

사고력을 키워주는 서·논술형 평가(중등)

해외 서술형 평가 사례:

LESSON 1.

국제학업성취도평가 - PISA, TIMSS

해외 논술형 평가 사례: 미국의 조지아 마일스톤 평가,

LESSON 2.

유럽의 IB 교육과정의 평가



1

해외 서술형 평가 사례:

국제학업성취도평가 - PISA, TIMSS

서술형 문항을 포함하는 가장 대표적인 해외 평가 사례로는 국제학업성취도 평가인 PISA와 TIMSS를 들 수 있습니다. 먼저 PISA(Programme for International Student Assessment)는 OECD에서 주관하는 3년 주기의 역량 중심 평가로서, 학생들이 사회 구성원으로서 삶을 살 준비가 되어 있는지를 점검하는 것을 목적으로 하며, 읽기, 수학, 과학 및 혁신 영역에 대해 평가하게 됩니다. TIMSS(Trends in International Mathematics and Science Study)는 국제 교육성취도 평가 협회(International Association for the Evaluation of Educational Achievement)에서 주관하는 평가로서, 4년을 주기로 학생들의 수학과 과학 과목에 대한 성취도를 국제적으로 비교하고, 여러 교육맥락변인과 성취도의 관계를 파악하는 것을 목적으로 실시되고 있습니다. PISA와 TIMSS 모두 평가의 목적에 따라 선다형 문항과 서술형(구성형) 문항을 적절히 선택하여 제작되고 있습니다. 해외 서술형 평가를 소개하는 이번 학습 목표에 따라 PISA와 TIMSS의 서술형 문항에 초점을 맞추어 각각의 문항 특성과 채점 기준 등에 대해 살펴보도록 하겠습니다.

1) PISA

(1) PISA란?

PISA는 Programme for International Student Assessment의 약자입니다. PISA는 OECD가 세계의 교육 시스템을 점검하기 위한 목적으로 2000년부터 시행하고 있는 역량 중심의 국제학업성취도평가이며, OECD의 회원국과 비회원국 모두 평가에 참여할 수 있습니다. PISA의 평가 대상은 의무교육이 종료되는 시기에 해당하는 만 15세 학생들입니다. PISA는 3년을 주기로 시행되는데, 최근에 시행된 PISA 2015 본검사에는 전 세계 72개국의 만 15세 학생 약 50만명이 참여하였습니다. PISA가 처음 시행된 2000년에는 32개국이 참여했던 것을 고려해보면, PISA에 대한 관심과 중요성에 대한 인식이 증가되었음을 짐작해 볼 수 있습니다.

(2) PISA의 평가틀 (Assessment Framework)

PISA는 읽기, 수학, 과학의 세 가지 역량 영역(competence field)에 대한 기본 소양(literacy)을 측정하며, 최근에는 미래 사회 시민으로서의 역량을 평가하기 위해 새로운 평가 영역과 평가 방법을 개발해오고 있습니다. PISA는 단순 지식의 습득 여부에 머무르지 않고, 실제 삶의 맥락에서 지식을 접목하여 문제를 해결하는 능력을 측정하는 것을 목표로 하는 역량 평가입니다. 따라서 암기된 지식보다는 실생활에 필요한 능력, 즉 지식을 상황과 목적에 맞게 활용할 수 있는 기본적인 소양을 강조합니다.

PISA의 기본 소양인 읽기, 수학, 과학 소양의 의미를 각각 알아보겠습니다.

먼저 읽기 소양은 “자신의 목적을 성취하고 지식과 잠재적 능력을 계발하며, 사회에 참여하기 위해 다양한 텍스트를 이해·활용하고 텍스트를 바탕으로 하여 성찰하며, 다양한 텍스트 읽기 활동에 참여하는 능력”을 의미합니다.

두 번째로 수학 소양은 “다양한 맥락에서 수학을 형식화하고, 이용하고, 해석하는 개인적인 능력이며, 현상을 기술하고 설명하며 예측하기 위해 수학적 추론과 수학적 개념, 절차, 사실, 도구를 사용하는 능력”을 의미합니다. 수학 소양은 개인이 실생활에서 수학이 차지하는 역할을 인식하도록 돕고, 건설적이며 참여적이고 사려 깊은 시민에게 필요한 사실에 근거한 판단과 결정을 지원합니다.

마지막으로 과학 소양은 “사려 깊은 시민으로서 과학적 사고(idea)를 가지고 과학과 관련된 문제에 참여하는 능력”을 의미합니다. 과학적인 소양이 있는 사람은 과학과 기술에 대한 조리 정연한 담론에 기여하고자 하는데, 이를 위해서는 현상을 과학적으로 설명하고 과학적인 탐구를 검증하고 설계하며, 데이터와 증거를 과학적으로 해석하는 능력이 요구됩니다.

한편 PISA는 주기별로 돌아가며 읽기, 수학, 과학 소양 중에서 ‘주영역’을 설정하여, 주영역의 문항을 새롭게 개발하여 심층적으로 평가하고 나머지 영역은 공통 문항을 중심으로 변화 추이만을 평가하게 됩니다. PISA가 처음으로 실시된 2000년의 주영역은 읽기였으며, 2003년은 수학, 2006년은 과학, 2009년은 다시 읽기가 주영역이 되는 식으로, PISA 2015에서는 과학이 주영역이었습니다.

또한 앞서 언급한 바와 같이 PISA는 주기별로 새로운 혁신 영역과 평가 방법을 시도하고 있습니다. 일례로 2003년에는 문제해결력을 혁신적 평가 영역으로 선정하여 평가를 시행하였으며, 2012년에는 컴퓨터 기반 문제해결력 평가, 2015년에는 협력적 문제해결력 평가, 2018년에는 글로벌 역량을 혁신적 평가 영역으로 선정하여 평가를 구성하였습니다. 또한 평가 방식도 기존의 지필평가 방식을 줄이고 컴퓨터 기반 평가를 늘리는 방식으로 발전하여, PISA 2015부터는 대부분의 평가를 컴퓨터 기반 평가로 시행하였습니다.

PISA 주기에 따른 평가 영역 구성

PISA 주기	2000	2003	2006	2009	2012	2015	2018
읽기	○	○	○	○	○	◎	◎
수학	○	○	○	○	○	◎	◎
과학	○	○	○	○	○	◎	◎
혁신적 평가 영역	-	문제해결력 ○	-	-	CBAPS ◎	CPS ◎	글로벌 역량 ◎

출처 : 조성민 외. (2018). OECD 국제학업성취도 평가문항자료집: PISA 2015 과학, 협력적 문제해결력

PISA의 각 영역별 평가들은 주기별로 기존의 평가들을 수정, 보완하여 개발됩니다. PISA 2015를 기준으로 최근까지의 각 영역별 평가들에 대해 알아보겠습니다. 우선 읽기 영역에 대한 평가들은 ‘상황’, ‘텍스트’, ‘양상’의 세 개 범주로 구성되며, 각 범주는 또 다시 세부 요소를 포함합니다.

읽기 영역의 평가틀



수학 영역의 평가들은 ‘맥락’, ‘수학적 내용’, ‘수학적 과정’의 세 개 범주로 구성되며, 각 범주는 다시 각각의 하위 세부 요소로 구성됩니다.

수학 영역의 평가틀



PISA 2015의 주영역인 과학에 대한 평가들은 ‘맥락’, ‘지식’, ‘역량’, ‘태도’, ‘인지적 요구’로 범주화되며, 각 범주는 다른 영역과 마찬가지로 세부 요소로 이루어져 있습니다.

과학 영역의 평가틀



종합해보면, 비록 읽기, 수학, 과학의 평가 영역의 평가틀을 구성하는 세부 범주의 구체적인 명칭은 구체적인 명칭은 조금씩 다르지만 대체적으로 내용, 과정, 상황의 세 차원으로 구성되어 있음을 알 수 있습니다. 이때 ‘내용’은 문제를 해결하기 위한 지식과 이해를 의미하며, ‘과정’은 문제해결을 위해 수행하는 인지적 과정을 의미합니다. 또한 ‘상황’은 지식과 인지과정이 적용되는 상황과 맥락을 뜻하는 것으로 이해할 수 있습니다. PISA는 학생이 가지고 있는 지식 자체를 평가하기 보다는 지식을 자신의 경험에 반영하여 실생활의 문제에 적용하는 능력에 대한 측정하는 평가입니다. 따라서 PISA의 평가 문항도 학생들이 다양한 맥락에서 개념을 이해하고 이를 적용하여 문제를 해결하는 능력을 측정할 수 있도록 제작됩니다.

(3) PISA 문항 구성과 채점의 특징

(가) PISA 문항 구성의 특징

PISA 문항은 하나의 상황에 대해 하나의 문항으로 제작되기도 하고, 하나의 상황에 여러 개의 문항이 함께 제시되는 세트형 문항으로도 제작됩니다. 또한 PISA 문항은 선다형 문항과 구성형 문항(서술형 문항 포함)으로 구분됩니다.

(나) PISA 구성형 문항 채점의 특징

구성형 문항의 채점은 채점 코드를 컴퓨터에 입력하여 이루어집니다. 채점 코드는 다양한 응답들을 범주화하기 위해 사용되며, 일반적으로 두 자리의 숫자 코드를 사용합니다. 앞자리의 숫자는 점수를 표기하고, 뒷자리 숫자는 답안의 유형을 표기합니다. 예를 들어, 어떤 학생의 답안에 대해 채점코드 12가 부여되었다면, 이 학생의 점수는 1점이며, 답안은 유형 2에 해당됨을 의미합니다. 이 때 답안의 유형을 나타내는 숫자는 PISA 국제본부에서 제시하는 채점표에 수록된 답안의 유형에 따라 정해집니다.

참고로 PISA에서는 ‘정답’ 대신 ‘만점’이라는 표현을 사용합니다. 학생의 답안이 ‘영점’에 해당될 때, 해당 답안이 무응답인 경우에는 9로 표기하여 답을 했으나 틀린

경우인 0과 구별합니다.

PISA는 정식 채점에 앞서 채점자 간에 일치된 판단이 확보될 수 있도록 채점자 연습 시간 부여를 규정하고 있습니다. 채점 시 고려 사항 몇 가지를 소개하면 다음과 같습니다. 먼저 답안에 나타난 철자나 문법적 실수는 의미 파악에 심각한 지장을 주지 않는 한 채점에 반영하지 않습니다. 또한 학생의 답이 부분적으로는 옳지만 틀린 부분도 있는 경우, 틀린 부분이 옳은 부분과 모순되지 않는다면 틀린 부분을 무시하고 옳게 쓴 부분에 해당하는 채점 코드를 부여하게 됩니다. 마지막으로, 답하는 방식이 잘못되었다고 해도 학생이 알고 있는 내용이라고 판단하면 그에 해당하는 채점 코드를 부여합니다.

(4) PISA 서술형 문항 예시

이제부터는 PISA의 서술형 문항과 채점기준을 예시를 통해 구체적으로 살펴보도록 하겠습니다. 읽기, 수학, 과학 각각에 대해 해당 영역이 주영역이었던 가장 최근의 문항을 살펴보겠습니다.

(가) 읽기 서술형 문항 예시 (PISA 2009)

[읽기가 주영역으로 시행된 평가는 PISA 2009와 PISA 2018임. PISA 2018의 경우 공개 문항이 아직 배포되지 않았기에 PISA 2009의 읽기 문항을 예시로 제시해야 함. PISA 2009 읽기 문항 예시는 기존의 연수 프로그램 자료에 이미 제시되어 있어, 본 원고에서는 추가적으로 다루지 않음]

(나) 수학 서술형 문항 예시 (PISA 2012)

다음은 PISA 2012의 수학 문항입니다.

PISA 2012의 수학 문항

다음 표는 다섯 국가에서 TV를 보유한 가구의 현황을 나타낸다.

이 표는 TV를 보유한 가구의 비율과 TV보유 가구 중 케이블 TV에 가입한 가구의 비율을 백분율로 보여주고 있다.



국가	TV를 보유한 가구 수 (단위: 만명)	전체 가구 중 TV를 보유한 가구의 백분율	TV보유 가구 중 케이블 TV에 가입한 가구의 백분율
일본	4800	99.8 %	51.4 %
프랑스	2450	97.0 %	15.4 %
벨기에	440	99.0 %	91.7 %
스위스	280	85.8 %	98.0 %
노르웨이	200	97.2 %	42.7 %

출처: ITU, World Telecommunication Indicators 2004/2005
ITU, World Telecommunication/ICT Development Report 2006

[문제]

경주는 이 표에서 프랑스와 노르웨이의 자료를 살펴보았다.

경주는 "전체 가구 중 TV를 보유하고 있는 비율은 두 나라가 거의 비슷하기 때문에, 케이블 TV에 가입한 가구 수는 노르웨이가 더 많다."고 말했다.

이 말이 왜 옳지 않은지 설명하시오. 그렇게 답한 이유를 함께 쓰시오.

.....
.....
.....

이 문항은 제시된 자료를 수학적으로 해석하여, 주어진 진술이 자료에 근거하여 타당한 내용인지를 판단하고, 자신의 판단에 대해 논리적으로 설명할 수 있는지를 묻는 문항입니다. 앞서 설명한 PISA 수학 영역 평가틀을 적용해 보면, 사회적 ‘맥락’에서 불확실성과 자료에 대한 ‘수학적 내용’을 ‘해석하기’라는 수학적 과정을 적용하여 제작된 서술형 문항임을 알 수 있습니다. 이 문제를 풀기 위해 학생들은 백분율의 개념과 TV를 보유한 가구 수와 TV를 보유한 가구 중 케이블 TV에 가입한 가구의 상관관계를 이해해야 합니다. 이러한 자료 해석 과정을 통해 학생들은 문제의 대상인 프랑스와 노르웨이에서 케이블 TV에 가입한 실제 가구 수를 산정하여, 경주의 진술이 타당한지를 판단할 수 있어야 합니다.

이 문항에 대해 PISA가 제시한 채점 기준입니다.

PISA 채점 기준

채점 코드		채점 기준
만점	코드 11	경주가 두 국가의 실제 TV 보유 가구 수를 고려했어야 한다(고려하지 않았다)는 것을 지적한 경우 [‘인구’ 는 ‘가구 수’ 와 같은 것으로 인정] [예시 답안] • TV를 보유한 가구 수는 프랑스가 2200만 가구 이상 많기 때문에 경주가 틀렸다 • 프랑스의 인구가 노르웨이의 약 10배 정도 많기 때문에 • 프랑스 인구가 더 많기 때문에 TV를 보유한 인구도 훨씬 더 많다. 따라서 케이블 TV의 가입자 수도 더 많다.
	코드 12	두 국가의 실제 가입자 수를 계산하여 답한 경우 [예시 답안] • 프랑스는 (2450만 x 0.154) = 370만 가구가 케이블 TV에 가입한 반면, 노르웨이는 (200만 x 0.427) = 80만 가구가 가입하였으므로, 프랑스의 케이블 TV 가입자가 더 많다.
영점	코드 00	기타
	코드 99	무응답

채점 기준을 살펴보면 코드 11은 두 국가의 케이블 TV 가입 가구 수의 차이를 함축적으로 비교하여 설명한 경우이며, 코드 12는 실제 백분율 수치를 적용하여 가입 가구 수를 계산하여 그 크기를 비교 설명한 경우입니다. 이와 같이 PISA에서는 문제를 통해 측정하고자 하는 학생의 역량을 고려하여, 이를 나타내 줄 수 있는 다양한 답을 인정하고 있습니다. 채점 기준에는 이러한 다양한 답에 대한 부분이 반영되어 있으며, 특히 다양한 예시 답안을 제공하여 채점의 신뢰도를 높이고 있습니다.

(다) 과학 서술형 문항 예시 (PISA 2015)

이 문항은 PISA 2015의 과학 문항입니다. PISA 2015부터 모든 영역이 컴퓨터 기반 평가로 시행되었기에, 문제의 모습도 기존의 PISA 문항과는 다르게 배치되어 있음을 알 수 있습니다. 이 문항은 하나의 문제 상황에 대해 두 가지 문항으로 구성된 세트형

문항입니다. ‘경사면 조사’ 라는 현장 과학 실험 상황을 상정한 공통 도입부의 설명을 읽고 다음 화면으로 넘어가면 순서대로 자료 수집과 관련된 첫 번째 문항과 자료 분석과 관련된 두 번째 문항이 제시됩니다.

PISA 2015의 과학 문항



(2) 문항 1 : 자료 수집



(3) 문항 2 : 자료 분석



우선 첫 번째로 자료 수집과 관련된 문항을 자세히 살펴보겠습니다. 이 문항에서 학생들은 제시된 화면의 정보를 참고하여 각 경사면에 설치된 여러 종류의 측정 기구가 2개씩 설치된 이유를 설명해야 합니다. 앞서 설명한 PISA 과학 영역 평가틀을 적용해 보면, 지역적/국가적 ‘맥락’에서 지구-우주계와 관련한 인식론적 ‘지식’을 적용하는 과학 탐구의 평가/설계 ‘역량’을 측정하도록 제작된 서술형 문항입니다. 따라서 이 문항은 학생들이 과학적 탐구 설계를 설명하기 위해 인식론적 지식을 적용하도록 요구하며, 특히 탐구하는 현상에 대해 복수의 독립된 측정 도구를 사용해야 하는 실험 절차의 근본적인 이유를 설명할 수 있도록 적절한 문제 상황을 제시하고 있습니다. 이 문항에 대해 PISA가 제시한 채점 기준은 다음과 같습니다. 아쉽게도 PISA 2015의 경우 현재 채점 코드별

구체적인 채점 기준은 공개되지 않았습니니다. 하지만 앞서의 다른 영역의 문항에 대한 채점 기준과 마찬가지로 다양한 예시 답안을 제공하여 채점의 신뢰도를 높이고 있음을 여전히 확인할 수 있습니다.

PISA 채점기준 1

점수	채점 기준
만점	<p>각 경사지에서 2개 이상의 측정 기구를 사용하는 것의 과학적인 이점을 하나 이상 설명하는 경우</p> <p>[예시 답안]</p> <ul style="list-style-type: none"> • 경사지 간의 차이가 의미가 있는지 판단하기 위해 • 동일 경사지 내에서 차이(variation)가 있을 수 있기 때문에 • 각 경사지에서의 측정 정확성을 높이기 위해 • 데이터가 보다 정확할 것이기에 • 두 개 중 하나가 오작동하는 경우를 대비하기 위해 • 경사지에서의 태양의 양적 차이를 비교하기 위해 [비교는 차이가 있다는 것을 함축함]

두 번째 문항은 자료 분석에 대한 문항입니다. 이 문항은 주어진 자료에 대한 해석을 바탕으로, 두 경사지 간 토양 습도의 차이가 발생하는 이유에 대한 두 학생의 의견을 검토하여 어느 의견이 옳은지 판단하고, 이러한 판단에 대한 이유를 설명하는 서술형 문항입니다. PISA 과학 영역 평가틀을 적용해 보면, 지역적/국가적 ‘맥락’에서 지구·우주계와 관련한 인식론적 ‘지식’을 적용하는 자료/증거의 과학적 해석 ‘역량’을 측정하도록 제작되었습니다. 학생들은 이 문항에 대답하기 위해 주어진 자료를 해석하여 학생 1과 학생 2의 주장을 평가해야 하는데, 주어진 자료에는 태양 복사량, 토양 습도, 강우량에 대한 각각의 측정치의 평균에 대한 신뢰 구간이 포함되어 있습니다. 이 문항은 학생들이 측정 오류가 어떻게 과학적 측정과 관련한 신뢰도에 영향을 미치는지를 이해하고, 이러한 이해를 자료 분석에 적용하는가를 묻고 있는데, 이러한 이해와 적용은 과학의 인식론적 지식의 주요 측면이라 할 수 있습니다. 이 문항에 대해 PISA가 제시한 채점 기준은 다음과 같습니다.

PISA 채점기준 2

점수	채점 기준
만점	<p>학생1을 선택함. 그리고 두 경사지 간 태양 복사량에 차이가 있음을 나타내는 설명을 제시하고/제시하거나 [and/or] 강우량에는 차이가 없다고 설명하는 경우</p> <p>[예시 답안]</p> <ul style="list-style-type: none"> • 경사지 B가 경사지 A에 비해 훨씬 많은 태양 복사량을 얻으나, 강우량은 동일하다. • 두 경사지 간 강우량의 차이는 없기 때문이다. • 경사지 A가 얻는 햇빛의 양은 경사지 B와 비교하여 큰 차이가 있다.

2) TIMSS

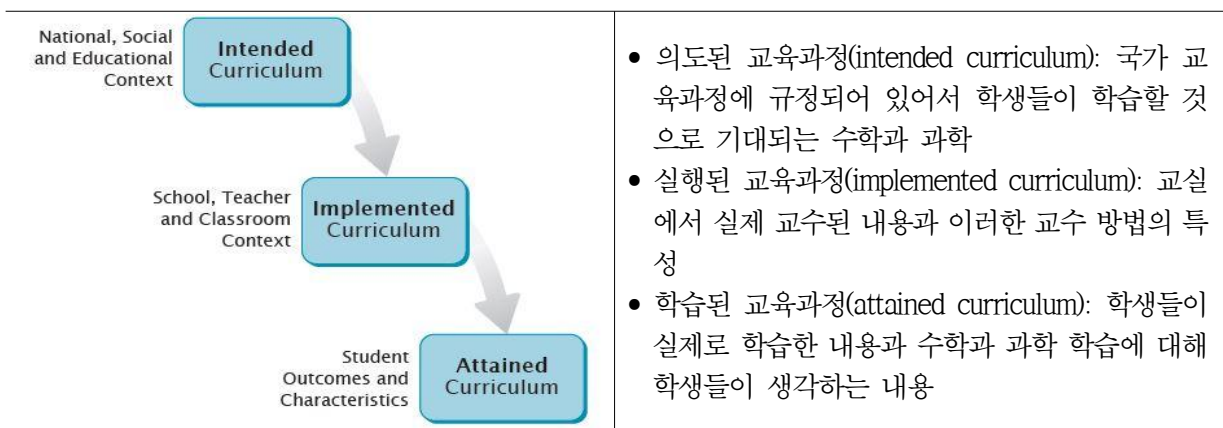
(1) TIMSS란?

TIMSS는 Trends in International Mathematics and Science Study의 약자입니다. TIMSS는 국제교육성취도 평가 협회(International Association for the Evaluation of Education Achievement, IEA)에서 주관하는 평가로, 우리나라의 초등학교 4학년과 중학교 2학년에 해당하는 학생들의 수학과 과학 과목에 대한 성취도를 파악하고, 성취도에 영향을 주는 요인을 분석하는 국제학업성취도평가의 하나입니다. TIMSS는 PISA와는 달리 참여 국가의 교육과정을 기반으로 학생들의 성취도를 평가하게 되며, 따라서 참여 국가의 교육과정 및 학교 교육의 효과성에 대한 정보를 직접적으로 보여주게 됩니다. TIMSS는 PISA보다 먼저 1995년부터 시작되었으며, 4년을 주기로 시행됩니다. 가장 최근에는 6번째 주기를 맞은 TIMSS 2015가 시행되었습니다. TIMSS 2015에는 초등학교 4학년 평가의 경우 총 49개국, 중학교 2학년 평가의 경우 총 39개국 참여하였습니다. 우리나라도 첫 번째 주기인 TIMSS 1995부터 평가에 지속적으로 참여하고 있으며, 최상위권의 성취도 평가 결과를 유지해 오고 있습니다.

(2) TIMSS의 평가틀

지금부터는 TIMSS의 평가틀을 포함해서 문항 제작과 관련된 내용을 살펴보겠습니다. TIMSS의 경우 수학과 과학 과목에 대해 학생들이 얼마나 성취했는가를 측정하는 국제적인 성취도평가로서, 국가 교육과정을 토대로 합니다. 즉, TIMSS는 참여국의 수학과 과학 교과 교육과정에 기초하여 평가 문항이 제작됩니다. 따라서 국가별로 교육과정이 다르기 때문에 수학과 과학 분야의 교육 전문가들이 각국의 교육과정을 조사하고, 핵심 내용을 추려 평가틀을 만듭니다. TIMSS의 교육과정 모형은 국가, 사회, 교육적 맥락에서의 ‘의도된 교육과정(intended curriculum)’, 학교, 교사, 교실 맥락에서의 ‘실행된 교육과정(implemented curriculum)’, 학생 성취 결과와 특성에 따른 ‘학습된 교육과정(attained curriculum)’으로 구분됩니다.

TIMSS의 평가틀



TIMSS의 평가들은 수학, 과학 과목별로 내용 영역과 인지 영역으로 구분됩니다. TIMSS의 평가들은 참여 국가의 교육과정 변경 사항을 반영하고, 성취도 추이 분석의 정확성을 확보하기 위해 매주기마다 소폭 수정됩니다. 가장 최근에 실시된 TIMSS 2015의 내용 영역과 인지 영역의 구성을 과목별로 살펴보겠습니다.

(가) 수학

중학교 2학년의 내용 영역은 ‘수, 대수, 기하, 자료와 가능성’의 4개 영역으로 구분됩니다. 인지 영역의 경우 ‘알기, 적용하기, 추론하기’의 3개 영역으로 구분됩니다. 이때 ‘알기’ 영역에는 ‘회상하기, 인식하기, 분류·정렬하기, 계산하기, 인출하기, 측정하기’가 하위 요소로 포함되고, ‘적용하기’ 영역에는 ‘결정하기, 표현·모델링하기, 실행하기’가 하위 요소로 포함됩니다. ‘추론하기’ 영역은 ‘분석하기, 통합·종합하기, 평가하기, 결론짓기, 일반화하기, 정당화하기’로 구성됩니다.

(나) 과학

중학교 2학년의 내용 영역은 ‘생물, 화학, 물리, 지구과학’의 4개 영역으로 구분됩니다. 인지 영역의 경우 ‘알기, 적용하기, 추론하기’의 3개 영역으로 구분됩니다. 이때 ‘알기’ 영역에는 ‘회상·인식하기, 정의하기, 기술하기, 예를 이용하여 설명하기, 과학적 도구의 사용법 설명하기’가 하위 요소로 포함되고, ‘적용하기’ 영역에는 ‘비교·대조·분류하기, 모델 사용하기, 관련짓기, 정보 해석하기, 해당찾기, 설명하기’가 하위 요소로 포함됩니다. ‘추론하기’ 영역은 ‘문제 분석하기, 통합·종합하기, 가설 설정·예상하기, 설계하기, 결론 도출하기, 일반화하기, 평가하기, 정당화하기’로 구성됩니다.

3) TIMSS 문항 구성과 채점의 특징

(1) TIMSS 문항 구성의 특징

TIMSS 역시 PISA와 마찬가지로 선다형 문항과 구성형 문항(서술형 문항)으로 구성되는데, 이 중 구성형 문항이 차지하는 비중은 약 45%에 이릅니다.

(2) TIMSS 구성형 문항 채점의 특징

TIMSS 구성형 문항의 채점은 채점 코드를 활용하여 이루어집니다. 문항의 복잡성 정도에 따라 완전한 정답일 때의 배점은 1점 혹은 2점으로 합니다. TIMSS 구성형 문항의 채점 코드는 PISA와 마찬가지로 학생들의 점수를 나타내는 숫자 코드와 학생들이 문제 해결에 사용한 방법에 대한 정보를 나타내는 숫자코드, 이렇게 두 자리 수로 구성됩니다. 따라서 두 자리 코드 중에서 첫 번째 자리는 답안의 정답 여부를 나타내며, 두 번째 자리는 학생의 사고 과정 분류나 오개념 추적에 사용됩니다.

TIMSS는 수학과 과학에 대한 학생의 성취 수준을 평가하는 것이기 때문에 쓰기 능력 자체는 평가되지 않습니다. 또한 학생들의 실제 답안에 기초하여 점수를 부여하고, 학생들이 보여주지 않은 사항에 대해 가정하여 점수를 부여하는 것을 금지하고 있습니다. 구성형 문항에 대한 채점 공정성과 신뢰성 확보를 위해 TIMSS에서는 채점 기준을 자세하고 분명하게 제시하며 채점자 연수도 실시합니다. 또한 본검사에 참여한 학생 수의 약 25%에 대해서는 복수 채점을 실시합니다.

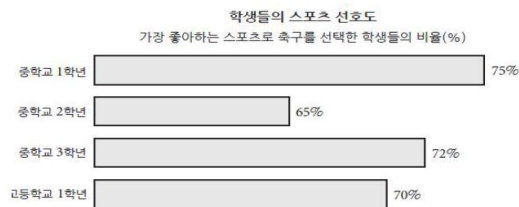
4) TIMSS 서술형 문항 예시

이제부터는 TIMSS의 서술형 문항과 채점기준을 구체적으로 알아보도록 하겠습니다. TIMSS의 최근 공개 문항 중 중학교 2학년을 대상으로 한 수학과 과학 과목의 예시 문항을 하나씩 살펴보겠습니다.

(1) 수학 과목 서술형 예시 문항

이 문항은 TIMSS 2015의 수학 문항입니다. 제시된 설문조사 결과 자려를 보고 이에 대한 해석이 잘못된 이유를 설명해야 하는 문항으로, ‘자료와 가능성’의 내용 영역 중 ‘자료 해석’의 하위요소와 ‘추론하기’의 인지 영역에 해당됩니다. 학생들은 문항에 함께 제시된 설문 조사 상황을 읽고, 설문 조사의 결과를 그래프로 표현한 자료를 토대로 ‘잘못된 해석’의 이유를 설명해야 합니다. 학생들은 이 문제에 대한 답을 찾기 위해 그림으로 주어진 막대그래프의 경우 기준점이나 단위 등이 표시되지 않았고 막대그래프의 값과 막대그래프의 길이가 정비례하지 않기 때문에, 그래프의 모양만 보고 수치로 주어진 %값을 고려하지 않는다면 그래프를 제대로 해석할 수 없다는 점에 유의해야 합니다. 이 문항은 우리나라의 중학교 수학 교육과정의 학습 내용 중 ‘주어진 자료를 표나 그래프로 나타내고, 이를 해석할 수 있다’에 해당합니다. 따라서 이 문항을 통해 자료와 그래프의 표현 및 해석과 관련한 수학 교과 교육과정 내용에 대해 학생들의 성취 여부를 판단할 수 있게 됩니다.

TIMSS 2015의 수학 문항



정호는 중학교 1학년에서 고등학교 1학년까지의 학생들을 대상으로 가장 좋아하는 스포츠를 조사하였다. 학년별로 100명이 응답하였다. 위의 그래프는 가장 좋아하는 스포츠로 축구를 선택한 학생들의 비율을 나타낸 것이다.

정호는 중학교 1학년과 2학년의 결과를 비교해 보았다. 정호는 중학교 1학년에서 축구를 선택한 학생 수가 2학년에서 축구를 선택한 학생 수의 두 배라고 생각하였다.

정호가 위의 그래프를 보고 잘못 생각하게 된 이유를 설명하시오.

이 문항에 대해 TIMSS가 제시한 채점 기준입니다. 채점 기준을 살펴보면 다양한 답안이 정답으로 인정되고 있음을 알 수 있습니다. 이러한 점을 통해, 학생들이 주어진 상황과 자료를 해석하고 추론하여 자신의 생각을 제시하도록 설계된 서술형 문항의 특성이 채점 기준에 잘 반영되고 있음을 확인할 수 있습니다.

TIMSS 채점 기준 1

채점 코드		채점 기준
정답	코드 10	<ul style="list-style-type: none"> • 중학교 1학년 막대 길이가 중학교 2학년 막대 길이의 두 배이다 또는 이와 동치인 표현 또는 • 시작 지점이 0이 아니다 또는 • 그래프에 척도(눈금)이 그려져 있지 않다.
오답	코드 79	오답(줄을 그어 지웠거나 의미 없는 표기, 답 표기란 외에 표기, 알아보기 어렵거나 문제와 상관없는 답을 한 경우 등)
무응답	코드 99	공란

(2) 과학 과목 서술형 예시 문항

이 문항은 TIMSS 2015의 과학 문항입니다. 내용 영역으로는 화학의 ‘물질의 성질’ 하위요소와 ‘적용하기’의 인지 영역에 해당합니다. 철 조각과 구리 조각이 섞여 있는 혼합물을 분리하는 방법을 제시된 두 가지 중에서 고르고 그 타당성을 설명해야 하는 서술형 문항으로, 질문이 ‘상황제시 - 선택 - 선택의 이유’의 형식으로 단계적으로 구성되어 있습니다. 이 문항의 경우 ‘선택한 방법이 적절한 이유’와 함께 ‘선택하지 않은 방법이 적절하지 않은 이유’, 즉 선택하지 않은 이유도 함께 묻고 있다는 점도 유의하여 살펴볼 필요가 있습니다. 이 문항을 풀기 위해 학생들은 혼합물의 분리와 관련하여 물질 간의 다른 성질, 즉 철과 구리가 각각 자석에 붙는지와 물에 녹거나 뜨는지를 모두 알아야 하고, 자신이 아는 내용에 대해 근거나 이유를 제시하는 방식으로 설명할 줄도 알아야 합니다. 자신이 고른 답안과 고르지 않은 답안에 대해 근거를 들어 논리적으로 설명하는 방식의 서술형 문항은 우리나라 학생들이 상대적으로 취약한 문항 형식입니다.

TIMSS 2015의 과학 문항

철 조각과 구리 조각이 섞여 있는 혼합물이 있다.

이 혼합물을 분리하기 위해 사용할 수 있는 방법은?

(하나만 고르시오.)

- 방법 1: 혼합물을 종이 위에 펼쳐 놓고 자석으로 훑는다.
- 방법 2: 혼합물을 물이 든 비커에 넣은 후 거른다.

1. 여러분이 선택한 방법이 적절한 이유를 설명하시오.

2. 여러분이 선택하지 않은 방법이 적절하지 않은 이유를 설명하시오.

이 문항에 대해 TIMSS가 제시한 채점 기준입니다. 채점 기준을 살펴보면, 이 문항은 2점이 만점이며, 부분답도 인정하여 1점을 부여하고 있습니다. 답안이 만점인 코드 20을 받기 위해서는 자신이 선택한 방법과 선택하지 않은 방법 모두에 대해 철과 구리가 지닌 물질의 성질에 근거하여 정확한 이유를 제시해야 합니다. 둘 중 한 가지 방법에 대해서만 정확한 이유를 제시했다면 부분 점수인 코드 10을 받게 됩니다.

TIMSS 채점 기준 2

채점 코드		채점 기준
정답	코드 20	방법 1을 선택하고, 아래의 두 가지 방법을 모두 언급한 경우 <ul style="list-style-type: none"> • 철이 자석에 붙기 때문에 방법 1은 제대로 작동할 것이다. (그리고 구리는 자석에 붙지 않는다) • 철과 구리는 물에 녹지 않기 때문에 그리고/또는 금속은 거름종이를 통과하지 못하기 때문에 방법 2는 제대로 작동하지 않을 것이다.
	코드 10	방법 1을 선택하고, 자석을 이용한 방법이나 여과 방법 중 하나에 대한 설명을 유효한 경우
오답	코드 79	오답(줄을 그어 지웠거나 의미 없는 표시, 답 표기란 외에 표기, 알아보기 어렵거나 문제와 상관없는 답을 한 경우 등)
무응답	코드 99	공란

2

해외 논술형 평가 사례: 미국의 조지아 마일스톤 평가, 유럽의 IB 교육과정의 평가

1) 조지아 마일스톤 평가

(1) 조지아 마일스톤 평가란?

조지아 마일스톤 평가(Georgia Milestones Assessment System, 이하 ‘마일스톤 평가’)는 미국의 조지주(州)에서 시행하고 있는 평가입니다. 마일스톤 평가는 주(州)에서 실시하는 일종의 학기말 종합시험이라고 볼 수 있으며, 영어(모국어), 수학, 과학, 사회 과목에 대한 학생의 지식과 기능 성취 여부를 평가합니다. 초등학교 3학년부터 우리나라의 중학교 2학년에 해당하는 8학년까지의 학생은 영어와 수학 과목에 대한 평가를 받게 되고, 5학년부터 8학년 학생은 여기에 더하여 과학과 사회 과목에 대한 평가를 추가로 받게 됩니다. 최종 학년말 성적의 최소 20%가 마일스톤 평가 결과를 반영하게 됩니다.

마일스톤 평가는 지난 2014학년도부터 기존의 선다형 문항 중심의 평가에서 벗어나 서·논술형 문항을 대폭 늘려 시행되고 있습니다. 특히 2017학년도 평가의 영어 과목(우리나라의 국어 과목에 해당)의 경우 문학과 쓰기 영역의 성취도가 대폭 향상되어 긍정적인 평을 받고 있습니다. 이번 학습에서는 마일스톤 평가의 영어 과목 논술형 문항을 중심으로 사례를 살펴보도록 하겠습니다.

(2) 마일스톤 평가의 구성

(가) 마일스톤 평가의 문항 체제

마일스톤 평가의 문항은 주어진 답안에서 답을 고르는 것에서 나아가, 어떻게 답을 찾아가는지, 학생이 응답한 답안을 뒷받침할 증거가 무엇인지를 묻는 것을 중시합니다. 마일스톤 평가에는 선다형 문항과 단답형 문항, 서술형 문항, 논술형 문항이 모두 포함되지만, 2014년의 평가 체제 개선을 통해 선다형 문항은 줄이고 논·서술형 문항을 확대하였습니다. 또한 영어 과목의 경우 읽기와 쓰기를 별도로 평가하던 것에서 벗어나 통합적으로 평가하게 되었습니다. 마일스톤 평가의 문항 체제를 문항 유형을 중심으로 정리해 보면 다음과 같습니다. 마일스톤 평가의 문항 구성은 매우 다양합니다. 흥미로운 것은 일반적인 서술형 문항과 논술형 문항의 중간 단계인 ‘확장된 서술형’ 문항을 두어 학생의 성취 수준을 다양하게 측정하고 있다는 점입니다.

마일스톤 평가의 문항 체제

문항 유형	문항 구성 특징	대상 과목	채점 관련 특징
선다형	<ul style="list-style-type: none"> • 사지선다형 	<ul style="list-style-type: none"> • 영어, 수학, 과학, 사회 	-
다중선택형, 다단계형	<ul style="list-style-type: none"> • 1개 이상의 정답 가능 • 첫 번째 문항에 대한 답안을 기반으로 두 번째 문항에 대한 답안을 작성하는 연계형 문항 	<ul style="list-style-type: none"> • 영어, 수학, 과학, 사회 	-
서술형	<ul style="list-style-type: none"> • 답을 직접 쓰거나 문제를 해결하는 문항으로 구성 • 선다형 문항 등과 세트형으로 제작 가능 	<ul style="list-style-type: none"> • 영어, 수학 	<ul style="list-style-type: none"> • 답안의 정확성, 학생 자신의 응답에 대한 근거 제시를 채점에 반영
확장된 서술형	<ul style="list-style-type: none"> • 서술형 문항과 유사하나, 	<ul style="list-style-type: none"> • 영어, 수학 	<ul style="list-style-type: none"> • 문항에서 언급한 조건

	서술형 문항보다 긴 답안을 요구하도록 제작		이 답안에 제시되어 있는지에 유의하여 채점
논술형	<ul style="list-style-type: none"> • 두 개의 지문을 활용하여 제작 • 답안 작성을 위해 복잡한 사고를 요구하는 문항으로 제작 • 읽기와 쓰기 능력을 통합적으로 측정할 수 있도록 제작 	• 영어	<ul style="list-style-type: none"> • 문항에서 언급한 조건이 답안에 제시되어 있는지와 글의 질(quality)을 고려하여 채점 • 지문에 제시된 정보를 근거로 주장이나 결론을 제시했는가에 유의하여 채점

마일스톤 평가의 논술형 문항은 영어 과목에만 해당하는 문항 유형입니다. 논술형 문항의 경우 서술형 문항과 달리 채점 기준에 ‘글의 질’이 포함된다는 점에서 큰 차이가 있습니다. 또한 긴 지문을 읽고, 지문에 제시된 정보를 활용하여 근거를 제시해야 하기 때문에 읽기와 쓰기의 통합적 성격을 지닌다는 점도 마일스톤 논술형 평가의 주요 특징입니다. 마일스톤 평가는 이와 같은 다양한 문항 유형의 조합으로 구성됩니다. 일례로 8학년 학생의 영어 과목 평가의 문항 구성은 다음과 같습니다. 문항 수와 점수 배점을 살펴보면, 서술형은 2문항 각 4점, 확장된 서술형은 1문항 4점, 논술형은 1문항 7점을 부여하여, 문항 유형별로 문항 수와 배점을 달리 하고 있음을 알 수 있습니다.

영어 과목 평가의 문항 구성

문항 유형	문항 수	CR 점수	NRT 피드백 점수
CR* 선다형	28	28	0
NRT 선다형**	20	10	20
CR 다단계형	1	2	0
CR 서술형	2	4	0
CR 확장된 서술형	1	4	0
CR 논술형	1	7	0
CR 개발 점검 문항(Field test items)	8	0	0
총문항수/점수	61	55	20

*CR: criterion-referenced (성취 기준에 따른 준거 참조 평가)

**NRT 선다형: norm-referenced selected-response items (규준 참조 평가; 여기서는 전국적인 비교를 위해 포함된 문항)

(나) 마일스톤 평가의 논술형 평가

① 마일스톤 논술형 문항 구성의 특징

마일스톤 평가의 논술형 문항에서 요구하는 글쓰기 장르는 주로 서사적인 글(narrative

writing), 정보전달적인 글(informational writing) 또는 설명적인 글(explanatory writing)과 설득하는 글(argumentative writing)입니다. 이러한 글쓰기 장르는 영어 과목의 쓰기 성취기준, 즉 ‘~글을 쓸 수 있다’ 와 같은 성취기준 기술에 대응하는 문항 유형이라고 할 수 있습니다. 마일스톤 평가의 논술형 문항은 제시된 지문을 읽고, 지문의 소재나 중심 내용과 관련된 글쓰기 주제에 대한 학생 자신의 의견이나 생각을 쓰도록 하는데, 이때 학생의 의견이나 생각에 대한 근거를 지문에서 찾아 쓰도록 요구하게 됩니다. 따라서 마일스톤 평가의 논술형 문항은 대부분 읽기 및 문학 영역과 쓰기 영역의 통합적 문항으로 구성됩니다.

② 마일스톤 논술형 문항 채점의 특징

이번에는 마일스톤 평가의 논술형 문항은 어떻게 채점되는지 알아보겠습니다. 먼저 마일스톤 평가 논술형 문항의 만점과 채점 기준은 글쓰기 장르에 따라 다르게 설정되어 있습니다. 먼저 서사적인 글은 4점 만점으로 총괄적 채점기준(holistic rubric)을 적용합니다. 정보전달적인 글 또는 설명적인 글과 설득하는 글은 7점 만점으로, ‘두 가지 특성 채점기준’ (two-trait rubric)을 적용합니다. 8학년의 경우 주로 설명하는 글이나 설득하는 글을 쓰게 되는 경우가 많기 때문에, 이번 학습에서는 이들 글쓰기 장르에 초점을 맞추어 살펴보겠습니다.

마일스톤 평가의 채점 기준은 다소 복잡한 체계로 구성되어 있습니다. 먼저 답안의 글쓰기 장르가 ‘정보전달적인 글 또는 설명적인 글’ 인지, ‘설득하는 글’ 인지에 따라 다른 채점 기준의 내용이 다릅니다. 두 번째로 두 가지 특성 채점 기준을 사용하여 일정 부분 분석적으로 진행됩니다. 이 때, 특성 1(Trait 1) 채점 기준은 논술의 사고적 측면과 관련한 기준이며, 특성 2(Trait 2) 채점 기준은 논술의 언어적 측면과 관련한 기준입니다. 즉, 특성 1의 채점 기준은 ‘내용의 전개와 구성 및 일관성(Idea development, Organization, & Coherence)’ 를 다루는데, 내용적 측면을 통합적으로 채점하여 최소 0점에서 최대 4점까지 부여할 수 있습니다. 특성 2의 채점 기준은 ‘언어 사용과 규범(Language usage & Conventions)’ 을 평가하게 되는데, 언어적 측면을 통합적으로 채점하여 최소 0점에서 3점까지 부여할 수 있습니다. 마일스톤 평가의 논술형 문항은 학생이 작성한 답안을 내용적 측면과 언어 사용의 측면으로 나누어 채점하는 방식으로, 일종의 간략한 분석적 채점을 사용한다고 볼 수 있습니다.

8학년(우리나라의 중학교 2학년) 대상 마일스톤 평가 논술형 문항의 채점 기준

논술 답안의 글쓰기 장르	채점기준		최대 점수
	분류	내용	
정보전달적인 글 또는 설명하는 글 (Informational/Explanatory Genre)	특성 1 채 점 기준	[내용의 전개와 구성 및 일관성] • 주어진 지문에 근거하여 자신의 생각(idea)을 효과적으로 표현하는 능력 • 주어진 지문에 제시된 근거를 이용하여 자신의 생각을 뒷	4

		받침할 수 있는 능력 • 자신의 생각을 예시, 묘사(illustration), 사실 및 기타 구체적인 근거를 들어 상술할 수 있는 능력 • 주어진 지문에서 찾은 정보를 자신의 언어로 통합하여, 정보전달적인/설명하는 글의 논리적 일관성을 유지하는 능력	
	특성 2 채점 기준	[언어 사용과 규범] • 문장 구성과 언어 사용 및 기술에 대한 전반적인 언어 통제 능력 • 학년에 따른 언어 능력 기준에 적합한 수준	3
	총점		7
설득하는 글 (Argumentative Genre)	특성 1 채점 기준	[내용의 전개와 구성 및 일관성] • 반대 주장에 대해 다루는 것을 포함하여 자신의 주장(claim)을 효과적으로 표현하는 능력 • 주어진 지문에 제시된 근거를 이용하여 자신의 주장을 뒷받침할 수 있는 능력 • 자신의 주장을 예시, 묘사(illustration), 사실 및 기타 구체적인 근거를 들어 상술할 수 있는 능력 • 주어진 지문에서 찾은 정보를 자신의 언어로 통합하여, 설득하는 글의 논리적 일관성을 유지하는 능력	4
	특성 2 채점 기준	[언어 사용과 규범] • 문장 구성과 언어 사용 및 기술에 대한 전반적인 언어 통제 능력 • 학년에 따른 언어 능력 기준에 적합한 수준	3
	총점		7

③ 마일스톤 평가의 논술형 문항 예시

이제부터는 예시를 통해 마일스톤 평가의 논술형 문항에 대해 보다 자세히 살펴보도록 하겠습니다. 이 문항은 8학년 학생들을 대상으로 시행되는 마일스톤 평가의 가이드북에 제시된 예시 문항입니다. 이 문항은 답안으로 설득적인 글을 작성해야 하는 논술형 문항으로, 학생들은 논란이 있는 어느 주제에 대해 어느 한 쪽의 편에 서서, 왜 자신의 주장이 타당한지를 상술하는 내용의 논술을 답안으로 작성해야 합니다. 즉, 이 문항에 대한 답안을 작성하기 위해 학생들은 유전자 변형(Genetically Modified, GM) 식품에 대한 옹호론과 비판론에 대해 모두 생각해 본 뒤, 유전자 변형 식품이 생산 및 소비되어야 하는가의 문제에 대해 어느 한 쪽의 편에서 설득적인 글을 써야 합니다.

마일스톤 평가의 논술형 문항 예시

WRITING TASK

There is an ongoing debate over the use of genetically modified (GM) foods. Think about BOTH sides of the debate. Should GM foods be produced and consumed? Write an **argumentative essay** in your own words supporting either side of the debate. Be sure to use information from BOTH passages in your **argumentative essay**.

Writer's Checklist

Be sure to:

- Introduce your claim.
- Support your claim with logical reasons and relevant evidence from the passages.
- Acknowledge and address alternate or opposing claims.
- Organize the reasons and evidence logically.
- Develop your ideas clearly and use your own words, except when quoting directly from the passages.
- Identify the passages by title or number when using details or facts directly from the passages.
- Use words, phrases, or clauses to connect ideas and to clarify the relationships among claims, counterclaims, reasons, and evidence.
- Establish and maintain a formal style.
- Use clear language and vocabulary.
- Provide a conclusion that supports the argument presented.
- Check your work for correct usage, grammar, spelling, capitalization, and punctuation.

Now write your **argumentative essay** on your answer document. Refer to the **Writer's Checklist** as you write and proofread your essay.

그런데 이 문항의 내용을 자세히 살펴보면 쓰기 과제(writing task)의 마지막 문장으로 ‘반드시 두 지문에서 제시된 정보를 사용하여 설득하는 논술을 작성할 것(Be sure to use information from BOTH passages in your argumentative essay)’이라는 문구가 포함되어 있음을 알 수 있습니다. 여기에서 관심 있게 봐야 할 것은 “Be sure to use information from BOTH passages” 라는 강조 문구입니다. 즉, 설득하는 논술을 작성할 때, 학생 자신의 의견이나 아이디어만으로 글을 쓰는 것이 아니라, 반드시 제시된 두 개의 지문에서 제시된 정보를 활용하여 글을 쓰라는 것입니다.

실제로 이 논술형 문항은 동일한 두 개의 지문을 기반으로 선다형-서술형-논술형의 3개 문항으로 구성된 세트형 문항의 마지막 문항입니다. 학생들은 처음에 나오는 선다형 문항이 시작되기 전에 다음과 같은 문항 관련 안내를 받게 됩니다. 안내에는 ‘지문을 읽으면서 유전자 변형 식품의 사용과 관련한 설득적인 논술에 사용할 수 있는 세부 정보에 대해 생각해 보라(As you read the passages think about details you may use in an argumentative essay about the use of genetically modified (GM) food)’는 내용이 담겨 있습니다. 이를 통해 학생들은 논술형 문항에 대한 답안을 작성하는데 필요한 지문의 세부 정보를 파악하는데 유의하여 지문을 읽을 수 있게 됩니다.

논술형 문항 예시 1

This section of the test assesses your skill to comprehend reading passages and use information from the passages to write an argumentative essay.

Before you begin writing your essay, you will read two passages and answer one multiple-choice question and one short constructed-response question about what you have read.

As you read the passages think about details you may use in an argumentative essay about the use of genetically modified (GM) food.

These are the titles of the passages you will read:

1. GM Food Saves Lives
2. What We Don't Know about GM Food

학생에게 설득하는 글을 쓰도록 요구하는 논술형 문항의 경우, 찬반이 나뉘어 논쟁이 되는 소재나 내용을 다루기 때문에 주어지는 지문 역시 찬성하는 주장과 반대하는 주장을 모두 포함하는 것이 중요합니다. 또한 문항이 요구하는 과업의 하나가 지금의 예시 문항과 같이 주어진 지문에서 필요한 근거를 제시하는 것이라면, 지문에서도 저자의 주장을 뒷받침하는 충분한 정보를 구체적으로 포함하는 글을 선정하는 것이 필요합니다.

논술형 문항 예시 2

GM Food Saves Lives by Rebecca Wilson	What We Don't Know about GM Food by Daniel McLeod
<p>Genetically modified (GM) food was introduced to the citizens of the United States in 1994. Since then, the use of genetics on produce and animals has become so widespread that each person in the United States is most likely eating GM food daily. A primary reason for its popularity is how beneficial it is to people and businesses.</p> <p>What is genetic modification?</p> <p>Plants and animals naturally go through a process of selection for survival. Features that make the plant or animal more likely to live are passed along, and features that are not advantageous are weeded out. These genetic mutations occur over generations, though, making improvement a slow-moving process. Scientists discovered that they could improve specific characteristics quickly by introducing foreign genes into an organism, such as those from plants, animals, and even viruses. For example, exposing a plant to a certain virus can make it more resistant to disease. Transferring genes from cows to pigs can help the pigs create more milk for larger litters of piglets. The targeting of genes allows scientists to bring out the specific traits of a product that will make it more successful.</p> <p>Uses of GM foods</p> <p>There are three main reasons for genetically modifying food: to produce more food at lower cost, to increase the health value of the food, and to make the food more desirable. When crops are modified to withstand disease and drought, it takes fewer resources to produce them, and fewer crops are lost. But altering food goes much further than this. Scientists are also able to make food more nutritious. For example, Golden Rice is infused with vitamin A in the hopes of saving the lives of children suffering from vitamin A deficiencies. However, the earliest uses of GM food are still the most popular. Genetic modification makes food look and taste better. Tomatoes stay ripe longer. Apples have fewer bruises. Strawberries grow larger.</p> <p>Safety</p> <p>Opponents of GM food say that changing an organism's genetic code is dangerous. They say that changes to a plant's durability can create superweeds that kill crops and that altering nutrition values could cause health problems for the people who eat the food. Yet thousands of research studies have shown no evidence that GM food causes harm, either to the environment or to people. It's safe, effective, and needed in a time when food shortages are skyrocketing.</p>	<p>Humans have a history of moving forward with great ideas—until they realize that those ideas weren't so great. Back in the 1940s, people around the world started using a miracle insecticide called dichlorodiphenyltrichloroethane (DDT). It killed every annoying insect out there! It was helping to eliminate malaria-carrying mosquitoes and life-threatening spiders. DDT was the best insecticide ever—until people realized the severe damage it was doing to the environment. It took over thirty years of using the chemical agent for scientists to verify the problems and for countries to ban DDT's use. Only now, seventy years since it became popular, are some of the species negatively affected by it finally regaining a foothold on life.</p> <p>Genetically modified (GM) food is our generation's DDT. Just as before, people have jumped headlong into the process of making food better, stronger, and different through changes to an organism's genetic code. Scientists are altering plants and animals at their most fundamental levels with no regard to the effects we might see in twenty, thirty, or even seventy years from now. True, this process is producing food at a lower cost and higher rate, something this world desperately needs, but at what cost?</p> <p>There have been documented cases of genetically altered crops affecting the durability of weeds that compete for the crops' resources. It's believed the genetic mutation of the crops spread to the weeds. These weeds, called superweeds, are aggressive and resistant to the chemicals used to kill them and now threaten the crops' growth. Another current problem is the reduction in insects such as butterflies and bees, which pollinate flowers. Crops designed to produce natural insecticides are killing off these important creatures. The ecosystem is thrown off balance without them.</p> <p>Those problems are nothing compared to the ones we don't know about yet. How will these modifications affect the humans who consume this food over a lifetime? How will unforeseen mutations affect the food? These questions can't be answered right now since we won't see the effects for decades.</p> <p>The biogenetics companies that produce GM food say the food has been tested by thousands of studies. What they don't say, however, is that they are the ones who funded the studies. Their financial interest in studies showing that GM food is safe compromises the believability of the studies. How might their corporate dollars have affected the results the scientists are reporting?</p> <p>The plain truth is that we don't know how GM food will affect humans, plants, and animals in the future. We shouldn't be risking our lives by eating altered food without knowing whether or not genetic modification is another DDT.</p>

종합해보면, 유전자 변형 식품의 사용과 관련된 학생의 의견을 논하라는 예시 문항은 앞서 설명한 마일스톤 평가의 논술형 문항의 특징을 잘 보여줍니다. 첫째로 마일스톤 평가의 논술형 문항은 학생의 의견이나 생각에 대한 근거를 지문에서 찾아 쓰도록 요구하여, 증거에 기반한 논리적인 사고와 이를 반영한 글쓰기 능력을 강조합니다. 예시 문항에서도 유전자 변형 식품의 사용을 찬성하는 지문(“GM Food Saves Lives”)과 반대하는 지문(“What We Don't Know about GM Food”)을 모두 읽고, 두 지문에서 제시된 정보를 모두 사용하여 논술형 답안을 작성할 것을 요구되고 있습니다. 둘째로 마일스톤

평가의 논술형 문항은 주어진 지문을 잘 읽고 관련된 정보를 이해할 수 있어야 문항이 요구하는 답안을 작성할 수 있는 읽기와 쓰기의 영역 통합적 성격을 강조합니다. 예시 문항에서도 유전자 변형 식품의 사용에 대한 찬반 지문의 정보를 정확하게 이해해야 논술에도 높은 점수를 얻을 수 있도록 읽기와 쓰기가 통합적으로 구성되어 있습니다. 예시 문항을 통해 살펴볼 수 있는 마일스톤 평가 논술형 문항 구성의 또 다른 특징은 문항 내에 ‘작성자 확인 사항(writer’s check list)’를 제시하여, 학생들이 답안 작성 시 반드시 고려해야 할 내용 구성과 언어 사용 측면의 주요 사항이나 유의점을 안내하고 있다는 것입니다. 이러한 작성자 확인 사항은 앞서 설명한 마일스톤 논술형 문항의 두 가지 특성 채점 기준(논술의 내용적 측면에 대한 채점인 특성 1 채점 기준과 논술의 언어적 측면에 대한 채점인 특성 2 채점 기준)과 상응하는 것입니다. 또한 문항의 맨 마지막에 ‘논술 답안을 작성하고 교정할 때 작성자 확인 사항을 참조하시오(Refer to the Writer’s Checklist as you write and proofread your essay)’라는 문장을 통해 작성자 확인 사항을 논술 답안 작성의 기준으로 참고할 것을 다시 한 번 강조하고 있습니다. 따라서 학생들은 이러한 문항 구성을 통해 자신의 논술 답안이 어떤 기준에 따라 채점될 것인지를 알 수 있게 됩니다.

마일스톤 평가 논술형 문항 구성 예시 1

WRITING TASK

There is an ongoing debate over the use of genetically modified (GM) foods. Think about BOTH sides of the debate. Should GM foods be produced and consumed? Write an **argumentative essay** in your own words supporting one side of the debate.

Be sure to use information from BOTH passages in your **argumentative** essay.

Writer’s Checklist

Be sure to:

- Introduce your claim.
- Support your claim with logical reasons and relevant evidence from the passages.
- Acknowledge and address alternate or opposing claims. *Organize the reasons and evidence logically.*
- Develop your ideas clearly and use your own words, except when quoting passages.
- Identify the passages by title or number when using details or facts directly from the passages.
- Use words, phrases, or clauses to connect ideas and to clarify the relationships among claims, counterclaims, reasons, and evidence.
- Establish and maintain a formal style.
- Use clear language and vocabulary.
- Provide a conclusion that supports the argument presented.
- Check your work for correct usage, grammar, spelling, capitalization, and punctuation.

Now write your **argumentative essay** on your answer document. Refer to the **Writer’s Checklist** as you write and proofread your essay.

예시 문항에 적용되는 채점 기준을 살펴보겠습니다. 예시 문항에서 요구하는 답안은 설득하는 글이기 때문에 채점 기준 역시 이에 따라 설득하는 글 장르에 해당하는 채점 기준을 사용하게 됩니다. 앞서 설명한 바와 같이 논술 답안의 내용적 측면과 관련한 4점 만점의 특성 1 채점 기준과 3점 만점의 특성 2 채점 기준을 별도로 적용하여 각각의 점수를 산출한 뒤, 두 점수를 합산하여 논술형 문항 답안에 대한 최종 점수를 산출하게 됩니다. 따라서 각 특성에서 모두 만점을 받은 답안의 최종 점수는 7점이 됩니다. 각

특성별 채점 기준은 0점에서 만점까지 점수별 답안의 특징을 기술하고 있는데, 점수별 답안의 종합적 측면을 먼저 제시하고 항목별 세부 내용을 구체적으로 기술하는 방식을 사용하고 있습니다. 예를 들어, 특성 1 채점 기준에서 ‘3점’을 받는 답안의 경우, “학생은 자신의 주장과 관련하여 지문에 제시된 근거를 일부 사용하여 온전히 논증하고 있다(The student’s response is a complete argument that relates and supports claims with some text-based evidence)”라는 종합적 특성을 먼저 제시하고, 도입(introduction), 구성 전략(organization strategy), 주장을 발전시키기 위한 지문 내용의 사용(use of sources to develop claims) 등에 대한 항목별 특성을 구체적으로 기술하고 있습니다. 따라서 채점자들은 이러한 구체적인 채점 기준을 통해 항목별 특성에 따른 학생 답안의 완성도를 평가하게 되며, 결과적으로 학생의 답안을 보다 객관적으로 채점할 수 있게 됩니다.

예시 문항에 적용되는 채점 기준

Trait 1 for Argumentative Genre		
Writing Trait	Points	Criteria
Idea Development, Organization, and Coherence This trait examines the writer’s ability to effectively establish a claim as well as to address counterclaims, to support the claim with evidence from the text(s) read, and to elaborate on the claim with examples, illustrations, facts, and other details. The writer must integrate the information from the text(s) into his/her own words and arrange the ideas and supporting evidence in order to create cohesion for an argument essay.	4	The student’s response is a well-developed argument that effectively relates and supports claims with clear reasons and relevant text-based evidence. <ul style="list-style-type: none"> Effectively introduces claim(s) Uses an organizational strategy to present reasons and relevant evidence logically Supports claim(s) with clear reasons and relevant evidence using specific, well-chosen facts, details, or other information from credible sources and demonstrates a good understanding of the topic or text Acknowledges and counters opposing claim(s), as appropriate Uses words, phrases, and/or clauses that effectively connect and show direct, strong relationships among claim(s), reasons, and evidence Establishes and maintains a formal style that is appropriate for the task, purpose, and audience Provides a strong concluding statement or section that logically follows from the argument presented
	3	The student’s response is a complete argument that relates and supports claims with some text-based evidence. <ul style="list-style-type: none"> Clearly introduces claim(s) Uses an organizational strategy to present some reasons and evidence Uses specific facts, details, definitions, examples, and/or other information from sources to develop claim(s) Attempts to acknowledge and/or counter opposing claim(s), as appropriate Uses words and/or phrases to connect ideas and show relationships among claim(s), reasons, and evidence Uses a formal style fairly consistently for task, purpose, and audience Provides a concluding statement or section that follows from the argument presented
	2	The student’s response is an incomplete or oversimplified argument that partially supports claims with loosely related text-based evidence. <ul style="list-style-type: none"> Attempts to introduce claim(s) Attempts to use an organizational structure which may be formulaic Develops, sometimes unevenly, reasons and/or evidence to support claim(s) Makes little, if any, attempt to acknowledge or counter opposing claim(s) Attempts to support claim(s) with facts, reasons, and other evidence sometimes, but logic and relevancy are often unclear Uses few words or phrases to connect ideas; connections are not always clear Uses a formal style inconsistently or an informal style that does not fit task, purpose, or audience Provides a weak concluding statement or section that may not follow the argument presented
	1	The student’s response is a weak attempt to write an argument and does not support claims with adequate text-based evidence. <ul style="list-style-type: none"> May not introduce claim(s)/claim(s) must be inferred May be too brief to demonstrate an organizational structure, or no structure is evident Has minimal support for claim(s) Makes no attempt to acknowledge or counter opposing claim(s) Uses minimal or no words or phrases to connect ideas Uses a very informal style that is not appropriate for task, purpose, or audience Provides a minimal or no concluding statement or section
	0	The student will receive a condition code for various reasons: <ul style="list-style-type: none"> Blank Copied Too Limited to Score/Illegible/Incomprehensible Non-English/Foreign Language Off Topic/Off Task/Offensive

Trait 2 for Argumentative Genre		
Writing Trait	Points	Criteria
Language Usage and Conventions This trait examines the writer’s ability to demonstrate control of sentence formation, usage, and mechanics as embodied in the grade-level expectations of the language standards.	3	The student’s response demonstrates full command of language usage and conventions. <ul style="list-style-type: none"> Effectively varies sentence patterns for meaning, reader/listener interest, and style Shows command of language and conventions when writing Any errors in usage and conventions do not interfere with meaning⁴
	2	The student’s response demonstrates partial command of language usage and conventions. <ul style="list-style-type: none"> Varies some sentence patterns for meaning, reader/listener interest, and style Shows some knowledge of language and conventions when writing Has minor errors in usage and conventions with no significant effect on meaning⁴
	1	The student’s response demonstrates weak command of language usage and conventions. <ul style="list-style-type: none"> Has fragments, run-ons, and/or other sentence structure errors Shows little knowledge of language and conventions when writing Has frequent errors in usage and conventions that interfere with meaning⁴
	0	The student will receive a condition code for various reasons: <ul style="list-style-type: none"> Blank Copied Too Limited to Score/Illegible/Incomprehensible Non-English/Foreign Language Off Topic/Off Task/Offensive

다음으로는 유전자 변형 식품의 사용과 관련한 예시 문항에 대해 마일스톤 평가에서 제시한 모범 답안을 살펴보겠습니다. 모범 답안은 앞에서 살펴 본 채점 기준을 적용했을 때 만점인 7점을 받을 수 있는 답안입니다. 찬반 논쟁이 있는 사안에 대해 어느 한쪽을 지지하는 설득적인 글을 작성하라는 논술 과제의 특성상, 마일스톤 평가에서 제공하는 모범 답안 역시 찬성인 경우와 반대인 경우 두 가지로 제시되고 있습니다.

먼저 학생이 유전자 변형 식품의 사용을 찬성하는 내용의 글을 쓴 경우의 모범 답안을 살펴보겠습니다. 우선 전체적으로 5개의 문단으로 구성되어 있습니다. 첫 번째 문단은 학생 자신의 의견은 “세상의 사람들이 음식을 먹고 건강을 유지하기 위해 유전자 변형 식품을 생산하고 필요하다”는 것임을 명확하게 밝히고 있으며, 이러한 주장에 대한 근거를

제시하여 자신의 주장을 뒷받침 하고 있습니다. 특히 두 번째와 세 번째 문단에서는 유전자 변형 식품을 찬성하는 지문으로 제시된 Rebecca Wilson의 글을 구체적으로 인용하여 자신의 주장에 대한 근거를 제공하고 있습니다. 또한 네 번째 문단에서는 유전자 변형 식품을 반대하는 주장에 대해 언급하면서(“People question how safe GM food is.”), 다시 Wilson의 글을 이용하여 이러한 반대 주장에 대한 반증을 제시합니다. 이러한 구성은 앞서 언급한 바와 같이 마일스톤 평가의 논술형 문항 구성에서 강조하고 있는 증거 기반 글쓰기(“Be sure to use information from BOTH passages”)를 반영한 것입니다.

유전자 변형 식품의 사용과 관련한 예시 문항

Producing and consuming genetically modified food is necessary to feed the people of the world and keep them healthy. Since GMOs have been introduced, farmers have been able to grow more food that uses fewer resources.

Rebecca Wilson states that we are all most likely eating genetically modified foods on a daily basis. She claims that “A primary reason for its popularity is how beneficial it is to people and businesses.” GM food grows bigger and tastier than regular food. This makes it more appealing to customers. The more customers want the food, the more they will buy it. This will help businesses succeed.

Modifying food also makes it stronger. According to Wilson, “When crops are modified to withstand disease and drought, it takes fewer resources to produce them, and fewer crops are lost.” Therefore, using genetic modification means there will be more food for more people. The food will also be healthier. An example of this is how vitamin A is being introduced to rice to help nourish children.

People question how safe GM food is. However, Wilson notes that thousands of studies have been done on it, and there is “no evidence that GM food causes harm, either to the environment or to people.” If it’s safe and effective, why not use it?

GM food is here to stay, and that is a good thing. The more we use genetic modification, the more we will be able to provide healthy food to more people.

유전자 변형 식품의 사용을 찬성하는 경우의 모범 답안

학생의 답안이 유전자 변형 식품의 사용을 반대하는 경우에 대한 모범 답안은 다음과 같습니다.

유전자 변형 식품의 사용을 반대 모범 답안

There’s no question that using genetic modification grows more food. The problem is that the food it produces has not been proven safe, so we shouldn’t be producing or consuming it.

As Daniel McLeod illustrates with the example of DDT, we often don’t know the negative effects a scientific discovery will have on us in the years to come. DDT went from being a miracle insecticide to a threat to the ecosystem. GM food could offer the same kind of threat to humans.

While GM foods may be bigger, stronger, and tastier, they also may be dangerous. The studies done to prove their safety were done by the companies selling the products themselves. McLeod rightfully asks, “How might their corporate dollars have affected the results the scientists are reporting?” We can’t trust results that are financially motivated.

Unfortunately, as Rebecca Wilson says, “The use of genetics on produce and animals has become so widespread that each person in the United States is most likely eating GM food daily.” Unless we take steps to avoid it, we have no choice but to consume something that could kill us. For that reason, modifying foods must stop.

유전자 변형 식품의 사용을 반대하는 경우의 모범 답안

지금까지 조지아 마일스톤 평가의 특징과 예시 문항을 살펴보았습니다. 마일스톤 평가의 논술형 문항은 학생의 의견이나 생각에 대한 근거를 지문에서 찾아 쓰도록 요구하여, 증거에 기반한 논리적인 사고와 이를 반영한 글쓰기 능력을 강조합니다. 또한 주어진 지문을 잘 읽고 관련된 정보를 이해할 수 있어야 문항이 요구하는 답안을 작성할 수 있는 읽기와 쓰기의 영역 통합적 성격을 강조합니다. 이러한 논술형 문항의 특징은 지금까지

살펴본 예시 문항과 채점 기준 및 모범 답안을 통해서도 그대로 구현되고 있음을 알 수 있었습니다.

2) IB 교육과정이란

(1) IB 교육과정이란?

IB(International Baccalaureate, IB) 교육과정은 스위스 제네바에 본부가 있는 비영리교육재단인 IB에서 운영하는 국제공인 교육과정으로, 150개 국가의 4,000여 학교에서 사용되고 있습니다. IB 교육과정은 암기를 기반으로 하는 주입식 교육이 아닌 토론과 논술을 중심으로 창의력과 비판적 사고력, 문제 해결 능력을 기르는 데 중점을 두고 있다는 점에서 최근 4차 산업혁명 시대 도래에 따른 미래형 인재 양성 노력과 관련하여 많은 관심을 받고 있습니다. IB 교육과정은 대상 학생의 나이와 진학 또는 취업 목표 여부에 따라 3세-12세의 초등 교육과정(Primary Years Programme, PYP), 11세-16세의 중등 교육과정(Middle Years Programme, MYP), 16-19세의 학위 교육과정(Diploma Programme, DP)과 직업 교육과정(Career-related Programme, CP)의 4개 교육과정으로 나뉩니다. 단위 학교에서 IB 교육과정을 도입하기 위해서는 IB 기관의 인증을 받아야 합니다.

평가의 측면에서 IB 교육과정이 세계적인 주목을 받는 이유는 기존의 선다형이나 단답형 평가의 틀에서 벗어나 모든 과목에서 학생 중심의 실행형 수행 평가와 확장된 서술형 및 논술형 평가를 시행하기 때문입니다. 이러한 학생 중심의 평가와 서술형 평가는 학생이 스스로 생각하고 비판적으로 판단하며, 적절한 질문을 형성할 수 있는 기술을 습득하도록 유도하는 데 효과적입니다. IB 교육과정의 평가는 과목별로 단위학교의 자체 평가(internal assessment)와 IB 기관 차원에서 전 세계적으로 시행하는 외부 평가(external assessment)로 구성됩니다. 외부 평가의 경우 IB 기관이 지정한 독립 채점관이 학생의 답안을 채점하게 됩니다. IB 교육과정의 평가는 이와 같이 엄격하게 관리되기 때문에, 전 세계적으로도 그 공신력을 인정받고 있습니다.

(2) IB 교육과정의 논술형 문항

IB 교육과정에서 에세이 형태의 논술형 평가가 주로 적용되는 과목은 어문 계열의 과목입니다. 과학 계열 과목의 경우, 수업에서 학습한 지식 개념을 토대로 실제로 실험을 계획하고 실행한 결과를 기술하는 보고서 형태의 평가나 과학적 개념을 적용하여 문제 해결을 유도하는 과제 중심형 서술형 평가를 주로 사용합니다. 이러한 점을 고려하여 이번 학습에서는 우리나라의 고등학교 교육과정에 해당하는 IB DP의 ‘Language A: Language and Literature(문학을 포함한 모국어 수준의 언어 과목)’의 표준 수준(Standard Level, SL) 과목의 평가가 어떻게 구성되었는지에 초점을 맞추어 살펴보도록 하겠습니다.

이 과목의 평가 개요를 살펴보면 앞서 설명한 바와 같이 외부 평가와 자체 평가로 구성되어 있습니다. 논술형 평가는 3가지의 외부 평가 중 하나의 평가 항목으로

수행됩니다.

과목 평가 개요

구분	평가 항목	내용	시간	점수	성적 반영 비율
외부 평가	시험1 (Paper 1): 자료 분석(textual analysis)	<ul style="list-style-type: none"> 시험은 두 개의 새로운 자료로 구성됨(two unseen texts) 학생들은 자료 중 하나에 대한 분석 작성 	90분	20	25%
	시험2 (Paper 2): 논술(essay)	<ul style="list-style-type: none"> 문학 관련 수업에서 학습한 작품에 대한 6개 질문 중 1개를 선택하여 답안 작성 	90분	25	25%
	서면 과제 (Written task)	<ul style="list-style-type: none"> 수업에서 학습한 자료에 근거하여 최소 2개의 서면 과제를 수행하여, 이중 한가지를 외부 평가를 위해 제출 각 과제는 800-1000 단어 수준으로 작성해야 하며, 200-300 단어를 근거(rationale) 제시를 위해 추가할 수 있음 	-	20	20%
	총 반영 비율				
자체 평가	개별 구두 평가(Individual oral commentary)	<ul style="list-style-type: none"> 문학 비평 수업 부분에서 읽은 문학 작품에 대한 학생 개인의 구두 논평 학생들은 평가 수행을 위해 2개의 안내 질문(guiding questions)을 받음 	30	15	
	추가 구두 활동(Further oral activity)	<ul style="list-style-type: none"> ‘문화적 맥락에서의 언어’와 ‘언어와 대중매체’ 수업 부분과 관련하여, 각각 최소 1개의 추가적인 구두 활동을 완료 	30	15	
	총 반영 비율				

논술형 평가의 경우 학생들은 수업에서 배운 문학 작품과 관련한 6개의 질문 중 하나를 선택하여 90분 안에 답안을 작성해야 합니다. 공개된 IB 교육과정의 과목 매뉴얼에는 논술형 문항의 실제 예시는 제공되지 않지만, 평가를 준비하기 위해 고려해야 할 쟁점들의 예시가 제시되어 있습니다. 따라서 교사와 학생은 이러한 쟁점들을 고려하여 문학 작품에 대한 토론 수업을 진행하게 됩니다.

IB 교육과정의 과목 매뉴얼

-
- How can we explain the continued interest in a particular work in different contexts and at different times?
 - What do you think of the assertion that the meaning of a text is fixed and does not change over time?
 - If beauty is a relative term, how do one or more of the works you have studied explore this idea?
 - How valid is the assertion that literature is a voice for the oppressed?
 - To what extent do male and female literary characters accurately reflect the role of men and women in society?
 - To what purpose do authors sometimes choose not to follow a chronological sequence of events in their literary works?
 - Do works of literary merit both reflect the spirit of the time and challenge it?
-

지금까지 살펴본 IB DP의 ‘Language A: Language and Literature’의 SL 과목 논술형 문항의 채점 기준을 살펴보겠습니다. 채점은 ‘지식과 이해(Knowledge and understanding)’, ‘질문에 대한 응답(Response to the question)’, ‘문체 표현의 사용과 효과에 대한 이해(Understanding of the use and effects of stylistic features)’, ‘(내용)구성과 발전(Organization and development)’, ‘언어(Language)’의 5가지 준거(criterion)에 따라 분석적으로 이루어집니다. 각 준거별 최대 점수는 5점으로, 이를 합산하면 논술형 문항의 총점인 25점이 됩니다. 각 준거에 대한 채점 기준은 표와 같습니다. 우선 준거 A는 ‘지식과 이해’에 대한 것으로, 학생의 답안에 학습한 문학 작품과 그 맥락에 대한 지식과 이해가 얼마나 나타나 있는가를 기준으로 점수를 부여하게 됩니다. 준거 B는 ‘질문에 대한 응답’에 대한 것으로, 질문의 주된 예상(main expectation of the question)에 대한 이해가 얼마나 제시되었는가와 답안이 이러한 예상에 얼마나 부합하는가를 기준으로 점수를 부여하게 됩니다. 준거 C는 ‘문체 표현의 사용과 효과에 대한 이해’에 대한 것으로, 작가가 문체 선택을 통해 어떻게 의미를 구성했는지에 대한 인식이 얼마나 나타나 있는가와 답안이 문체의 효과에 대한 이해를 얼마나 보여주는가를 중심으로 점수를 부여하게 됩니다. 준거 D는 답안의 ‘구성과 전개’에 대한 것으로, 논술의 논거가 얼마나 조리 있고 효과적인가와 논술의 형태적 구조가 얼마나 효과적인가를 기준으로 점수를 부여하게 됩니다. 마지막으로 준거 E는 ‘언어 사용’과 관련한 것으로, 답안 작성에 사용된 언어가 얼마나 명확하고 다채로우며 정확한가와 사용역(register)의 선택이 얼마나 적절한가를 기준으로 점수를 부여하게 됩니다.

SL 과목 논술형 문항 채점 기준

항목	채점 기준														
<p>준거 A : 지식과 이해</p> <ul style="list-style-type: none"> ● 학생의 답안에 학습한 문학 작품과 그 맥락에 대한 지식과 이해가 얼마나 나타나 있는가? 	<table border="1" style="width: 100%; border-collapse: collapse;"> <thead> <tr style="background-color: #d9d9d9;"> <th style="width: 15%;">Marks</th> <th style="width: 85%;">Level descriptor</th> </tr> </thead> <tbody> <tr> <td style="text-align: center;">0</td> <td>The work does not reach a standard described by the descriptors below.</td> </tr> <tr> <td style="text-align: center;">1</td> <td>Little knowledge of the part 3 works is demonstrated.</td> </tr> <tr> <td style="text-align: center;">2</td> <td>Some knowledge of the part 3 works and their context is demonstrated, but understanding is limited.</td> </tr> <tr> <td style="text-align: center;">3</td> <td>Knowledge of the part 3 works and the way in which context affects their meaning is adequately demonstrated and shows a general understanding.</td> </tr> <tr> <td style="text-align: center;">4</td> <td>Knowledge of the part 3 works and the way in which context affects their meaning is substantially demonstrated, and the understanding shown is good.</td> </tr> <tr> <td style="text-align: center;">5</td> <td>Knowledge of the part 3 works and the way in which context affects their meaning is thoroughly demonstrated, and the understanding shown is very good.</td> </tr> </tbody> </table>	Marks	Level descriptor	0	The work does not reach a standard described by the descriptors below.	1	Little knowledge of the part 3 works is demonstrated.	2	Some knowledge of the part 3 works and their context is demonstrated, but understanding is limited.	3	Knowledge of the part 3 works and the way in which context affects their meaning is adequately demonstrated and shows a general understanding.	4	Knowledge of the part 3 works and the way in which context affects their meaning is substantially demonstrated, and the understanding shown is good.	5	Knowledge of the part 3 works and the way in which context affects their meaning is thoroughly demonstrated, and the understanding shown is very good.
Marks	Level descriptor														
0	The work does not reach a standard described by the descriptors below.														
1	Little knowledge of the part 3 works is demonstrated.														
2	Some knowledge of the part 3 works and their context is demonstrated, but understanding is limited.														
3	Knowledge of the part 3 works and the way in which context affects their meaning is adequately demonstrated and shows a general understanding.														
4	Knowledge of the part 3 works and the way in which context affects their meaning is substantially demonstrated, and the understanding shown is good.														
5	Knowledge of the part 3 works and the way in which context affects their meaning is thoroughly demonstrated, and the understanding shown is very good.														
<p>준거 B : 질문에 대한 응답</p> <ul style="list-style-type: none"> ● 질문의 주된 예상(main expectation of the question)에 대한 이해가 얼마나 제시되었는가? ● 답안이 이러한 예상에 얼마나 부합하는가? 	<table border="1" style="width: 100%; border-collapse: collapse;"> <thead> <tr style="background-color: #d9d9d9;"> <th style="width: 15%;">Marks</th> <th style="width: 85%;">Level descriptor</th> </tr> </thead> <tbody> <tr> <td style="text-align: center;">0</td> <td>The work does not reach a standard described by the descriptors below.</td> </tr> <tr> <td style="text-align: center;">1</td> <td>There is little awareness of the main expectations of the question.</td> </tr> <tr> <td style="text-align: center;">2</td> <td>There is some awareness of the main expectations of the question; the response is mainly unsubstantiated generalization.</td> </tr> <tr> <td style="text-align: center;">3</td> <td>There is adequate awareness of the main expectations of the question, with a generally relevant response.</td> </tr> <tr> <td style="text-align: center;">4</td> <td>There is good understanding and awareness of the main expectations of the question, with a mostly relevant response.</td> </tr> <tr> <td style="text-align: center;">5</td> <td>There is very good understanding and awareness of the expectations of the question, with a consistently relevant response.</td> </tr> </tbody> </table>	Marks	Level descriptor	0	The work does not reach a standard described by the descriptors below.	1	There is little awareness of the main expectations of the question.	2	There is some awareness of the main expectations of the question; the response is mainly unsubstantiated generalization.	3	There is adequate awareness of the main expectations of the question, with a generally relevant response.	4	There is good understanding and awareness of the main expectations of the question, with a mostly relevant response.	5	There is very good understanding and awareness of the expectations of the question, with a consistently relevant response.
Marks	Level descriptor														
0	The work does not reach a standard described by the descriptors below.														
1	There is little awareness of the main expectations of the question.														
2	There is some awareness of the main expectations of the question; the response is mainly unsubstantiated generalization.														
3	There is adequate awareness of the main expectations of the question, with a generally relevant response.														
4	There is good understanding and awareness of the main expectations of the question, with a mostly relevant response.														
5	There is very good understanding and awareness of the expectations of the question, with a consistently relevant response.														
<p>준거 C : 문체 표현의 사용과 효과에 대한 이해</p> <ul style="list-style-type: none"> ● 작가가 문체 선택을 통해 어떻게 의미를 구성했는지에 대한 인식이 얼마나 나타나 있는가? ● 답안이 문체의 효과에 대한 이해를 얼마나 보여주는가? 	<table border="1" style="width: 100%; border-collapse: collapse;"> <thead> <tr style="background-color: #d9d9d9;"> <th style="width: 15%;">Marks</th> <th style="width: 85%;">Level descriptor</th> </tr> </thead> <tbody> <tr> <td style="text-align: center;">0</td> <td>The work does not reach a standard described by the descriptors below.</td> </tr> <tr> <td style="text-align: center;">1</td> <td>There is little awareness or illustration of the use of stylistic features.</td> </tr> <tr> <td style="text-align: center;">2</td> <td>There is some awareness and illustration of stylistic features.</td> </tr> <tr> <td style="text-align: center;">3</td> <td>There is adequate awareness and illustration of stylistic features, with some understanding of their effects.</td> </tr> <tr> <td style="text-align: center;">4</td> <td>There is good awareness and illustration of stylistic features, with adequate understanding of their effects.</td> </tr> <tr> <td style="text-align: center;">5</td> <td>There is very good awareness and illustration of stylistic features, with good understanding of their effects.</td> </tr> </tbody> </table>	Marks	Level descriptor	0	The work does not reach a standard described by the descriptors below.	1	There is little awareness or illustration of the use of stylistic features.	2	There is some awareness and illustration of stylistic features.	3	There is adequate awareness and illustration of stylistic features, with some understanding of their effects.	4	There is good awareness and illustration of stylistic features, with adequate understanding of their effects.	5	There is very good awareness and illustration of stylistic features, with good understanding of their effects.
Marks	Level descriptor														
0	The work does not reach a standard described by the descriptors below.														
1	There is little awareness or illustration of the use of stylistic features.														
2	There is some awareness and illustration of stylistic features.														
3	There is adequate awareness and illustration of stylistic features, with some understanding of their effects.														
4	There is good awareness and illustration of stylistic features, with adequate understanding of their effects.														
5	There is very good awareness and illustration of stylistic features, with good understanding of their effects.														

<p>준거 D : 구성과 전개</p>	<ul style="list-style-type: none"> • 논술의 논거가 얼마나 조리있고 효과적인가? • 논술의 형태적 구조가 얼마나 효과적인가? 	<table border="1"> <thead> <tr> <th>Marks</th> <th>Level descriptor</th> </tr> </thead> <tbody> <tr> <td>0</td> <td>The work does not reach a standard described by the descriptors below.</td> </tr> <tr> <td>1</td> <td>There is little focus, structure and development.</td> </tr> <tr> <td>2</td> <td>There is some focus, structure and development.</td> </tr> <tr> <td>3</td> <td>There is adequate focus, structure and development.</td> </tr> <tr> <td>4</td> <td>There is good focus, structure and development.</td> </tr> <tr> <td>5</td> <td>There is very good focus, structure and development.</td> </tr> </tbody> </table>	Marks	Level descriptor	0	The work does not reach a standard described by the descriptors below.	1	There is little focus, structure and development.	2	There is some focus, structure and development.	3	There is adequate focus, structure and development.	4	There is good focus, structure and development.	5	There is very good focus, structure and development.		
Marks	Level descriptor																	
0	The work does not reach a standard described by the descriptors below.																	
1	There is little focus, structure and development.																	
2	There is some focus, structure and development.																	
3	There is adequate focus, structure and development.																	
4	There is good focus, structure and development.																	
5	There is very good focus, structure and development.																	
<p>준거 E : 언어 사용</p>	<ul style="list-style-type: none"> • 언어가 얼마나 명확하고 다채로우며 정확한가? • 사용역(register)의 선택이 얼마나 적절한가? 	<table border="1"> <thead> <tr> <th>Marks</th> <th>Level descriptor</th> </tr> </thead> <tbody> <tr> <td>0</td> <td>The work does not reach a standard described by the descriptors below.</td> </tr> <tr> <td>1</td> <td>Language is rarely clear and appropriate; there are many errors in grammar, vocabulary and sentence construction and little sense of register and style.</td> </tr> <tr> <td>2</td> <td>Language is sometimes clear and carefully chosen; grammar, vocabulary and sentence construction are fairly accurate, although errors and inconsistencies are apparent; the register and style are to some extent appropriate to the task.</td> </tr> <tr> <td>3</td> <td>Language is clear and carefully chosen with an adequate degree of accuracy in grammar, vocabulary and sentence construction despite some lapses; register and style are mostly appropriate to the task.</td> </tr> </tbody> </table> <table border="1"> <thead> <tr> <th>Marks</th> <th>Level descriptor</th> </tr> </thead> <tbody> <tr> <td>4</td> <td>Language is clear and carefully chosen, with a good degree of accuracy in grammar, vocabulary and sentence construction; register and style are consistently appropriate to the task.</td> </tr> <tr> <td>5</td> <td>Language is very clear, effective, carefully chosen and precise, with a high degree of accuracy in grammar, vocabulary and sentence construction; register and style are effective and appropriate to the task.</td> </tr> </tbody> </table>	Marks	Level descriptor	0	The work does not reach a standard described by the descriptors below.	1	Language is rarely clear and appropriate; there are many errors in grammar, vocabulary and sentence construction and little sense of register and style.	2	Language is sometimes clear and carefully chosen; grammar, vocabulary and sentence construction are fairly accurate, although errors and inconsistencies are apparent; the register and style are to some extent appropriate to the task.	3	Language is clear and carefully chosen with an adequate degree of accuracy in grammar, vocabulary and sentence construction despite some lapses; register and style are mostly appropriate to the task.	Marks	Level descriptor	4	Language is clear and carefully chosen, with a good degree of accuracy in grammar, vocabulary and sentence construction; register and style are consistently appropriate to the task.	5	Language is very clear, effective, carefully chosen and precise, with a high degree of accuracy in grammar, vocabulary and sentence construction; register and style are effective and appropriate to the task.
Marks	Level descriptor																	
0	The work does not reach a standard described by the descriptors below.																	
1	Language is rarely clear and appropriate; there are many errors in grammar, vocabulary and sentence construction and little sense of register and style.																	
2	Language is sometimes clear and carefully chosen; grammar, vocabulary and sentence construction are fairly accurate, although errors and inconsistencies are apparent; the register and style are to some extent appropriate to the task.																	
3	Language is clear and carefully chosen with an adequate degree of accuracy in grammar, vocabulary and sentence construction despite some lapses; register and style are mostly appropriate to the task.																	
Marks	Level descriptor																	
4	Language is clear and carefully chosen, with a good degree of accuracy in grammar, vocabulary and sentence construction; register and style are consistently appropriate to the task.																	
5	Language is very clear, effective, carefully chosen and precise, with a high degree of accuracy in grammar, vocabulary and sentence construction; register and style are effective and appropriate to the task.																	

IB 교육과정의 논술형 평가는 이처럼 매우 정교하고 세부적으로 구성된 채점 기준을 적용하고 있습니다. 앞서 소개한 조지아 마일스톤 평가의 논술형 평가의 경우, 답안의 내용적 측면과 언어 사용 측면의 2가지 특성에 대해서만 각각의 채점 기준을 적용하고, 각각 4점과 3점을 만점으로 하여 총 7점 만점의 점수를 부여한 것과 비교해서도, IB 논술형 평가의 채점 기준이 매우 정교함을 알 수 있습니다. 이는 IB 교육과정이 전 세계에서 공통으로 적용되는 국제적인 교육과정인 동시에, 교육과정의 핵심적 요소를 지키되 개별 학교의 자율성을 부여하는 특징이 있다는 점에서 공통적인 외부 평가를 통해 과목 이수에 대한 학점의 질 관리를 철저하게 하여 교육과정의 공신력을 확보하기 위한 것으로 이해될 수 있습니다.