩을 제공받은 학생을 대상으로 추가 설문을 진행하였다. Table 5는 1차 형성평가에서 과학 개념을 충분히 이해하지 못해 문제를 틀린 학생들을 대상으로, ChatGPT 프롬프트 스캐폴딩이 이후 2차 형성평가에서 개념 이해와 문제 해결에 실제로 도움이 되었는지를 조사한 설문 결과를 나타낸 것이다.

Table 5. Students' Perceived Helpfulness of ChatGPT Prompt Scaffolding in Understanding Scientific Concepts

Survey Question:
Among the students who answered incorrectly or guessed in the first gravity formative assessment, only those who later answered correctly in the second assessment were asked:
"Did the ChatGPT question-and-answer activity help you understand scientific concepts that you previously misunderstood or could not fully grasp?"
Response Options(color):
Strongly agree(blue), Agree(red), Neutral(yellow), Disagree(green), Strongly disagree(purple)



응답 결과, 네 개의 문항 모두에서 대다수 학생이 ChatGPT와의 질의응답 활동이 과학 개념을 이해하고 문제 해결 과정을 점검하는 데 도움이 되었다고 인식하였다. '매우 그렇다'와 '그렇다'로 응답한 비율은 문항별로 60~80% 수준으로 나타났으며, '보통이다' 이상의 긍정적 응답이 전체의 약 90% 이상을 차지하였다. 이는 ChatGPT가 단순히 정답을 제공하는 도구가 아니라, 학생 스스로 개념을 점검하고 사고 과정을 구체화하도록 돕는 인지적 지원자(cognitive supporter) 역할을 수행했음을 보여준다.

특히 1번과 3번 문항에서는 각각 71.8%, 72.3%의 학생이 ChatGPT의 피드백이 도움이 되었다고 응답했으며, 2번과 4번 문항에서도 각각 67.7%, 68.1%가 긍정적인 반응을 보였다. 이는 ChatGPT가 제시하는 후속 질문이나 추가 설명이 학습자의 오류 인식과 개념 재구성에 기여했음을 시사한다. 또한 일부 학생들은 ChatGPT의 응답을 통해 자신이 잘못 이해했던 부분을 스스로 인식하고, 이를 교정하는 과정에서 문제 해결 방법을 명확히 할 수 있었다고 서술하였다. 이러한 결과는 ChatGPT 프롬프트 스캐폴딩이 학습자의 자기 주도적 탐구를 촉진하고 학습 몰입도를 향상시키는 교육적 잠재력을 지니고 있음을 의미한다.

결과적으로, 학생들은 ChatGPT 프롬프트 스캐폴딩을 학습의 보조적 도구로 인식했으며, 과학 개념을 재구성하거나 문제 해결 과정을 점검하는 과정에서 실제적인 도움을 받았다고 평가하였다. 이는 생성형 AI가 교사의 개입이 제한

된 상황에서도 학습자의 사고를 유도하고 개념적 이해를 지원할 수 있는 효과적인 디지털 스캐폴딩 도구로 기능할 수 있음을 보여준다.

5. 결론 및 제언

5.1 결론

본 연구는 생성형 인공지능인 ChatGPT를 중학교 과학 수업의 문제 풀이 학습에 적용하여, 인공지능이 학습자의 사고 과정을 지원하는 스캐폴딩(scaffolding)으로 기능할 수 있는지에 대한 가능성을 탐색하였다. 연구 결과, ChatGPT 프롬프트 스캐폴딩을 활용한 학생들은 '중력'과 '부력' 개념 형성평가 모두에서 정답률이 향상되었으며, 이러한 결과는 사람 스캐폴딩과 비교했을 때 통계적으로 유의한 차이가 없었다. 특히 중력 단원에서는 ChatGPT 스캐폴딩이 일부 문항에서 사람 스캐폴딩보다 더 높은 향상 폭을 보였고, 부력 단원에서는 두 스캐폴딩 간 유사한 수준의 효과가 확인되었다.

설문 조사에서도 60~80%의 학생들이 ChatGPT와의 질의응답 활동이 과학 개념의 이해와 문제 해결력 향상에 도움이 되었다고 응답하였다. 이러한 결과는 ChatGPT가 단순한 정보 제공을 넘어, 학습자가 스스로 개념을 점검하고 사고를 재구성할 수 있도록 돕는 인지적 촉진자(cognitive facilitator)로서의 역할을 수행할 수 있음을 보여준다. 학생들 역시 ChatGPT를 인지적 지원자(cognitive supporter)이자 개념적 촉진자(cognitive facilitator)로 인식했다고 여겨진다.

특히 교사나 또래에게 질문하기 어려워하는 학생들에게 ChatGPT는 부담 없이 의견을 표현하고 즉각적인 피드백을 받을 수 있는 환경을 제공함으로써, 학습 참여와 개념 이해를 동시에 향상시키는 긍정적인 효과를 나타냈다. 따라서 ChatGPT 프롬프트 스캐폴딩은 과학 문제 풀이 학습에서 학생들의 개별적 사고를 확장하고, 자기주도적 개념 형성을 지원하는 새로운 형태의 AI 기반 학습 지원 도구로 활용될 수 있다. 다만, 이와 같은 효과는 순수한 AI의 능력이라기보다는, 교사의 전문적인 프롬프트 설계 및 지도 전략인 '프롬프트 스캐폴딩'이 필수적으로 적용된 결과로 해석해야 한다. ChatGPT 스캐폴딩이 사람 스캐폴딩과 비교하여 열등하지 않다고 결론을 내리는 것은 타당하지만, 생성형 AI가 사람의 도움을 "부분적으로 대체(partially substitute)"할 가능성을 논하기 위해서는 사람과의 질적 차이에 대한 심층적인 분석이 보완될 필요가 있다. 즉 생성형 AI의 효율성과 정보의 정확성, 학습자의 사고력 유지 사이의 균형 잡힌 접근 및 교사의 생성형 AI 활용 역량 강화가 중요하다. 본 연구의 비교 설계는 사람 스캐폴딩과 생성형 AI 스캐폴딩의 우열을 판단하기 위한 목적보다는, 기존 스캐폴딩 이론에서 제시한 핵심 요소들이 생성형 AI 기반 학습 환경에서도 유사하게 구현될 수 있는지를 탐색하는 데 연구의 초점을 두었다. 즉,

본 연구는 생성형 AI가 교사나 또래의 도움을 완전히 대체할 수 있는지를 논하기보다는, 실제 교실 수업 맥락에서 학습자의 문제 해결 과정을 보완하는 스캐폴딩 도구로 기능할 수 있는 가능성을 검증하고자 하였다.

이러한 관점에서 ChatGPT 프롬프트 스캐폴딩은 질문 기반 상호작용을 통해 학습자의 사고를 촉진하고, 기존 연구에서 강조한 단계적 질문, 오개념 인식, 자기 설명 유도 전략을 AI 환경에서 실현한 사례로 해석될 수 있다.

5.2 제언

첫째, 생성형 인공지능을 수업에 적용할 때 발생할 수 있는 오답 생성이나 할루시네이션(hallucination) 현상에 대비해야 한다. 교사는 학습자가 인공지능의 응답을 비판적으로 검토하고, 정보의 신뢰성을 판단할 수 있도록 안내해야 하며, 이를 위해 사실 검증 능력과 비판적 사고력 함양을 위한 추가 교육이 필요하다. 이러한 AI의 신뢰성 저하 문제는 교육적 활용의 근본적인 한계로, 결과 해석 단계에서 이 부분이 충분히 비판적으로 다루어졌는지에 대한 보완이 필요하다.

둘째, 본 연구는 중학생을 대상으로 한 45분 단위의 실제 수업 맥락에서 형성평가, 오답 점검, 스캐폴딩 제공, 개념 재이해 과정을 하나의 차시 안에 포함하도록 설계되었다. 이에 따라 학습자의 인지적 부담과 수업 시간의 제약을 고려할 때, 다수의 형성평가 문항을 적용하는 데에는 현실적인 한계가 있었다. 이러한 점은 본 연구의 제한점으로 작용하며, 연구 결과를 일반화하는 데 주의가 필요하다. 향후 연구에서는 더 많은 문항 수와 다양한 과학 개념 영역을 포함한 실험 설계를 통해, 생성형 AI 기반 스캐폴딩의 효과를 장기적·확장적으로 검증할 필요가 있다. 또한 제한된 차시 내 수업 맥락에서 형성평가를 중심으로 분석이 이루어졌기 때문에, 집단 간 또는 시점 간 효과를 정량적으로 비교할 수 있는 효과크기(effect size)를 제시하지 못하였다. 향후 연구에서는 Cohen's d 또는 η^2 와 같은 효과크기 지표를 포함한 분석을 통해 생성형 AI 기반 스캐폴딩의 효과를 보다 정밀하게 검증할 필요가 있다.

셋째, 학습자의 성향과 인지적 특성에 따라 생성형 AI 스캐폴딩의 효과가 다르게 나타날 수 있으므로, 학습 수준과 참여 태도를 고려한 맞춤형 스캐폴딩 설계가 요구된다. AI의 피드백 강도, 설명 방식, 질문 유형을 개별 학습자 특성에 맞게 조정하는 것이 중요하다. 이를 위해 향후 논의에는 현재 데이터 내에서 학습 수준(예: 1차 형성평가 점수)에 따른 AI 스캐폴딩의 효과 차이를 추가적으로 분석하거나 논의에 포함하여 맞춤형 스캐폴딩 설계의 중요성을 구체화할 필요가 있다. 또한 향후 연구에서는 학습자의 사전 성취도 수준에 따른 생성형 AI 기반 스캐폴딩의 효과 차이를 보다 정밀하게 분석할 필요가 있다. 예를 들어, 상위 수준 학습자에게서는 사람 스캐폴딩과 ChatGPT 스캐폴딩 간의 차이가 크지 않게 나타나는 반면, 중·하위 수준 학습자에게서는 질문 기반 피드백과 단계적 추론을 제공하는 ChatGPT 스캐폴딩

이 보다 의미 있는 학습 지원으로 작용할 가능성이 있다. 이러한 수준별 분석은 생성형 AI 스캐폴딩이 어떤 학습자 집단에서 효과적인지를 구체적으로 밝히는 데 기여할 수 있으며, 향후 맞춤형 AI 스캐폴딩 설계를 위한 중요한 근거를 제공할 것이다.

넷째, 향후 연구에서는 본 연구의 접근 방식을 실제 학교 현장에 장기적으로 적용하여, AI 기반 스캐폴딩이 학습 성취도, 개념 유지력, 인공지능 수용 태도 등에 미치는 영향을 심층적으로 탐구할 필요가 있다. 이를 통해 생성형 AI가 학습 지속성과 개념 전이에 미치는 장기적 효과를 검증할 수 있을 것이다. 특히 본 연구는 단기적인 형성평가 결과를 비교하는 데 그쳤으므로, 향후 연구에서는 AI 기반 스캐폴딩이 학습 성취도, 개념 유지력, 개념 전이에 미치는 장기적인 효과를 탐구할 필요가 있다.

마지막으로, 교육 현장에서 생성형 AI를 성공적으로 도입하기 위해서는 교사의 AI 활용 역량 강화와 전문성 개발이 필수적이다. 교사들이 AI의 작동 원리를 이해하고 수업 맥락에 적합하게 통합할 수 있도록, 실제 수업 사례 중심의 연수 프로그램과 교육 자료가 체계적으로 개발, 보급되어야 한다.

참고문헌

- [1] Cho, Y., Lee, J., Lim, G., Jeong, H., & Han, I. (2023). Future education with generative AI: From machine to collaborative partner. *Journal of Educational Technology*, 39(4), 1449-1478. DOI: <http://dx.doi.org/10.17232/KSET.39.4.1449>
- [2] Kim, S. (2023). Development of a programming education code generation prompt using generative AI. *Journal of the Korean Association of Computer Education*, 26(5), 107-117. DOI: <http://dx.doi.org/10.32431/kace.2023.26.5.009>
- [3] Kim, J., Kang, D., & Ko, Y. (2023). A Study on Educative Utilization of Generative AI - Focusing on ChatGPT Utilization. *Journal of the Korean Association of Information Education*, 27(6), 691-704. DOI: <http://dx.doi.org/10.14352/jkaie.2023.27.6.691>
- [4] Gui, J., Sun, Z., Wen, Y., Tao, D., & Ye, J. (2021). A review on generative adversarial networks: Algorithms, theory, and applications. *IEEE Transactions on Knowledge and Data Engineering*. DOI: <https://doi.org/10.1109/tkde.2021.3130191>
- [5] Jo, E., & Lee, S. (2023). A study on elementary students' use patterns and attitudes toward generative artificial intelligence in AI-integrated classes. *Journal of the Korean Association of Computer Education*, 27(4), 21-39. DOI: <http://dx.doi.org/10.32431/kace.2024.27.4.003>
- [6] Jang, J. (2024). The effect of ChatGPT prompt scaffolding on argumentative writing skills. Master's thesis. Seoul National University. url:<https://hdl.handle.net/10371/209648>
- [7] Park, E. (2024). The impact of learners' question types using generative artificial intelligence on problem-solving

- outcomes. Master's thesis. Ewha Womans University.
- [8] Baidoo-Anu, D., & Owusu Ansah, L. (2023). Education in the era of generative artificial intelligence (AI): Understanding the potential benefits of ChatGPT in promoting teaching and learning. Available at SSRN 4337484. DOI: <https://doi.org/10.61969/jai.1337500>
- [9] Adamson, V., & Bagerfeldt, J. (2023). Assessing the effectiveness of ChatGPT in generating Python code. DOI: <http://urn.kb.se/resolve?urn=urn:nbn:se:his:diva-22860>
- [10] Liu, C., Bao, X., Zhang, H., Zhang, N., Hu, H., Zhang, X., & Yan, M. (2023). Improving ChatGPT Prompt for Code Generation. DOI: <http://arxiv.org/abs/2305.08360>
- [11] Cho, H. (2023). Understanding of Generative Artificial Intelligence Based on Textual Data and Discussion for Its Application in Science Education. *Journal of the Korean Association for Science Education*. 43(3). 307-319 DOI: <http://dx.doi.org/10.14697/jkase.2023.43.3.307>
- [12] Kim, J. & Yoo, H. (2023). Exploring applications of ChatGPT for Physics Education: Focusing on High School and General Physics Class. *School Science Journal*. 17(3). 216-239 DOI : <http://dx.doi.org/10.15737/ssj.17.3.202308.216>
- [13] Rojas-Drummond, S. et al. (2013). 'Dialogic scaffolding': Enhancing learning and understanding in collaborative contexts. *Learning, Culture and Social Interaction*. 2(1), 11-21. DOI: <https://doi.org/10.1016/j.lcsi.2012.12.003>
- [14] Choi, S. & Kim, G. (2025). Development of a generative AI-based online platform prototype using a scaffolding strategy. *Proceedings of the Winter Conference of the Korea Computer Information Society*, 33(1).



김미나

- 2008년 제주대학교 과학교육학과 (교육학석사)
- 2024년 제주대학교 교육대학원 인공지능융합교육 전공(교육학석사)
- 2026년 제주대학교 일반대학원 과학교육학부 컴퓨터교육과 박사과정 수료
- 2015년~현재 제주특별자치도교육청 중등 과학 교사

⊕ 관심분야 : 인공지능 교육, 과학 교육, 융합 교육
 ✉ kimna1214@naver.com



김성백

- 1989년 서울대학교 컴퓨터공학과(공학사)
- 1991년 서울대학교 컴퓨터공학과(공학석사)
- 1995년 서울대학교 컴퓨터공학과(공학박사)
- 2020년~현재 제주대학교 교육대학원 인공지능 융합교육전공 교수
- 1996년~현재 제주대학교 컴퓨터교육과 교수

⊕ 관심분야 : 컴퓨터시스템, SWAI 교육, 원격 교육, 컴퓨터교육
 ✉ sbkim@jejunu.ac.kr