

생성형 AI의 비판적 활용을 위한 토론 기반 수업 설계 연구

안윤미*

국문초록 본 연구는 생성형 AI 시대에 요구되는 비판적 활용 역량 함양을 목표로 한, 국어과 토론 기반 수업을 설계하는 데 목적이 있다. 생성형 AI는 정보 탐색, 글쓰기, 토론 준비 등의 다양한 학습 활동을 지원하지만, 환각 현상과 정보 신뢰성 문제를 내포하고 있다. 이에 본 연구에서는 생성형 AI의 비판적 활용 역량을 생성형 AI의 생성 원리와 한계에 대한 이해, 정보의 사실성과 신뢰성 검증, 근거의 비교·평가 및 관점의 재구성 능력을 포함하는 개념으로 정의하였다. 이를 바탕으로 생성형 AI의 이해, 정보 검증, 비판적 활용의 과정을 단계적으로 경험할 수 있도록 총 3차시의 토론 기반 수업을 설계하였다. 1차시에서는 생성형 AI의 개념과 특성을 이해하고 응답의 가변성을 탐색하도록 하였으며, 2차시에서는 환각 현상과 정보 검증 방법을 학습하도록 하였다. 3차시에서는 기후 위기를 주제로 생성형 AI가 제시한 논거를 검증하고 이를 토론 활동에 활용하도록 구성하였다. 본 연구는 생성형 AI를 비판적 검토의 대상으로 설정하고, 이해·검증·토론·성찰 등의 과정을 연계한 토론 기반 수업 모형과 차시별 지도 방안을 설계·제안하였다는 점에 의의가 있다.

주제어: 생성형 AI, 비판적 활용 역량, 비판적 문식성, 정보 검증, 토론 수업, 국어과 교육과정

목차

- I. 머리말
- II. 이론적 배경
- III. 생성형 AI의 비판적 활용을 위한 토론 기반 수업 설계
- IV. 생성형 AI의 비판적 활용을 위한 토론 기반 수업의 실제
- V. 결론

논문접수일: 2026.06.18.

논문수정일: 2026.07.16.

게재확정일: 2026.08.01.

I. 머리말

최근 사회 전반에 확산된 생성형 인공지능(Generative Artificial Intelligence)은 교육 현장에서의 교수·학습 환경에도 큰 변화를 가져오고 있다. 생성형 AI는 사용자의 질문에 자연어로 응답하고 다양한 형태의 결과물을 생성할 수 있다는 점에서 기존의 디지털 기술과 차별화된다. 특히 정보 탐색, 자료 요약, 글쓰기, 발표 준비, 토론 자료 수집 등 다양한 학습 활동에 활용되면서 교육적 가능성에 대한 관심이 높아지고 있다.

최근 연구들에서도 생성형 AI가 학습자의 사고를 지원하고 학습 경험을 확장하는 교수·학습 도구로 활용될 수 있음을 보여주고 있다. 생성형 AI는 학습자의 수준과 요구에 맞는 정보를 제공하고 아이디어

* 서울청원여자고등학교 국어 교사 / yoonmia@hanmail.net

생성과 문제 해결 과정을 지원함으로써 자기주도적 학습을 촉진할 수 있다. 이러한 긍정적 영향으로 인해 학교 현장에서도 생성형 AI를 활용한 다양한 수업 사례와 교수·학습 방안이 제안되고 있다.

그러나 생성형 AI의 확산은 새로운 교육적 과제를 제기하기도 한다. 생성형 AI는 사실을 이해하거나 판단하는 존재가 아니라 확률적으로 가장 적절한 언어를 생성하는 기술이므로 환각(hallucination), 편향, 정보 신뢰성 등의 문제를 내포하고 있다. 따라서 생성형 AI가 생산한 정보는 검증 과정을 필요로 하며, 이를 무비판적으로 수용할 경우 잘못된 지식 형성으로 이어질 가능성이 있다. 이러한 점에서 생성형 AI 시대의 교육은 생성형 AI를 활용하는 방법을 가르치는 것과 더불어, 생성형 AI가 생산한 정보의 정확성을 검토하고 평가하는 비판적 활용 역량을 함양하도록 하는 방향으로 나아가야 할 필요가 있다. 즉, 생성형 AI를 효과적으로 활용하기 위해서는 생성형 AI의 응답을 그대로 수용하는 것이 아니라 다양한 자료와 비교·검토를 통해 정확성과 타당성을 판단할 수 있어야 한다. 특히 최근에는 생성형 AI가 생산한 텍스트가 학습자의 새로운 읽기 대상이 되기도 한다는 점에서 생성형 AI의 응답을 비판적으로 분석하고 평가하는 능력은 무엇보다 중요한 교육적 과제로 보인다.

이러한 관점에서 토론 활동은 생성형 AI의 비판적 활용 역량을 함양할 수 있는 효과적인 교수·학습 방법이 될 수 있다. 토론은 다양한 관점을 비교·검토하고 주장과 근거의 타당성을 평가하는 과정을 포함한다는 점에서, 생성형 AI의 응답을 비판적으로 분석하고 검증하는 활동과 밀접하게 연결된다. 따라서 생성형 AI가 생산한 정보를 토론의 대상으로 활용할 경우, 학습자는 정보의 신뢰성과 타당성을 검토하며 비판적 사고력을 기르게 된다.

본 연구는 생성형 AI의 비판적 활용 역량 함양을 위한 교수·학습 방법으로 토론 활동을 제안하고자 한다. 생성형 AI의 생성 결과물은 환각 현상과 정보의 신뢰성 문제를 내포하고 있으므로, 학습자는 생성형 AI의 응답을 무비판적으로 수용하기보다 그 정확성과 타당성을 검토하고 평가할 수 있어야 한다. 이에 본 연구는 생성형 AI의 확률적 문장 생성 특성과 환각 현상에 대한 이해를 바탕으로 학습자가 생성형 AI 응답의 오류와 한계를 비판적으로 검토할 수 있는 교육 방안을 모색하고자 한다.

이를 위해 ‘기후 위기’와 같은 사회적 쟁점을 중심으로 생성형 AI가 제시한 다양한 관점과 논거를 비교·분석하고, 그 정확성과 타당성을 검증하는 토론 기반 수업 모형을 설계하고자 한다. 생성형 AI가 생산한 정보를 검토하고 교차 검증하는 과정을 통해 학습자가 정보의 신뢰성을 판단하며, 토론 과정에서 제시된 주장과 근거를 비판적으로 평가할 수 있도록 한다. 또한 토론 이후에는 생성형 AI의 응답과 토론 과정에서 도출된 쟁점을 바탕으로 자신의 관점을 재구성하는 성찰적 활동을 수행함으로써 주체적인 판단 능력을 형성하도록 한다.

본 연구의 목적은 학생들이 생성형 AI를 정보 제공 도구에서 나아가 비판적 검토와 성찰의 대상으로 인식하고, 토론 활동을 통해 학습자가 생성형 AI의 생성물을 검증·평가하며 이를 주체적으로 활용할 수 있도록 하는 수업을 설계하는 데 있다. 또한 토론 활동의 기반이 되는 단계별 학습 과정을 제시하고, 교사가 학습자의 검증·평가 활동을 확인할 수 있는 기준을 함께 제안하고자 한다. 아울러 이를 통해, 생성형 AI 시대에 요구되는 비판적 활용 역량 함양을 위한 교육의 방향을 모색해 보고자 한다.

II. 이론적 배경

1. 생성형 AI의 교육적 활용과 한계

생성형 AI의 등장 이후 교육 현장에서는 생성형 AI를 교수·학습 도구로 활용하려는 시도가 확대되고 있다. 생성형 AI는 정보 탐색, 자료 요약, 글쓰기, 토론 자료 수집 등 다양한 학습 활동을 지원할 수 있으며, 학습자의 자기주도적 학습을 촉진하는 교육적 가능성을 지닌다. 이러한 가능성은 여러 선행 연구를 통해 확인된다. 조용림(2025)은 생성형 AI가 글쓰기, 발표, 토론 등 언어 활동 전반에 영향을 미칠 수 있으며, 사고의 조직과 표현, 소통 과정을 지원하는 교육적 도구가 될 수 있다고 보았다. 또한 오수경, 신미성, 임다미(2025)는 국내 대학 교육에서의 생성형 AI 활용 연구를 분석한 결과, 생성형 AI가 학습 효과와 학습 효율성 향상에 긍정적인 영향을 미친다고 보았다. 이러한 연구 결과는 학교 현장의 활용 양상에서도 확인된다. 안윤미, 김유하, 이서희(2025)의 중·고등학교 생성형 AI 활용 실태 연구에서는 생성형 AI가 학교 현장에 빠르게 확산되고 있으며, 교사들은 수업 자료 제작, 교수·학습 활동 지원 등 다양한 목적으로 생성형 AI를 활용하고 있는 것으로 나타났다. 반면, 정보의 신뢰성, 표절과 저작권 문제, 평가의 공정성 확보 등은 생성형 AI 활용 과정에서의 주요 우려 사항으로 나타났다.

이처럼 생성형 AI는 교육적으로 다양한 가능성을 지니고 있지만, 동시에 활용 과정에서 해결해야 할 과제도 안고 있다. 생성형 AI는 사실을 이해하거나 판단하는 존재가 아니라 대규모 언어 모델에 기반하여 확률적으로 언어를 생성하는 기술이다. 따라서 생성형 AI의 응답은 사실적 정확성보다 문장의 개연성과 유창성이 우선될 수 있으며, 존재하지 않는 정보를 사실처럼 제시하는 환각 현상을 보이기도 한다. 따라서 생성형 AI를 교육적으로 활용하기 위해서는 생성된 정보를 그대로 수용하기보다 정보의 정확성과 신뢰성을 검토하고, 다양한 자료를 통해 그 타당성을 확인하는 과정이 필요하다.

2. 생성형 AI 시대의 비판적 활용 역량

생성형 AI의 확산과 함께 문식성의 개념도 새롭게 논의되고 있다. 생성형 AI 시대에는 인간이 작성한 텍스트뿐만 아니라, 생성형 AI가 생산한 텍스트를 읽고 해석하며 그 정확성과 신뢰성을 판단하는 능력이 중요해지고 있다. 생성형 AI가 인간의 언어 생산 과정에 개입하면서 기존의 읽기와 쓰기 개념이 변화하고 있으며, 생성형 AI가 생산한 텍스트를 비판적으로 읽고 해석하는 능력이 중요해지고 있다(오규설, 2024).

최근에는 AI가 생성한 텍스트 자체를 새로운 문식성의 대상으로 바라보려는 논의도 나타나고 있다. 생성형 AI가 생산한 텍스트는 인간 저자가 작성한 전통적 텍스트와 구별되는 새로운 읽기 대상으로

이해되며, 이를 ‘AI 텍스트’로 개념화하기도 한다(오선주, 이경화, 2025). AI 텍스트의 특징으로는 비저작성, 불투명성, 환각이 제시된다. 이에 따르면 AI 텍스트는 특정 저자의 의도에 의해 생산된 것이 아니라 방대한 데이터의 통계적 패턴에 기반하여 생성되며, 생성 과정 역시 블랙박스 형태로 작동하기 때문에 독자가 그 과정을 직접 확인하기 어렵다. 또한 존재하지 않는 사실을 실제 정보처럼 제시하는 환각 현상을 내포한다(오선주, 이경화, 2025).

이러한 특성을 통해 생성형 AI 시대의 문식성은 사실성 검증과 정보 평가 능력을 포함해야 함을 알 수 있다. 생성형 AI 활용 수업에서는 생성형 AI 응답에 대한 비판적 검증과 학습 과정에 대한 성찰이 핵심 원리로 제시되고 있으며(오유나, 김도연, 윤영현, 황혜지, 2025), 교사 역시 생성형 AI의 응답을 그대로 수용하기보다 수정·재구성하는 것으로 나타났다(백민경, 박찬홍, 2025). 이는 생성형 AI의 응답이 정답이 아니라 검토와 판단의 대상임을 보여주는 결과라 할 수 있다.

이상의 논의를 바탕으로 본 연구에서는 생성형 AI의 비판적 활용 역량을 생성형 AI의 생성 원리와 한계에 대한 이해, 정보의 사실성과 신뢰성 검증, 근거의 비교·평가 및 관점의 재구성 능력을 포괄하는 통합적 능력으로 보고, 이를 인식적 기반, 분석적 수행, 논증적 수행의 세 가지 차원으로 구체화하였다.

인식적 기반은 생성형 AI의 생성 원리와 한계를 이해하여 생성형 AI의 응답을 절대적인 지식이 아니라 확률적으로 생성된 결과물로 인식하고, 환각 가능성을 고려하며 비판적으로 접근하는 태도를 의미한다. 이는 AI 리터러시에서 강조하는 생성형 AI의 생성 원리와 한계에 대한 이해, 그리고 비판적 문식성에서 강조하는 비판적 수용에 관한 논의와 관련된다. 분석적 수행은 생성형 AI 응답의 오류와 편향을 식별하고, 출처 확인, 교차 검증, 최신성 점검 등을 통해 정보의 사실성과 신뢰성을 평가하는 능력을 의미한다. 이는 정보의 출처와 신뢰성을 평가하고 다양한 자료를 통해 정보를 검증하는 정보 리터러시의 논의와 관련된다. 논증적 수행은 검증된 정보와 근거를 비교·평가하여 자신의 입장을 구성하고, 토론 과정에서 다양한 관점을 검토하며 자신의 견해를 성찰하고 재구성하는 능력을 의미한다. 이는 국어과의 논증 교육에서 강조하는 근거의 타당성 평가와 논증의 재구성에 관한 논의와 관련된다.

이와 같이 본 연구의 비판적 활용 역량은 생성형 AI의 생성 특성을 이해하는 인식적 기반을 바탕으로 정보를 분석·검증하고, 이를 토론과 성찰 과정에서 자신의 판단으로 재구성하는 단계적이고 통합적인 역량으로 설정하였다.

3. 토론 활동과 생성형 AI의 비판적 활용

토론은 다양한 관점을 비교·검토하고 주장과 근거의 타당성을 평가하는 과정에서 비판적 사고력을 함양하는 대표적인 교수·학습 방법이다. 생성형 AI 기반 토론 수업 역시 학습자의 비판적 사고력, 문제 해결력, 의사 결정력을 신장시키는 데 효과적인 교수·학습 방법으로 제안되고 있다(이세미, 이동형, 박창균, 2025).

최근에는 생성형 AI를 토론 수업의 보조 도구로 활용하는 방안도 제안되고 있다. 생성형 AI는 토론

준비 과정에서 자료 탐색과 논거 생성을 지원하고 AI 리터러시 향상에 긍정적인 영향을 미치는 것으로 나타났으며(박광길, 2025), 자료 탐색과 논증 구성을 지원하는 도구로도 활용될 수 있다(허선아, 2025).

한편 ‘기후 위기’와 같은 사회적 쟁점은 생성형 AI의 비판적 활용을 위한 토론 수업에 적합한 주제라 할 수 있다. 토의·토론 기반 기후 변화 교육은 학생들의 비판적 사고력 향상에 긍정적인 영향을 미치며, 기후 변화와 같은 복합적 사회 문제를 다룰 때 토론 활동은 정보 분석과 가치 판단 능력을 함양하는 데 효과적이다(최은경, 이용섭, 2025).

지금까지의 연구들을 살펴보면, 생성형 AI를 토론 준비를 지원하거나 토론 수행을 보조하는 도구로 활용하는 데 주로 초점을 두고 있다. 그러나 생성형 AI가 생산한 정보 자체를 비판적으로 검토하고 평가하는 과정을 토론 활동과 연계하여 체계적으로 다룬 연구는 상대적으로 부족한 편이다. 생성형 AI의 비판적 활용 역량은 학습자가 생성형 AI를 활용하는 과정에서 자연스럽게 형성되기 어렵다. 또한 생성형 AI를 비판적으로 활용해야 한다는 안내 수준의 지도만으로는 학습자가 그 필요성과 방법을 내면화하기 어렵다. 따라서 학습자가 생성형 AI의 응답 생성의 특성을 이해하고, 생성형 AI가 제시한 정보의 신뢰성을 검증하며, 검증된 정보를 바탕으로 다양한 관점과 근거를 비교·평가할 수 있도록 하는 교육이 필요하다.

토론 활동은 이러한 과정을 자연스럽게 포함한다. 학습자는 토론 과정에서 생성형 AI가 제시한 주장과 근거를 검토하고 상대방의 논증을 비판적으로 평가하게 되며, 이를 통해 생성형 AI의 비판적 활용 역량을 함양할 수 있다. 이에 본 연구에서는 생성형 AI의 생성 원리에 대한 이해, 정보 검증, 토론, 성찰의 과정을 단계적으로 구성한 토론 기반 수업을 설계하고자 한다.

III. 생성형 AI의 비판적 활용을 위한 토론 기반 수업 설계

1. 수업 설계의 방향

생성형 AI는 정보 탐색, 자료 요약, 글쓰기, 발표 준비 등 다양한 학습 활동에 활용되며, 이미 교육 현장의 주요 디지털 도구로 자리 잡고 있다. 그러나 생성형 AI는 인간처럼 사실을 이해하거나 판단하는 존재가 아니라 대규모 언어 모델을 기반으로 다음에 올 단어를 확률적으로 예측하여 문장을 생성하는 기술이다.

생성형 AI는 확률론적 언어 모델을 사용해 입력한 일련의 글 다음에 올 단어의 확률을 이전 단어들을 기반으로 추정하는 조건부 확률을 사용한다. 또한 자동회귀 자기 지도학습으로 이전에 만든 단어들의 정보를 사용해 다음에 올 단어를 예측하며, 예측한 값과 실제 값 간에 존재하는 오차를 확률적 경사

하강법(Stochastic Gradient Descent) 등으로 최소화한다(임종섭, 2024). 따라서 생성형 AI는 사실과 다른 정보를 그럴듯하게 제시하는 환각 현상을 보일 수 있으며, 논리적으로 설득력 있어 보이는 내용이라 하더라도 그 정확성과 타당성이 항상 보장되는 것은 아니다.

그럼에도 불구하고 학교 현장에서 생성형 AI 활용 교육은 프롬프트 작성이나 결과물 생성과 같은 기능적 활용에 초점을 두는 경우가 많다. 학생들은 생성형 AI를 활용하여 정보를 수집하고 글을 작성하지만, 생성형 AI가 생산한 정보의 신뢰성을 검토하거나 논리적 타당성을 평가하는 것에 대한 교육은 구체적이지 않거나 단발적 안내 수준인 경향이 있다. 그러나 생성형 AI 시대에 필요한 역량은 AI의 활용 능력이 아니라 생성형 AI가 생산한 정보를 검증하고 평가하여 주체적으로 활용할 수 있는 능력이라는 점에서 이에 대한 교육이 필요하다.

본 연구는 고등학생을 대상으로 생성형 AI의 비판적 활용 역량을 함양하기 위한 토론 기반 수업을 설계하였다. 생성형 AI의 비판적 활용은 생성형 AI가 제시한 정보 검증뿐만 아니라, 다양한 관점과 근거를 비교·평가하여 자신의 판단을 수정하고 재구성하는 과정을 포함한다. 이러한 활동은 학습자 간 상호작용을 바탕으로 주장과 근거를 검토하는 토론 활동을 통해 효과적으로 이루어질 수 있다.

본고에서 토론 기반 수업을 생성형 AI 비판적 활용 교육에 적용하는 이유는 다음과 같다.

첫째, 토론 활동은 국제기구가 강조하는 인간 중심의 AI 활용 역량을 실천할 수 있는 교수·학습 방법이 될 수 있다. UNESCO(2024)는 미래 학습자가 갖추어야 할 핵심 역량으로 인간 주체성(Human Agency), 비판적 사고(Critical Thinking), AI 윤리(AI Ethics)를 제시하였다. 즉 학습자는 AI가 생성한 정보를 수동적으로 수용하는 존재가 아니라 그 결과를 검토하고 평가하며 책임 있게 활용할 수 있는 주체로 성장해야 한다는 것이다. 토론 활동은 생성형 AI의 응답을 그대로 받아들이는 것이 아니라, 동료 학습자들과 함께 검증하고 논의하는 과정을 제공한다는 점에서 인간 중심의 AI 활용 역량을 실천하는 교육적 방안이 될 수 있다.

둘째, 토론 활동은 생성형 AI 응답에 대한 검증의 필요성을 경험하게 한다. 생성형 AI가 제시한 정보는 외형적으로 완결성과 설득력을 갖추고 있어 사실로 받아들여지기 쉽다. 그러나 토론에서는 주장과 근거의 타당성을 입증해야 하므로 생성형 AI가 제시한 통계 자료나 사례, 인용 자료의 사실 여부를 확인하는 과정이 필수적으로 요구된다. 이 과정에서 학습자는 생성형 AI의 응답을 정답이 아니라 검증의 대상으로 인식하게 된다.

셋째, 토론 활동은 생성형 AI 응답의 오류와 한계를 발견할 수 있는 사회적 검증 환경을 제공한다. 개별적으로 생성형 AI의 응답을 검토할 경우에는 자신의 배경지식이나 관점에 따라 특정 오류를 발견하지 못할 수 있다. 반면 토론 과정에서는 서로 다른 관점과 경험을 가진 학습자들이 동일한 정보를 함께 검토하고 의견을 교환하므로 생성형 AI 응답의 오류와 편향을 보다 효과적으로 발견할 수 있다.

넷째, 토론 활동은 생성형 AI가 제시한 정보를 근거 중심으로 평가하는 태도를 형성하게 한다. 생성형 AI는 설득력 있는 문체와 논리적 구조를 갖춘 응답을 생성하기 때문에 학습자는 정보의 출처보다 표현의 완성도에 영향을 받을 수 있다. 그러나 토론에서는 주장 자체보다 그것을 뒷받침하는 근거의

신뢰성과 타당성이 중요하게 다루어진다. 따라서 학습자는 생성형 AI가 생성한 자료의 출처와 논거의 타당성을 평가하면서 비판적 사고력을 함양할 수 있게 된다.

다섯째, 토론 활동은 생성형 AI 활용 과정에 대한 성찰적 사고를 촉진한다. 학습자는 토론 이후 자신의 판단 과정을 되돌아보며 어떠한 기준으로 정보를 선택하고 검증하였는지를 성찰하게 된다. 또한 생성형 AI의 정보가 지닌 한계와 가능성을 함께 검토함으로써 생성형 AI의 정보를 수동적으로 받아들이는 것이 아니라 비판적으로 활용해야 할 대상으로 인식하게 된다.

이와 같은 토론 활동은 생성형 AI의 생성 특성에 대한 이해를 바탕으로 정보를 검증하고, 다양한 관점과 근거를 비교·평가하며, 자신의 판단을 성찰하는 활동을 통합적으로 수행할 수 있다. 또한, 생성형 AI의 이해에서 출발하여 정보 검증, 근거 중심 판단, 성찰적 활용으로 이어지는 단계적 학습 과정을 실현할 수 있다는 점에서 생성형 AI에 대한 비판적 활용 역량 함양을 위한 효과적인 교수·학습 방법이 될 수 있다.

2. 제재 및 성취기준

1) 제재

본 수업은 생성형 AI의 비판적 활용 역량 형성을 목적으로 하므로, 특정 교과 지식의 습득보다는 생성형 AI의 응답 생성 방식과 정보 활용 과정을 이해하는 데 중점을 두었다. 이에 따라 1차시에서는 특정 주제를 교사가 일괄적으로 제시하기보다 학생들이 일상생활, 진로, 학교생활, 사회 문제 등 자신이 관심 있는 주제를 자유롭게 선정하도록 한다. 다만 모둠 내에서는 동일한 질문을 설정하고, 모둠 간에는 서로 다른 질문을 설정하여 생성형 AI의 응답에 대해 다양하게 비교·분석할 수 있도록 한다. 예를 들면, 학생들은 선정한 질문을 ChatGPT, Gemini, Claude 등 서로 다른 생성형 AI에 입력하고 생성된 응답의 내용과 표현 방식을 비교한다. 이후 모둠 내에서 응답의 공통점과 차이점을 분석하고, 다른 모둠의 결과와 공유하면서 생성형 AI가 동일한 질문에 대해서도 다양한 응답을 생성할 수 있음을 확인한다. 이러한 활동은 학생들이 생성형 AI의 응답을 절대적인 정답이 아니라 확률적으로 생성된 결과물로 이해할 수 있게 한다.

2차시에서는 1차시 활동 과정에서 생성된 응답을 활용하여 생성형 AI의 환각 현상과 정보 검증 방법을 학습하도록 한다. 학생들은 생성형 AI가 제시한 정보 가운데 사실 여부를 확인할 필요가 있는 내용을 선정하고, 출처 확인과 교차 검증을 하면서 생성형 AI 응답의 신뢰성을 판단하는 경험을 하게 된다. 이를 통해 생성형 AI의 응답을 그대로 수용하기보다 검토와 검증의 대상으로 인식하도록 한다.

3차시의 토론 활동에서는 사회적 쟁점을 제재로 선정한다. 2022 개정 교육과정 총론에서는 핵심역량으로 다양한 영역의 지식과 정보를 비판적으로 탐구하고 활용하는 지식정보처리 역량과, 지역·국가·세계 공동체의 지속 가능한 발전에 책임감 있게 참여하는 공동체 역량을 제시하고 있다(교육부,

2022a: 6). 이에 본 연구에서는 다양한 사회적 쟁점 가운데 기후 위기를 토론 제재로 선정하였다.¹⁾ 기후 위기는 환경 문제만이 아닌 경제, 산업, 에너지 정책, 세대 간 책임, 국제 협력 등 다양한 관점이 복합적으로 얽혀 있는 대표적인 사회적 쟁점으로서 찬반 논의와 가치 판단이 활발하게 이루어질 수 있는 주제이다. 또한 인터넷과 생성형 AI를 통해 방대한 양의 정보와 논거를 쉽게 수집할 수 있는 반면, 정보의 정확성과 신뢰성을 검증해야 할 필요성도 큰 주제이기도 하다. 따라서 학생들은 생성형 AI가 제시한 기후 위기 관련 정보와 논거를 검토하고 사실 여부를 확인하는 과정에서 정보의 신뢰성과 근거의 타당성을 평가할 수 있으며, 토론 과정에서는 서로 다른 관점과 근거를 비교·검토하며 자신의 입장을 형성하고 재구성할 수 있게 된다. 이러한 점에서 기후 위기는 생성형 AI의 비판적 활용 역량과 토론 능력을 함께 함양하기에 적절한 제재라고 볼 수 있다.

이와 같이 본 수업의 제재는 1차시의 자유로운 질문 생성 활동에서 출발하여 2차시의 정보 검증 활동, 3차시의 기후 위기 토론으로 연결되도록 구성하였다. 이러한 단계적 구성은 생성형 AI의 특성을 이해하는 단계에서 시작하여 생성형 AI가 생산한 정보를 평가하고 활용하는 단계로 학습을 심화하기 위한 것이다.

2) 성취기준

본 연구의 수업은 고등학교 2022 개정 교육과정 「공통국어1」의 성취기준과 교수·학습 방향을 바탕으로 설계하였다. 특히 듣기·말하기 영역의 [10공국1-01-02], 읽기 영역의 [10공국1-02-01], 쓰기 영역의 [10공국1-03-01]을 중심으로 생성형 AI의 비판적 활용 과정을 토론 활동과 연계하였다. 본 연구는 「공통국어1」의 듣기·말하기, 읽기, 쓰기 성취기준을 핵심 기준으로 설정하였다. 또한 「독서 토론과 글쓰기」와 「주제 탐구 독서」의 성취기준은 본 수업 모형을 관련 선택 과목으로 확장하여 적용할 수 있는 참고 기준으로 함께 제시하였다. 이는 <표 1>과 같다.

먼저 본 수업은 생성형 AI가 제시한 정보를 검토하고 이를 토대로 찬반 논거를 구성하며 토론하는 과정에서 [10공국1-01-02]의 내용을 반영하였다. 학생들은 기후 위기라는 사회적 쟁점을 중심으로 다양한 관점을 비교하고 상대방의 주장과 근거를 평가하면서 논증의 타당성을 판단하게 된다.

또한 생성형 AI의 응답과 다양한 외부 자료를 비교·검토하고 정보의 신뢰성을 확인하는 과정은 [10공국1-02-01]과 관련된다. 학생들은 생성형 AI가 제시한 정보를 그대로 수용하지 않고 출처 확인과 교차 검증을 통해 정보의 타당성을 판단하며, 이를 바탕으로 자신의 입장을 수정하거나 보완하게 된다.

1) 최은경·이용섭(2025)은 토의·토론 기반 기후 변화 교육의 필요성과 관련하여, 미래 세대가 기후 위기의 심각성을 인식하고 기후 행동 및 기후 적응 조치에 참여하도록 해야 한다는 유네스코(2024)의 교육 방향과 학생들이 실생활의 복잡한 문제를 해결할 수 있도록 다양한 지식과 경험을 통합·융합하도록 강조한 교육부(2022c)의 교육과정 방향을 설명하였다. 본고에서 설계한 수업은 생성형 AI의 비판적 활용에 대한 교육을 주된 목적으로 하지만, 기후 변화를 토론 주제로 활용함으로써 기후 변화에 대한 문제의식을 형성하고 대응 방안을 탐색하는 학습도 함께 이루어질 수 있다.

〈표 1〉 생성형 AI의 비판적 활용 토론 수업 관련 성취기준(교육부, 2022b)

영역		성취기준
공통 국어	듣기·말하기	[10공국1-01-02] 논제의 필수 쟁점별로 논증을 구성하고 논증이 타당하지 평가하며 토론한다.
	읽기	[10공국1-02-01] 다양한 글이나 자료를 읽으며 논증의 타당성을 평가하고 자신의 관점을 바탕으로 논증을 재구성한다.
	쓰기	[10공국1-03-01] 내용 전개의 일반적 원리를 고려하여 사회적 쟁점에 대한 자신의 견해를 정교하게 표현하는 글을 쓴다.
독서 토론과 글쓰기		[12독토01-02] 대화, 토의, 토론 등 적절한 방법을 활용하여 서로 다른 생각과 관점을 존중하며 독서 토론을 한다.
		[12독토01-03] 독서 토론의 내용을 바탕으로 쓰기 목적, 독자, 매체를 고려하여 글을 쓰고 공유한다.
		[12독토01-06] 사회적 현안이나 쟁점이 담긴 책을 읽고 독서 토론하고 글을 쓰며 공동체 문제를 해결하고 사회적 담론에 참여한다.
주제 탐구 독서		[12주탐01-04] 주제와 관련된 책이나 자료를 탐색하면서 신뢰할 수 있고 가치 있는 정보를 선정하여 분석하며 읽는다.
		[12주탐01-05] 주제에 관련된 책과 자료를 종합하여 읽으며 자신의 관점과 견해를 형성한다.

토론 이후에는 생성형 AI 활용 과정에서 확인한 문제점과 자신의 생각을 정리하는 성찰 활동을 실시하도록 한다. 이는 사회적 쟁점에 대한 자신의 견해를 표현하는 [10공국1-03-01]과 연계된다. 학생들은 토론 과정에서 형성된 자신의 관점을 정리하고 생성형 AI를 활용할 때 필요한 태도와 유의점을 자신의 언어로 표현하게 된다.

이러한 설계는 공통국어 교육과정이 강조하는 비판적·창의적 사고 역량, 디지털·미디어 역량, 의사소통 역량, 공동체·대인 관계 역량, 자기 성찰·계발 역량을 통합적으로 반영한 것이다. 학생들은 생성형 AI라는 디지털 매체가 생산한 정보를 비판적으로 검토하고, 토론을 통해 다양한 관점을 이해하며, 성찰 활동을 통해 자신의 판단 과정을 점검하게 된다. 「공통국어」 과목 외에 「독서 토론과 글쓰기」와 「주제 탐구 독서」의 성취기준은 사회적 쟁점을 중심으로 한 토론 활동과 정보 탐색·검증 활동을 포함하고 있어 본 수업 모형을 관련 선택 과목으로 확장 적용할 때 참고할 수 있다.

이와 같이 본 수업은 「공통국어1」의 듣기·말하기, 읽기, 쓰기 성취기준을 핵심 설계 기준으로 하여 생성형 AI가 생산한 정보를 비판적으로 검토하고 이를 토론과 글쓰기에 활용하도록 구성하였다.

3. 차시별 수업 구성

이상의 설계 원리를 바탕으로 본 연구는 생성형 AI에 대한 이해를 바탕으로 정보를 검증하고, 이를 토대로 토론과 성찰이 이루어질 수 있도록 총 3차시의 수업을 구성하였다.

1차시는 생성형 AI의 개념과 특징을 이해하고 동일한 질문에 대한 응답의 가변성을 탐색하는 단계이다. 학생들은 서로 다른 생성형 AI에 동일한 질문을 입력하고 응답을 비교함으로써, 생성형 AI의 응답이 고정된 정답이 아닌 확률적으로 생성된 결과물이라는 점을 이해하게 된다. 이를 통해, 생성형 AI의 응답을 비판적으로 검토하기 위한 기초를 마련하도록 하였다.

2차시는 생성형 AI의 한계와 정보 검증 방법을 이해하는 단계이다. 학생들은 생성형 AI의 확률적 생성 원리와 환각 현상을 학습하고 실제 오류 사례를 분석한다. 또한 출처 확인, 교차 검증, 사실

확인 방법을 적용하여 생성형 AI 응답의 신뢰성을 판단한다. 이를 통해 토론에서 활용할 자료의 사실성과 타당성을 검토할 수 있도록 하였다.

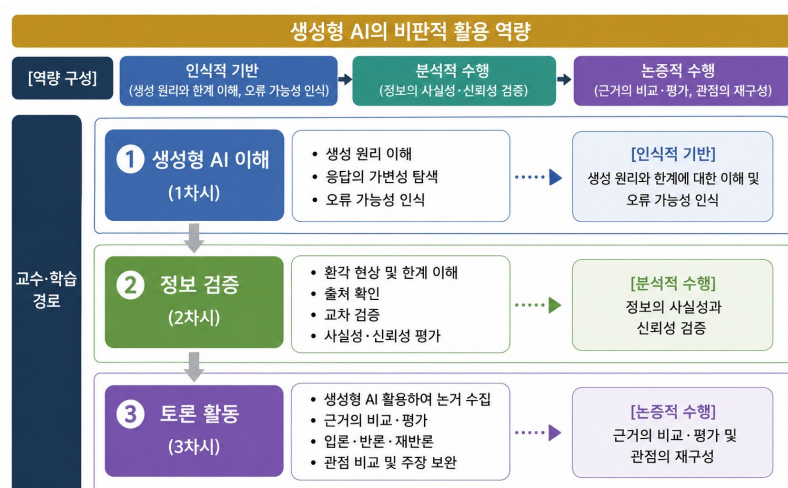
3차시는 생성형 AI의 비판적 활용 단계이다. 학생들은 생성형 AI를 활용하여 토론 논거를 수집한 후 해당 정보의 사실성과 신뢰성을 검증한다. 이후 검증된 자료를 바탕으로 토론을 수행하고 성찰 활동을 통해 생성형 AI 활용 과정에서의 판단과 태도를 점검한다.

이를 통해 학생들이 생성형 AI를 정보 제공 도구만이 아닌 비판적 검토와 평가의 대상으로 인식하고 활용하며, 생성형 AI 시대에 요구되는 비판적 문식성과 주체적 정보 활용 역량을 함양할 수 있도록 설계하였다.

다만, 학교 현장에서는 수업 여건과 학습자의 수준에 따라 한 차시 안에 논거 수집, 검증, 토론, 성찰 활동을 모두 수행하기 어려운 경우가 있을 수 있다. 이에 학교의 수업 여건과 학습자의 수준에 따라 논거 수집을 사전 과제 또는 3차시 도입 전 활동으로 운영하거나, 토론 및 성찰 활동을 별도의 차시로 확장하여 운영하는 것도 가능하다.

각 차시의 활동은 앞서 정의한 생성형 AI 비판적 활용 역량의 구성 요소와 대응한다. 즉, 1차시의 응답 비교 활동은 동일한 질문에 대한 생성형 AI의 응답이 고정된 정답이 아니라 확률적 생성의 결과임을 확인하는 활동으로서, 생성 원리와 한계에 대한 이해 및 오류 가능성의 인식이라는 인식적 기반의 형성과 대응한다. 2차시의 검증 활동은 출처 확인과 교차 검증을 통해 생성형 AI 응답의 사실 여부를 판단하는 활동으로서, 사실성과 신뢰성의 검증이라는 분석적 수행과 대응한다. 3차시의 토론 활동은 검증된 자료를 바탕으로 상대방의 근거를 평가·반박하고 자신의 관점을 수정·보완하는 활동으로서, 근거의 비교·평가 및 관점의 재구성이라는 논증적 수행과 대응한다.

이상의 내용을 반영한 생성형 AI의 비판적 활용을 위한 토론 기반 수업 모형은 [그림 1]과 같다.



[그림 1] 생성형 AI의 비판적 활용을 위한 토론 기반 수업 모형

IV. 생성형 AI의 비판적 활용을 위한 토론 기반 수업의 실제

1. 수업 내용 체계

1) 생성형 AI의 이해와 응답의 가변성 탐색 (1차시)

1차시는 생성형 AI의 개념과 특징을 이해하고 생성형 AI 응답의 가변성을 인식하는 데 초점을 두었다. 최근 학생들은 생성형 AI를 일상적으로 활용하고 있으나, 생성형 AI의 응답을 검색 결과나 정답과 동일한 것으로 인식하는 경우가 많다. 따라서 생성형 AI의 비판적 활용 역량을 함양하기 위해서는 생성형 AI가 어떠한 방식으로 응답을 생성하는지 이해하고, 생성형 AI의 응답이 항상 동일하거나 절대적인 결과가 아니라는 점을 인식할 필요가 있다.

도입 단계에서는 동일한 질문에 대해 서로 다른 생성형 AI가 생성한 응답 사례를 제시하였다. 학생들은 같은 질문임에도 불구하고 서로 다른 내용과 표현이 생성된다는 사실을 확인하면서 생성형 AI의 응답이 하나의 정답이 아니라 다양한 가능성 가운데 생성된 결과물일 수 있음을 인식하게 된다. 이러한 활동은 생성형 AI 응답에 대한 문제의식을 형성하고 생성형 AI 응답을 비판적으로 검토할 필요성을 인식하도록 하기 위해 구성하였다.

전개 단계의 활동 1에서는 생성형 AI의 개념과 특징을 이해하도록 하였다. 학생들은 ChatGPT, Gemini, Claude와 같은 생성형 AI의 사례를 살펴보고 활용 사례를 탐색한다. 이를 통해 생성형 AI가 다양한 분야에서 활용되고 있는 디지털 기술임을 이해하고, 향후 토론 활동에서 활용할 도구의 특성을 이해하도록 하였다.

활동 2에서는 모듈별로 동일한 프롬프트를 서로 다른 생성형 AI에 입력하고 결과를 비교하도록 하였다. 학생들은 생성된 답변의 핵심 키워드, 강조 내용, 표현 방식을 분석하면서 생성형 AI마다 서로 다른 응답을 생성할 수 있음을 확인하게 된다. 이 활동은 생성형 AI 응답의 가변성을 직접 경험하도록 하기 위한 것으로, 학생들이 생성형 AI의 응답을 객관적 사실과 동일한 것으로 인식하지 않도록 하는 데 목적이 있다.

활동 3에서는 생성형 AI의 작동 원리를 학습한다. 학생들이 기술적 원리를 쉽게 이해할 수 있도록 계산기와 생성형 AI를 비교하였다. 계산기는 동일한 입력값에 대해 동일한 결과를 산출하지만 생성형 AI는 확률적으로 단어를 선택하여 문장을 생성한다. 학생들은 이러한 차이를 통해 동일한 질문에 대해서도 다양한 응답이 생성될 수 있는 이유를 이해하게 된다.

정리 단계에서는 생성형 AI 응답의 공통점과 차이점을 정리하고, 생성형 AI의 답변이 고정된 정답이 아니라 생성 결과물임을 확인하도록 하였다. 또한 다음 차시에 학습할 환각 현상과 정보 검증의 필요성을 안내하여 학습 내용을 연결하였다. 이를 통해 학생들이 생성형 AI 응답에 대한 비판적 검토의 필요성을 인식하도록 하였다. 이러한 1차시 수업을 정리하면 <표 2>와 같다.

〈표 2〉 1차시 수업 지도안 (1/3)

학습 주제	생성형 AI의 이해와 응답의 가변성 탐색		
학습 목표	① 생성형 AI의 개념과 특징을 설명할 수 있다. ② 동일한 질문에 대한 생성형 AI의 다양한 응답을 비교할 수 있다. ③ 생성형 AI의 응답이 변화할 수 있는 생성물임을 이해할 수 있다.		
단계(시간)	학습 요소	교수학습활동	자료 및 유의점
도입(10분)	문제 인식 및 동기 유발	■ 동일한 질문에 대한 생성형 AI의 서로 다른 응답 사례를 제시한다. ■ 응답의 차이점을 비교하며 그 이유를 생각해 본다. ■ 학습 목표와 활동 내용을 확인한다.	ChatGPT, Gemini, Claude 등 화면 제시
전개(35분)	활동 1. 생성형 AI 이해하기	■ 생성형 AI의 개념과 특징을 이해한다. ■ ChatGPT, Gemini, Claude 등 대표적인 생성형 AI 사례를 살펴본다. ■ 생성형 AI의 활용 사례를 탐색한다.	PPT 자료
	활동 2. 동일 질문에 대한 다른 답변 결과 확인하기	■ 모듈별로 동일한 프롬프트(“인공지능 시대에 학생에게 필요한 능력은 무엇인가?”)를 서로 다른 생성형 AI에 입력한다. ■ 생성된 답변의 핵심 키워드와 표현 방식을 비교한다. ■ 공통점과 차이점을 정리하여 공유한다.	■ 스마트 기기(디벗) ■ 생성형 AI 서비스
	활동 3. 생성형 AI의 작동 원리 이해하기	■ 생성형 AI의 확률적 생성 원리를 이해한다. ■ 계산기와 생성형 AI를 비교하며 결정론적 도구와 확률적 생성 도구의 차이를 이해한다. ■ 생성형 AI 응답이 항상 동일하거나 절대적인 결과가 아님을 이해한다.	■ 시각 자료 ■ 기술적 설명보다 가변성 이해에 초점
정리 (5분)	학습 내용 정리	■ 생성형 AI 응답의 공통점과 차이점을 정리한다. ■ 생성형 AI의 답변은 고정된 정답이 아니라 생성 결과물임을 확인한다. ■ 다음 차시의 환각 현상과 정보 검증 학습을 예고하고, 생성형 AI 응답을 비판적으로 검토할 필요성을 확인한다.	

2) 생성형 AI의 환각 현상 이해와 정보 검증 (2차시)

2차시는 생성형 AI의 환각 현상을 이해하고, 생성형 AI가 제공한 정보를 검증하는 방법을 익히는 데 학습 목표를 두었다. 학생들이 동일한 질문에 대해서도 서로 다른 응답이 생성될 수 있음을 확인하도록 한 1차시 수업에 이어, 2차시에서는 이러한 응답의 차이가 표현상의 차이뿐만이 아닌, 정보의 정확성과 신뢰성 문제로 이어질 수 있음을 이해하도록 하였다. 또한 생성형 AI가 제시한 정보를 토론에서 활용하기 위해서는 사실성과 신뢰성을 검증하는 과정이 필요함을 이해하도록 하였다.

도입 단계에서는 1차시 활동 결과를 다시 확인하고 “생성형 AI의 답변은 모두 믿을 수 있을까?”와 같은 질문을 통해 동기를 유발한다. 학생들은 생성형 AI가 제공한 정보의 신뢰성에 대해 자신의 생각을 공유하면서 정보 검증의 필요성을 인식하게 된다.

전개 단계의 활동 1에서는 생성형 AI의 확률적 생성 원리를 살펴본다. 학생들이 기술적 내용을 쉽게 이해할 수 있도록 다음 단어를 예측하는 간단한 예시를 활용한다. 학생들이 생성형 AI는 정보를 검색하여 제시하는 것이 아니라, 학습된 언어 패턴을 바탕으로 다음에 올 가능성이 높은 단어를 생성한다는 점을 이해하도록 한다. 이를 통해 생성형 AI의 응답이 사실 그 자체가 아니라 확률적 예측의 결과라는 점을 파악하도록 한다.

활동 2에서는 생성형 AI의 환각 현상을 분석한다. 생성형 AI가 실제 존재하지 않는 정보나 잘못된

사실을 그럴듯하게 제시한 사례를 검토하고 오류가 발생한 부분을 찾도록 하였다. 이은선과 진은진(2024)에 제시된 환각 사례로 ‘안중근 의사 관련 오류 사례’와²⁾ 같이 학생들이 이미 알고 있는 역사적 사실을 활용함으로써 환각 현상을 보다 쉽게 이해하도록 하였다. 학생들은 오류를 발견하는 과정에서 생성형 AI가 제공한 정보라 하더라도 사실 여부를 확인해야 함을 인식하게 된다. 또한 생성형 AI의 오류를 AI라는 기술적 결함으로 이해하기보다 생성형 AI의 특성에서 비롯되는 현상으로 이해하도록 지도한다.

활동 3에서는 생성형 AI 응답을 검증하는 방법을 학습하도록 한다. 학생들은 생성형 AI가 제시한 정보의 출처를 확인하고, 다른 자료와 비교하며, 최신 정보를 검색하는 방법을 실습한다. 학술 논문 자료, 공공기관 자료, 언론 기사, 통계 자료 등을 활용하여 생성형 AI 응답의 사실 여부를 확인하도록 한다. 이를 통해 생성형 AI가 제시한 정보의 신뢰성을 스스로 판단할 수 있도록 한다.

정리 단계에서는 생성형 AI가 학습과 정보 탐색에 유용한 도구이지만 오류 가능성을 내포하고 있으므로, 생성형 AI의 응답을 활용할 때에는 반드시 검증 과정을 거쳐야 함을 강조한다. 이를 통해 학생들은 생성형 AI를 사용할 때 필요한 기본적인 비판적 태도를 갖도록 지도한다. 마지막으로 3차시 수업에 진행될 토론 활동에 대해 사전 안내한다.³⁾ 2차시 수업을 정리하면 <표 3>과 같다.

<표 3> 2차시 수업 지도안 (2/3)

학습 주제	생성형 AI의 환각 현상 이해와 정보 검증		
학습 목표	① 생성형 AI의 확률적 생성 원리를 설명할 수 있다. ② 생성형 AI의 환각(Hallucination) 현상을 이해하고 오류 사례를 분석할 수 있다. ③ 생성형 AI가 제공한 정보의 신뢰성을 검증하는 방법을 활용할 수 있다.		
단계(시간)	학습 요소	교수학습활동	자료 및 유의점
도입(5분)	답변의 신뢰성 탐색	■ 동일한 질문에 대한 생성형 AI의 서로 다른 답변을 다시 살펴본다. ■ ‘생성형 AI의 답변은 모두 믿을 수 있을까?’를 주제로 생각을 나눈다. ■ 학습 목표와 활동 내용을 확인한다.	■ PPT 자료 비판적 태도보다 문제 의식 형성에 초점
전개(40분)	활동 1. 생성형 AI는 어떻게 답을 만들까?	■ 생성형 AI의 확률적 생성 원리를 이해한다. ■ 다음 단어 예측 방식을 간단한 예시를 통해 살펴본다. ■ 계산기와 생성형 AI를 비교하며 결정론적 도구와 확률적 생성 도구의 차이를 이해한다.	■ 시각 자료 ■ 기술적 설명은 최소화
	활동 2. 환각(Hallucination) 찾아보기	■ 생성형 AI의 환각 현상을 이해한다. ■ 안중근 의사 사례 등 실제 오류 사례를 분석한다. ■ 어떤 정보가 사실과 다르며 왜 오류가 발생했는지 토의한다.	■ 오류 사례 자료 ■ AI를 부정적으로 인식 하지 않도록 유의

2) 이은선과 진은진(2024)은 환각 예시 찾기 사례로 ‘안중근 의사님이 환자를 구한 사례들을 알려줘.’라는 사용자의 질문에 대한 ‘안중근 의사는 대한민국의 독립운동가이며, 의사로도 활약했습니다.’라는 ChatGPT 3.5의 답변을 들고 있다. 그러나 이와 같은 정보에 대한 학습이 이루어진 최근 ChatGPT의 이 질문에 대한 답변은 ‘안중근 의사님께서 환자들을 치료하거나 의료 활동을 하신 분이 아닙니다. 여기서 ‘의사’는 의료인을 뜻하는 의사(醫師)가 아니라 나라와 민족을 위해 헌신한 의로운 사람이라는 뜻의 의사(義士)를 의미합니다.’로 제시된다. 이러한 답변 차이의 예는 학생들에게 생성형 AI의 성능 향상 과정을 보여주는 예로도 가능하다.

3) 2차시의 차시 예고 단계에서 3차시 토론 주제를 안내하고, 생성형 AI를 활용하여 찬반 입장의 기초적인 논거를 수집해 오도록 과제를 제시할 수 있다. 이를 통해 3차시에는 생성형 AI가 제시한 논거를 바탕으로 토론을 진행하면서 상대 측 논거의 출처와 사실 여부를 확인·검증하는 활동에 보다 집중할 수 있다.

전개(40분)	활동 3. 생성형 생성형 AI 답변 검증하기	■ 생성형 AI 답변을 검증하는 방법을 학습한다. ■ 출처 확인, 교차 검증, 최신 정보 확인 방법을 익힌다. ■ 모듈별로 AI 답변을 검증하고 결과를 정리한다.	■ 스마트 기기 ■ 공신력 있는 사이트 활용
정리(5분)	학습 내용 정리	■ 생성형 AI는 유용하지만 오류 가능성이 있음을 정리한다. ■ AI 답변을 검증해야 하는 이유를 정리한다. ■ 다음 차시 토론 활동을 안내한다.	■ 3차시 토론 주제를 미리 안내하고 그에 대한 근거 자료 찾기 과정을 제시할 수 있음.

3) 생성형 AI를 활용한 토론 활동과 성찰 (3차시)

3차시는 생성형 AI의 비판적 활용 역량을 실제 토론 상황에 적용하는 단계로 구성하였다. 앞선 차시에서는 학습한 생성형 AI의 특성과 정보 검증 방법을 바탕으로 생성형 AI가 제공한 논거를 분석하고 검증하며, 이를 활용하여 토론을 수행하도록 하였다. 본 차시는 생성형 AI를 토론 준비를 위한 보조 도구로 활용하는 데 목적이 있는 것이 아니라, 생성형 AI가 생산한 정보의 신뢰성과 타당성을 검토하는 경험을 제공하는 데 목적이 있다.⁴⁾ 즉, 생성형 AI를 활용하여 토론하는 것보다 생성형 AI가 제시한 정보를 비판적으로 검토하고 활용하는 데 중점을 두었다. 이를 통해 생성형 AI를 비판적으로 활용하는 과정과 토론 능력을 함께 신장할 수 있도록 하였다.

도입 단계에서는 ‘기후 위기’와 관련된 토론 주제를 제시하고 찬성·반대 입장을 배정한다. 예를 들어 ‘기후 위기 대응을 위해 내연기관 자동차 판매를 금지해야 한다.’나 ‘기후 위기 대응을 위해 탄소세(Carbon Tax)를 도입해야 한다.’와 같이 찬성과 반대의 입장이 가능한 하위 주제를 활용할 수 있다. ‘기후 위기’와 관련한 하위 토론 주제는 하나의 주제에 대해 찬·반 토론을 진행할 수도 있고, 여러 모듈인 경우 <표 4>와 같이 모듈별로 다른 하위 주제를 제시하여 진행할 수 있다.

<표 4> ‘기후 위기’ 토론 주제 활용의 예

구분	하위 토론 주제	찬성 조	반대 조
주제 1	기후 위기 대응을 위해 내연기관 자동차 판매를 금지해야 한다.	1조	2조
주제 2	기후 위기 대응을 위해 탄소세(Carbon Tax)를 도입해야 한다.	3조	4조

이러한 주제는 앞서 설명한 바와 같이 경제적, 사회적, 윤리적 관점이 복합적으로 작용하는 사회적 쟁점으로 다양한 입장과 논거가 존재한다. 또한, 생성형 AI가 풍부한 정보를 제공하는 주제이기 때문에 생성형 AI가 제시한 정보의 신뢰성을 검증하고 토론하기에 적합한 주제라고 판단하였다.

전개 단계의 활동 1에서는 생성형 AI를 활용하여 찬성과 반대 입장의 논거를 수집하도록 한다. 학생들은 생성형 AI가 제시한 주장, 통계 자료, 사례 등을 정리하고 토론에 활용할 수 있는 정보를

4) 본 수업의 주된 목표가 토론 활동에서 생성형 AI를 활용하면서 정보에 대한 정확성을 검증하는 데 있다. 그러나 이러한 과정에서 학습자는 자신이 갖추어 놓은 자료나 논리적 근거를 수정·보완할 수 있고 생성형 AI를 통하여 예상 질문을 받아 주장을 강화할 수 있다(박광길, 2025). 이를 통해 학습자는 토론 능력이 함양될 수 있다.

선별하였다. 이 과정에서 학생들은 생성형 AI가 짧은 시간 안에 다양한 논거를 제공할 수 있다는 점을 확인하게 된다. 동시에 생성형 AI가 제공한 자료가 실제 토론의 근거로 활용될 수 있는지 검토할 필요가 있음을 인식하게 된다.

활동 2에서는 생성형 AI가 제시한 논거를 검증하도록 한다. 학생들은 생성형 AI가 제시한 통계 수치, 기관명, 보고서, 연구 결과 등의 출처를 확인하여, 출처 자료와 비교한다. 또한 사실과 다른 정보나 불명확한 근거가 포함되어 있는지 검토하고 필요한 경우 자료를 수정하거나 보완하도록 한다. 이 활동은 생성형 AI의 정보를 무비판적으로 활용하는 것이 아니라 근거의 신뢰성을 검토하는 경험을 제공하는 데 목적이 있다. 학생들은 검증 과정을 통해 토론의 질이 자료의 양이 아니라 자료의 신뢰성에 의해 결정될 수 있음을 이해하게 된다.

활동 3에서는 검증된 자료를 활용하여 토론을 실시한다. 토론은 찬성과 반대 입론, 반론, 질의응답, 제반론, 최종 발언의 순으로 진행하며, 학생들은 상대방의 주장과 근거를 검토하고 그 결과를 바탕으로 자신의 주장을 수정·보완하도록 한다. 특히 상대방이 제시한 통계 자료나 사례의 출처와 신뢰성, 논리적 타당성을 확인하며 질문과 반론을 제기하도록 지도한다. 또한 토론에 참여하지 않는 학생은 청중의 입장에서 주장과 근거의 적절성을 기록하고 토론 후 피드백을 제공하도록 하여 학급 전체가 근거를 검토하는 과정에 참여하도록 한다. 이를 통해 학생들은 주장 자체보다 근거의 적절성과 신뢰성을 중심으로 논증을 평가하고, 토론 과정에서 다양한 관점을 수용하며 자신의 입장을 논리적으로 조정하는 능력을 기를 수 있다.

활동 4에서는 성찰 활동을 실시한다.⁵⁾ 학생들은 토론 과정에서 활용된 생성형 AI 정보 가운데 오류가 있었던 사례를 다시 확인하고, 생성형 AI를 사용할 때 필요한 태도에 대해 성찰지를 작성한다. 또한 생성형 AI의 장점과 한계를 정리하고 앞으로 생성형 AI를 활용할 때 어떤 점을 유의해야 하는지 자신의 생각을 정리하도록 한다. 이 활동을 통해 학생들은 학습 과정에서 경험한 내용을 자신의 언어로 재구성하고 내면화할 수 있다.

정리 단계에서는 생성형 AI의 정보가 검증의 대상이 될 수 있음을 다시 확인하고 정보 검증의 중요성을 정리한다. 또한 생성형 AI를 활용할 때에는 정보의 정확성과 근거의 신뢰성을 스스로 판단해야 함을 강조한다. 3차시는 생성형 AI의 이해와 정보 검증을 바탕으로 토론 활동을 수행하고 성찰에 이르기까지의 과정을 통합적으로 경험하도록 함으로써 생성형 AI 시대에 요구되는 비판적 문식성과 주체적 정보 활용 역량을 함양하도록 구성하였다. 수업 운영 시에는 학습자의 수준과 토론 진행 상황에 따라 토론 시간과 성찰 활동의 비중을 조정하거나 일부 활동을 과제로 연계하여 운영할 수 있다. 3차시 수업을 정리하면 <표 5>와 같다.

5) 토론 및 성찰 활동을 보다 심도 있게 운영하고자 하는 경우에는 토론과 성찰 활동을 별도의 차시로 구성하여 운영할 수도 있다.

〈표 5〉 3차시 수업 지도안 (3/3)

학습 주제	생성형 AI를 활용한 토론 활동과 성찰		
학습 목표	① 생성형 AI를 활용하여 토론에 필요한 논거를 수집할 수 있다. ② 생성형 AI가 제시한 논거의 신뢰성과 타당성을 검증할 수 있다. ③ 검증한 자료를 바탕으로 토론하고 생성형 AI의 비판적 활용 과정을 성찰할 수 있다.		
단계(시간)	학습 요소	교수학습활동	자료 및 유의점
도입 (5분)	토론 준비	■ 토론 주제를 확인한다. ■ 찬성·반대 입장을 배정한다. ■ 학습 목표와 활동 내용을 확인한다.	■ 토론 주제 제시 ■ 역할 분담
전개 (40분)	활동 1. AI로 논거 수집하기	■ 생성형 AI를 활용하여 찬성·반대 논거를 수집한다. ■ 관련 통계 자료와 사례를 탐색한다. ■ 생성형 AI가 제시한 주요 논거를 정리한다.	■ 스마트 기기 ■ 생성형 AI 서비스
	활동 2. 논거 검증하기	■ 생성형 AI가 제시한 논거와 자료를 검증한다. ■ 통계 수치, 기관명, 보고서 등을 확인한다. ■ 검증 결과를 정리하고 수정·보완한다.	■ 팩트체크 활동지 ■ 검증 과정 자체를 중시
	활동 3. 토론하기	■ 검증된 자료를 바탕으로 토론을 실시한다. ■ 찬성 입론, 반대 입론, 반론, 질의응답, 재반론, 최종 발언 순으로 진행한다. ■ 상대방이 제시한 근거의 출처와 신뢰성, 논리적 타당성을 확인하며 질문하고 반론을 제기한다. ■ 토론 과정에서 자신의 주장과 근거를 수정·보완한다.	■ 토론 규칙 안내 ■ 근거 중심 토론 유도 ■ 청중은 발표자의 주장과 근거의 적절성을 기록하고 토론 후 피드백을 제공
	활동 4. 성찰하기	■ 토론 과정을 되돌아보며 성찰지를 작성한다. ■ AI가 제시한 정보 중 오류가 있었는지 확인한다. ■ 토론 과정에서 자신의 주장과 근거가 수정·보완된 과정을 돌아본다. ■ 생성형 AI를 사용할 때 필요한 태도를 정리한다.	■ 성찰 활동지
정리 (5분)	학습 내용 정리	■ 생성형 AI의 정보는 비판적으로 검토되어야 함을 확인한다. ■ 정보 검증의 중요성을 정리한다. ■ 생성형 AI를 활용할 때 필요한 태도를 정리한다.	

2. 생성형 AI 응답 검증 기준 및 평가 기준

생성형 AI의 비판적 활용 역량 교육을 위한 수업에서 학생들이 생성형 AI의 응답을 검증할 때 어떤 기준에 따라 검증할 것인지, 그리고 토론 과정에서 교사가 학습자의 비판적 활용 역량을 어떠한 요소를 중심으로 확인할 것인지를 계획할 필요가 있다. 이를 위해 학생들이 생성형 AI의 응답을 검토할 때 활용할 검증 기준으로는 <표 6>, 교사가 토론 및 성찰 활동을 평가할 때 활용할 평가 기준으로는 <표 7>을 활용할 수 있다. 먼저, 생성형 AI의 응답 검증 기준은 2차시 정보 검증 활동에서 활용할 수 있으며, 토론 및 성찰 평가 기준은 3차시 토론 활동과 성찰 활동에서 활용할 수 있다.

<표 6>은 생성형 AI가 제시한 정보를 그대로 수용하지 않고 사실성과 신뢰성을 확인하기 위한 기준을 제시한 것이다. 학생들은 출처 확인, 교차 검증, 최신성 확인, 근거의 적절성, 논리적 타당성 등을 중심으로 생성형 AI 응답을 검토함으로써 정보의 신뢰성을 판단하고 토론에 활용할 자료를 선별할 수 있다.

〈표 6〉 생성형 AI의 응답 검증 내용

검증 요소	검증 내용
출처 확인	생성형 AI가 제시한 기관명, 연구, 통계, 인용 자료 등이 실제 존재하는지 확인한다.
교차 검증	공공기관, 학술자료, 언론 기사 등 다른 자료와 비교하여 정보가 일치하는지 확인한다.
최신성 확인	제시된 통계와 정책, 사례가 현재에서도 유효한지 확인한다.
근거의 적절성	제시된 근거가 주장과 관련성이 있는지 확인한다.
논리적 타당성	근거가 주장을 충분히 뒷받침하는지 검토한다.

〈표 7〉 토론 및 성찰 평가 기준

평가 요소	상	중	하
근거의 신뢰성 검증	출처와 교차 검증 결과를 근거로 신뢰성을 판단한다.	검증을 시도하였으나 일부만 제시한다.	검증 없이 생성형 AI 응답을 활용한다.
근거 중심 토론	상대방 근거의 타당성을 근거를 들어 평가·반박한다.	근거를 일부 활용한다.	자신의 주장만 반복한다.
관점의 재구성	검증과 토론을 거쳐 자신의 견해를 수정·보완한다.	일부 수정한다.	기존 입장을 그대로 유지한다.
성찰	생성형 AI 활용 과정에서의 한계와 유의점을 구체적으로 성찰한다.	한계와 유의점을 일부 제시하며 성찰한다.	한계와 유의점을 형식적으로 작성한다.

〈표 7〉은 생성형 AI의 비판적 활용 역량이 토론 활동에서 어떻게 나타나는지를 확인하기 위한 평가 기준이다. 교사는 생성형 AI 응답의 신뢰성을 검증하였는지, 검증한 자료를 토대로 토론을 수행하였는지, 토론 과정에서 자신의 관점을 수정·보완하였는지, 생성형 AI 활용 과정에 대해 성찰하였는지를 중심으로 학습 과정을 평가할 수 있다.

V. 결론

본 연구는 생성형 AI의 확산에 따라 요구되는 비판적 활용 역량을 함양하기 위한 국어과 토론 수업 방안을 설계하는 데 목적을 두었다. 생성형 AI는 정보 탐색, 글쓰기, 발표 준비 등 다양한 학습 활동에 활용되고 있으나, 확률적 언어 생성 기술이라는 특성으로 인해 환각 현상과 정보 신뢰성 문제를 내포하고 있다. 이에 따라 생성형 AI 시대의 교육은 생성형 AI를 활용하는 기능적 능력과 함께 생성형 AI가 생산한 정보를 검토·평가하고 적절하게 활용할 수 있는 역량을 함양하는 방향으로 이루어질 필요가 있다.

본 연구에서는 생성형 AI의 개념과 특성을 이해하고, 생성형 AI 응답의 신뢰성을 검증하며, 생성형 AI가 생산한 정보를 토론과 성찰에 활용하는 3차시 수업을 설계하였다. 1차시에서는 생성형 AI의 개념과 특성을 이해하고 동일한 질문에 대한 다양한 응답을 비교함으로써 생성형 AI 응답의 가변성을

인식하도록 하였다. 2차시에서는 환각 현상과 정보 검증 방법을 학습하며 생성형 AI 응답의 신뢰성을 분석·평가하도록 하였다. 3차시에서는 생성형 AI가 제시한 논거를 검증한 후 토론에 활용하고 성찰하는 활동을 통해 생성형 AI가 생산한 정보를 비판적으로 수용하고 재구성하도록 하였다. 이를 통해 생성형 AI의 비판적 활용 역량을 단계적으로 함양할 수 있는 국어과 토론 수업 모형을 제안하였다.

본 연구는 생성형 AI의 교육적 활용 전반을 포괄하는 수업 모형을 제안하기보다, 생성형 AI가 생산한 정보와 논거를 검증·평가의 대상으로 삼는 국어과 토론 수업을 설계하는 데 초점을 두었다는 점에 의의가 있다. 또한 생성형 AI 시대의 국어과 교육에서 정보 생산 능력과 함께 정보의 타당성과 신뢰성을 평가하는 비판적 문식성 교육의 필요성을 구체적인 수업 모형으로 제시하였다는 점에서도 의미가 있다. 특히 본 연구에서 설계한 수업은 고등학교 국어과 2022 개정 교육과정이 강조하는 비판적·창의적 사고 역량, 디지털·미디어 역량, 의사소통 역량, 공동체·대인 관계 역량 및 자기 성찰·계발 역량 등을 수업 설계에 반영하였다는 점에서 교육과정적 의미를 지닌다.

그러나 본 연구는 수업 설계와 지도 방안 제안에 초점을 둔 연구로서 실제 수업 적용을 통한 학습자 변화와 교육적 효과를 검증하지 못하였다는 한계를 가진다. 따라서 후속 연구에서는 본 연구에서 설계한 수업 모형을 학생들에게 적용하여 생성형 AI의 비판적 활용 역량이 어떠한 양상으로 형성되는지 분석할 필요가 있다. 특히 생성형 AI 응답에 대한 학생들의 신뢰도 변화, 정보 검증 과정에서 활용하는 판단 기준, 토론 과정에서 나타나는 논증 능력의 변화 등을 다각적으로 살펴볼 필요가 있다.

또한 생성형 AI의 비판적 활용에 대한 교육은 특정 교과와 과제가 아니라 미래 사회의 디지털 시민성을 함양하기 위한 공통의 교육 과제라는 점에서 다양한 교과와의 연계 연구가 요구된다. 국어과에서는 생성형 AI의 비판적 활용 교육을 토론뿐 아니라 읽기, 쓰기, 매체 등의 영역과 연계한 수업 모형을 개발할 필요가 있으며, 사회·과학·정보 교과 등과 연계한 융합적 접근 역시 가능할 것이다. 아울러 생성형 AI를 가상의 토론자나 반론 생성기로 활용하는 수업 모형을 개발하고 그 교육적 가능성과 적용 방안을 검토하는 후속 연구도 이루어질 필요가 있다.

참고문헌

- 교육부 (2022a). 『초·중등학교 교육과정 총론』. 교육부 고시 제2022-33호 [별책 1].
- 교육부 (2022b). 『고등학교 교육과정』. 교육부 고시 제2022-33호 [별책 4].
- 교육부 (2022c). 『과학과 교육과정』. 교육부 고시 제2022-33호 [별책 9].
- 박광길 (2025). 생성형 AI를 활용한 대학 토론 수업 사례 연구. **어문논집**, 103, 267-302.
- 백민경·박찬홍 (2025). 교사-AI 협업에 기반한 평가 문항 설계 탐색: 가정 교과 교사의 ChatGPT 활용 사례를 중심으로. **교육연구**, 93, 259-281.

- 안윤미 · 김유하 · 이서희 (2025). 『생성형 AI의 교육적 활용 실태 및 요구 조사 연구』. 서울: 서울특별시교육청교육연구정보원.
- 오규설 (2024). 생성형 인공지능 시대의 문식성 교육 방향 탐색: 행위자 네트워크 이론을 중심으로. **한글**, 85(3), 717-751.
- 오선주 · 이경화 (2025). AI 시대의 비판적 문식성 재개념화. **청람어문교육**, 108, 67-95.
- 오수경 · 신미성 · 임다미 (2025). 국내 대학 교육에서의 생성형 AI 활용 연구 동향 분석. **멀티미디어 학회 논문지**, 28(2), 353-365.
- 오유나 · 김도연 · 윤영현 · 황혜지 (2025). 비판적 사고 함양을 위한 PBL 기반 생성형 AI 활용 질문 활동 수업 모형 개발. **교육공학연구**, 41(1), 157-203.
- 이세미 · 이동형 · 박창균 (2025). 생성형 인공지능 기반 토론 수업 모형 개발 연구. **국어교육연구**, 87, 83-109.
- 이은선 · 진은진 (2024). 대학 글쓰기 수업에서의 ChatGPT ‘환각(Hallucination)’ 교육 사례 연구. **돈암어문학**, 45, 211-243.
- 임종섭 (2024). AI 시대 뉴스룸... 생성형 AI는 무엇인가?. **관훈저널**, 171, 76-82.
- 조용림 (2025). 생성형 AI 기반 교양국어 교육 적용 탐구. **대학교양교육연구**, 10(2), 161-185.
- 최은경 · 이용섭 (2025). 토의 · 토론 기반 기후변화 교육이 초등학생의 기후 소양 및 비판적 사고력에 미치는 효과. **대한지구과학교육학회지**, 18(2), 154-165.
- 허선아 (2025). AI 시대 화법 교육에서 챗GPT의 교육적 활용과 한계: 고등학교 토론 수업 사례 연구. **새국어교육**, 143, 49-87.
- UNESCO. (2024). *AI competency framework for students*. Paris: UNESCO.

Abstract

A Study on a Discussion-Based Instructional Model for the Critical Use of Generative AI

An, Yoonmi¹

¹*Korean Language Teacher, Seoul Cheongwon Girls' High School*

The purpose of this study is to design a discussion-based instructional model for Korean language education aimed at fostering students' critical use of generative artificial intelligence (AI). Although generative AI supports various learning activities, such as information retrieval, writing, and discussion preparation, it also poses challenges related to AI hallucinations and the reliability of generated information. Accordingly, this study defines the critical use of generative AI as the ability to understand its underlying principles and limitations, verify the factual accuracy and reliability of AI-generated information, compare and evaluate supporting evidence, and reconstruct one's perspectives. Based on this framework, a three-session discussion-based instructional model was designed to provide students with opportunities to experience the sequential processes of understanding generative AI, verifying information, and engaging in its critical use. The first session focuses on understanding the concepts and characteristics of generative AI and exploring the variability of AI-generated responses. The second session is designed to help students examine AI hallucinations and learn methods for verifying information. In the third session, students verify AI-generated arguments on the topic of climate change and use verified evidence in classroom discussions. This study is meaningful in that it conceptualizes generative AI as an object of critical examination and proposes a discussion-based instructional model and lesson plans that integrate the processes of understanding, verification, discussion, and reflection.

Key Words: generative AI, critical utilization competency, critical AI literacy, information verification, discussion-based instruction, Korean language curriculum