

컴퓨터교육학회 논문지 2026년 제29권 제8호
<https://doi.org/10.32431/kace.2026.29.8.002>



국내 SW 교육이 K—12 학습자의 컴퓨팅 사고력에 미치는 효과 : 메타분석적 접근을 중심으로*

The Effects of Software Education on K-12 Learners' Computational Thinking: A Meta-Analytic Approach in Korea

백승애[†] · 김미량^{††} 
 Seungae Baek[†] · Mi Ryang Kim^{††}

요약

본 논문은 2015-2025년 국내 K-12 소프트웨어(SW) 교육이 컴퓨팅 사고력(CT)에 미치는 효과를 종합적으로 검증하기 위해 111편의 논문에서 추출한 356개의 효과크기를 대상으로 메타분석을 수행하였다. PRISMA 2020 기반 문헌 선별과 무선희과모형과 하위 집단분석을 통해 전체 및 조절효과를 추정한 결과, 전체 효과크기는 $\hat{\Delta} = 0.689$ 로 중간 이상의 CT 향상 효과가 나타났다. AI-assisted 활용 환경($\hat{\Delta} = 0.543$)이 미활용 환경($\hat{\Delta} = 0.740$)보다 낮은 효과를 보인 것은 인지적 부하 가중과 AI 자동 생성 기능이 학습자의 문제 해결 과정에 영향을 미쳤을 가능성을 시사하며, AI-assisted 활용 기제와 학습 성취 간 인과관계를 규명하는 후속 연구가 필요하다. 본 연구는 K-12 학습자로 범위를 한정하고 최신 교육 변인을 투입함으로써 국내 SW 교육의 효과성을 실증하였다.

주제어 컴퓨팅 사고력, 소프트웨어 교육, K-12, 메타분석, AI-assisted 활용, 교육정책

ABSTRACT

This study conducted a meta-analysis on 356 effect sizes extracted from 111 studies to comprehensively examine the effects of K-12 software (SW) education on computational thinking (CT) in South Korea from 2015 to 2025. Following PRISMA 2020 guidelines for literature screening, overall and moderating effects were estimated using a random-effects model and subgroup analyses. The results revealed an overall effect size of $\hat{\Delta} = 0.689$, indicating a moderate-to-strong improvement in CT. The finding that AI-assisted learning environments ($\hat{\Delta} = 0.543$) yielded lower effects than non-AI environments ($\hat{\Delta} = 0.740$) suggests that increased cognitive load and AI-generated automation may have interfered with learners' problem-solving processes, highlighting the need for future research to investigate the causal relationships between AI-assisted instructional mechanisms and learning outcomes. By restricting the scope to K-12 learners and incorporating contemporary educational variables, this study provides empirical evidence for the effectiveness of SW education in South Korea.

Keywords Computational thinking, Software education, K-12, Meta-analysis, AI-assisted learning, Education policy

†정회원 성균관대학교 대학원 컴퓨터교육학과 석사
 ††중신회원 성균관대학교 사범대학 컴퓨터교육과 교수(교신저자)
 논문투고 2026년 02월 20일
 심사완료 2026년 06월 17일
 게재확정 2026년 06월 30일
 발행일자 2026년 08월 04일

* 본 논문은 제1저자의 성균관대학교 일반대학원 석사학위논문 일부를 발췌하여 요약, 정리한 것임.

1. 서론

1.1 연구의 필요성

21세기 사회는 인공지능(AI)과 디지털 전환(Digital Transformation)으로 대표되는 급격한 기술 변화 속에서 문제 해결력과 창의적 사고를 갖춘 인재 양성을 요구하고 있다. 이러한 시대적 흐름 속에서 컴퓨팅 사고력(Computational Thinking, CT)은 디지털 리터러시와 함께 핵심 역량으로 부상하고 있으며[1, 2], 국내에서도 2015 개정 교육과정을 통해 초·중등학교 대상 필수 정보 교육이 본격적으로 시행되었다[3]. 이어 2022 개정 교육과정에서는 AI 및 데이터 활용 역량이 정보 교과와 핵심 역량으로 명시되며, 디지털 환경 변화에 주체적으로 대응할 수 있는 능력이 더욱 강조되고 있다[3]. 이러한 변화에 따라 K-12 교육현장에서는 단순한 프로그래밍 교육을 넘어 창의적 문제 해결 중심의 컴퓨팅 사고력 교육이 빠르게 확산되고 있다[4, 5].

또한 교육 현장에서는 컴퓨팅 사고력을 기반으로 한 ‘인공지능 사고력(AI Thinking)’ 교육이 강조되고 있으며[6], 실생활 문제를 해결하는 AI 메이커 교육 등 고차원적인 SW·AI 융합 교육이 활발히 시도되고 있다[7]. 최근 생성형 AI의 등장은 SW 교육 패러다임에 근본적인 변화를 불러일으키고 있다. 생성형 AI는 텍스트와 코드 등 복잡한 산출물을 생성하며 학습 성취도에 긍정적인 영향을 미치는 것으로 보고되고 있으나[8], 동시에 알고리즘 설계와 같은 핵심적인 사고 과정을 자동화함으로써 학습자의 주도적인 CT 함양을 저해할 수 있다는 우려도 공존한다. 따라서 AI-assisted 활용 교육이 학습자의 컴퓨팅 사고력에 대한 메타분석적 검증은 부족한 실정이다.

SW 교육의 컴퓨팅 사고력 향상 효과를 검증하기 위한 실증연구는 국내외에서 꾸준히 누적되어 왔다. 국내 1세대 메타분석[9]은 2006-2017년 연구를 분석하여 SW 교육이 CT 향상에 중간 수준의 효과($\Delta=0.314$)를 보인다고 보고하였다. 그러나 이 연구는 ① 도입기 중심의 시기(2006-2017)에 국한되어 최근 10년간 학습 환경 변화를 반영하지 못했으며, ② 초등-중등-대학을 모두 포함한 분석으로 인해 K-12 발달단계에 따른 효과를 명확히 파악하기 어려웠다. 또한 ③ 생성형 AI 활용 교수 등 2018년 이후 등장한 새로운 교육 방식을 조절 변수로 포함하지 못한 한계가 있다.

한편 국내 연구 동향을 살펴보면, 컴퓨팅 사고력 교육은 실행 방안과 프로그램 개발 중심의 연구가 주를 이루고 있으며[10] CT 평가 도구 개발 및 체계적인 메타분석 연구는 여전히 제한적인 것으로 나타났다[11]. 이는 CT 개념의 명확한 조작적 정의와 표준화된 평가 체계 구축이 필요함을 시사한다[12].

따라서 현 시점에서 K-12 수준의 SW 교육이 컴퓨팅 사고력 향상에 미치는 효과를 다시 검증하는 연구는 필수적이다. K-12는 구체적 조작기에서 형식적 조작기로 이행하는 발달적 전환기[13]로, 대학생 학습자와는 학습 동기, 선행 지식, 자기조절 등에서 본질적으로 차이가 존재한다는 연구

결과가 보고되고 있다[14]. 또한 국가 교육과정이 적용되는 범위이기 때문에, K-12 대상 CT 교육 효과를 분석하는 연구는 교육정책 수립 및 교육과정 개선에 직접적으로 기여할 수 있다. 국제적 연구 동향 또한 최신 CT 메타분석이 대부분 K-12에 집중되고 있다는 점에서[15, 16], 본 연구의 분석 범위는 학문적·정책적 타당성을 가진다.

1.2 연구의 목적

본 연구는 2015-2025년 사이 국내에서 수행된 K-12 소프트웨어 교육 관련 연구들을 대상으로 컴퓨팅 사고력 향상에 관한 효과를 메타분석 방법으로 종합적으로 검토하고자 한다. 구체적인 연구 목적은 다음과 같다. 첫째, 해당 기간 국내 K-12 SW 교육이 학습자의 컴퓨팅 사고력 향상에 미치는 전체 효과크기를 산출한다. 2015년 정보 교육 필수화 이후 관련 연구가 꾸준히 축적되어 왔음에도 불구하고, 최근 10년간의 효과를 통합적으로 분석한 연구는 충분하지 않은 실정이다. 이에 개별 연구들을 체계적으로 종합함으로써 국내 SW 교육의 전반적인 교육 성과를 파악하고자 한다.

둘째, 교수·학습 방식, 프로그래밍 활동 유형, AI-assisted 활용 여부 등 최신 교육 변인을 조절 변수로 포함하여 SW 교육의 조건적 효과 차이를 보다 정교하게 검증하고자 한다. 특히 AI-assisted 활용 여부를 조절 변인으로 설정한 이유는 신기술 도입이 CT 형성에 미치는 실질적 성과와 한계를 비판적으로 검토함으로써, 향후 SW 교육과정 설계 및 교수학습 모델 개발을 위한 기초 자료로 활용하고자 한다.

셋째, SW 교육 도입기를 중심으로 수행된 1세대 메타분석[9]과의 비교를 통해 지난 10년간 SW 교육 효과의 변화 양상을 살펴보고자 한다. 기존 메타분석에서는 분석 대상에 대학생이 포함되어 K-12 학습자만의 효과를 세밀히 구분하는 데 한계가 있었다. 본 연구는 이러한 점을 보완하여 K-12에 한정된 분석을 수행함으로써, 학교급별 발달 단계를 고려한 위계적 교육과정 구성 및 관련 정책 개선을 위한 실증적 근거를 제시하고자 한다. 이를 통해 급변하는 교육 환경 속에서 근거에 기반한 교육적 시사점을 도출하는 것을 목표로 한다.

1.3 연구 문제

[연구 문제 1] 2015-2025년 국내 K-12 SW 교육이 학습자의 컴퓨팅 사고력 향상에 미치는 전체 효과크기는 어떠한가?

[연구 문제 2] 컴퓨팅 사고력 향상 효과는 교육 수준, 교수·학습 방식, 프로그래밍 유형, AI-assisted 활용 여부 등 조절 변수에 따라 차이를 보이는가?

[연구문제 3] 기존 1세대 메타분석(2006-2017)과 비교할 때, SW 교육 효과는 시기적 변화를 보이는가?

2. 이론적 배경

2.1 컴퓨팅 사고력

컴퓨팅 사고력(Computational Thinking, CT)은 Wing(2006)이 “모든 학습자가 갖추어야 할 기초 소양(a fundamental skill for everyone)”으로 제시하면서[1], 학계에 본격적으로 논의되기 시작하였다. Wing은 컴퓨팅 사고력을 컴퓨터 과학의 근본 개념을 활용해 문제를 이해하고 해결하는 인지적 과정으로 정의하며, 읽기·쓰기·셈하기와 같이 학교 교육의 필수 소양으로 포함되어야 한다고 강조하였다[1, 17, 18]. 이러한 관점은 컴퓨팅 사고가 단순한 프로그래밍 능력 이상의 폭넓은 문제해결 체계를 포함한다는 점을 시사한다[19].

그러나 컴퓨팅 사고력의 국제적 중요성에도 불구하고 합의된 정의는 아직 존재하지 않으며, 국가·학문 분야마다 다양한 해석이 공존한다[20]. 일부 교육 현장에서는 컴퓨팅 사고력이 코딩 또는 로봇 활동과 동일한 개념으로 혼용되는 사례도 나타난다[20]. 최근 문헌에서는 컴퓨팅 사고력을 복잡한 문제를 체계적으로 해결하기 위해 컴퓨터 과학 개념·알고리즘 원리·컴퓨팅 시스템을 적용하는 절차적 사고 능력으로 정의하는 경향이 확산되고 있다[21].

Brennan & Resnick(2012)은 컴퓨팅 사고력을 개념(Concepts), 실행(Practices), 관점(Perspectives)의 세 차원으로 구조화하였다[22]. 개념 영역은 반복, 조건, 데이터 표현 등 프로그래밍 수행을 위한 기초 원리를 의미하며, 실행 영역은 디버깅, 코드 재사용, 시스템 설계와 같은 실천적 수행 과정을 포함한다[22]. 관점 영역은 기술을 활용해 자신과 세계를 이해하고 표현하는 태도를 의미하며 협업·창의성·기술적 관점 형성을 포함한다[22].

국제 기준에서도 유사한 흐름이 나타난다. ISTE(2016)는 컴퓨팅 사고력을 알고리즘적 사고와 자동화를 중심으로 정의하였으며[23], NRC(2012)는 수학·과학 학습의 핵심 실행으로 컴퓨팅 사고력을 제시하였다[24]. 국내 교육과정에서도 CT는 “컴퓨터 과학의 개념·원리·컴퓨팅 시스템을 활용해 실생활 문제를 창의적으로 해결하는 능력”으로 규정되어 왔다[25]. 특히 추상화(abstraction)와 모델링 과정이 핵심 요소로 강조되지만, 각 학문 분야에서 추상화 개념이 혼재되어 있어 명확한 개념 정립이 필요하다는 지적이 이어지고 있다[12].

컴퓨팅 사고력은 기능적·인지적·정의적 관점에서 다층적으로 해석될 수 있으며, 초등 SW 교육 현장에서는 이러한 관점에 기반한 통합적 역량 함양이 강조되고 있다[26]. 이는 컴퓨팅 사고가 단순한 프로그래밍 기술 습득이 아니라 문제 해결 과정 전반을 구성하는 사고 체계임을 반영한다.

2.2 컴퓨팅 사고력의 하위 영역 및 조작적 정의

본 연구는 컴퓨팅 사고력(Computational Thinking, CT)의 구성요소를 메타분석의 효과크기 추출 기준으로 활용하고자, 주요 선행 프레임워크를 비교·분석하여 분석틀을

재구성하였다. Wing(2006), Brennan과 Resnick(2012), ISTE(2016) 등의 선행연구에서는 CT를 문제 해결을 위한 사고 과정으로 개념화하고, 추상화, 알고리즘적 사고, 분해, 패턴 인식, 평가 등을 하위 구성요소로 제시하고 있다. 그러나 연구자에 따라 구성요소의 범위와 명칭이 상이하여, 메타분석에서 일관된 분류 기준을 적용하는 데 어려움이 따른다.

이를 해결하기 위해 본 연구에서는 CSTA·ISTE(2011), 교육부(2015), Selby와 Woollard(2013), Computing At School(2011) 등 주요 프레임워크를 종합적으로 검토하였다(Table 1)[27]. 비교 분석 결과, 추상화(Abstraction), 문제분해(Decomposition), 알고리즘적 사고(Algorithmic Thinking)는 모든 프레임워크에서 공통적으로 확인된 핵심 요소로 나타났다. 반면, 일반화(Generalization)와 평가(Evaluation)는 일부 프레임워크에만 포함되어 있었으며, 코딩·시뮬레이션·디버깅 등 실행 중심의 활동은 독립적인 요소로 구분되기보다 알고리즘적 사고 또는 적용 과정에 통합되는 경향이 확인되었다. 분석틀 구성 시에는 주요 프레임워크 간 공통 요소를 우선적으로 반영하였으며, 측정 가능성과 요소 간 독립성을 주요 선정 기준으로 삼았고 Table 2와 같이 설정하였다.

이러한 원칙에 따라 인지적 영역은 추상화, 문제분해, 알고리즘적 사고, 자동화의 네 가지 요소로 구성하였다. 자동화(Automation)는 기존 프레임워크에서 독립적인 구성요소로 명시되지 않는 경우가 있으나, 실제 교육 연구에서 코딩·시뮬레이션·디버깅 등 구현 활동이 주요 측정 변인으로 빈번히 활용되고 있다는 점을 고려하여 별도의 요소로 포함하였다. 일반화와 평가는 알고리즘 설계 및 문제 해결 과정 전반에 걸쳐 통합적으로 나타나는 특성을 지니므로, 독립적인 분류 범주로 구분하지 않았다. 또한, 컴퓨팅 사고력을 하위 요소로 세분화하지 않고 통합적인 개념으로만 다룬 경우에는 ‘Comb’로 표기하였다.

Table 1. Comparison of Computational Thinking Components Across Key Educational Frameworks

Construct	ISTE & CSTA (2011)	MOE, South Korea (2015)	Selby & Woollard (2013)	CAS (2011)
Data Collection	O			
Data Analysis	O			
Representation	O	O		
Decomposition	O	O	O	O
Abstraction	O	O	O	O
Algorithmic Thinking	O	O	O	O
Modeling	O	O		
Automation	O	O		
Parallelization	O			
Simulation	O			
Generalization		O	O	O
Evaluation			O	O

Table 2. Operational Definitions of Cognitive Components of Computational Thinking (CT)

Construct	Example
Abstraction (Ab)	Identifying key features, simplifying complex problems
Problem Decomposition (PD)	Breaking tasks into subtasks, analyzing sub-problems
Algorithmic Thinking (AT)	Designing algorithm flow, constructing solution procedures
Automation (Au)	Coding, simulation, debugging

Table 3. Operational Definitions of Affective Components of Computational Thinking

Construct	Example
Attitude (ATD)	Positive attitude toward programming, valuing coding activities
Motivation (MTV)	Task engagement, sustained learning effort
Interest (INT)	Enjoyment of coding tasks, curiosity about computing
Self-efficacy (SEF)	Confidence in problem-solving, belief in coding ability
Satisfaction (SAT)	Satisfaction with learning results, positive classroom experience

본 연구에서는 학습 성과를 인지적 영역과 정의적 영역으로 구분하였다. 인지적 영역은 Table 2와 같이 추상화, 알고리즘적 사고, 자동화, 문제분해 및 자료조직 등의 하위 요소로 구성되며, 선행 연구를 바탕으로 이를 코드화하여 각각 1~4의 값으로 분류하였다. 정의적 영역은 Table 3와 같이 학습자의 태도, 동기, 흥미, 자기효능감, 만족감 등을 포함하는 정서적 특성으로 정의하였으며, 이를 바탕으로 각 변인을 1~5으로 세분화하여 코딩하였다. 이러한 구분 기준은 메타분석을 위한 효과크기 추출 및 하위집단 분석에 활용되었다.

3. 연구 방법

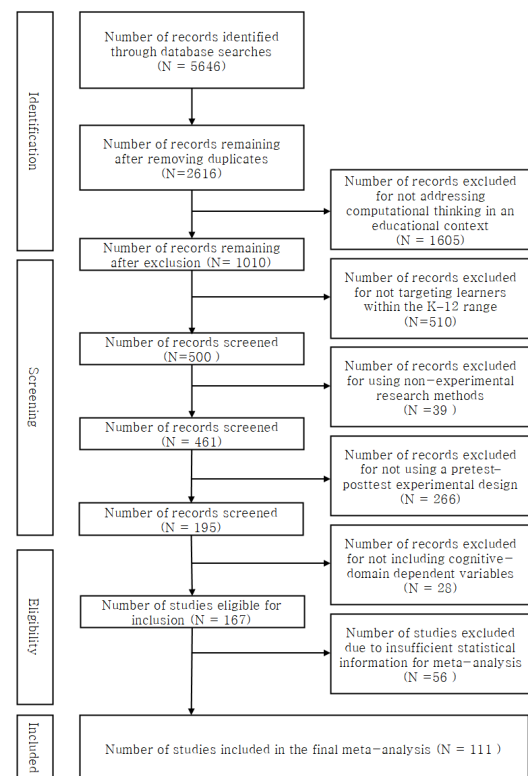
3.1 연구 대상 및 자료 수집

본 연구는 2015 개정 교육과정 이후 국내 K-12 SW교육이 컴퓨팅 사고력(CT)에 미치는 효과를 종합적으로 파악하기 위해, 2015년 1월부터 2025년 10월까지 발표된 국내 학술지 논문 및 학위논문을 대상으로 메타분석을 수행하였다. 자료 검색은 RISS, KCI, 교보문고 스콜라, KISS, DBpia 등 5개 학술 데이터베이스를 활용하였으며 “컴퓨팅 사고력”, “SW 교육” 키워드를 조합하여 실시하였다.

검색 및 선별 절차는 PRISMA 2020 가이드라인을 따랐으며, 중복 제거와 주제 및 연구 설계 적합성, 통계 정보 보고 여부 등 단계적으로 검토하여 최종 분석 대상을 선정하였다. 그 결과, 총 5646편의 문헌 중 111편의 연구가 포함되었고, 이들로부터 356개의 효과크기가 산출되었다.

문헌 선정 과정의 신뢰도와 객관성을 확보하기 위해 3단계 검증 절차를 수행하였다. 먼저, 연구자가 PRISMA 체크리스트를 기준으로 1차 분류를 실시하고, 포함·제외 기준 적용이 모호한 사례에 대해서는 교육공학 전공 박사의 자문을 받아 최종 판단하였다.

이와 병행하여 연구자의 주관적 편향을 최소화하고 대규모 검색 결과를 일관성 있게 검토하기 위해 ChatGPT 기반 AI도구를 활용한 교차 검증을 실시하였다. 본 절차는 PRISMA의 문헌 선별 절차에 맞추어 동일한 포함·제외 기준을 프롬프트(Prompt)로 구조화하여 적용하였다. 구체적으로 해당 프롬프트에는 AI에게 시스템적 문헌고찰 연구자로서의 역할을 부여하고, 연구 대상, 중재, 결과 변수 등 명확한 포함·제외 기준을 제시하여 각 문헌의 초록을 ‘포함’, ‘제외’, ‘보류’로 분류하도록 지시하는 내용이 포함되었다. 대량의 문헌 초록을 피로도 없이 동일한 기준으로 빠르게 스크리닝을 하는 데 그 활용 범위를 두었다. 다만, AI가 실제 존재하지 않는 정보를 생성하는 ‘환각(Hallucination)’ 현상이나 교육적 맥락의 미묘한 차이를 간과할 수 있는 맥락 해석의 한계를 인지하여, AI의 분석 결과를 최종 결정 근거로 단독 사용하지 않았다. AI는 연구자와의 분류 일치 여부를 확인하는 보조적 병행 검증 도구로만 활용되었으며, 불일치하거나 판단의 모호성이 남는 사례에 대해서는 연구자와 전문가가 해당 원문을 전수 재검토하고 심층 논의를 거쳐 최종 포함 여부를 결정함으로써 분석 데이터의 정확성을 담보하였다.

**Figure 1.** PRISMA Table

3.2 자료 코딩 및 변수 구성

최종 선정된 111편의 논문에 대해 연도, 출판 유형, 표본 크기, 학습 기간, CT 하위 영역, 통계치 등 분석에 필요한 정보를 코딩하였다. 종속변수인 컴퓨팅 사고력은 연구자가 설정한 조작적 정의에 근거하여 인지적 영역(추상화, 문제분해, 알고리즘화, 자동화, 통합 CT)과 정의적 영역(태도, 동기, 흥미, 자기효능감, 만족감)으로 분류하였다.

조절변수는 학습자 관련 변수와 교수·학습 관련 변수로 나누어 코딩하였다. 학습자 관련 변수로는 교육 수준(초등 저학년, 초등 고학년, 중학교, 고등학교), 영재 여부를 포함하였으며, 교수·학습 관련 변수로는 프로그래밍 언어 활용 유형(블록코딩, 언플러그드 활동, 명령어 기반, 피지컬 컴퓨팅, 혼합), 협력학습 여부(협력학습 및 개인학습), AI-assisted 활용 여부(AI-assisted 활용 및 AI-assisted 미활용)를 설정하였다. 이때, AI-assisted 여부를 도구의 명칭이 아닌 기능적 특성을 기준으로 정의하였다. 포함 사례로는 학습자가 챗봇(예: ChatGPT)과 직접 상호작용하며 코드를 디버깅하거나, 지능형 튜터링 시스템(ITS)으로부터 개별화된 오답 피드백을 받은 경우, 혹은 기계학습 모델(예: Teachable Machine)을 활용해 자신의 데이터로 모델을 직접 훈련하고 결과를 확인하는 연구 등이 해당된다. 반면 제외 사례로는, 문헌 제목에 'AI'가 포함되어 있더라도 교사가 수업 설계 목적으로만 AI를 활용하고 학습자가 직접 다루지 않은 경우, 또는 정해진 규칙에 따라 구문 오류만을 지적하는 단순 정적 컴파일러나 비적응형 평가 도구를 사용한 연구를 제외하였다. 분류가 모호한 경계 사례로 'AI 기반 코딩 플랫폼'을 활용했다고 명시되어 있으나 학습자의 구체적인 상호작용 수준이 기술되지 않은 연구에는 '해당 도구가 학습자의 알고리즘 설계 및 문제 해결 과정에서 인지적이고 적응적인 피드백을 실질적으로 제공하였는가'를 핵심 판단 기준으로 삼았다. 이러한 사례들은 연구자 단독으로 판단하지 않고, 교육공학 전공 박사 자문자와 함께 원문을 심층적으로 재검토한 후 포함 여부를 최종 결정함으로써 코딩의 신뢰성을 높이고자 하였다. 즉, 학습자의 입력에 대해 인공지능 모델이 예측, 분류, 생성 또는 적응적 피드백을 제공하는 경우를 AI-assisted로 분류하였으며, 단순 정보 제공이나 비적응적 기능만을 수행하는 디지털 도구는 포함하지 않았다. 또한 AI-assisted의 적용 범위를 명확히 하기 위해, 생성형 AI, 지능형 튜터링 시스템, 그리고 코드 추천 및 오류 수정 기능을 포함한 코딩 보조 도구 등 다양한 유형을 포괄하는 것으로 기준을 설정하였다.

코딩은 교육공학 석사과정 1인이 수행하였다.

3.3 효과크기 산출

본 연구의 효과크기 산출 및 분산 계산은 나청수 외(2018)의 메타분석 연구 절차를 따랐다. 사전-사후 설계를 사용한 연구의 경우, 사전 및 사후 검사 데이터를 활용하여 표준화된 평균 변화량(standardized mean-change

difference)을 기준으로 Becker의 효과크기를 산출 방식으로 계산하였다[28]. 개별 연구에서 사전-사후 점수 간 상관관계수가 제시되지 않은 경우, 선행연구의 기준 및 메타분석의 보수적 추정 원칙을 준용하여 실험집단의 상관관계수는 $r=.50$, 통제집단의 상관관계수는 $r=.60$ 으로 설정하였다[9]. 이는 단일집단 사전-사후 설계에서 변화량의 분산을 추정할 때 효과크기가 과대 추정되는 것을 방지하기 위한 설정이며, 1세대 메타분석 연구[9]에서 채택한 기준과 동일하게 설정함으로써 연구 간 비교의 일관성을 확보하였다.

실험집단-통제집단 설계의 경우, 각 집단의 사전-사후 차이를 표준화한 값을 Hedges g 로 보정하여 효과크기를 산출하였고, 두 집단 간 차이를 통해 표준화된 평균 변화차(Δ)를 구하였다. 표본 크기에 따른 편향을 줄이기 위해 Hedges와 Olkin(1985)의 가중치를 적용하였다.

일부 연구에서는 동일 논문에서 여러 효과크기가 보고되어 독립성 가정 위배 가능성이 존재하므로, 본 연구는 Cooper(1998)의 '분석단위의 전환(shifting unit of analysis)' 방법을 적용하였다. 전체 효과크기 산출 시에는 연구(study) 단위를 기준으로, 하위집단분석에서는 개별 효과크기를 분석 단위로 사용하여 정보 손실을 최소화하였다.

3.4 분석 절차 및 통계 모형

포함된 연구들 간 수업 맥락, 학습자 수준, 교수·학습 방식, 평가 도구 등이 상이하다는 점을 고려하여, 본 연구는 무선효과모형(random-effect model)을 적용하였다. 고정효과모형을 가정하여 실시한 동질성 검정 결과는 Table 4과 같이 나타났다. 중간 이상의 이질성이 확인되었으며, 이에 따라 무선효과모형을 사용하여 전체 및 하위집단 효과를 추정하였다.

Table 4. Test of Homogeneity Using the Fixed-Effects Model

Q	I^2	r^2	p	df	ES	95% CI	
						LB	UB
29.69	65.4	0.176	.002	11	1.11	0.815	1.404

메타분석 절차는 다음 세 단계로 구성하였다. 첫째, 전체 CT 및 인지적·정의적 하위 영역에 대한 평균 효과크기를 산출하였다. 둘째, 효과크기의 변동을 설명하기 위해 교육 수준, 영재 여부, 프로그래밍 유형, 협력학습 여부, 측정 방식, AI-assisted 활용 여부 등을 조절 변수로 한 하위집단분석(subgroup analysis)을 실시하고, Qb 검정을 통해 통계적 유의성을 확인하였다. 셋째, 분석 결과의 안정성을 확보하기 위해 출판편향 검증 및 민감도 분석을 수행하였다. 특히 통제집단 변화값의 대표치()를 활용하여 추정한 효과크기의 타당성을 검토하기 위해 추가적인 민감도 분석을 병행하였다.

3.5 출판편향 검증

출판편향 검증을 위해 funnel plot을 통한 시각적 검토, Egger's 회귀분석, Duval과 Tweedie(2000)의 trim-and-fill 방법을 사용하였다. Egger's 회귀분석에서 절편의 p 값이 0.222로 유의하지 않았으며, trim-and-fill 결과 추가로 보정된 연구 수는 0편, 보정 전후 전체 효과크기($\Delta = 0.689$)와 95% 신뢰구간이 동일하게 유지되어 출판편향 가능성은 크지 않은 것으로 판단하였다.

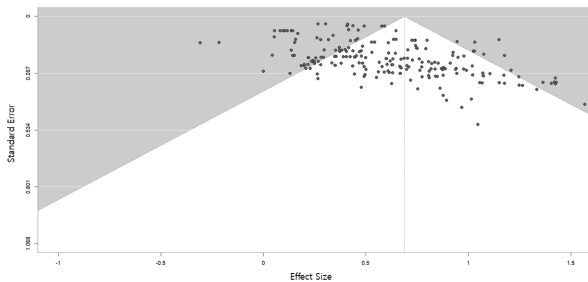


Figure 2. Trim-and-Fill funnel plot

Table 5. Results of Egger's regression test for publication bias

	Coef.	SE	t	p	95% CI	
					LB	UB
β	3.685	0.225	16.391	<.001	3.244	4.126
α	-0.059	0.048	-1.220	0.222	-0.154	0.036

Table 6. Results of trim-and-fill analysis for publication bias

	k (add)	Δ	Q	p	95% CI	
					LB	UB
Observed	0	0.689	1549.209	<.001	0.631	0.746
Adjusted	0	0.689	1549.209	<.001	0.631	0.746

*k=Number of effect sizes included in the meta-analysis, Δ = Estimated difference in effect sizes between groups

3.6 효과크기의 계산과 해석

모든 분석은 R 4.5.1 환경에서 수행되었으며, metafor(4.8-0) 패키지와 dplyr, tidyr, purrr 등을 활용하였다. 효과크기의 해석은 Cohen(1988)의 기준(0.20: 작은 효과, 0.50: 중간 효과, 0.80 이상: 큰 효과)에 따라 이루어졌다[29].

4. 연구 결과

4.1 전체효과크기

무선효과모형을 적용하여 전체 효과크기를 산출한 결과, Table 7과 같이 SW 교육이 K-12 학습자의 컴퓨팅 사고력에 미치는 평균 효과크기는 $\Delta = 0.689$ (95% CI [0.631, 0.746])로 나타났다. Cohen(1988)의 기준에 따르면 이는 중간 이상 수준의 효과에 해당하며[29], 국내 SW 교육이 학

습자의 컴퓨팅 사고력 신장에 중간 이상 수준의 기여를 하고 있음을 확인할 수 있다. 이는 1세대 메타분석이 보고했던 효과크기($\Delta = 0.314$) 기준과 대비해 볼 때, 최근 K-12 연구에서 컴퓨팅 사고력 향상 효과가 지속적으로 유지 및 강화되는 경향을 띠고 있음을 보여준다.

고정효과모형 기반 동질성 검정 결과, $Q(11)=29.69$, $p=0.002$, 이질성 지수 $I^2=65.4\%$, $\tau^2 = 0.176$ 으로 나타나. 연구 간 변동성이 존재하는 것으로 확인되었다. 이에 따라 본 연구는 무선효과모형에서 하위집단 분석을 추가로 수행하였다.

이러한 효과 향상은 다음과 같은 변화에 기인한 것으로 해석된다. 첫째, 교육과정의 안착과 표준화이다. 1세대 연구 시기(2006-2017)가 SW 교육의 도입 및 탐색기였다면, 본 연구의 분석 기간(2015-2025)은 개정 교육과정을 통해 정보교육이 공교육 필수 역량으로 자리 잡은 시기에 해당한다. 국가 수준의 교육과정 가이드라인이 보급되면서 교수 내용의 표준화와 상향 평준화가 이루어진 결과로 예측할 수 있다. 둘째, 연구 간 변동성의 수용이다. 동질성 검정 결과 확인된 높은 이질성($I^2=65.4\%$)은 국내 SW 교육이 다양한 학교급과 교수 환경에서 고유한 변동성을 지니며 시행되고 있음을 보여준다. 이는 특정 환경에서만 유효한 결과가 아니라, 다양한 교육 맥락 속에서도 컴퓨팅 사고력 향상이라는 공통의 목표를 향해 SW 교육이 유연하게 시행되고 있음을 시사할 수 있다.

Table 7. Overall Effect Size Using the Random-Effects Model

k	Effect Size		Z	p	95% CI	
	Δ	SE			LB	UB
356	0.689	0.029	23.475	<.001	0.631	0.746

*k=Number of effect sizes included in the meta-analysis, Δ =Estimated difference in effect sizes between groups

인지적 영역의 세부 요소별 효과크기는 Table 8 과 같이 나타났다. 문제분해, 추상화, 알고리즘화, 자동화, 통합 CT 등 모든 하위 요소에서 중간 이상의 안정적 효과가 확인되었다. 각 효과의 95% 신뢰구간은 모두 0.50 이상으로, 인지적 CT 영역은 SW 교육의 가장 강력한 효과가 나타나는 영역임을 보여준다.

이질성 검사에서 $Qb = 1.19$, $p = 0.879$ 로 나타나, 인지적 하위요인 간 효과 차이는 통계적으로 유의하지 않았으며, 이는 SW 교육이 인지적 CT의 전 영역을 균형 있게 함양시키고 있음을 시사한다.

이러한 결과는 1세대 메타분석 연구[9]와 비교할 때, 최근 K-12 교육 현장에서 각 하위 요소들이 고르게 발달하는 긍정적인 시기적 추세를 관찰할 수 있다. 1세대 메타분석[9]에서는 수행 영역($\Delta = 0.409$)과 개념 영역($\Delta = 0.259$)에서 상대적으로 낮은 효과크기가 보고되었으며, 특히 기초 지식 이해에 해당하는 개념 영역의 효과는 중간 수준에 미치지 못하는 것으로 나타났다. 반면 본 연구에서는 알고리즘화

($\Delta^*=0.790$), 문제분해($\Delta^*=0.704$), 추상화($\Delta^*=0.673$), 자동화($\Delta^*=0.662$) 등 전 영역에서 효과크기가 크게 상승하였을 뿐만 아니라, 요소 간 편차 없이 고른 발달이 이루어지고 있음을 확인할 수 있다.

Table 8. Effect Sizes by Cognitive Domain of Computational Thinking

		PD	Ab	AT	Au	Comb
	k	34	37	39	40	68
	n	2074	2479	2600	2418	4335
	$\Delta^{\wedge}$	0.704	0.673	0.790	0.662	0.753
	SE	0.091	0.081	0.107	0.073	0.061
95% CI	LB	0.522	0.515	0.581	0.519	0.634
	UB	0.908	0.831	0.998	0.806	0.872
Test of Heterogeneity		Qb	df		p	
		1,19	4		0.879	

k=Number of effect sizes included in the meta-analysis, n=Sample size of each group, Δ^ =Estimated difference in effect sizes between groups, Qb=Between-group heterogeneity statistic used for subgroup analysis

정의적 영역의 효과크기는 Table 9 과 같이 나타났다. 효과크기는 전반적으로 작음에서 중간 수준의 크기로 인지적 영역보다 효과는 다소 낮다. 각 정의적 요소별 효과는 모두 $p < .05$ 으로 유의하였으며, 95% 신뢰구간 또한 양의 범위에 있다. 하위 요소 간 차이는 $Qb = 0.941$, $p = 0.919$ 로 유의하지 않았다.

Table 9. Effect Sizes by Affective Domain of Computational Thinking

		ATD	MTV	INT	SEF	SAT
k		7	7	10	8	3
n		956	159	1216	1888	87
Δ°		0.544	0.539	0.642	0.476	0.654
SE		0.080	0.146	0.124	0.080	0.215
95% CI	LB	0.387	0.253	0.398	0.320	0.233
	UB	0.702	0.825	0.885	0.632	1.076
Test of Heterogeneity		Qb	df		p	
		0.941	4		0.919	

k=Number of effect sizes included in the meta-analysis, n=Sample size of each group, Δ^ =Estimated difference in effect sizes between groups, Qb=Between-group heterogeneity statistic used for subgroup analysis

4.2 조절변수에 따른 효과크기

Table 10 과 같이 교육 수준에 따른 효과는 $Qb = 647.90$, $p < .001$ 으로 통계적으로 유의미하다. 특히, 초등 고학년(UE, Upper Elementary)에서 가장 높은 효과가 나타나, SW 교육은 발달 초기의 상위 초등 단계(LE, Lower Elementary)가장 효과적으로 작동함을 볼 수 있다.

Table 10. Effect Sizes by Educational Level of Korean Learners

		LE	UE	MS	HS	Comb
k		66	132	18	29	7
n		3,260	7,461	1,543	1,031	4,907
$\Delta^{\wedge}$		0.505	0.889	0.503	0.653	0.329
SE		0.053	0.054	0.041	0.059	0.027
95% CI	LB	0.401	0.784	0.422	0.537	0.276
	UB	0.609	0.995	0.583	0.769	0.382
Test of Heterogeneity		Qb		df	p	
		647.90		5	< .001	

k=Number of effect sizes included in the meta-analysis, n=Sample size of each group, Δ^ =Estimated difference in effect sizes between groups, Qb=Between-group heterogeneity statistic used for subgroup analysis

프로그래밍 유형에 따른 효과는 Table 11 과 같이 나타났다. 전체적으로 중간 이상의 효과를 보였으며, 하위집단 차이는 $Qb = 0.856$, $p = 0.931$ 으로 통계적으로 유의하지 않았다. 특히 혼합형 활동(Hybrid Type)($\Delta^* = 0.968$)에서 높았지만 정확히 어떤 활동과 혼합되었는지는 알 수 없다.

Table 11. Effect Sizes by Programming Language Type

		LEUnplugged	Block	Text-Based	Physical Computing	Hybrid Type
k		35	33	10	124	49
n		1330	3450	425	6139	6591
Δ^*		0.703	0.619	0.687	0.680	0.968
SE		0.085	0.048	0.094	0.043	0.145
95% CI	LB	0.536	0.526	0.502	0.595	0.684
	UB	0.870	0.712	0.873	0.765	1.252
Test of Heterogeneity		Qb		df		p
		0.856		4		0.931

k=Number of effect sizes included in the meta-analysis, n=Sample size of each group, Δ^ =Estimated difference in effect sizes between groups, Qb=Between-group heterogeneity statistic used for subgroup analysis

협력학습 여부에 따른 효과는 Table 12와 같이 나타났다. 두 조건의 효과 차이는 $Qb = 19.703$, $p < .001$ 로 나타났다. 통계적으로 유의하며, 개인 학습 조건(Individual)에서 약간 더 큰 효과가 나타났다. 이는 협력학습 과정(Collaborative)에서 인지적 부담이나 조율 부담이 발생할 수 있음을 시사한다.

Table 12. Effect Sizes by Collaborative Learning Condition

	Individual	Collaborative
k	40	200
n	1036	15273
Δ^c	0.736	0.719
SE	0.061	0.034
95% CI	LB	0.617
	UB	0.854
Test of Heterogeneity	Qb	df
	19.703	1
		p
		< .001

*k=Number of effect sizes included in the meta-analysis, n=Sample size of each group, Δ^c =Estimated difference in effect sizes between groups, Qb=Between-group heterogeneity statistic used for subgroup analysis

AI-assisted 여부에 따른 효과크기는 Table 13 과 같이 나타났다. 조절효과는 Qb = 7.17, p = 0.007으로 통계적으로 유의하다. 예상과 달리 AI-assisted 미활용 조건에서 더 큰 효과가 나타났는데, 이는 AI-assisted 활용 수업이 도입 초기 단계에 있어 연구 축적 및 수업 안정성 연구가 충분히 확보되지 않았을 가능성을 시사한다. 본 메타분석은 문헌들의 효과크기 경향성을 확인한 것으로, 조절효과 분석 결과를 AI-assisted 활용이 컴퓨팅 사고력 향상을 직접적으로 저해한다는 인과적 결론으로 단정 지어 해석하는 데에는 주의가 필요하다.

Table 13. Effect Sizes by AI-assisted vs. Non-AI-assisted Conditions

	AI-assisted	Non-AI-assisted
k	66	187
n	8117	10095
Δ^c	0.543	0.740
SE	0.048	0.035
95% CI	LB	0.448
	UB	0.638
Test of Heterogeneity	Qb	df
	7.17	1
		p
		0.007

*k=Number of effect sizes included in the meta-analysis, n=Sample size of each group, Δ^c =Estimated difference in effect sizes between groups, Qb=Between-group heterogeneity statistic used for subgroup analysis

4.3 민감도 분석

본 연구에서는 통제집단의 평균 변화값 d^c 추정치에 대한 의존성을 확인하기 위해, 추정치의 최솟값 Min(d^c), 기준값 d^c , 최대값 Max(d^c)을 적용한 민감도 분석을 수행하였다. 하위 영역별 효과크기와 조절효과 결과는 Table 14-16 에 제시하였다.

컴퓨팅 사고력의 하위 영역별 추정된 Min(d^c), d^c , Max(d^c) 조건에 따른 효과크기의 변화는 Table 13 과 같이 나타났다. 조건에 따른 효과크기의 변화는 소수점 셋째 자리 수준의 매우 미미한 차이에 그쳤다. 이는 하위 영역별 효

과가 특정 추정치 변화에 민감하게 흔들리지 않으며, 매우 안정적인 효과 구조를 갖고 있음을 의미한다.

Table 14. Estimated Min(d^c), d^c , Max(d^c) for Sub-Domains of Computational Thinking

	Min(d^c)	d^c	Max(d^c)
PD	0.698	0.704	0.710
Ab	0.667	0.673	0.679
AT	0.787	0.790	0.792
Au	0.659	0.662	0.664
CT	0.748	0.753	0.758
ATD	0.525	0.544	0.565
MTV	0.537	0.539	0.541
INT	0.629	0.642	0.654
SEF	0.461	0.476	0.493
SAT	0.646	0.654	0.661

* d^c = mean effect size for the control group

전체 효과크기 또한 Min(d^c), Max(d^c) 조건 모두에서 거의 동일하게 나타났다(Table 15). 95% 신뢰구간 역시 모든 조건에서 0.631-0.746 수준으로 동일하였으며, Z값 또한 23.47 전후로 일관된 유의미성 (p = < .001)을 유지하였다. 또한 전체 이질성(Q) 역시 세 조건에서 약 1546-1553으로 큰 차이를 보이지 않아, 전체 효과 추정값이 추정치 변화에 대해 매우 강건함을 확인할 수 있다.

Table 15. Heterogeneity Tests and Effect Size Estimates Using the Estimated Min(d^c), d^c , Max(d^c)

		Min(d^c)	d^c	Max(d^c)
Q		1546.214	1549.209	1553.697
	p	< .001	< .001	< .001
Effect Size	Δ^c	0.688	0.689	0.689
	SE	0.029	0.029	0.029
Z		23.477	23.475	23.472
P		< .001	< .001	< .001
95%	LB	0.631	0.631	0.631
	CI	0.746	0.746	0.747

추정된 Min(d^c), d^c , Max(d^c)에 따른 조절변수 별 통계적 유의성 변화 여부를 Table 16 과 같이 검토하였다.

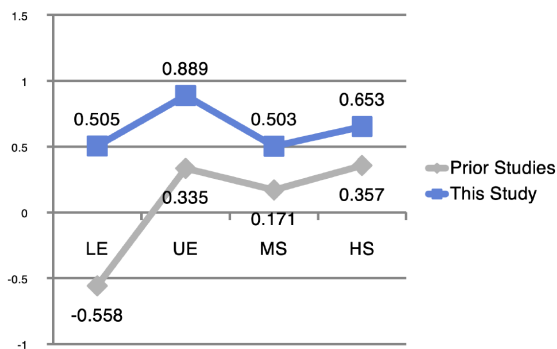
프로그래밍 언어 유형에서는 모든 조건에서 비유의하였다. 즉, 프로그래밍 언어의 조절효과는 추정치 변화와 무관하게 p값이 유의하지 않은 패턴을 유지하였다. 협력학습 여부에서는 세 조건 모두에서 p값이 유의하였다. 협력학습 여부는 추정치 변화와 상관없이 일관된 조절효과를 갖는다. AI-assisted 활용 여부의 조절효과는 세 조건 모두에서 유의한 차이가 확인되었다. AI-assisted 활용 여부가 효과크기 변동에 구조적으로 영향을 미치는 독립적 조절 변수임을 나타낸다.

Table 16. Moderator Analysis Using the Estimated $\text{Min}(d^c)$, d^c , $\text{Max}(d^c)$

		Edu Level	PL Type	Collaboration	AI-assisted
Min (d^c)	Qb	34.678	1.467	19.735	7.253
	df	5	6	2	1
	p	<.001	0.962	<.001	0.007
d^c	Qb	34.709	1.462	19.703	7.170
	df	5	6	2	1
	p	<.001	0.962	<.001	0.007
Max (d^c)	Qb	34.738	1.458	19.663	7.076
	df	5	6	2	1
	p	<.001	0.962	<.001	0.008

4.4 선행 메타분석과 본 연구 결과 비교

본 연구에서 1세대 국내 메타분석인 나청수 외(2018)의 분석 결과와 비교하여, 최근 10년간(2015-2025) K-12 SW 교육의 효과가 시기적으로 변화하였는지를 검토하였다 (Figure 3-5). 본 연구의 전체 효과크기는 기존 연구 결과와 비교하여 전반적으로 그 효과가 유지·강화되는 경향으로 나타났다. Fig. 3에서 제시한 바와 같이, 모든 학교급에서 본 연구의 효과크기가 선행 연구보다 높게 나타났다. 특히 초·중·고학년의 효과크기 증가가 가장 두드러지며, 이는 SW 교육이 발달 초기 고학년 단계에서 더욱 효과적으로 작용하고 있음을 알 수 있다.

**Figure 3.** Comparison of Prior Meta-Analyses and the Present Study on Effect Sizes Across Educational Levels

프로그래밍 언어별 효과 또한 기존 연구보다 크게 나타났다. 특히 혼합형 언어 활용이 상승폭이 매우 크다.

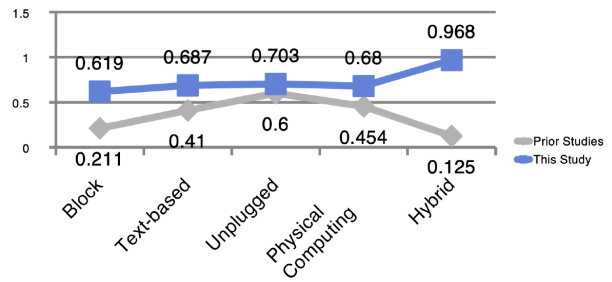
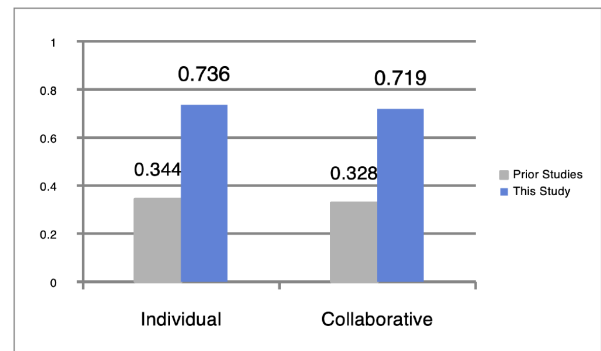
**Figure 4.** Comparison of Prior Meta-Analyses and the Present Study on Effect Sizes by Programming Language

Fig. 5 과 같이, 협력학습 여부에서도 본 연구가 기존 연구보다 높은 효과를 보였다.

**Figure 5.** Comparison of Prior Meta-Analyses and the Present Study on Effect Sizes by Collaborative Learning Condition

본 연구의 2015-2025의 CT 향상 효과는 1세대 메타분석(2006-2017)의 결과와 비교하여 최근 K-12 연구에서 관찰되는 긍정적인 시기적 추세를 반영한다. 이는 교육과정 개정 및 다양한 SW 교육 활동의 도입 등 최근 교육 환경 변화의 영향이 효과크기의 강화에 기여했을 가능성을 시사할 수 있다.

하지만 두 연구 간 직접적인 비교에는 몇 가지 한계가 존재한다. 첫째, 분석 기간이 일부 중첩되어 있으며, 둘째 기존 연구는 K-12에서 대학생까지 포함한 폭넓은 학습자 집단을 분석 대상으로 삼았다. 본 연구는 K-12 학습자만을 대상으로 범위를 한정하였다. 따라서 두 연구 간 효과크기의 차이는 연구 설계나 맥락의 차이에서 비롯될 수 있어 직접적인 우열을 단정하기 어렵다. 이에 본 연구에서는 수치 자체의 절대적 비교보다는, 최근 K-12 환경에서 유사한 효과가 유지되거나 다소 강화되는 경향이 관찰된다는 시기적 추세에 주목하여 결과를 해석하고자 하였다.

그럼에도 불구하고, AI-assisted 학습 도구의 등장, 피지컬 컴퓨팅 및 프로젝트 기반 SW 교육의 확산 등 최근 10년간의 교육 환경 변화가 CT 향상 효과를 전반적으로 확대해왔다는 점은 일관되게 나타난다. 이와 같은 분석은 K-12 교육 현장에서 SW 교육이 어떻게 진화해왔는지 보여주는 중

요한 근거로 기여할 수 있다.

5. 결론

5.1 SW 교육 효과에 대한 주요 결과

본 연구는 2015-2025년 국내 K-12 SW 교육 연구에서 보고된 356개의 효과크기를 대상으로 메타분석을 수행하였다. 분석 결과, 전체 효과크기는 $\Delta = 0.689$ 로 나타났으며, 이는 SW 교육이 학습자의 컴퓨팅 사고력 향상에 중간 이상의 안정적 효과를 갖는다는 점을 보여준다. 출판편향 검증에서도 유의한 편향이 확인되지 않아 연구결과의 신뢰성이 확보되었다. 본 연구의 핵심적 기여는 다음과 같다.

첫째, 도입기 중심의 1세대 메타분석($\Delta = 0.314$)과 비교하여 국내 SW 교육이 양적 확대를 넘어 질적 안착 단계에 진입하며 그 효과가 유지·강화되는 시기적 추세에 있음을 확인하였다. 이는 정보 교육 필수화와 교육과정 표준화가 학교 현장의 교수·학습 실행력을 상향 평준화시켰음을 시사한다. 하지만, 1세대 메타분석 연구[9]에 대학생이 포함되어 있다는 점에서 직접적인 비교에는 한계가 있으므로 결과 해석에 주의가 필요하다.

둘째, 인지적 영역이 정의적 영역보다 상대적으로 높은 효과를 보인 점은 국내 SW 교육이 논리적 문제 해결 도구로서의 CT 함양에 최적화되어 있음을 보여준다. 특히 알고리즘화($\Delta = 0.790$)와 문제분해($\Delta = 0.704$)에서 나타난 높은 성과는 프로그래밍의 기술적 습득이 실제적인 사고 체계로 전이되고 있음을 뒷받침한다. 반면 정의적 영역(흥미·동기·자기효능감 등)의 상대적으로 낮은 효과는 국내 SW 교육이 인지적 CT 요소 중심으로 설계되어 왔음을 시사하며, 향후 정의적 영역을 포괄하는 통합적 설계의 필요성을 제기한다.

셋째, 초등 고학년에서 관찰된 최대 효과크기($\Delta = 0.889$)는 해당 시기가 CT 교육의 최적 발달 단계임을 확인하며, 발달 단계에 따른 정보 교육과정의 위계적 설계가 필요함을 나타냈다. 반면 중학교 단계에서 상대적으로 낮은 효과가 나타난 것은 2022 개정 교육과정의 구성 체계와 연계될 가능성이 있다. 초등 고학년은 '디지털 사회와 인공지능' 영역에서 특정 프로그래밍 언어의 문법 습득보다 놀이와 체험 중심의 문제 해결 과정을 강조하고 있으며, 프로그래밍의 결과물보다 과정을 통한 사고력 향상을 측정하도록 권장하는 과정 중심 평가 기제가 해당 단계에서 효과적으로 작용했을 가능성이 있다. 반면 중학교 단계의 상대적으로 낮은 효과크기($\Delta = 0.503$)는 교육과정의 급격한 심화와 학문적 위계성에서 기인한 것으로 보인다. 중학교 정보과 교육과정은 운영 체제의 기능, 퍼지컬 컴퓨팅 시스템의 구성, 데이터의 구조화 등 초등 단계에 비해 높은 수준의 지식 이해와 과정 기능을 요구하며, 학습자에게 인지적 부하를 야기할 수 있다. 초등 단계의 놀이·체험 중심 활동에서 시스템적 이해와 논리적 엄밀성을 요구하는 중학교 교육과정으로 전환되는 과정에서 인지적 격차가 컴퓨팅 사고력의 즉각적인 향상을 억제

하는 요인으로 작용했을 가능성이 있다.

5.2 AI-assisted 학습의 조절효과와 교육적 해석

조절효과 분석에서 교육 수준($Q_b = 647.90, p < 0.001$)과 AI-assisted 활용 여부($Q_b = 7.17, p = 0.007$)가 유의한 조절 변인으로 확인되었다. 특히 AI-assisted 활용 여부에 대해, AI-assisted 미활용 환경($\Delta = 0.740$)이 AI-assisted 활용 환경($\Delta = 0.543$)보다 더 큰 효과를 보인 이유에는 두 가지 가능성이 있다. 첫째, 교수 실행의 안정성이 부족한 점에서, 현재 AI-assisted 수업은 기술 도입 초기 단계에 머물러 있기 때문에 AI 도구가 학습자에게 추가적인 인지적 부하로 작용했을 가능성을 제기해 볼 수 있다. 둘째, AI-assisted 환경이 학습자가 스스로 문제를 분석하고 알고리즘을 설계하는 과정을 생략하게 만들어 CT의 발달을 저해했을 가능성이 있다.

하지만 이러한 결과를 해석함에 있어 신중한 접근이 필요하다. 본 메타분석 결과는 보고된 효과크기의 집단 간 평균적 차이를 확인하는 데 그치며, AI-assisted 활용 방식과 학습자 성취 간의 직접적인 인과관계를 입증한 것은 아니다. 따라서 본 연구 결과가 AI-assisted 활용 자체의 교육적 효과가 부족하거나 컴퓨팅 사고력에 부정적인 영향을 미친다고 단정하기보다는, 각 연구에서 채택된 교수 설계의 적절성이 부족하거나 현재의 평가 도구가 생성형 AI-assisted 활용 환경에서 발생하는 변화를 충분히 포착하지 못했기 때문일 수 있다. 이 결과는 생성형 AI 도구를 구체적인 활용 방식과 교육 성과 간의 관계를 정교하게 규명해야 할 필요가 있음을 시사한다.

5.3 본 연구의 제한점 및 제언

첫째, 본 연구는 국내 연구만을 대상으로 하였기 때문에 결과의 일반화 범위가 한국 교육 상황으로 제한된다. 향후에는 OECD 국가나 AI 교육을 선도하는 나라들과 비교하는 국제 메타분석이 필요하다.

둘째, 포함된 연구들의 설계 및 보고 방식의 이질성이 존재하였다. 일부 연구는 통계 데이터가 충분하지 않아 별도의 추정 과정이 필요했으며, 이는 효과크기 계산의 정밀도에 영향을 미칠 수 있다. 후속 연구에서는 메타회귀(meta-regression) 또는 개별 참가자 데이터(IPD) 기반 메타분석을 통해 설계 변인을 정교하게 통제할 필요가 있다.

셋째, CT 평가도구의 표준화가 미흡하여 정의적 영역 측정 척도의 부족과 도구 간 이질성은 결과 해석에 제한을 초래할 수 있다. 향후 연구에서는 인지적·정의적 영역을 통합한 타당화된 CT 평가도구 개발이 요구된다.

넷째, 본 연구에서는 AI-assisted 활용 여부에 따른 조절 효과를 확인하였으나, 실제 학습 환경에서의 AI 활용 방식과 CT 성장 간의 상관관계를 심층적으로 분석하는 데는 한계가 있었다. 따라서 후속 연구에는 AI 기반 피드백 구조, 상호작용 패턴, 학습자의 AI 리터러시 수준 등 세부 요소를

통제된 실험 연구를 축적함으로써, 생성형 AI 도구의 구체적인 활용 방식과 교육 성과 사이의 정교한 상관관계를 규명할 필요가 있다.

다섯째, 초등 고학년에서 중학교로 전이할 때 교육 효과를 높이기 위해 연계 교육 및 교수 설계 개선에 대한 연구가 필요하다. 본 연구에서는 중학교 단계에서 다루는 교과 내용이 초등 고학년 교육과정에 비해 보다 고차원적 개념을 담고 있어 학습자의 인지적 부담이 증가할 가능성이 있음을 확인하였다. 따라서 이러한 인지적 부하를 최소화할 수 있는 교수 설계에 대한 보완 연구가 요구된다.

결론적으로 본 연구는 국내 K-12 SW 교육이 컴퓨팅 사고력 향상에 미치는 효과를 체계적으로 검증하고, 향후 SW 교육의 설계·평가·정책 방향을 제시하였다는 점에서 의미가 있다. 또한 연구 과정에서 확인된 국제 연구 데이터의 선별 경험은 향후 단계적 확장 연구의 기반을 제공한다.

참고문헌

- [1] Wing, J. (2006). Computational thinking. *Communications of the ACM*, 49(3), 33–35. <https://doi.org/10.1145/1118178.1118215>
- [2] Yang, X., & Zhang, Y. (2020). International comparative study on computational thinking education in K-12. *The Educational Review*, 4(8), 166–175. <https://doi.org/10.26855/er.2020.08.002>
- [3] Ministry of Education. (2021). *Major Issues of the 2022 Revised National Curriculum (Draft)*. Ministry of Education. <https://www.moe.go.kr>
- [4] Hong, S., Lee, Y., & Kim, S. (2024). Effect of artificial intelligence convergence education using ChatGPT on computational thinking of high school students in Korea. *International Journal on Advanced Science, Engineering and Information Technology*, 14(6), 1982–1990. <https://doi.org/10.18517/ijaseit.14.6.10092>
- [5] Kim, J., Lee, H., & Cho, Y. (2022). Learning design to support student-AI collaboration: Perspectives of leading teachers for AI in education. *Education and Information Technologies*, 27, 6069–6104. <https://doi.org/10.1007/s10639-021-10831-6>
- [6] Choe, H. (2021). Study of AI Thinking Education based on Computational Thinking. *The Journal of Korean Association of Computer Education*, 24(3), 57 - 65 <https://doi.org/10.32431/kace.2021.24.3.006>
- [7] Lee, J., Kim, D., & Lee, S. (2021). An analysis of the impact of AI maker coding education on improving computational thinking. *Journal of the Korean Association of Information Education*, 25(5), 779–790. <https://doi.org/10.14352/jkaie.2021.25.5.779>
- [8] Baek, D., Son, W., Song, J., & Yoo, M. (2024). Meta-analysis of learning effectiveness using generative AI. *Journal of Korean Association for Educational Information and Media*, 30(4), 1261–1285. <https://doi.org/10.15833/KAFEIAM.30.4.1261>
- [9] Na, C., Joo, H., Lee, J., & Kim, D. (2018). *Inducing computational thinking in Korean SW education: Synthesizing standardized mean changes through meta-analysis*. *Journal of Educational Technology*, 34(3), 775–815. <https://doi.org/10.17232/KSET.34.3.775>
- [10] Kim, D., & Lee, T. (2019). Review of Concept of Abstraction in Computational Thinking. *Proceedings of the Korean Society of Computer Information Conference*, Gyeongsangbuk-do, Republic of Korea.
- [11] Han, J., Jung, U., & Lee, Y. (2017). Analysis on Research Trends related Computational Thinking in Korea. *Proceedings of the Korean Association of Computer Education Conference*, Seoul, South Korea, 21(2), 3–5.
- [12] Choi, S. (2018). A Study on the Digital Competency for the Fourth Industrial Revolution. *The Journal of Korean Association of Computer Education*, 21(5), 25–35. <https://doi.org/10.32431/kace.2018.21.5.003>
- [13] Ghazi, S., Khan, U., Shahzada, G., & Ullah, K. (2014). Formal Operational Stage of Piaget's Cognitive Development Theory: An Implication in Learning Mathematics. *Journal of Educational Research*, 17, 71.
- [14] Vosniadou, S. (2020). Bridging secondary and higher education: The importance of self-regulated learning. *European Review*, 28(S1), S94–S103. <https://doi.org/10.1017/S1062798720000939>
- [15] Merino-Armero, J., González-Calero, J., & Cózar-Gutiérrez, R. (2022). Computational thinking in K-12 education: An insight through meta-analysis. *Journal of Research on Technology in Education*, 54(3), 410–437. <https://doi.org/10.1080/15391523.2020.1870250>
- [16] Chen, Y., Wang, Y., & Li, Y. (2023). The effectiveness of teaching approaches in computational thinking education: A meta-analysis. *Proceedings of the 2022 ACM Conference on International Computing Education Research*, 156–169. <https://doi.org/10.1145/3572549.3572611>
- [17] Settle, A., & Perkovic, L. (2010). *Computational Thinking across the Curriculum: A Conceptual Framework*. Technical Reports, DePaul University. <https://via.library.depaul.edu/tr/13>
- [18] Kim, T. (2015). *STEAM education program based on programming to improve computational thinking ability* [Doctoral dissertation]. Jeju National University
- [19] Choi, S. (2011). An Analysis of 'Informatics' Curriculum from the Perspective of 21st Century Skills and Computational Thinking. *Proceedings of the Korean Association of Computer Education Conference*, 15(2), 99–103. <https://doi.org/10.32431/kace.2011.14.6.003>
- [20] Lee, S. (2018). A Study on Designing a Class of Convergence Thinking based on Computational Thinking. *The Korean Society of Science & Art*, 36, 255–263. <https://doi.org/10.17548/ksaf.2018.12.30.255>
- [21] Choi, J., & Lee, Y. (2014). The Design of Method for Evaluating Computational Thinking. *Proceedings of the Korean Society of Computer Information Conference*, 22(2),

177-178.

- [22] Brennan, K., & Resnick, M. (2012). New frameworks for studying and assessing the development of computational thinking. *Proceedings of the 2012 Annual Meeting of the American Educational Research Association*, 1-25.
- [23] International Society for Technology in Education. (2016). *ISTE standards for students*. <https://www.iste.org/standards/for-students>
- [24] National Research Council. (2012). *A framework for K-12 science education: Practices, crosscutting concepts, and core ideas*. National Academies Press. <https://doi.org/10.17226/13165>
- [25] Kang, H., Cho, J., & Kim, H. (2019). A Study on Software Analysis and Design Education Model based on Computational Thinking. *Journal of Digital Contents Society*, 20(5), 947-954. <https://doi.org/10.9728/dcs.2019.20.5.947>
- [26] Kim, C., Ahn, J., Noh, D., Moon, J., & Kim, S. (2022). A Study on Core Competencies for Elementary and Secondary Education in the Era of the Fourth Industrial Revolution. *Journal of Digital Contents Society*, 23(11), 2159-2170.
- [27] Csizmadia, A., Curzon, P., Dorling, M., Humphreys, S., Ng, T., Selby, C., & Woollard, J. (2015). *Computational thinking: A guide for teachers*. Computing At School. <http://computingatschool.org.uk/computationalthinking>
- [28] Becker, B. (1988). Synthesizing standardized mean-change measures. *British Journal of Mathematical and Statistical Psychology*, 41, 257-278. <https://doi.org/10.1111/j.2044-8317.1988.tb00901.x>
- [29] Cohen, J. (1988). *Statistical power analysis for the behavioral sciences* (2nd ed.). Lawrence Erlbaum Associates.

**백승애**

· 2023년 충북대학교 산림학과(농학사)
· 2026년 성균관대학교 일반대학원 컴퓨터교육학 전공(컴퓨터교육학석사)

✚ 관심분야 : 정보교육, 교육과정평가, 데이터기반 분석, 에듀테크

✉ bsa100na@naver.com

**김미량**

· 1987년 서울대학교 영어영문학과(문학사)
· 1988년 Lehigh University 교육공학전공(이학 석사)
· 1998년 서울대학교 대학원 교육학과 교육방법·공학 전공(교육학박사)
· 1999년~현재 성균관대학교 사범대학 컴퓨터교육과 교수

✚ 관심분야 : 컴퓨터교과교육, 온라인교수설계, 데이터 기반분석

✉ mrkim@skku.edu