

논단

가톨릭대학교 인간학연구소

『인간연구』 제58호

(2026/봄)

제1저자. ckwcho@naver.com

초등 3-4 도덕 교과서에 담긴 인공지능(AI) 윤리교육 요소 고찰

2022 개정 교육과정을 기반으로

조광원(서울사당초등학교 · 도덕교육 · 교사)

<https://doi.org/10.21738/JHS.2026.4.58.287>

I. 서론

■ 인공지능(AI) 기술은 챗봇, 추천 알고리즘, 자율주행 자동차 등 우리 생활 속에서 영향력이 날로 커지고 있다. 이러한 기술의 발전은 편리함과 효율성을 가져다주지만, 동시에 데이터 편향성, 사생활 침해, 딥페이크, 인간 소외 등 많은 윤리적 문제를 야기하고 있다. 이에 따라 인공지능(AI) 기술을 올바르게 이해하고 윤리적으로 활용할 수 있는 능력인 ‘인공지능(AI) 윤리 소양’의 함양이 주목받고 있다.

특히, 가치관과 도덕적 판단력의 기초가 형성되는 초등학교 시기부터 인공지능(AI) 윤리교육을 체계적으로 실시하는 것은 매우 중요하다. 초등학생들은 이미 일상에서 다양한 인공지능(AI) 기술을 접하고 있으며, 이 과정에서 무비판적인 기술 수용으로 잠재적인 윤리적 문제에 노출될 수 있다.

이러한 시대적 요구에 부응하여, 2022 개정 교육과정은 ‘디지털 기초 소양’을 모든 교과를 통해 갖출 수 있도록 하며 지식과 정보를 비판적으로 탐구하고 활용하는 ‘지식정보처리 역량’을 설정하였고,¹ 도덕과 교육과정에서는 ‘인공지능 및 디지털 윤리’의 교육을 명시적으로 강조하였다.² 이는 인공지능(AI)과의 공존을 위한 윤리적 규범

1 교육부, 『초·중등학교 교육과정 총론』(교육부 고시 제2022-33호[별책 1]), 교육부, 2022, 5~6쪽.

2 교육부, 『도덕과 교육과정』(교육부 고시 제2022-33호[별책 6]), 교육부, 2022, 3쪽.

교육을 공교육의 중요한 책무로 인식하기 시작했음을 의미한다.

초등학교 3-4학년은 지적·도덕적 발달 특성상 타인의 관점을 이해하고 규칙의 필요성을 인식하기 시작하는 시기이다.³ 도덕적 판단에 있어서 자기중심성에서 벗어나 점차 사회적 규범과 관계를 고려하기 때문에 인공지능(AI)에 대한 윤리적 관점을 탐구할 수 있는 시기라 할 수 있다.

이에 본 연구는 초등학교 3-4학년 도덕 교과서에 인공지능(AI) 윤리교육 관련 요소가 어떻게 포함되어 있는지 고찰하고자 한다. 먼저, 인공지능(AI) 윤리를 고찰하고, 초등 도덕 교육과 인공지능(AI) 윤리교육을 주제로 한 선행 연구를 살핀 뒤, 초등학교 3-4학년 도덕 교과서 및 교사용 지도서를 대상으로 텍스트, 삽화, 활동 자료 등에 ‘인공지능(AI)’이 들어간 내용을 고찰하고자 한다. 이는 향후 인공지능(AI) 윤리교육의 내실화를 위한 개선 방안을 모색하는 데 기초 자료를 제공하는 것을 목적으로 한다.

본 연구의 목적을 달성하기 위한 연구 문제는 다음과 같다. 첫째, 2022 개정 교육과정 기반 초등학교 3-4학년 도덕 교과서에 인공지능(AI) 윤리 관련 내용은 어떻게 제시되어 있는가? 둘째, 교과서에 나타난 인공지능(AI) 윤리교육 내용은 인공지능(AI) 윤리의 가치 및 핵심 요건을 어떻게 반영하고 있는가? 셋째, 교과서의 인공지능(AI) 윤리 관련 내용(자료 및 활동)은 학생들이 인공지능(AI)을 비판적으로 탐구하며 활용할 수 있는 능력을 함양하도록 구성되어 있는가?

3 김봉제, 「도덕판단력 발달단계를 활용한 학급 인성교육 목표설정 방법 탐색」, 『초등도덕교육』, 51(2016), 8~9쪽.

II. 이론적 배경

1. 인공지능(AI) 윤리

유네스코한국위원회에서는 인공지능(AI) 윤리를 인공지능과 관련된 다양한 쟁점에 대한 선악 판단의 기준보다 사회에서의 핵심적 가치와 이를 균형 있게 존중하는 방식으로 인공지능(AI)을 개발 활용함에 있어 주의점 및 제도적 장치 등에 대한 통합적인 탐색으로 보았다.⁴ 또한, 인공지능(AI) 윤리를 인공지능을 개발하고 운영하고 사용하는 것에 있어 개발자와 소비자에게 요구되는 윤리의식으로 보기도 한다.⁵ 즉, 인공지능(AI) 윤리란 인공지능(AI) 그 자체에 대한 윤리라기보다는 인공지능(AI)을 개발하고 이를 활용하는 사람들이 가져야 할 윤리의식을 말한다.

따라서 인공지능(AI) 윤리는 그 자체가 무엇인지에 대한 구체적인 개념보다는 수많은 원칙이나 가이드라인의 모습으로 제시되었다. 이는 단순히 기술적인 문제를 넘어, 인공지능(AI)이 인간의 삶과 사회에 미치는 영향을 도덕적 관점에서 성찰하고 바람직한 방향을 모색하는 것을 포함함을 뜻한다.

이와 관련하여 유네스코(UNESCO)에서는 AI 윤리 예비연구에 나타난 교육, 과학, 문화 등 과 관련된 AI 시스템의 광범위한 윤리 문제에 주목하며, 인공지능(AI) 윤리에 대한 초안을 작성하고 여러 나라들과 협의한 끝에 2021년에 권고안을 제시하였다.⁶ 이러한 시도는

4 이상욱, 『유네스코 인공지능(AI) 윤리 권고 해설서: 인공지능 윤리 이해하기』, 유네스코한국위원회, 2021, 15쪽.

5 신현화 · 신태수 · 신현우, 「초등학교 도덕과 수업에서 인공지능 윤리교육 방안 탐색 연구」, 『학교와 수업 연구』, 7(2022), 77쪽.

6 UNESCO, 「AI 윤리에 대한 권고 초안_국문번역본」, 제2차 정부 간 회의, 2021년 6월 25일, 파리.

인공지능(AI)의 근본적인 윤리적 우려를 제기하면서 인간의 존엄성과 권리를 비롯하여 젠더 평등, 사회경제적 정의와 발전, 신체적·정신적 안녕, 다양성, 상호연결성, 포용성, 환경 및 생태계 보호와 관련된 윤리적 논제를 다루기 위해 국제적 기준을 제시함으로써 AI 기술을 책임감 있는 방향으로 이끌고자 하였으며, 결국 193개의 모든 회원국이 만장일치로 채택하여 글로벌 표준 규범이 되었다. 여기서 제시한 4개의 가치는 다음과 같다.⁷

〈표 1〉 유네스코에서 제시한 4대 가치(Values)

순	가치
Value 1 (V1)	인권, 기본적 자유, 인간 존엄성의 존중·보호·증진 (Respect, protection and promotion of human rights, fundamental freedoms and human dignity)
Value 2 (V2)	환경 및 생태계의 번영 (Environments and ecosystem flourishing)
Value 3 (V3)	다양성 및 포용성 보장 (Ensuring diversity and inclusiveness)
Value 4 (V4)	평화롭고 정의로우며 상호연결된 사회 구축 (Living in peaceful, just and interconnected societies)

이 4개의 가치는 유네스코의 회원국들이 인공지능(AI)을 개발·활용함에 있어 궁극적으로 추구해야 할 가치이자, 규범으로 인공지능(AI) 윤리기준에 해당한다. V1은 인간의 존엄성과 권리, 신체적·정신적 안녕의 윤리적 논제를 바탕으로 인공지능(AI)이 기술 이전에 인간을 최우선에 두어야 하며 모든 개인의 존엄성, 인권, 기본적 자유

7 UNESCO, “Recommendation on the Ethics of Artificial Intelligence”, 41st session of the General Conference, 2021.11.23., Paris.

를 침해해서는 안 되며, 이를 보호하고 증진하는 방향으로 사용되어야 한다고 한다. 즉, 인공지능(AI)이 모든 개인의 인권과 존엄성을 최우선으로 존중하고 보호해야 함을 말하고 있다. V2는 환경 및 생태계 보호의 윤리적 논제를 바탕으로 인공지능(AI) 기술이 개발부터 활용까지 전 과정에서 환경에 미치는 영향을 고려하며 지속가능성에 기여해야 함을 뜻한다. 즉, 인공지능(AI)은 기술을 개발하고 운영하는 전 과정에서 발생하는 환경적 영향을 고려해야 하며, 자원을 고갈시키는 것이 아닌 지속가능한 환경과 생태계의 번영에 기여하도록 해야 한다. V3는 젠더 평등, 다양성, 포용성의 윤리적 논제를 바탕으로 인공지능(AI)은 성별, 인종, 문화, 언어 등에서 발생하는 편견을 답습하거나 강화하지 않고, 모든 사람이 공평하게 혜택을 누릴 수 있도록 다양성과 포용성을 증진해야 함을 말한다. 인공지능(AI)은 학습의 방향에 따라 특정 집단의 편견을 학습하거나 차별을 강화하게 할 수 있다. 따라서 인공지능(AI)은 성별, 인종, 문화, 언어 등 사회의 다양한 가치를 반영하고, 모든 사람이 소외되지 않고 인공지능(AI)의 혜택을 누릴 수 있도록 포용적으로 설계되어야 한다. V4는 사회경제적 정의와 발전, 상호연결성의 윤리적 논제를 바탕으로 인공지능(AI)이 사회적 갈등이 아닌 평화와 정의, 공동체 간의 상호 이해와 연결을 촉진하는 방향으로 사용되어야 함을 말한다. 인공지능(AI)을 잘못된 방향으로 개발하고 활용하게 되면 사회적 갈등이나 불의를 조장할 수 있다. 따라서 인공지능(AI)은 다양한 공동체 간의 평화, 정의, 상호 이해를 돕는 도구로 개발하고 활용해야 한다.

한국은 기술 오용, 데이터 편향성과 같은 인공지능 윤리 이슈를 대비하여 인공지능 윤리기준을 2020년에 이미 발표하였다. 이른바 ‘사람이 중심이 되는 AI 윤리기준’은 한국 정부 주도의 AI 윤리기준

으로,⁸ 인간성(Humanity)을 최고 가치로 설정하여 사람 중심의 인공지능을 지향하기 위해 인간 존엄성, 사회의 공공선, 기술의 합목적성의 3대 기본원칙을 제시하였으며, 이를 요약하면 다음과 같다.⁹

〈표 2〉 3대 기본원칙

기본원칙	내용
인간 존엄성	- 인간은 생명체로 인공지능과는 교환 불가능한 가치가 있음 - 인공지능은 안전성을 갖춰 인간의 생명, 정신적 및 신체적 건강 등 인간에게 해가 되지 않는 범위에서 개발 및 활용
사회의 공공선	- 인공지능은 사회적 약자와 취약 계층의 접근성을 보장하도록 개발 및 활용 - 공익 증진을 위한 인공지능 개발 및 활용은 인류의 보편적 복지를 향상
기술의 합목적성	- 인공지능은 인류의 삶에 필요한 도구라는 목적과 의도에 부합되게 윤리적으로 개발 및 활용 - 인류의 삶과 번영을 위한 인공지능 개발 및 활용 장려

또한, 이를 실현하고자 하는 세부 요건으로 인공지능(AI)을 개발하고 활용하는 모든 과정에서 필요한 10대 핵심 요건을 제시하였는데, 이는 다음과 같다.¹⁰

〈표 3〉 10대 핵심 요건

요건 명	내용
인권보장	- 모든 인간의 권리 존중 및 보장 - 인간의 권리와 자유 침해 금지
프라이버시 보호	- 개발 및 활용의 전 과정에서 프라이버시 보호 - 개인 정보의 오용 최소화

8 이인재, 「AI 윤리 기준을 활용한 AI 윤리교육 방안 연구」, 『초등도덕교육』, 76(2022), 163쪽.

9 과학기술정보통신부, 『사람이 중심이 되는 「국가 인공지능 윤리기준」』, 과학기술정보통신부, 2020, 2쪽.

10 위의 글, 2~4쪽.

요건 명	내용
다양성 존중	- 개인 특성에 따른 편향과 차별 최소화 - 사회적 약자를 비롯한 모든 사람에게 골고루 분배
침해금지	- 직간접적인 피해를 주는 목적으로 활용 금지 - 인공지능의 위험과 부정적 결과에 대응 방안 마련
공공성	- 사회적 공공성 증진과 인류 공동 이익을 위해 활용 - 긍정적 사회 변화를 이끄는 방향으로 활용 및 교육
연대성	- 다양한 집단 간의 관계 연대성 유지 - 다양한 주체들의 공정한 참여 기회 보장
데이터 관리	- 각각의 데이터를 목적에 부합하게만 활용 - 데이터 편향성은 최소화되도록 품질과 위험 관리
책임성	- 모든 과정에서 책임 주체를 설정하여 피해 최소화 - 다양한 주체 간의 책임 소재를 명확히 함
안전성	- 잠재적 위험 방지 및 안전 보장 - 명백한 오류나 침해 발생 시 제어 기능을 갖춘
투명성	- 활용 시 적합한 수준의 투명성과 설명 가능성 높임 - 서비스 제공 시 유의사항 사전 고지

이렇듯 인공지능(AI) 윤리란 인공지능(AI)을 개발 및 활용하는 모든 사람이 갖추고 지켜야 할 가치, 요건 등을 모두 포괄하는 윤리 기준이다. 윤리적 가치와 원칙은 윤리적으로 건전한 인공지능(AI) 체계를 개발하고 활용하는 데 필요하며, 인공지능(AI) 윤리 체계 구성이나 인공지능(AI) 윤리교육을 할 때는 이러한 인공지능(AI) 윤리기준 속에 들어있는 핵심 가치를 반영해야 한다.¹¹⁾

본 연구에서는 인공지능(AI) 윤리의 수많은 가치와 요건 중 유네스코(UNESCO)에서 제시한 최초의 인공지능(AI) 윤리 글로벌 표준 규범인 인공지능(AI) 윤리 권고에 담긴 4대 가치와 국내 인공지능

11 이인재, 앞의 글, 167쪽.

(AI) 윤리기준의 10대 핵심 요건을 고찰 기준으로 설정한다. 유네스코에서 제시한 인공지능(AI) 윤리 권고의 가치는 한국을 포함한 유네스코에 속한 모든 회원국의 인공지능(AI)과 관련한 정책 및 법적 규범 형성에 있어 바람직한 행동을 고무하는 원칙의 근본이 되기 때문에 그 어떤 기준보다 보편성을 갖고 있다.¹² 아울러, 국내 인공지능(AI) 윤리기준의 10대 핵심 요건을 고찰 기준으로 설정한 이유는 이 역시 한국 정부 주도로 만들어 한국 사회의 구성원이 기본적으로 참고하는 보편성과 범용성을 가진 일반 원칙으로, 이후 세부 규범이 유연하게 발전해 나갈 수 있는 기반을 제공하고 있기 때문이다.¹³

2. 초등 도덕 교육과 인공지능(AI) 윤리교육

교육부에서는 미래 교육의 필수 교양으로 인공지능(AI)의 개발에 따라 인공지능(AI)과 소통하고 인공지능(AI)을 윤리적인 판단을 바탕으로 활용하는 능력을 제시하였다.¹⁴ 많은 선행 연구에서 초등 도덕 교육에서 지향해야 할 인공지능(AI) 윤리교육을 제시하고 있으나, 몇 가지만 알아보고자 한다.

먼저, 초등 인공지능(AI) 윤리교육을 제시한 연구가 있다. 규범적 인공지능(AI) 윤리 가이드라인을 바탕으로 초등 인공지능(AI) 윤리교육을 제시한 연구에서는 규범적 인공지능(AI) 윤리 가이드라인을 주제어와 질문으로 연결하고, 이를 규범적 인공지능(AI) 윤리교육의 목표, 주제, 내용으로 제시하였다.¹⁵ 생성형 인공지능(AI) 윤리교육

12 UNESCO, 「AI 윤리에 대한 권고 초안_국문번역본」, 제2차 정부 간 회의, 2021년 6월 25일, 파리.

13 과학기술정보통신부, 위의 글, 1쪽.

14 교육부, 『인공지능시대 교육정책방향과 핵심과제』, 교육부, 2020, 5쪽.

15 박형빈, 「규범적 AI윤리 가이드라인과 초등 AI윤리교육 적용 방안 모색」, 『한국초등교육』, 121(2024), 88~91쪽.

을 제시한 연구에서는 할루시네이션과 같은 위험성을 고려하여 생성형 인공지능(AI)의 결과물을 비판적으로 감독하고 판단할 수 있도록 16개의 주제를 초등 인공지능(AI) 윤리교육으로 제시하였다.¹⁶ 이 두 연구는 직접적으로 도덕 교육으로 제시하지는 않았으나, 도덕 교육의 내용 요소로 충분히 구성할 수 있도록 구체적인 내용을 제시하였다.

다음으로, 인공지능(AI) 윤리교육을 초등 도덕과로 직접 제시한 연구들이다. 교육과정에서부터 ‘인간 이외의 존재’를 넣어 인공지능(AI) 윤리교육을 목표, 방향, 내용 체계 및 성취기준을 제시한 연구가 있고,¹⁷ 초등 도덕과 교육에서 인공지능(AI) 윤리교육으로 디지털 리터러시 추구하고 인간성 추구를 제시하여 각각에 맞는 교수학습 과정을 제시한 연구도 있다.¹⁸ 또한, 2022 개정 도덕과 교육과정에서 인공지능(AI) 윤리교육을 로봇에 대한 구체적인 맥락을 제시하고 인공지능(AI) 로봇에 대한 지나친 의인화를 사용하지 않은 뒤 구체적인 사례에 대한 윤리적 사용 원칙을 만들어보도록 구체적인 방향을 제시한 연구가 있으며,¹⁹ 초등 도덕과의 내용으로 뇌-컴퓨터 인터페이스(Brain-Computer Interface)를 활용한 인공지능(AI) 윤리교육을 구성하여 글쓰기, 역할놀이, 교과 통합 수업, 프로젝트 학습 등 다양한 활동을 제시한 연구도 있다.²⁰ 그리고 인공지능(AI) 윤리교육과정을 설계하고 프로그램 모형을 개발 및 적용하여 투명성, 책임성, 공공성,

16 박형빈, 「생성 AI(Generative AI)의 할루시네이션(hallucination) 이슈와 윤리적 과제: 초등 AI 윤리교육에서 활용 가능한 주제를 중심으로」, 『한국초등교육』, 119(2023), 32~33쪽.

17 홍현주, 「인공지능 윤리교육의 초등도덕 교육과정 적용 방안」, 『초등도덕교육』, 75(2021), 191~196쪽.

18 신현화 · 신태수 · 신현우, 위의 글, 84~93쪽.

19 김하민, 「2022 개정 초등 도덕 교육과정에 대한 비판적 고찰: 인간과 인공지능 로봇 관계 내용 서술을 중심으로」, 『도덕윤리과교육』, 79(2023), 45~49쪽.

20 박형빈, 「뇌-컴퓨터 인터페이스(BCI)의 인간 증진 및 윤리적 쟁점과 초등도덕 AI윤리교육 활용 방안」, 『초등도덕교육』, 91(2025), 260~264쪽.

다양성 존중 측면에서 효과가 있음을 검증한 연구도 있으며,²¹ 초등 도덕과에서 인공적 도덕 행위자(AMA) 논의를 중심으로 윤리적 행위 원칙을 바탕으로 인공적 도덕 행위자(AMA)를 적용한 인공지능윤리 교육 방안을 제시한 연구도 있었다.²²

이와 같은 선행 연구들은 인공지능(AI) 윤리교육이 초등 도덕 교육과 밀접한 연관이 있으며, 초등 도덕 교육에 반영될 필요가 있다는 점을 드러낸다. 다만, 선행 연구에서 현 도덕 교과서에 담긴 인공지능(AI) 윤리를 살펴본 연구는 없으므로, 이러한 연구는 필요하다.

Ⅲ. 초등 3-4학년 도덕 교과서 고찰

■ 본 장에서는 초등 3-4학년 도덕 교과서에 나타난 인공지능(AI) 윤리교육 내용은 무엇인지, 인공지능(AI) 윤리 권고에 담긴 4개 가치와 국내 인공지능 윤리기준의 10대 핵심 요건을 모두 반영하는지, 학생들이 인공지능(AI)을 비판적으로 탐구하며 활용할 수 있는 능력을 함양하기에 적절한 활동인지, 텍스트, 삽화, 활동 자료 등 어떤 유형으로 반영되어 있는지를 고찰한다.

21 이지형 · 변순용, 「인공지능 윤리교육 프로그램의 효과 검증: 초등 도덕과 교육 적용을 중심으로」, 『초등 도덕교육』, 84(2023), 288~305쪽.

22 권누리, 「초등 도덕과에서의 인공지능윤리교육 방안 연구: 인공적 도덕 행위자(AMA) 논의를 중심으로」, 『초등도덕교육』, 87(2024), 87~108쪽.

1. 3-4학년 도덕 교과서 내용

1) 3학년 도덕 교과서 고찰 결과

3학년에서는 자율형 단원인 4단원 ‘인공지능 로봇 연구소에 가요!’에서만 인공지능(AI) 윤리를 다루고 있다. 이 단원은 인공지능(AI) 로봇을 통해 앞의 1, 2, 3단원의 가치인 자아 존중, 성실, 효·우애를 바탕으로 학습을 심화시키도록 구성하여 인공지능(AI) 윤리에 대한 장벽을 낮췄다.²³ 이를 고찰한 결과는 다음과 같다.

〈표 4〉 3학년 4단원 고찰

차시명	내용	가치	핵심 요건	반영 유형	함양 여부
나와 같은 인공지능 로봇 연구실	나와 같은 인공지능 로봇으로 나에 대해 탐구하기	V1	인권보장, 프라이버시 보호, 다양성 존중, 침해금지	삽화, 4컷 만화, 활동 자료	○
꿈 성장 인공지능 로봇 연구실	인공지능 로봇이 나의 일을 모두 대신 한다면 어떻게 생각하기		공공성, 책임성	텍스트, 활동 자료	○
가족 행복 인공지능 로봇 연구실	가족의 행복을 위한 인공지능 로봇 기능 설정하기	V3, V4	프라이버시 보호, 다양성 존중, 침해금지, 공공성, 연대성, 안전성, 투명성	텍스트, 삽화, 활동 자료	○

23 교육부, 『도덕 3 교사용 지도서』, 비상교육, 2025, 95쪽.

차시명	내용	가치	핵심 요건	반영 유형	함양 여부
인공지능 로봇의 올바른 사용	인공지능 로봇을 활용하는 올바른 태도에 대한 마음 다지기	V1, V3, V4	인권보장, 프라이버시 보호, 다양성 존중, 침해금지, 공공성, 데이터 관리, 연대성, 책임성, 안전성, 투명성	활동 자료	O
반영할 점	인공지능 로봇을 활용함에 있어 환경 고려하기	V2			

이 단원에서는 인공지능(AI) 로봇을 소재로 하여 나의 모습을 되돌아보고 인공지능(AI) 로봇과 함께하는 세상에 대해서도 진지하게 고민해 보도록 한다.²⁴ 1차시에서는 나와 같은 인공지능(AI) 로봇으로 나에 대해 탐구하면서 인공지능(AI) 로봇이 나를 대체할 수 없음을 인식하도록 한다.²⁵ 인공지능(AI) 로봇과 ‘나’를 비교하며 ‘나’의 고유성을 탐구하는데, 인공지능(AI)이 나를 ‘대체할 수 없음’을 인식하는 것 자체가 곧 인간이 ‘인간 존엄성’을 지닌 존재임을 깨닫는 과정이다. 따라서 V1을 담고 있음을 확인할 수 있다. 이러한 내용은 인권보장, 프라이버시 보호, 다양성 존중, 침해금지의 핵심 요건을 갖추고 있다. 먼저, ‘나’는 인공지능(AI) 로봇과 달리 고유한 감정과 생각을 가진 대체할 수 없는 존엄한 존재임을 깨닫는 것은 인간 존엄성의 핵심이므로 인권보장에 해당한다. 다음으로, ‘나와 같은 인공지능(AI) 로봇’을 만들기 위해 나에 대한 정보를 써보는 활동에서는 ‘나에 대해 알리고 싶지 않은 정보’도 생각해 보도록 한다. 예를 들어, “로봇이 ‘나’처럼 되려면, 내가 뭘 좋아하는지, 어떻게 말하는지, 내 비밀이

24 위의 책, 198쪽.

25 위의 책, 202쪽.

나 생각까지 다 알아야 할까?”와 같은 질문은 프라이버시 보호를 확인할 수 있다. 또한, ‘나’라는 존재의 고유성을 인정하는 것은 곧 나와 다른 타인의 고유성을 존중하는 첫걸음이다. 이는 모든 사람이 인공지능(AI)으로 대체될 수 없는 각자 고유한 가치를 지님을 인식하므로 다양성 존중에 해당한다. 마지막으로, 인공지능(AI) 로봇과 관련해 ‘진짜 나’와 다른 점을 써보며 ‘나’를 대체할 수 없음을 배우는 것은 인공지능(AI) 로봇이 인간 고유의 영역인 감정, 생각, 자율성 등을 함부로 침해해서는 안 된다는 윤리적 경계선을 배우는 것이기도 하므로, 침해금지에 해당한다. 이러한 활동들은 각 활동에서의 활동 자료와 ‘나와 같은 인공지능(AI) 로봇’이 있다면 어떤 일이 일어날지 상상해 보며 친구들과 이야기를 나누는 삽화의 내용을 통해 제시되며, 학생들로 하여금 인공지능(AI) 로봇을 비판적으로 탐구하며 활용할 수 있는 능력을 함양하기에 적합하다.

2차시에서는 인공지능(AI) 로봇의 활용이 성실한 생활에 미치는 영향을 알고, 자기 성장을 위한 인공지능(AI) 로봇의 활용 방법을 모색한다.²⁶ 활동으로는 인공지능(AI)이 내가 할 일을 대신 해주고 도와줄 때, 자신의 꿈을 이루는 데 도움이 되는지 생각해 보고, 인공지능(AI) 로봇에게 시키고 싶은 일과 시켜서는 안 되는 일을 구분한 뒤, 인공지능(AI) 로봇 사용 원칙을 만들어보도록 한다. 이러한 활동은 자기 성장과 성실한 생활을 위한 도구로 인공지능(AI) 로봇을 올바르게 활용하도록 하고 있다.²⁷ 특히, ‘성실’이라는 가치를 반영하여 인간의 존엄성을 해치지 않도록 인공지능(AI) 로봇을 활용하도록 하므로 V1이 담겨 있다. 핵심 요건으로는 공공성, 책임성을 반영하고 있는데, 먼저, 공공성은 인공지능(AI) 로봇을 자신의 성실한 생활에 도움

26 위의 책, 208쪽.

27 위의 책, 208쪽.

이 되도록 사용하므로 긍정적인 사회 변화를 이끄는 방향으로 활용하는 것에서 드러난다. 다음으로, 책임성은 인공지능(AI) 로봇을 올바르게 활용하려는 원칙을 만들어보는 것에서 드러난다. 올바른 활용하려는 태도 자체가 곧 주도적으로 인공지능(AI) 로봇을 활용하는 책임이 자신에게 있음을 배우기 때문이다. 이러한 내용들은 각 활동에서의 활동 자료와 ‘꿈 성장 로봇 연구실의 인공지능 로봇이 나를 도와준다면?’이라는 텍스트를 통해 제시되며, 학생들이 인공지능(AI) 로봇을 비판적으로 탐구하며 활용할 수 있도록 한다.

3차시에서는 가족의 행복에 도움이 되는 인공지능(AI) 로봇의 기능을 탐색한다.²⁸ 활동으로는 인공지능(AI) 로봇이 우리 가족의 행복에 어떤 도움을 줄지 생각하기, 도움을 주는 기능을 상상하여 적어 보기, 상상한 가족 행복 인공지능(AI) 로봇에 대해 친구들과 이야기 나누기, 로봇이 아닌 내가 직접 하고 싶은 일 말하기로 구성되어 있다. 이 차시는 ‘가족’이라는 학생이 경험하는 가장 기본적인 공동체의 행복을 위해 인공지능(AI) 로봇을 설계하여 가족 구성원 간의 관계를 연결하고 화목하게 만드는 측면에서 V4인 ‘평화롭고 정의로우며 상호연결된 사회 구축’이 담겨 있고, 로봇의 도움이 필요한 대상으로 할머니, 부모님, 동생 등 모든 가족 구성원을 고려한다는 측면에서 V3인 ‘다양성 및 포용성 보장’이 담겨 있다. 핵심 요건으로는 프라이버시 보호, 다양성 존중, 침해금지, 공공성, 연대성, 안전성, 투명성을 갖추고 있다. 먼저, 프라이버시 보호는 가족과 함께 생활하는 인공지능 로봇(AI)을 설계하기 때문에 가족의 사적인 대화나 생활을 보호할지 고민해야 하는 것에서 드러나고, 다양성 존중은 가족의 행복을 위해 다양한 가족 구성원의 필요와 생각을 골고루 존중하는 것에서

28 위의 책, 212쪽.

드러난다. 침해금지 는 가족의 행복에 도움이 되는 로봇을 만들기 때 문에 가족 간의 대화를 방해하거나 해를 끼치지 않도록 하는 것에서 드러나며, 공공성과 연대성은 ‘가족’이라는 작은 공동체의 공공 이익 이자 공동의 목표인 ‘가족의 행복’을 위해 인공지능(AI) 로봇을 설계 하는 것에서 드러난다. 안전성은 학생들이 로봇을 설계할 때 가족과 함께 지내기에 안전한지를 고려하는 것에서 드러나고, 투명성은 가 족들이 로봇의 기능과 행동을 쉽게 알도록 설계하는 것에서 드러난 다. 이러한 내용들은 ‘가정용 인공지능 로봇’이라는 텍스트, 인공지능 (AI) 로봇이 가족의 생활에 활용되고 있는 삽화, 각 활동에서의 활동 자료에서 볼 수 있으며, 각각의 활동들은 학생들이 가족의 행복을 위 해 인공지능(AI) 로봇을 비판적으로 탐구하며 활용하고자 고민해 보 는 경험을 제공한다.

4차시에서는 인공지능(AI) 로봇의 장단점을 생각해 보고 올바 르게 인공지능(AI) 로봇을 사용하는 태도를 기르도록 한다.²⁹ 활동으 로는 인공지능(AI) 로봇을 올바르게 사용해야 하는 까닭을 생각해 보 고, 인공지능(AI) 로봇의 장단점을 생각하며 이전 3차시 동안 배운 인 공지능(AI) 로봇들 중 내가 가지고 싶은 로봇을 고르고 까닭을 생각 한 뒤, 인공지능(AI) 로봇을 올바르게 사용하는 방법을 이야기 나눈 다. 단원 마무리에서도 인공지능(AI) 로봇을 올바르게 사용하는 방법 을 생각하며 정리하며 자신의 모습을 되돌아보도록 한다. 이 차시는 앞서 배운 내용들을 바탕으로 인공지능(AI) 로봇을 비판적으로 평가 하게 하여 V1, V3, V4를 담고 있다. 인공지능(AI)을 잘못 사용한 예 로는 인공지능(AI)에 너무 의존하여 ‘나’의 고유한 생각이나 판단력 이 약해지거나 가족 관계에 문제가 생길 수도 있고, 집안일을 엄마와

29 위의 책, 218쪽.

같은 특정인에게만 시키는 등 편견을 가질 수도 있다. 따라서 인공지능(AI) 로봇을 올바르게 사용하는 방법을 생각 및 정리하는 것은 V1, V3, V4를 침해하지 않는 것을 포함한다. 핵심 요건으로는 10대 핵심 요건을 모두 포함하고 있는데, 인공지능(AI) 로봇을 올바르게 사용하는 방법을 생각하는 것에서 인간의 존엄성과 인권을 침해하지 않는 사용(인권보장)과 사용자의 올바른 태도 기르기(책임성), 인공지능(AI) 로봇의 장단점을 파악하기 위해 인공지능(AI) 로봇이 어떻게 작동하고 어떤 영향을 미치는지 알기(투명성), 장점으로는 사회 전체와 공동체에 도움(공공성), 단점으로는 잘못 사용했을 때 불평등 심화 문제(연대성), 프라이버시 침해 문제(프라이버시 보호), 편견을 가지는 차별 문제(다양성 존중), 사람에게 해를 끼치는 문제(침해금지), 고장이나 위험해질 수 있는 안전 문제(안전성), 데이터 유출 문제(데이터 관리) 등에서 드러난다. 이러한 내용들은 각 활동에서의 활동 자료에서 볼 수 있으며, 각각의 활동들은 인공지능(AI) 로봇을 올바르게 사용하기 위해 비판적으로 탐구하도록 한다.

2) 4학년 도덕 교과서 고찰 결과

4학년에서도 자율형 단위인 8단원 ‘비무장 지대, 평화와 생명의 땅을 지켜요!’에서만 인공지능(AI) 윤리 내용을 다루고 있다. 이 단원은 5, 6, 7단원의 디지털 윤리, 통일, 자연의 가치를 융합하여 인공지능 윤리 원칙의 필요성을 생각하고 디지털 윤리에 대한 학습도 심화할 수 있도록 구성하였다.³⁰ 이 단원을 고찰한 결과는 다음과 같다.

30 교육부, 『도덕 4 교과용 지도서』, 비상교육, 2025, 95쪽.

〈표 5〉 4학년 8단원 고찰

차시명	내용	가치	핵심 요건	반영 유형	함양 여부
비무장 지대, 평화와 생명의 땅	비무장 지대가 평화와 생명의 땅인 까닭 탐색하기	X	X	X	X
비무장 지대를 지키는 인공지능 로봇을 설계해 봐요	비무장 지대를 지키는 인공지능 로봇 설계하기	V2, V4	침해금지, 공공성, 안전성	텍스트, 삽화, 활동 자료	O
비무장 지대를 지키는 인공지능 로봇에 필요한 행동 규칙은?	비무장 지대를 지키는 인공지능 로봇에게 행동 규칙이 필요함을 인식하기	V1, V2	인권보장, 침해금지, 책임성, 안전성	6컷 만화, 활동 자료	O
비무장 지대를 지키는 인공지능 로봇을 소개해 봐요	비무장 지대를 지키는 인공지능 로봇 개발 시 가져야 할 올바른 태도 내면화하기	V1, V2, V4	인권보장, 침해금지, 공공성, 책임성, 안전성, 투명성	활동 자료, 삽화, 텍스트	O
반영할 점	인공지능 로봇이 필요한 이유에 다양성 및 포용성 고려하기	V3	프라이버시 보호, 다양성 존중, 연대성, 데이터 관리		

1차시에서는 비무장 지대의 의미와 중요성을 인식하고, 비무장 지대가 평화와 생명의 땅인 까닭을 탐색하도록 한다.³¹ 비무장 지대의 가치인 ‘평화’와 ‘생명’을 탐색하고 이를 바탕으로 뒷차시에서 만들 인공지능(AI) 로봇이 지향해야 할 윤리적 목표를 제시한다고 한다.³² 활동으로는 ‘비무장 지대는 어떤 곳일까?’라는 영상을 통해 비무장 지대의 의미와 중요성을 파악하고, 글을 통해 비무장 지대를 평

31 위의 책, 310쪽.

32 위의 책, 306쪽.

화와 생명의 땅으로 부르는 이유를 파악하도록 하고, 평화 통일 후 생태환경이 풍부한 비무장 지대의 모습을 상상해보도록 한다. 이 차시는 살펴보면, 인공지능(AI) 로봇이나 윤리를 직접적으로 다루지 않는다.

2차시에서는 1차시에서의 비무장 지대를 바탕으로 비무장 지대에 필요한 인공지능(AI) 로봇의 생김새를 그려보고 설계하도록 한다.³³ 활동은 텍스트와 삽화를 통해 비무장 지대에서 어떤 로봇이 필요한지 떠올려보고, 비무장 지대를 지키는 인공지능(AI) 로봇을 설계하여 발표하도록 구성되어 있다. 여기서 비무장 지대를 평화롭게 지키기 위한 기능으로 멸종위기 동물을 식별하고 보호하는 기능, 동식물을 위협하는 산불을 감시하는 기능 등을 고려하는 것은 V2를 담고 있고, 포탄, 지뢰, 철조망을 안전하게 탐지하는 기능, 군사 활동을 감시하는 기능 등을 고려하는 것은 V4를 담고 있다. V2는 환경 및 생태계 보호의 윤리적 논제를 담고 있고, V4는 평화적인 방향으로 인공지능(AI)을 사용하도록 하기 때문이다. 이 차시에서는 침해금지, 공공성, 안전성의 핵심 요건을 갖추고 있다. 침해금지는 인공지능(AI) 로봇의 기능이 오히려 비무장 지대의 생태계를 해치거나 군사적 긴장을 유발하지 않도록 설계하는 것에서 드러나고, 공공성은 인공지능(AI) 로봇의 기능이 1차시에 배운 비무장 지대의 평화와 생명이라는 공공의 목적에 부합하도록 설계하는 것에서 드러나며, 안전성은 비무장 지대의 생태환경과 인간에게 안전한 로봇을 설계하는 것에서 드러난다. 이러한 가치와 핵심 요건을 반영한 활동들은 학생들이 비판적인 탐구를 통해 인공지능(AI) 윤리에 맞는 로봇을 설계하도록 한다.

3차시에서는 비무장 지대를 지키는 인공지능(AI) 로봇에게 행동 규칙이 필요함을 이해하고, 이에 따라 인공지능(AI) 로봇이 지켜야 할 행동 규칙을 탐구한다.³⁴ 활동은 6컷 만화를 통해 인공지능(AI) 로봇에게 행동 규칙이 필요함을 이해하고, 인공지능(AI) 로봇이 지켜야 할 행동 규칙을 활동 자료를 통해 탐구한 뒤, 친구들과 논의하여 수정하거나 보완하도록 구성되어 있다. 이 차시에는 앞서 설계된 인공지능(AI) 로봇에게 윤리적 제약과 원칙을 부여하므로, V1과 V2가 담겨 있다. 로봇이 비무장지대에서 인간(민간인, 연구원, 군인 등)을 만났을 때는 인간의 존엄성과 생명을 최우선으로 보호해야 한다는 규칙에서는 V1을, 로봇이 비무장지대에서 동식물을 해치지 않도록 하는 규칙에서는 V2를 담고 있기 때문이다. 핵심 요건으로는 인권보장, 침해금지, 책임성, 안전성을 갖추고 있는데, 인간을 다치게 하지 않는다는 규칙에서 인권보장이 드러나고, 비무장 지대의 자연과 평화를 침해하지 않도록 하는 규칙에서 침해금지가 드러난다. 또한, 행동 규칙은 인공지능(AI) 로봇이 잘못된 판단을 했을 때 책임의 근거가 되므로 책임성이 드러나며, 행동 규칙이 로봇의 안전한 임무 수행을 보장하는 것에서 안전성이 드러난다. 이러한 가치와 핵심 요건을 반영한 활동들은 학생들이 인공지능(AI) 로봇에게 필요한 행동 규칙을 설정하기 위해 비판적인 탐구를 하게 한다.

4차시에서는 비무장 지대를 지키는 인공지능(AI) 로봇을 소개하는 홍보물을 제작하고 이러한 로봇에게 행동 규칙이 중요한 까닭을 설명하도록 한다.³⁵ 활동은 앞 차시부터 만들었던 인공지능(AI) 로봇을 기능, 행동 규칙, 좋은 점을 활동 자료에 적어 소개 자료를 만드는 활동, 삽화와 같이 인공지능(AI) 로봇을 학급 친구들에게 소개하

34 위의 책, 318쪽.

35 위의 책, 326쪽.

는 활동, 내가 만든 로봇이 비무장 지대를 누비는 모습을 상상하여 신문 기사를 써보는 활동으로 구성되어 있다. 단원 마무리로는 인공지능(AI) 윤리를 생각해볼 수 있도록 활동 자료를 이용하여 로봇 개발자와 사용자에게 각각 올바른 개발과 사용에 대해 편지글을 써보고, 생각 놀이터에서는 ‘튜링 테스트’ 텍스트를 읽어보며 학생들이 인공지능(AI)의 사고에 대해 생각해보도록 하였다. 이 차시는 1, 2, 3차시의 모든 활동을 아우르고 정당화하는 종합 단계이므로, V1, V2, V4를 담고 있다. 이러한 내용은 핵심 요건으로 인권보장, 침해금지, 공공성, 책임성, 안전성, 투명성을 갖추고 있다. 홍보를 통해 이 인공지능(AI) 로봇이 비무장 지대에서 ‘평화와 생명’이라는 공공의 이익을 위해 일한다는 것을 널리 알리는 것에서 공공성이 드러나고, 인공지능(AI) 로봇의 행동 규칙을 설명하면서 이에 대한 책임을 갖는 것에서 책임성이 드러나며, 사람들에게 홍보물을 제작하여 인공지능(AI) 로봇에 대해 설명하고 알리는 것에서 투명성이 드러난다. 그 외 인권보장, 침해금지, 안전성은 3차시에서와 같이 인공지능(AI) 로봇에 정한 행동 규칙에서 드러난다. 이러한 가치와 핵심 요건을 반영하면서, 학생들이 내가 만든 인공지능(AI) 로봇을 홍보하면서 이를 활용할 수 있는 능력을 함양하고, 다른 친구들의 인공지능(AI) 로봇의 소개를 들으며 비판적 탐구 능력을 함양한다.

2. 특징 및 시사점

먼저, 3학년 고찰 결과, 가장 두드러지는 특징은 인공지능(AI) 윤리교육의 출발점을 학생들에게 가장 친숙한 주제인 ‘나’로 삼았다는 점이다. 1차시에서 ‘나와 같은 로봇’을 상상하며 ‘대체 불가능성’을 탐구하게 한 것은 인공지능(AI) 윤리의 가장 근본적 가치인 V1을

학생들 자신의 이야기로 풀어낸 것이다.

4단원의 1차시(나)에서는 ‘나’의 고유성과 존엄성, 프라이버시를 탐구하고, 2차시(나의 성장)에서는 ‘나의 일’, ‘나의 꿈’과 인공지능(AI) 로봇과의 관계를 설정하며 책임성을 배운다. 3차시(가족)에서는 ‘나’에서 ‘가족 공동체’로 윤리적 고려 범위를 확장하고, 4차시(사회)에서는 인공지능(AI) 로봇의 ‘올바른 사용’이라는 사회적 맥락에서 인공지능(AI) 로봇의 장단점을 비판적으로 종합한다. 이러한 나 → 나의 성장 → 가족 → 사회로 이어지는 관계의 확장은 3학년 학생들에게 인공지능(AI) 윤리에 대한 사고를 자연스럽게 심화시킨다. 이 단원에서 학생들은 능동적인 설계자이자 규칙 제정자이다. 2차시의 시켜도 되는 일과 시키면 안 되는 일에 대해 구분하기, 로봇 사용 원칙 만들기, 3차시의 가족의 행복을 위한 인공지능(AI) 로봇 기능 설정하기는 학생들이 인공지능(AI) 로봇의 바람직한 역할과 한계를 스스로 정의하며 윤리적 주체성을 기르도록 유도한다.

이 단원은 유네스코의 가치(V1, V3, V4)와 국가 10대 핵심 요건의 모든 요건을 자연스럽게 융합하였다. 이는 ‘나’와 ‘가족’이라는 친숙한 맥락 안에서 인공지능(AI)이 야기할 수 있는 복합적인 윤리 문제들을 3학년 학생들도 충분히 탐색할 수 있음을 보여준다. 다만, 이 단원에서 V2가 반영되어 있지 않음은 아쉬울 수 있는 부분이다. 물론, 자아 존중, 성실, 효·우애의 가치를 V2와 연결하기는 어렵겠지만, 3학년 학생의 수준에서 환경을 생각하면서 위의 가치들을 바탕으로 인공지능(AI) 활용할 수 있도록 한다면 더욱 좋을 것으로 보인다.

다음으로, 4학년 고찰 결과, 가장 두드러지는 특징은 ‘비무장 지대(DMZ)’라는 구체적이고 복합적인 맥락을 활용한다는 점이다. 인공지능(AI) 윤리를 추상적으로 가르치는 것이 아니라, ‘평화’와 ‘생명’이라는 가치 안에서 기술의 역할을 탐색하게 한다. 이는 인공지능

(AI) 윤리가 기술 자체의 문제가 아니라 인간과 사회의 가치를 실현하기 위한 도구라는 점을 명확히 한다.

이 단원은 유네스코의 가치(V1, V2, V4)와 국가 10대 핵심 요건(인권보장, 침해금지, 공공성, 책임성, 안전성, 투명성)을 자연스럽게 융합한다. 학생들은 비무장 지대를 지키는 인공지능(AI) 로봇을 설계하는 과정에서 이 두 가지 수준의 윤리 규범을 동시에 고려하게 된다.

다만, 8단원의 4개의 차시가 갖춘 기본적인 구조 단계인 Why-What-How-So What에서 아쉬운 부분이 있다. 1차시(Why)에서는 인공지능(AI) 로봇을 활용하는 ‘윤리적 목표’를 먼저 설정한다고는 하지만, 비무장지대의 내용은 직접적으로 인공지능(AI)을 다루는 것은 아니기 때문에 오히려 통일의 가치를 다루는 6단원인 ‘통일로 가는 한 걸음’에서 구성할 필요가 있다. 그 외 2차시(What)에서 설정된 가치를 구현하기 위한 인공지능(AI) 로봇의 기술적 기능을 설계하고, 3차시(How)에서 인공지능(AI) 기술이 가치를 침해하지 않도록 윤리적 제약인 행동 규칙을 부여하며, 4차시(So What)에서 완성한 윤리적 인공지능(AI) 기술을 사회와 소통하고 사회적 책임을 내면화하는 구성은 바람직하다. 따라서 1차시가 조금 아쉬운 4단계 구성으로 볼 수 있으므로, 1차시는 비무장지대에서 인공지능(AI) 로봇이 왜 필요한지와 이에 대한 다양한 사람들의 입장을 다루어 V3의 가치와 프라이버시 보호, 다양성 존중, 연대성, 데이터 관리의 핵심 요건도 반영할 필요가 있다. 활동 자료로는 다양한 사람들의 역할을 직접 맡아서 얘기할 수 있는 역할놀이 자료, 관련된 6컷 만화, 매체 자료 등을 제시할 수 있다.

이러한 3-4학년 도덕 교과서 고찰 결과의 특징을 살펴보았을 때 다음과 같은 시사점을 제시할 수 있다. 첫째, 인공지능(AI) 윤리교육은 초등 3, 4학년에서도 충분히 가능하다. 3학년에서는 학생들에게

가장 친숙한 소재를 활용하여 인공지능(AI) 윤리의 가장 근본적이고 기초적인 개념을 도입하였으며, 4학년에서는 ‘비무장 지대’라는 구체적인 사회·생태적 맥락에서 인공지능(AI) 윤리를 다룰 수 있음을 보여주었다. 둘째, 발달 단계를 고려한 체계적인 인공지능(AI) 윤리 설계 모델을 제시하고자 하였다. 3학년에서는 나, 나의 성장, 가족, 사회로 이어지는 관계 확장에 따른 인공지능(AI) 모델로 제시하였고, 4학년에서는 목표 정립(Why), 기능 설계(What), 행동 규칙(How), 사회적 책임(So What)로 이어지는 인공지능(AI) 윤리 설계 모델로 제시하였다. 단, 4학년에서는 1차시의 내용을 보완할 필요가 있었다. 셋째, 인공지능(AI) 윤리교육은 학생들에게 비판적 사용자이자 윤리적 설계자로서의 역량을 함양하도록 한다. 학생들은 단순히 인공지능(AI)의 장단점을 아는 것을 넘어, 인공지능(AI)의 윤리적 문제를 비판적으로 탐구하고, 인공지능(AI) 개발자로서 기술의 바람직한 역할과 한계를 고민하는 능동적인 윤리적 설계자로 참여한다. 넷째, 3학년, 4학년의 인공지능(AI) 윤리 교육은 내용과 가치를 보완할 필요가 있다. 3학년에서는 인공지능(AI) 로봇을 활용하는 환경을 고려하여 V2를 반영해야 하고, 4학년에서는 인공지능 로봇을 설계하고 활용하는 데 있어 V3를 반영해야 하고, 프라이버시, 다양성 존중, 연대성, 데이터 관리의 핵심 요건을 반영할 필요가 있다. 다섯째, 1개 학년에 1개 단원으로만 국한되어 있다. 인공지능(AI) 윤리교육은 2022 개정 도덕과 교육과정에 처음 도입되었다. 학생들에게 친숙하지 않은 주제라 어려울 수 있지만, 충분히 수업이 가능하다. 따라서 인공지능(AI) 윤리교육은 일회성보다는 지속적이고 통합적인 교육으로 확장될 필요가 있다.

IV. 결론

■ 본 연구는 2022 개정 교육과정에 따라 2025년부터 적용된 초등학교 3-4학년 도덕 교과서의 자율형 단원에 나타난 인공지능(AI) 윤리교육 요소를 고찰하였다. 유네스코(UNESCO)의 4대 가치와 국내 인공지능 윤리기준 10대 핵심 요건을 기준으로 텍스트, 삽화, 활동 자료를 고찰하였으며, 주요 결론 및 제언은 다음과 같다.

첫째, 2022 개정 교육과정 3-4학년 도덕 교과서는 인공지능(AI) 윤리교육을 학생들의 발달 단계에 맞춰 체계적으로 제시하고자 하였다. 3학년은 ‘나’로부터 시작하여 ‘나의 성장’, ‘가족’으로 이어지는 관계 확장 모델을 통해 인공지능(AI) 윤리의 기본 가치인 인간 존엄성(V1)과 프라이버시, 책임성 등을 탐구하도록 하였다. 4학년은 ‘비무장 지대(DMZ)’라는 구체적인 사회·생태적 맥락을 제시하고, Why→What→How→So What의 ‘윤리적 설계 모델’ 4단계를 적용하였다. 이는 초등학생 수준에서도 인공지능(AI) 윤리의 가치(V1, V2, V3, V4)와 조건들을 충분히 다룰 수 있음을 보여준다. 물론, 이에 대한 내용, 가치, 핵심 요건에 대한 보완은 필요하다.

둘째, 교과서는 학생들을 AI 기술의 수동적인 ‘소비자’나 ‘사용자’가 아닌, 기술의 목적과 한계를 비판적으로 탐구하고 바람직한 역할과 규칙을 스스로 정의하는 ‘윤리적 설계자’의 역할로 제시한다. 3학년의 ‘로봇 사용 원칙 만들기’, 4학년의 ‘로봇 행동 규칙 탐구’ 등의 활동은 학생들이 인공지능(AI)과 관련된 윤리적 딜레마 상황에서 주체적으로 판단하고 실천하는 역량을 기르도록 유도한다.

셋째, 고찰 결과 3-4학년의 인공지능(AI) 윤리교육은 각각 하나의 자율형 단원에 국한되어 제시되었다. 이는 인공지능(AI) 윤리교육을 공교육에 처음 도입하는 단계에서는 의미가 있으나, 자칫 일회성

교육으로 그칠 우려가 있다. 인공지능(AI)이 학생들의 일상에 깊숙이 스며든 만큼, 인공지능(AI) 윤리교육이 특정 단원에만 머무르지 않고 ‘존중’, ‘배려’, ‘정의’, ‘책임’의 도덕과 핵심 가치와 연관하여 다양한 단원에서 지속적이고 통합적으로 다루어질 필요가 있다.

넷째, 본 연구는 교과서의 내용 고찰에 중점을 두었다. 향후 이 교과서가 실제 교육 현장에 적용될 때, ‘교사들이 이 단원을 어떻게 재구성하여 가르치는지, 이러한 수업이 학생들의 인공지능(AI) 윤리 의식과 비판적 탐구 능력에 실질적인 변화를 가져오는지에 대한 수업 실행 연구 및 효과성 검증 연구’가 후속으로 이루어져야 할 것이다. 또한, 2026년부터 적용되는 5-6학년 도덕 교과서에서는 인공지능(AI) 윤리교육이 어떻게 연계되고 심화되는지, 또는 타 교과와 어떻게 통합적으로 다루어지는지에 대한 후속 연구를 기대해 본다.

- 과학기술정보통신부, 『사람이 중심이 되는 「국가 인공지능 윤리기준」』, 과학기술정보통신부, 2020.
- 교육부, 『인공지능시대 교육정책방향과 핵심과제』, 교육부, 2020.
- _____, 『도덕과 교육과정』 (교육부 고시 제2022-33호[별책 6]), 교육부, 2022.
- _____, 『도덕 3 교사용 지도서』, 비상교육, 2025.
- _____, 『도덕 4 교사용 지도서』, 비상교육, 2025.
- _____, 『초·중등학교 교육과정 총론』 (교육부 고시 제2022-33호[별책 11]), 교육부, 2022.
- 권누리, 「초등 도덕과에서의 인공지능윤리교육 방안 연구: 인공적 도덕 행위자(AMA) 논의를 중심으로」, 『초등도덕교육』, 87(2024), 87~108쪽.
- 김하민, 「2022 개정 초등 도덕 교육과정에 대한 비판적 고찰: 인간과 인공지능 로봇 관계 내용 서술을 중심으로」, 『도덕윤리과교육』, 79(2023), 45~49쪽.
- 박형빈, 「규범적 AI윤리 가이드라인과 초등 AI윤리교육 적용 방안 모색」, 『한국초등교육』, 121(2024), 88~91쪽.
- _____, 「뇌-컴퓨터 인터페이스(BCI)의 인간 증진 및 윤리적 쟁점과 초등도덕 AI 윤리교육 활용 방안」, 『초등도덕교육』, 91(2025), 260~264쪽.
- _____, 「생성 AI(Generative AI)의 할루시네이션(hallucination) 이슈와 윤리적 과제: 초등 AI 윤리교육에서 활용 가능한 주제를 중심으로」, 『한국초등교육』, 119(2023), 32~33쪽.
- 신현화·신태수·신현우, 「초등학교 도덕과 수업에서 인공지능 윤리교육 방안 탐색 연구」, 『학교와 수업 연구』, 7(2022), 77~94쪽.

이상욱, 『유네스코 인공지능(AI) 윤리 권고 해설서: 인공지능 윤리 이해하기』, 유네스코한국위원회, 2021.

이인재, 「AI 윤리 기준을 활용한 AI 윤리교육 방안 연구」, 『초등도덕교육』, 76(2022), 157-180.

이지형·변순용, 「인공지능 윤리교육 프로그램의 효과 검증: 초등 도덕과 교육 적용을 중심으로」, 『초등도덕교육』, 84(2023), 288~305쪽.

홍현주, 「인공지능 윤리교육의 초등도덕 교육과정 적용 방안」, 『초등도덕교육』, 75(2021), 191~196쪽.

UNESCO, 「AI 윤리에 대한 권고 초안_국문번역본」, 제2차 정부 간 회의, 2021년 6월 25일, 파리.

_____, “Recommendation on the Ethics of Artificial Intelligence”, 41st session of the General Conference, 2021.11.23., Paris.



■ 본 연구는 2025년부터 적용되는 2022 개정 교육과정 기반 초등 학교 3-4학년 도덕 교과서에 인공지능(AI) 윤리교육 요소가 어떻게 구현되었는지를 글로벌 표준 규범인 유네스코(UNESCO)의 4대 가치와 국내 ‘인공지능 윤리기준’의 10대 핵심 요건을 활용하여 고찰하였다.

고찰 결과, 2022 개정 도덕 교과서는 초등학생의 발달 단계를 고려한 체계적인 접근을 보여주하고자 하였다. 3학년은 ‘관계 확장 모델’을, 4학년은 ‘비무장 지대(DMZ)’라는 구체적 맥락에서 ‘윤리적 설계 모델’을 적용하여 가치 및 핵심 요건을 탐구하게 하였으나, 보완할 부분이 있었다. 그럼에도 가장 큰 특징은 학생들을 인공지능(AI) 기술의 목적과 한계를 비판적으로 탐구하고 규칙을 직접 정해보는 ‘윤리적 설계자’로 제시한다는 점이다.

2022 개정 초등 3-4 도덕 교과서의 인공지능(AI) 윤리교육은 일부 보완할 점이 있으나, 자연스럽게 도입했음을 보여준다. 다만, 인공지능(AI) 윤리교육이 학년별 1개의 자율형 단원에 국한되어 있어 자칫 일회성 교육에 그칠 우려가 있다. 따라서 향후 인공지능(AI) 윤리교육이 도덕과 핵심 가치와 연계되어 지속적이고 통합적으로 다루어져야 함을 기대한다.

주제어: 인공지능(AI) 윤리, 2022 개정 교육과정, 3-4 도덕 교과서, 초등 도덕 교육



A Study on the Elements of Artificial Intelligence (AI) Ethics Education in Elementary Grade 3–4 Ethics Textbooks: Based on the 2022 Revised Curriculum

Cho, Kwang Won

(Seoul Sadang Elementary School · Moral Education · Teacher)

■ This study examines how elements of artificial intelligence (AI) ethics are implemented in moral education textbooks for Grades 3 and 4 under the 2022 revised curriculum, which is scheduled for implementation in 2025. The analysis draws on UNESCO’s four core values as a global standard, along with the core requirements of *Korea’s Artificial Intelligence Ethics Standards*.

The findings indicate that the revised 2022 textbooks attempt a systematic approach that reflects the developmental stages of elementary school students. In Grade 3, the “expanded relationship model” is applied, whereas in Grade 4, the “ethical design model” is introduced within the specific context of the Demilitarized Zone (DMZ) to explore key values and requirements. Despite

these strengths, several areas for improvement remain. Notably, the textbooks position students as “ethical designers” who critically examine the purposes and limitations of AI technologies and actively participate in establishing ethical guidelines.

Overall, AI ethics education in the revised Grades 3–4 textbooks is introduced naturally, although some limitations persist. In particular, the content is confined to a single autonomous unit per grade, raising concerns that it may function as a one-off lesson rather than a sustained component of the curriculum. Accordingly, there is a need for AI ethics education to be more consistently and systematically integrated across the curriculum, with stronger connections to broader moral values and ethical competencies.

Keywords: Artificial Intelligence (AI) Ethics, 2022 Revised Curriculum, Grades 3–4 Moral Textbooks, Elementary Moral Education

- ◎ 투고 접수 2026. 03. 09.
- ◎ 심사 완료 2026. 04. 20.
- ◎ 게재 결정 2026. 04. 23.