



## 세계의 학생 평가 International student assessment

### 성적이 아닌 학습을 위한 평가

여러 교육 선진국은 학생 평가를 단순한 성적 산출이나 선발 도구가 아닌 학습 지원 수단으로 활용하는 방향으로 전환하고 있다. 평가를 통해 학생이 무엇을 알고 무엇을 할 수 있는지 파악하여 개별 학습을 돋고, 다음 학습 단계에 대한 피드백을 제공하는 역할이 강조된다. OECD 등 국제기구들도 평가가 학습 성과 향상에 기여하도록 지식 암기 위주의 평가에서 역량 중심 평가로의 변화를 강조하고 있다.

[자료 출처: OECD]

### 전인적 역량을 반영한 평가 지표의 개발

OECD Education 2030 프로젝트는 지식뿐 아니라 협업, 감정조절, 시민성과 같은 전인적 역량을 포함하는 핵심역량 중심 평가체계를 제안하고 있다. 이에 발맞춰 세계 각국에서는 미래 사회에 필요한 핵심 역량을 평가하기 위해 필기 평가뿐 아니라 프로젝트, 포트폴리오, 수행평가 등 다양한 방식을 도입하고 있다. 국제기구들 역시 새로운 역량 평가의 중요성을 강조하며, OECD는 국제학업성취도평가(PISA)에 글로벌 역량(2018)과 창의적 사고력(2022) 영역을 신설하는 등 평가 범위를 확대했다.

[자료 출처: OECD]

### 새로운 평가 도구의 등장

AI 기반 평가 도구는 학생의 학습 진행 상황을 확인하고, 오류를 파악하며, 개인화된 피드백을 제공하는 등 교사의 평가 업무의 부담을 크게 줄여준다. 인도에서는 방과 후에 제공되는 개인형 평가 소프트웨어의 활용이 특히 성적이 낮은 학생들의 학습 향상에 긍정적인 영향을 미치는 것으로 나타났다. 그러나 미국에서는 AI 기반 학습 및 평가 시스템을 사용하더라도 전통적인 방법보다 학업 성과를 더 크게 개선했다고 보기 어렵다는 결과가 나타나기도 하였다. 이렇듯 새로운 평가 도구의 등장은 그 자체로 의미가 있다기보다, 학생 성장이라는 평가의 본래 목적에 얼마나 기여할 수 있느냐에 따라 그 가치가 결정될 것이다.

[자료 출처: UNESCO]



## 독일

### ▶ 독일의 역량 핵심 중심 평가 방안

- 학생 평가는 역량 중심 절대 평가 시행
- 수업 과정에서 협력 중시, 진로의 다양성 인정, 성장 중심 평가 지향
- 교육 표준에 따른 평가 모니터링으로 교육의 질 개선, 학생 역량 함양 추구

### ▶ 교육 평가

국제  
학업  
성취도  
비교



피사인턴셔널  
(PISA-International)



팀스  
(TIMSS)



PIRLS  
/IGLU

교육  
표준에  
대한  
검토  
및 적용



Institut zur Qualitätsentwicklung  
im Bildungswesen

연방정부 차원(IQB)



학교 수준(VERA)

주정부  
차원  
평가



바덴뷔르템베르크주의 나침반 4(Kompass 4)



핀란드의 학생 평가는 단순한 결과 확인을 넘어 학습 과정과 성장을 지원하는 데 초점을 두며, 전인적 역량까지 평가에 포함시킨다. 이러한 접근은 평가를 위한 평가가 아닌 '학습을 위한 평가(assessment for learning)'라는 핀란드의 학생 평가 철학이 반영되어 있다.

### ▶ 핀란드 학생 평가의 특징



교사 평가권의 강조



표준화 시험의  
최소화



과정중심평가를  
통한 상시 평가



목표 설정 및 평가  
과정에서의 학생 참여

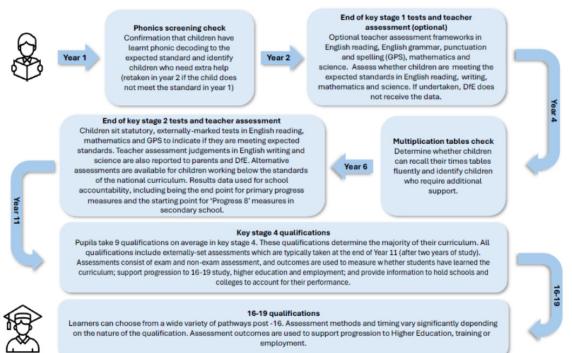


전인적 역량 평가



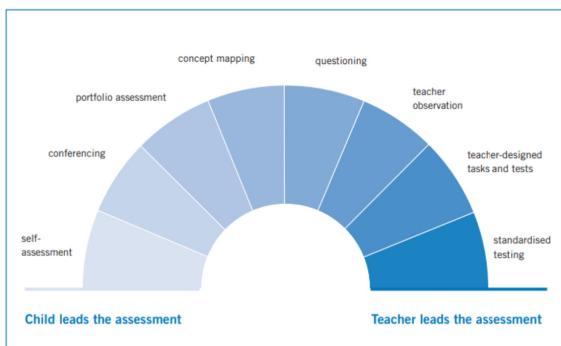
### ▶ 국가 표준 시험(SATs) 공정성 확보

Figure 4. National assessments and exams in England's education system



### ▶ 교사평가(Teacher Assessment) 개인의 성장 중심

Figure 2: A continuum of assessment methods

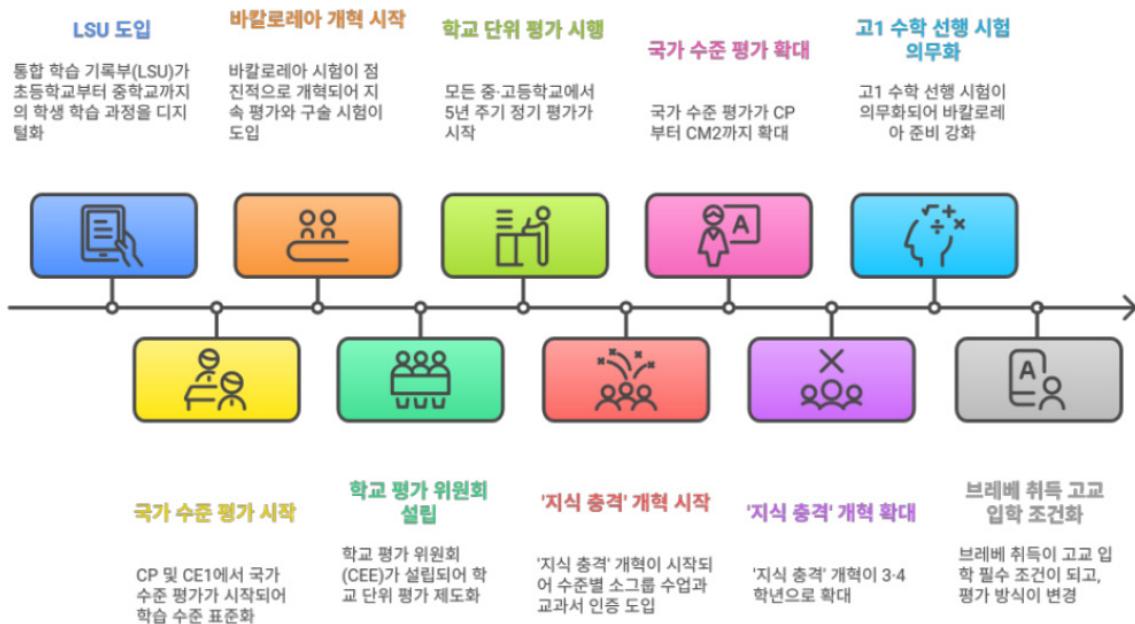




## 프랑스

지난 10년간 프랑스 교육 평가 시스템은 통합 학습 기록부(LSU) 도입, 국가 표준화 평가 확대, 바칼로레아 개혁, 학교 단위 평가, '지식 충격(Choc des savoirs)' 등 주요 변화를 겪었다. 이들 개혁은 중앙집중화, 데이터 기반 관리, 결과 중심 교육 전환을 보여준다. 그러나 교육학적 목표와 제도적 목표 간 긴장, 평등·공정성 확보의 어려움, 교사 자율성 대 표준화 갈등, 정책의 현장 수용 및 실행 문제 등 도전 과제는 여전하다. 특히 '지식 충격' 개혁은 공화주의 이상과 충돌하며 심각한 논쟁을 야기했다. 이러한 과제가 반복되는 것은 단일 정책을 넘어 시스템 내부의 구조적·문화적 문제를 시사한다고 하겠다.

### ▶ 프랑스 교육계의 평가 방식 개혁의 주요 이정표





## 미국

### ▶ 학생의 가능성에 주목하는 평가 혁신, 미국의 방향성

#### • 진단평가

(1) STAAR (State of Texas Assessments of Academic Readiness)

#### • 수행평가(Performance-based Assessment): 비판적 사고와 창의성의 무대

(1) Massachusetts Consortium for Innovative Education Assessment (MCIEA)

(2) 캘리포니아 STEM 고등학교: 실제 문제 중심 설계 과제 도입

#### • 형성평가(Formative Assessment)

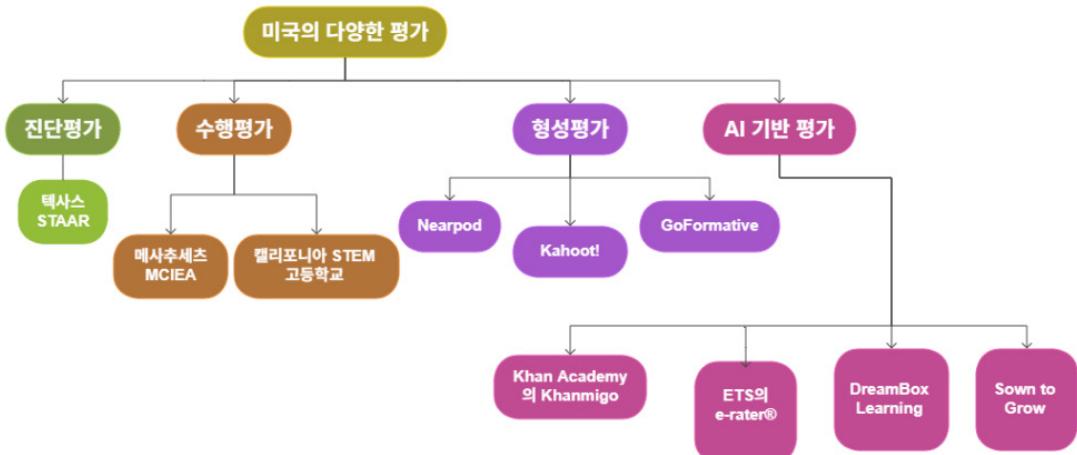
(1) Nearpod (2) Kahoot! (3) GoFormative (Formative)

#### • AI 기반 평가: 자동화된 및 춤형 피드백과 학습 지원

(1) ETS의 e-rater® (2) Khan Academy의 Khanmigo

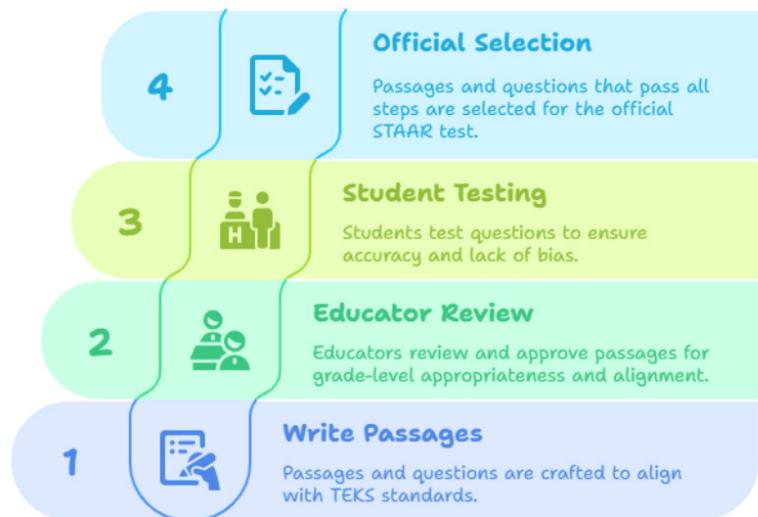
(3) DreamBox Learning (4) Sown to Grow

### 미국의 평가 혁신 방향성



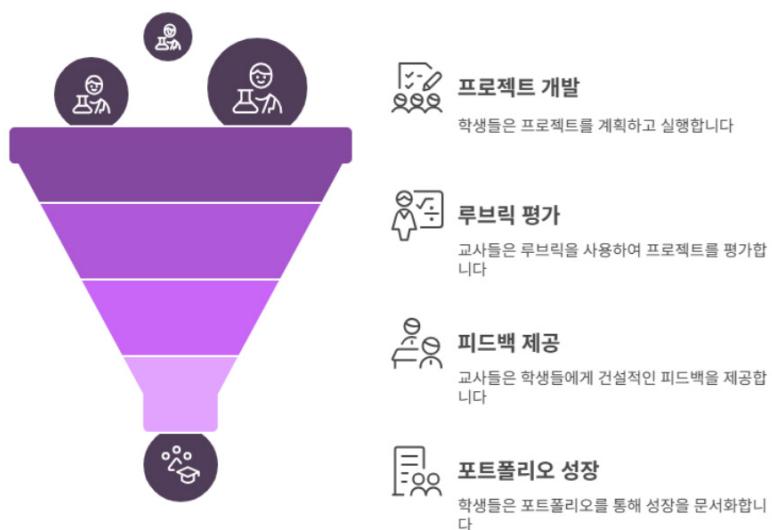
## ▶ 진단평가(Diagnostic Assessment): STAAR

1. Write Passages
2. Educator Review
3. Student Testing
4. Official Selection



## ▶ 수행평가(Performance-based Assessment): MCIEA

### MCIEA 평가 프로세스





## 캐나다

### ▶ 교원 전문성에 기반한 AfL와 AaL 중심 평가 체계

- AfL(Assessment for Learning, 형성평가) : 과정 중심  
학습 중 학습 진행 상황 파악 및 피드백 제공, 학습 방향 조정 목적
- AaL(Assessment as Learning, 자기평가) : 메타인지 중심  
학습자 스스로 학습 과정 성찰 및 점검, 자기주도 학습 촉진 목적
- AfL과 AaL 실현 : 교원 지율성 및 전문성 위주 평가 체계

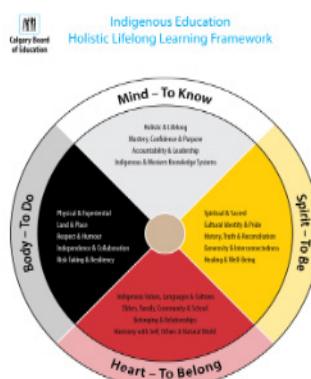
### ▶ 표준화 평가 확대 시도



- 2024 알버타 주의 표준화 평가 확대 발표
- 평가 연령 확대 : 유치원생 대상 평가 사상 최초 도입
- 평가 횟수 확대 : 초등 졸업까지 치는 평가 횟수 최대 10회 → 32회
- 교원단체의 강한 반발 "학생들에게 필요한 것은 평가가 아닌 지원"

### ▶ 학생 평가의 형평성과 포용성

- 다문화 배경 학습자  
온타리오 주 CCRP) 학습자 문화를 고려한 반응적인 평가
- ESL, ELL 학습자  
영어, 프랑스어를 모국어로 사용하지 않는 학생 대상 평가 가이드라인 제시
- 원주민 학습자  
원주민 학생 문화를 존중하는 평생 학습 구조와 인정 기반 평가

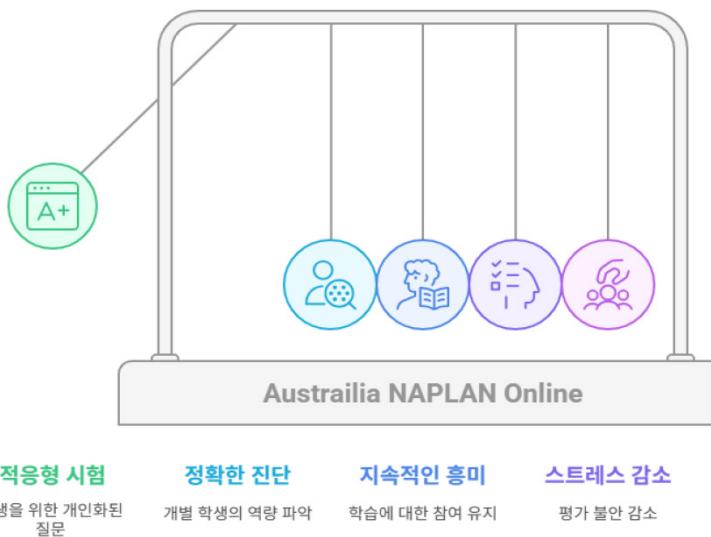




## 호주와 뉴질랜드

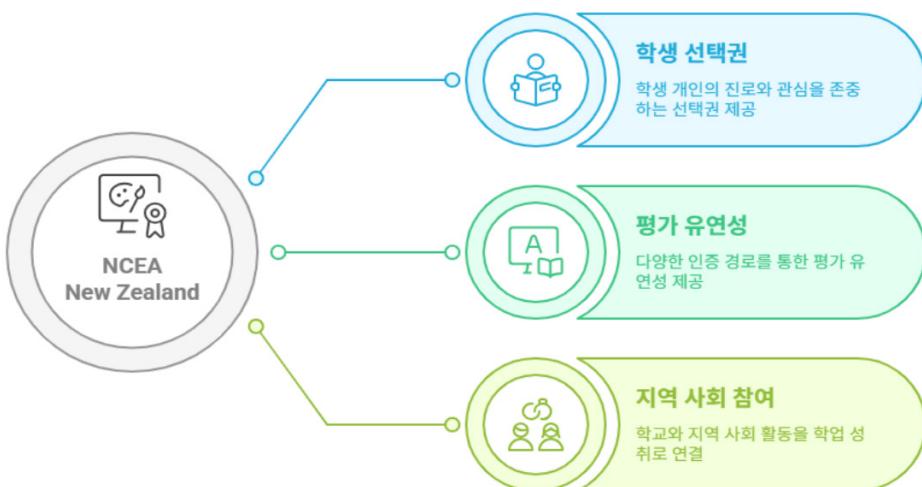
### ▶ 호주의 학생 평가: NAPLAN

“모두에게 같은 시험이 아닌 각자에게 맞는 배움으로 학생 한 명 한 명의 가능성을 여는 평가”



### ▶ 뉴질랜드의 학생 평가: NCEA

“학생의 선택을 존중하는 유연한 평가를 통해 학습과 진로, 삶을 연결하는 미래 지향적 교육 운영”





## ▶ 변화의 배경

2022년부터 시행된 새로운 교육과정에서 '지식·기능', '사고·판단·표현', '주체적으로 학습에 임하는 태도·인간성'이라는 세 가지의 관점에서 평가하는 '관점별 평가'를 강조함에 따라 각 학교에서는 정기 평가를 중심으로 한 학습 평가 방식의 개선 필요성을 느끼게 됨

## ▶ 변화 내용

변화	내용	모습
정기 평가 폐지 움직임	<ul style="list-style-type: none"><li>2018년 도쿄도 치요다구립 코지마치 중학교는 일본에서 최초로 정기 평가를 폐지함</li><li>정기 평가를 단원 평가로 바꾸고 학생이 원하는 경우 단원 평가 재도전을 허용함</li><li>'평가를 어떻게 학력 향상으로 연결할 것인가?'라는 교육의 본질적인 목표 도달을 위한 시도</li></ul>	
진단 평가에서의 CBT 도입	<ul style="list-style-type: none"><li>일본의 전국 학력·학습 실태 조사는 매년 시행되는 전국 단위의 진단 평가임</li><li>2023년부터 CBT(Computer-Based Testing) 도입이 추진됨</li><li>반응형 문항 구성, 자동 채점 및 데이터 수집, AI 활용 서술형 자동 채점, 장애 학생을 위한 확대 문자 제공 등이 가능</li></ul>	
등교 거부 (不登校) 학생 평가	<ul style="list-style-type: none"><li>온라인 수업 출석 인정: 등교 거부 학생이 비대면으로 실시간 수업 참여 또는 과제 제출을 할 수 있도록 방안 마련</li><li>1인 1기기 정책 연계 학습 추적: 디지털 과제와 학습 로그를 바탕으로 개별 평가가 가능</li></ul>	

**기획 및 편집** 경상북도교육청연구원 정책연구부 교육연구사 안주연 및 발간위원 8명

**표지 디자인** 이천초등학교 교사 최희도

**주소** 054-840-2276

**주소** 경상북도 안동시 강남로 152

**WEB** [www.gbe.kr/gber](http://www.gbe.kr/gber)



# 세계의 학생 평가

International student assessment



## 독일의 학생 평가

발간위원 : 최신영 (영주가흥초등학교 교사)

독일은 16개의 연방 주로 구성된 국가이고, 각 연방 주는 기본법에 따라 독립된 교육정책을 펼친다. 연방 주별 교육의 질을 유지하고 연방 주 간 협력을 위해, 교육문화부 장관회의(Kultusministerkonferenz, KMK)에서는 각 연방주의 교육 모니터링을 실시하고 있다.

KMK의 교육 모니터링은 크게 국제 학업성취도 비교, 교육 표준에 대한 검토 및 적용, 학교 수준의 모니터링으로 분류할 수 있다.

### 1. 국제 학업성취도 비교

1997년 10월, 독일연방공화국의 KMK에서는 과학적 연구의 틀 안에서 독일의 학교 시스템을 국제적으로 비교하기를 결정했다(콘스탄츠 결의안). 국가 간 경쟁이 교육의 질적 발전을 위한 필수 전제 조건이라 강조하고, 목표는 핵심 역량 분야에서 학생들의 강점과 약점에 대한 신뢰할 수 있는 결과를 얻는 것이었다.

#### 가. 피사 인터내셔널(PISA-International)



[그림 1] OECD PISA

PISA(Programme for International Student Assessment, 국제 학생 평가프로그램)는 경제협력개발기구(OECD)가 관리하며 2000년부터 3~4년마다 15세 학생의 읽기, 수학, 과학 능력을 평가해 왔다. 전 세계적으로 81개국, 약 69만 명의 학생이 참여한 이 평가에 독일은 2002년에 8번째로 참여했다. 읽기, 수학, 과학 분야에서 2시간 동안 컴퓨터로 시험을 치르며, 학생·학부모·교사·학교 행정부는 학생 배경 특성과 교수·학습 조건에 대한

설문지를 작성한다.

2023년에 발표된 ‘PISA 2022’ 결과에서 독일 15세 청소년의 학업성취도는 수학, 읽기, 과학 분야에서 눈에 띄게 떨어지는 것으로 나타났다. KMK에서는 결과 분석을 통해 독일 내 학교 운영에 대한 장기간의 제한, 팬데믹 관련 제한 조치, 학생 구성의 다종화 등을 그 원인으로 들었다. 사회 경제적으로 유리한 청소년들의 성적이 우수하며, 이민자 배경의 청소년들이 상대적으로 낮은 성적을 보였다.

이에 따라 독일에서는 사회적 배경과 교육적 성공 사이의 연관성이 강한 것으로 보고, 사회적으로 취약한 학생을 위한 맞춤형 지원을 도입했다. 연방 정부와 주정부가 협력하여 스타트찬센(Startchancen) 프로그램을 통해 사회적 취약 비율이 높은 약 4,000개 학교를 지원하고 있다. 연방 정부와 주정부는 10년간 총 200억 유로를 투자하여 읽기, 쓰기, 셈하기에서 최소 기준을 충족하지 못하는 학생 비율을 절반으로 줄이는 걸 목표로 하고 있다.

#### 나. 팀스(Trends in International Mathematics and Science Study, TIMSS)



TIMSS(수학·과학 성취도 추이 변화 국제 비교 연구)는 4년 주기로 4학년 말에 학생들의 기본적인 수학 및 과학 이해력을 측정한다. 전 세계적으로 총 58개국이 ‘TIMSS 2023’에 참여했는데, 이 중 22개국은 EU 회원국이고 29개국은 OECD 국가이다. 국제 수준에서 교육 성취 평가를 위한 국제 협회(IEA)에서 만든 이 평가에 독일에서는 2023년에 참여했고, TIMSS 2027에도 참여하기로 했다.

TIMSS 2023의 결과(2024년 12월 4일 발표됨), 2020년~2021년 팬데믹으로 인한 학교 운영 제한에도 불구하고, 2019년과 비교하여 수학과 과학에서 일관된 성취도를 보였다. 국제적으로 비교했을 때, 독일은 전체 순위의 중간 정도에 있다.

KMK에서는 결과 분석을 통해, 전반적으로 코로나19 영향이 나타나지 않았으나, 교육 성공이 사회적 배경과 밀접한 연관이 있음을 설명했다. 이민 경험의 유무에 따라 학생의 학업성취도에 차이가 있어, 교육 시스템이 다문화 사회에 걸맞게 노력해야 함을 표명했다. 스타트찬센(Startchancen) 프로그램, STEM 프로젝트 집중 지원 등으로 기초 역량 강화에 힘쓸 것을 발표했다.

#### 다. PIRLS/IGLU(Progress in International Reading Literacy Study bzw. Internationale Grundschul-Lese-Untersuchung)

PIRLS/IGLU(국제 독해 능력 향상 연구)에서는 2001년부터 5년마다 4학년 학생들의 읽기 능력을 조사했다. 독해 이해력, 독서 태도, 독서 습관을 측정하며, 2021년에 처음으로 독해 능력도 디지털 방식으로 측정되었다. 교육 성취 평가를 위한 국제협회(IEA)에서 실시하는 이 평가에 독일은 2021년

에 참여했고, 2026년에도 참여를 결정했다.

2021년 결과는 2023년에 발표되었는데, 독일 4학년 학생들의 읽기 능력은 2016년 이후 내림세를 보임으로 나타났다. 국제적인 비교에서 독일의 순위는 중간 정도이다. 독일 4학년 학생의 4분의 1은 읽기 학습 능력이 최소 능력 수준 미달이며, 코로나19 팬데믹 동안 학교가 문을 닫으면서 독서 성과에 상당한 영향을 미쳤다고 분석했다. 이러한 결과 분석을 통해 이민 배경을 가진 학생과 사회적으로 취약한 지역의 학생에게 스타트찬센(Startchancen) 프로그램을 비롯한 맞춤형 지원을 지속해서 제공하기로 했다.

## 2. 교육 표준에 대한 검토 및 적용(연방 정부에서 실시하는 학업 성취도 평가)

2004년부터 KMK에서는 독일 각 연방 주마다 국제적 학업성취도 평가(PISA, TIMSS, IGLU) 외에도 독일 전역에 적용할 교육 표준을 채택해 학생들의 학업성취도를 평가하도록 하였다. 교육 표준은 학습 성과의 형태로 교육의 목표를 설정한다. 학생들이 학교생활의 특정 단계까지 어떤 역량을 개발해야 하는지 과목별로 명시하였다. 역량은 각 과목의 지식과 기술을 적용해 문제를 해결하는 능력이라고 볼 수 있다.

### 가. 교육 품질 개발 연구소(Istitut zur Qualitätsentwicklung im Bildungswesen, IQB)

IQB는 이 국가 교육 표준을 기반으로 각 연방주의 교육 기준(학생들이 학교생활의 특정 시점까지 개발해야 하는 역량) 달성을 정기적으로 모니터링한다. IQB가 담당하는 주요 업무 중 하나는 초등 및 중등 교육의 교육 기준에 따라 시험 항목을 개발하고, 원하는 역량 목표가 어느 정도 달성되었는지 확인하는 것이다.

〈표 1〉 연방 정부 실시 학업 성취도 평가(IQB)

평가 대상	내용
초등학교 4학년	-과목: 독일어, 수학 -주기: 5년마다
중학교 9학년	-과목: 언어(독일어, 영어, 프랑스어), 수학 및 과학(수학, 생물, 화학, 물리) -주기: 3년마다 언어 과목과 수학·과학 과목을 번갈아 가며 평가 시행 (언어는 2015년, 2022년 평가 시행/ 수학·과학은 2018년, 2024년 시행) ※2028년 언어 평가 예정, 2030년 수학·과학 평가 예정

2020년, KMK에서는 초등 및 중등 교육의 교육 기준을 발전시키길 결정했고, 새로운 시험 항목을 개발하였으며, 향후 개발될 교육 표준에 대한 역량 수준 모델 제시 개발 일정을 잡았다.

2029년 말까지 독일어와 수학(초등학교·중학교 수준), 제1 외국어(영어, 프랑스어/중학교 수준), 과학(생물·화학·물리학/중학교 수준) 과목에 대한 새로운 과제 풀을 개발할 예정이다.

#### 나. 2024년 9학년 IQB 교육 동향

독일 모든 연방주의 9학년 학생들이 2024년 4월부터 7월 사이(구체적인 기간은 연방 주별로 조금씩 다름) IQB에 참여하였다. 과목별 역량 평가 영역은 다음과 같다(표 2). 이 역량 평가 연구 자료 분석 결과는 2025년 가을에 발표될 예정이다. 보고서 발표 후, 모든 참여 학교는 달성한 결과에 대한 별도의 피드백 요약을 받게 된다.

〈표 2〉 2024년에 실시한 9학년 IQB 평가 영역

과목	평가 역량
수학	<p>-<b>일반 수학 역량(6가지)</b>: 수학적 추론, 수학적 문제 해결, 수학적 모델링, 수학적 표현 사용, 수학의 상징적·형식적·기술적 요소 다루기, 의사소통</p> <p>-<b>내용 관련 수학 역량(핵심 아이디어 5가지)</b>: 숫자, 측정, 공간과 형태, 기능적 관계, 자료와 우연</p>
과학	<p>-<b>공통 역량(4가지)</b>: 전문 지식, 지식 습득, 평가, 소통</p> <p>-<b>공통 역량이 과목별(생물학, 화학, 물리학) 제시되어 있음</b></p>

#### 3. 학교 수준의 비교 연구(Vergleichsarbeiten in der 3. 8. Jahrgangsstufe, VERA)

VERA는 KMK의 교육 모니터링 중 하나로, 2006년에 채택되고 2015년에 개정되었다. VERA는 독일의 모든 주에서 3학년과 8학년이 의무적으로 치러야 하는 필기시험이다. 일부 주에서는 VERA를 시험이라 하지 않고 ‘학습 평가(헤센, 노르트라인베스트팔렌)’, ‘KERMIT-역량 결정(함부르크)’, ‘역량 시험(작센, 튜링겐)’이라고 한다. 2010년부터 이탈리아의 남티롤 지방, 벨기에의 독일어권 지역도 VERA에 참여하고 있다. VERA에서 사용하는 과제는 IQB의 중재하에 전국 교사들이 개발하고, 해당 과목 교수자들의 검토를 거쳤으며, 많은 학생을 대상으로 여러 차례 테스트를 거친다. VERA 3(3학년) 및 VERA 8(8학년)의 구현 방식은 각주마다 다르다. 구체적인 규정은 현지 상황을 고려해야 하기 때문이다.

〈표 3〉 VERA 2025 일정

평가 대상	내용
VERA 3 (3학년)	-과목: 독일어, 수학 -시험 일정: 2025년 4월 22일 ~ 5월 28일
VERA 8 (8학년)	-과목: 독일어, 수학, 제1 외국어(영어, 프랑스어) -시험 일정: 2025년 2월 11일 ~ 3월 28일

※각 연방 주는 시험지를 어떻게 편찬할지, 매년 학생들이 어떤 역량 영역에서 시험을 치러야 할지 독립적으로 결정한다.

VERA의 핵심 목적은 각 학교의 교육과 학교 발전을 지원하는 것이다. VERA는 교사들에게 학급 및 학교 간 학급 성취 수준에 대한 객관적인 정보를 제공한다. 결과를 통해 학교와 교사에게 학생들이 역량 목표를 어느 정도 달성했는지, 초등학교(보통 4학년)와 중등학교(9학년)에서 교육 기준을 달성해야 하는 시기 1~2년 전에 어떤 분야에 더 많은 관심을 기울여야 하는지에 대한 피드백을 제공한다. VERA는 채점하여 평가하는 것이 아니라 결과 피드백에 대한 분석과 해석에 중점을 둔다.

#### 4. 나침반 4(Kompass 4)-바덴뷔르템베르크주

2024년 5월, 바덴뷔르템베르크주 정부는 초등학교에서 중학교로의 전환이 관련된 내용을 포함한 교육 패키지를 발표했다. ‘나침반 4(Kompass 4)’ 역량 평가는 2024/2025학년도부터 바덴뷔르템베르크주의 모든 공립 초등학교와 초등학교 교육을 제공하는 특수 교육 및 상담 센터에서 의무적으로 실시된다. 나침반 4는 독일어와 수학 성취 수준을 평가한다. 평가 목표는 전환기 동안 중등학교 진학에 대한 조언을 강화하고, 교사의 추천과 학부모/보호자의 중등학교 유형에 관한 결정을 지원하는 것이다.<sup>1)</sup>

독일어는 독해, 철자, 언어 사용을 평가하고, 수학에서는 숫자와 연산, 공간과 모양, 크기와 측정, 데이터, 빈도와 확률 영역을 평가한다. 이와 더불어, 교사가 학생의 행동이나 작업, 학습을 평가하는 것이 포함된다. ‘나침반 4’ 평가에는 반영되지 않지만, 2025년 1월 31일까지 학부모/보호자와 담임교사가 나누는 정보 및 협의도 중요한 부분이라 볼 수 있다.

1) 독일의 초등학교는 4년제이고, 4학년 2학기에 진학을 희망하는 중등학교 유형(김나지움, 하웁트슐레, 레알슐레 등)을 결정해야 한다.

이 역량 평가는 바덴뷔르템베르크주의 표준화된 지침과 평가 기준에 따라 각 과목 교사에 의해 평가되며, 지정된 시간에 시행된다(오전 9시, 과목별 평가 시간은 45분으로 통일).

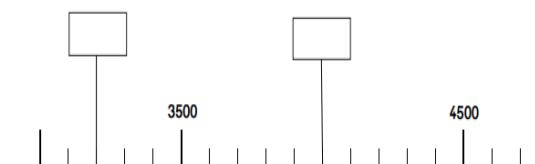
1. Beantworte die Fragen zum Text. Kreuze an.

/ 1,5 Punkte

- a) Wie sorgen die Menschen dafür, dass Wölfe sich wieder in Deutschland ansiedeln?  
 Wölfe werden regelmäßig gefüttert.  
 Wölfe stehen unter Schutz und dürfen nicht gejagt werden.  
 Es werden spezielle Schutzgebiete für Wölfe in Städten gebaut.  
 Wölfe werden trainiert, um Menschen zu meiden.

1. Ergänze die fehlenden Zahlen in den Kästchen.

/ 1 Punkt



[그림 3] 'Kompass 4' 2024 출제 문제(독일어)

[그림 4] 'Kompass 4' 2024 출제 문제(수학)

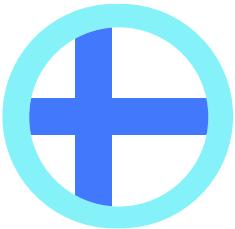
청각, 언어, 시각 장애가 있는 학생들에게는 수정된 과제가 제공된다. 사립·국립 초등학교는 2024/2025학년도에 ‘나침반 4’를 제공할 수 있고, 2025/2026학년도부터는 법적으로 의무화되어 실시해야 한다.

## 5. 맷음말

독일에서는 학생 평가로 역량 중심 절대 평가를 시행한다. 수업 과정에서 협력을 중시하며 진로의 다양성을 인정하고 성장 중심 평가를 지향한다. 각 연방 주별 교육정책 실시로 교육 표준에 따라 모니터링을 해서 피드백을 통해 교육의 질 개선, 학생 역량 함양을 꾀하는 점이 특징이다.

### 【참고 자료】

- ▶ Bildungsstandards der Kultusministerkonferenz, <https://www.kmk.org/themen/qualitaetssicherung-in-schulen/bildungsstandards.html>
- ▶ IQB-Bildungstrends, <https://www.kmk.org/themen/qualitaetssicherung-in-schulen/bildungsmonitoring/ueberpruefungumsetzung-der-bildungsstandards.html>
- ▶ Kompass 4, <https://ibbw-bw.de/Lde/Startseite/Kompetenzmessung/kompass4>
- ▶ PIRLS/IGLU, <https://www.kmk.org/themen/qualitaetssicherung-in-schulen/bildungsmonitoring/internationale-schulleistungsvergleiche/pirlsiglu.html>
- ▶ PISA-International, <https://www.kmk.org/presse/pressearchiv/mitteilung/pisa.html>
- ▶ TIMSS, <https://www.kmk.org/presse/pressearchiv/mitteilung/grundschuelerinnen-und-schueler-halten-ihr-leistungen-in-mathematik-und-naturwissenschaften-und-das-trotz-pandemie.html>
- ▶ VERA – Ein Überblick, <https://www.iqb.hu-berlin.de/vera/ueberblick/>



## 핀란드의 학생 평가

발간위원 : 최휘도 (아천초등학교 교사)

핀란드 교육은 오랫동안 교육적 평등의 추구와 혁신성 그리고 학생 중심 접근으로 여러 교육 선진국 중에서도 주목을 받아왔다. 특히 다수의 국가들이 전국 단위 표준화 시험과 이에 기반한 평가를 강조할 때 핀란드는 지속적이고 개인화된 형성평가를 통해 학생의 전인적 발달과 학습 과정을 평가하는 독특한 방식을 꾸준히 견지하고 있다. 이에 본 기사에서는 핀란드 학생 평가의 철학적 토대와 실제 운영 방식을 살펴보고 구체적인 사례의 확인을 통해 핀란드 평가 방식의 특징과 그 교육적 의미를 알아보고자 한다.

### 1. 핀란드 학생 평가의 교육 철학적 기반

핀란드 학생 평가의 기반에는 평가를 학습의 일부로 보는, 더 나아가 학습을 촉진하는 도구로 보는 시각이 자리 잡고 있다. 핀란드 교육 당국은 ‘학생 평가의 주요 목적은 학생들로 하여금 자신의 학습에 대해 성찰하고 자기평가를 할 수 있도록 격려하는 데 있다.’라고 명시적으로 선언한다. 이는 과거처럼 평가를 단순히 성적을 매기기 위한 수단이 아니라 학생이 스스로 자신의 학습 과정을 확인하고 학습하는 방법을 익힐 수 있도록 돋는 과정으로 인식한다는 것을 의미한다. 실제로 핀란드에서는 평가가 수업 활동과 구분되어 운영되는 것이 아니라 일상적인 수업에 통합되어 이뤄지고 있으며, 학생들이 자신의 성장 상황을 확인하고 지속적으로 성찰 및 개선할 수 있도록 돋는다. 이렇듯 핀란드에서 학생 평가는 동료와의 비교나 석차 산출을 의미하는 것이 아닌 학생 개개인의 성장에 초점을 맞춤으로써 학생이 비교와 석차에서 나오는 두려움과 압박감에 휘둘리지 않고 학생 자신의 자신감, 호기심, 내적 동기를 키우는 것을 목표로 한다. 즉, **핀란드에서의 학생 평가란 곧 학습을 위한 평가(assessment for learning)를 의미하며 경쟁보다는 성찰과 성장을 중시하는 철학적 기반 위에 구축되어 있다.**

## 2. 핀란드 학생 평가의 특징

서열을 위한 평가가 아닌 ‘학습을 위한 평가’라는 핀란드의 교육 철학은 사실 그들만의 독특한 철학은 아니다. 이와 같은 평가에 대한 시각의 변화는 교육 선진국을 중심으로 많이 자리 잡았다. 우리나라로 지난 10여 년 동안 학습과 성장을 위한 평가가 강조되었으며 ‘과정 중심 평가’의 도입은 그 대표적인 예라고 할 수 있을 것이다. 하지만 교육 현장에서 정말 학생 성장을 위한 평가가 이뤄지고 있는가? 라는 질문에는 선뜻 대답하기 쉽지 않다. 그렇다면 핀란드의 학생 평가는 어떤 특징들을 가지고 있을까?

### 가. 교사 평가권의 강조

핀란드에서는 교사가 학생 평가의 핵심 주체가 된다. 전국 단위의 일제고사나 외부 평가기관의 개입 없이 모든 학생 평가를 각 학교의 교사가 책임지고 수행한다. 이와 같은 교사 평가권의 강조는 교사들에게 상당한 평가 권한과 재량을 부여하는데 이는 교사에 대한 사회적 신뢰를 바탕으로 한다. 핀란드 교육 당국은 넓은 범위에서의 기준과 목표만 제시할 뿐 각 학교와 교사가 자체적인 평가 방법을 개발한다. 교사는 이러한 자율성 속에서 학생들의 강점, 약점, 창의성, 성장 과정 등을 총체적으로 평가 하며 평가 결과에 대한 책임도 스스로 진다. 즉, **핀란드의 교사들은 교육 전문가로서 전문성과 소신을 갖고 학생 평가를 주도할 수 있으며 외부 간섭 없이 학생들을 가장 잘 평가할 수 있는 방법을 선택할 권한을 가지고 있는 것이다.**

### 나. 표준화 시험의 역할 최소화

핀란드 학생 평가의 가장 두드러진 특징 중 하나는 표준화 시험의 최소화이다. **핀란드는 종합학교 단계<sup>2)</sup>에서 전국적으로 시행되는 시험이 전혀 없다.** 이렇듯 핀란드에서는 학생과 학교를 서열화할 수 있는 국가 단위의 표준화 시험을 실시하지 않는다. 유일하게 존재하는 표준화 시험은 우리나라의 수능에 해당하는 대학 입학시험이다. 따라서 대학 진학을 목표로 하지 않는 학생이나 직업 교육으로 진로를 정한 학생들은 학교 교육과정 내에서는 어떠한 전국 공통 시험의 압박도 받지 않고 학업을 마칠 수 있다. 그렇다면 왜 핀란드는 이렇듯 표준화 시험을 지양하는 것일까? 우선 그들은 **단일 시험으로 모든 학생의 학습 성과를 정확히 평가할 수 없다고 인식한다.** 또한 **시험 성적은 학생이 시험을 잘 보는 능력만을 반영할 뿐 개인의 다양한 강점과 역량을 보여주지 못한다고 본다.** 대신 학생 평가를 통해 각자의 강점과 기술에 주목하여, 학생이 무엇을 알고 무엇을 할 수 있는지를 다각도로 파악하고자 한다. 마지막으로 국가나 외부 기관에서 출제한 시험보다 **교사가 일년 내내 관찰하고 평가한 결과가 더 공정하고 유의미하다고 인식하고 있다.**

2) Comprehensive School, 우리나라의 초중등 교육에 해당

## 다. 과정 중심 평가

핀란드의 평가는 매일의 수업 활동 속에서 이루어지며, 교사는 일상적으로 학생에게 피드백을 제공하며 도움을 준다. 교사는 학생들의 참여도, 과제 수행, 토의 기여도, 문제 해결 과정 등을 관찰하고 기록함으로써 시험 점수만으로는 파악할 수 없는 학습 과정의 질을 평가한다. 이러한 과정 중심 평가는 학생에게 상시적인 피드백을 제공하여 학습 개선을 촉진한다. 하지만 과정 중심 평가 그 자체보다 더 중요한 것은 핀란드가 어떻게 과정 중심 평가를 실천하고 있는가이다. 핀란드의 과정 중심 평가에는 몇 가지 특징이 있다. **첫째, 상시적이고 낮은 진장도의 피드백이 일상적으로 제공된다.** 교사는 학기 내내 학생들에게 구체적 피드백을 제공하여 학생이 자신의 장점과 개선점을 스스로 확인할 수 있도록 유도한다. **둘째, 학생들이 학습 목표 설정에 적극적으로 참여한다.** 평가 과정에서 학생은 자신의 학습을 확인하고 다음 목표를 스스로 설정함으로써 학습과 평가의 연속적인 과정을 자신의 성장 과정으로 이해할 수 있게 된다.

**Pupil set goals for the new school year**
**Goals for the academic year 20xx-20xx**

name \_\_\_\_\_

I'm planning my own school work in the 5th grade.

I set goals for myself that we follow with the teacher and parents.

**Working skills at school**
Independent work goes well with me.  
My goal:Working with a pair goes well with me.  
My goal:Working in a group is easy for me.  
My goal:I let the others work in peace.  
My goal:
**Working skills at home**

I can almost always do my homework myself. \_\_\_\_\_

I need help with my homework. \_\_\_\_\_

**Behavior**

I behave towards others \_\_\_\_\_

In case of a dispute, I \_\_\_\_\_

I always obey adults  
I don't always obey adults \_\_\_\_\_

My goal: \_\_\_\_\_

**Goals for mother tongue and literature study**

I think my reading ability is \_\_\_\_\_  
 I write different kind of texts  
 completely error-free \_\_\_\_\_  
 almost without error \_\_\_\_\_  
 there will still be quite a few errors \_\_\_\_\_

Writing my own stories and texts is something I think is for me  
 very easy \_\_\_\_\_  
 quite easy \_\_\_\_\_  
 a bit challenging \_\_\_\_\_  
 really challenging \_\_\_\_\_

**Reading books**

I'm going to read a whole thick book properly \_\_\_\_\_  
 Reading a whole book is still quite challenging for me \_\_\_\_\_  
 I'm not really excited about reading whole books \_\_\_\_\_

**Goals for studying mathematics**

I can do basic calculations (+, -, ×, ÷) from 0 to 100000  
 very well \_\_\_\_\_  
 I can't yet \_\_\_\_\_

I recognize and can compare and calculate numbers from 0 to 1000000  
 very well \_\_\_\_\_  
 pretty good \_\_\_\_\_  
 I still need exercise \_\_\_\_\_

I can do calculations with fractions  
 very well \_\_\_\_\_  
 pretty good \_\_\_\_\_  
 I still need exercise \_\_\_\_\_

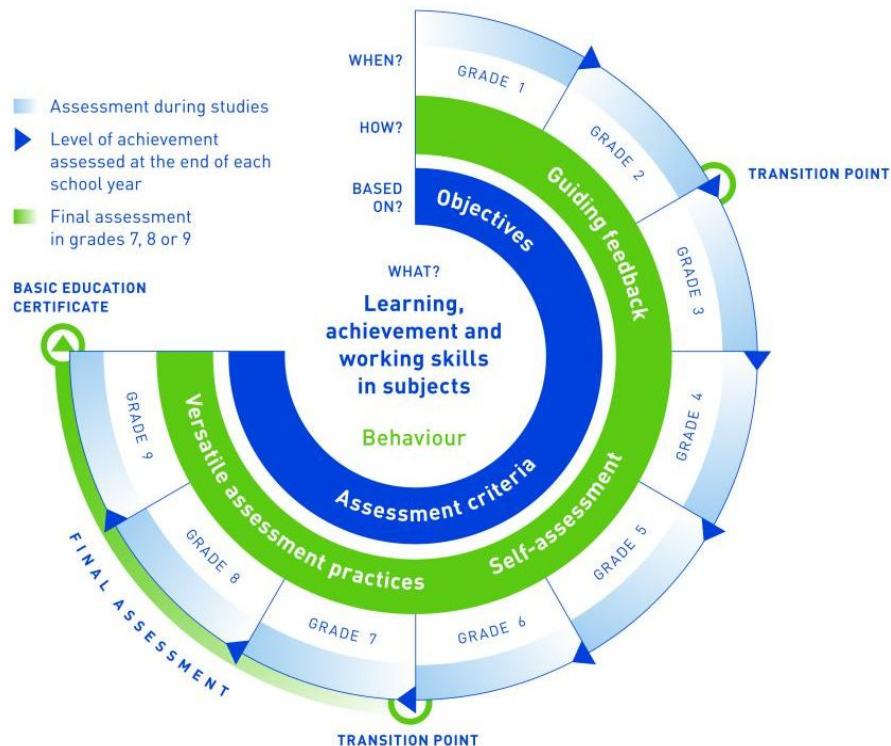
For me verbal or written tasks are  
 very easy \_\_\_\_\_  
 pretty easy \_\_\_\_\_  
 challenging \_\_\_\_\_  
 I can't solve them alone \_\_\_\_\_

[그림 1] 학생의 학습 목표 설정 자료<sup>3)</sup>

**셋째, 핀란드에서 평가의 범위는 학업 역량 이외의 다양한 기술과 소양까지 포함한다.** 예를 들어 교사는 친구와의 협동 능력, 창의적 문제해결력, 감성 지능 등의 비인지적 역량도 함께 관찰하여 평가함으로써 평가가 학생의 전인적 성장을 반영하도록 한다. 이처럼 핀란드의 과정 중심 평가 체계는 ‘학

3) 5학년 학생의 학급 목표 설정 자료로 자신의 학교생활, 가정생활, 행동 특징 및 국어, 독서, 수학 학습 수준을 스스로 평가하고 자신의 목표를 설정할 수 있도록 구성되어 있다.

습을 위한 평가’라는 핀란드의 철학인 그대로 반영되어 있으며, 학생들로 하여금 실패를 두려워하기보다는 실수를 통한 학습을 받아들이도록 지도한다. 다만 핀란드가 과정 중심 평가만을 운영하는 것은 아니며, 과정 중심 평가를 기반으로 매 학기 말에 총괄평가 역시 실시하고 있다.



[그림 2] 핀란드 종합학교의 평가 과정

#### 라. 수치 척도가 함께 제공되는 평가 리포트

과정 중심 평가와 총괄평가를 종합하여 핀란드의 종합학교에서는 매 학년말 ‘평가 리포트(School year report)’를 학생들에게 배부하는데, 해당 리포트에는 학생의 학업 성취 즉, 교과별 성취 수준과 진급 여부가 기재된다. 이를 바탕으로 학생과 학부모는 학습 진행 상황을 전체적으로 파악할 수 있으며, 진로와 진급을 위한 자료로 활용한다. 흥미로운 점은 이와 같은 평가 리포트에서 술형 피드백뿐만 아니라 수치 평가도 함께 제시된다는 점이다. 종합학교 4학년부터는 서술형 피드백과 함께 학생들의 평가 척도가 4에서부터 10까지 제시된다. 과정 중심 평가를 누구보다 강조하는 핀란드가 학생들에게 수치 평가도 함께 제공하고 있다는 사실은 현재 초등학교 과정에서 수치 평가를 전혀 제시하고 있지 않은 우리나라에서도 생각할 거리를 제공한다고 여겨진다. 각 수치 척도의 의미는 다음과 같다.

- ▶ 4 Fail(낙제)
- ▶ 5 Adequate(거우 통과함)
- ▶ 6 Moderate(보통)
- ▶ 7 Satisfactory(만족스러움)
- ▶ 8 Good(우수함)
- ▶ 9 Very good(매우 우수함)
- ▶ 10 Excellent(탁월함)

### 3. 교육 현장에서의 학생 평가: 펠른캉가스 학교(Pöllönkangas School)의 사례

핀란드는 주지하다시피 교육과정 운영에 지역과 학교가 큰 재량권을 가지고 있다. 학생 평가에 있어도 다르지 않다. 각 지역과 학교는 자신들이 처한 학생 수준과 주변 환경을 고려하여 가장 적절한 학생 평가를 제공하고 있다. 따라서 학생 평가의 방법은 학교마다 조금씩 차이를 보인다. 이에 실제 학교에서 앞서 살펴본 핀란드 학생 평가의 일반적인 특징들이 어떻게 실천되고 있는지 확인할 필요가 있다. 핀란드 오울루시에 위치한 펠른캉가스 학교의 사례를 통해 이를 확인해 보고자 한다.

펠른캉가스 학교는 일단 평가 목표를 설정할 때 학생의 의견을 청취한다. 학년 초에 모든 학생들과 함께 평가 과정 개발 토론회를 개최하여, 학생들의 참여하에 해당 학년도에 학습하게 될 학습 목표를 설정한다. 학생뿐만 아니라 학부모 또한 해당 토론회에 참가할 수 있으며, 학습 목표의 달성을 여부는 다음 해에 개최되는 평가 토론회에서 확인할 수 있다. 또한 학생들은 평가 리포트를 제공받는다. 일반적으로 봄 학기가 끝날 때 학생들은 리포트를 받게 되며, 1~3학년 학생들은 서술형 리포트를 4~9학년 학생들은 수치 자료가 첨부된 리포트를 제공받는다. 만약 해당 학생이 일정 과목에서 낙제 점수를 받게 되면 학부모에게 통보되며, 학생은 학년이 끝나기 전에 학부모와 함께 낙제 사유에 대한 소명의 기회를 얻을 수 있다.

펠른캉가스 학교는 학생들에게는 교과목 평가 외에도 학생 행동에 대한 평가도 제공하고 있다. 학생 행동은 핀란드 국립 교육청과 지역 교육과정에서 설정된 목표와 지침에 따르며, 학생 행동은 학생이 다른 사람과 주변 환경을 어떻게 인식하며, 협력 작업에서 어떻게 자신과 타인을 평가하는지 그리고 규칙과 예절을 얼마나 존중하고 따르는지에 중점을 두고 평가된다.

〈표 1〉 펠른캉가스 학교의 학생(4~9학년) 행동 평가 수치 척도

척도	의미
10	- 타인과 환경을 고려하여 긍정적인 모범이 되며, 다른 학생과 학습 상황에 영향을 미침. - 자신의 그룹과 학교 공동체의 이익을 위해 행동하고 문제를 처리하는 데 주도성을 보임.
9	- 타인과 환경을 고려하여 합의된 규칙과 지시에 따라 책임감 있게 행동하며, 친절하고 도움이 되는 방식으로 행동하고 긍정적 분위기 형성을 위해 노력함.
8	- 상황에 적합한 방식으로 행동하고, 일반적으로 학교의 규칙을 준수하며 예의 있게 행동하여 주도적으로 갈등 상황을 해결할 수 있으며, 협력하는 방법을 알고 있음.
7	- 일반적으로 상황에 적합한 방식으로 행동하고, 학교의 규칙을 알고 수용하며, 교사의 지도하에 갈등 상황을 해결할 수 있고, 지원을 받는 그룹에서 역할을 할 수 있음.
6	- 종종 동료, 학업 및 학교 환경의 안정에 무관심하고, 반복적으로 학교 규칙을 어기고, 태도와 행동으로 부정적인 분위기를 조성하고, 행동에 대한 지속적인 지도와 조언이 필요함.
5	- 동료, 학업 및 학교 환경에 대해 매우 무관심한 행동으로 학습의 안전을 위해 특별한 조치가 필요함.
4	- 행동이 정상적인 학교 공동체에 완전히 부적절함.

1~3학년 학생의 경우 서술형 평가를 통해 자신의 행동에 대한 평가를 받으며, ‘매우 우수, 우수, 만족, 안내 필요’ 4단계의 등급이 있다.

#### 4. 맷음말

핀란드의 학생 평가는 단순히 결과를 확인하는 수단을 넘어 학생의 학습 과정을 살피고 성장을 지원하는 데 초점을 두고 있다. 학생을 줄 세우는 평가가 아닌 학생이 자신의 학습을 성찰하고 스스로 의미를 찾을 수 있도록 돋는 평가 방식인 것이다. **교사의 자율성과 전문성을 바탕으로 한 평가, 학생 간 비교보다는 자기 성장을 중심에 둔 문화 그리고 학생이 학습 목표를 직접 세우고 평가에 적극 참여하는 구조**는 핀란드 평가 체제의 핵심이라 할 수 있다.

함께 주목할 점은 교과 지식이나 기능에 국한되지 않고 협력, 자기조절, 감정 조절 능력 등 **전인적 역량까지 평가의 범위에 포함**하고 있다는 것이다. 이는 평가를 단순한 측정 도구가 아닌 학습을 이끄는 촉진자로 바라보는 핀란드 교육의 철학을 잘 보여주는 대목이다. 결과보다 과정을 중시하고 수치보다 피드백을 우선시하는 이러한 접근은 학생들에게 실수와 실패마저도 하나의 학습 기회로 받아들일 수 있도록 여유를 제공하며 배움 그 자체에 대한 내적 동기를 자극한다.

우리 교육에서도 평가의 역할을 재정립하려는 시도가 이어지고 있다. 그러나 여전히 소위 명문 학교 진학을 위한 성적과 등수에 집중하는 현실은 평가의 본래 목적과는 거리가 있다. 평가의 방향을 바꾸는 일은 결국 교육의 본질을 다시 묻는 일이다. 무엇을 어떻게 교육할 것이며 그에 걸맞은 평가는 어떻게 이뤄져야 하는지를 고민하게 만든다. 핀란드 또한 수치 척도에 대한 찬반 등 학생 성장을 위한 평가에 대한 논의가 치열하게 이어지고 있다. 학생 평가와 관련된 핀란드의 경험은 교육과 평가에 대한 여러 물음에 일관되며 설득력 있는 하나의 대답이 되어주고 있다.

#### 【참고 자료】

- ▶ techclass, Finland's Approach to Student Assessment: What We Can Learn  
<https://www.techclass.com/education-insights/finlands-approach-to-student-assessment-what-we-can-learn>
- ▶ Finish National Agency for Education, Basic information about primary and lower secondary education  
<https://www.oph.fi/en/education-and-qualifications/basic-information-about-primary-and-lower-secondary-education>
- ▶ Learning together, Formative Assessment: Examples of Forms Filled in by Finnish Teachers, Children and Their Parents  
<https://nushub.org.ua/lt/en/news/formative-assessment/>
- ▶ OULU, Pöllönkankaan koulu Arvointi  
<https://www.ouka.fi/pollonkankaan-koulu/arvointi?accordion=accordion-25647>



## 영국의 학생 평가

발간위원 : 정기업 (청송중학교부동분교장 교사)

### 1. 들어가며

영국 교육에서 ‘평가’를 이해하려면, 먼저 영국이 경쟁 중심의 교육을 지향한다는 점을 명확히 인식할 필요가 있다. 영국 사회에서 학교는 단순히 지식을 전달하는 공간이 아니라, 학생들이 다양한 경험을 통해 자신의 소질과 재능을 계발할 수 있는 기회의 장으로 여겨진다. 이러한 교육 철학 아래에서는 학생들이 폭넓은 교과목을 자유롭게 탐색하고, 여러 활동에 자발적으로 참여할 수 있도록 성적 중심의 서열화된 평가가 배제되어야 한다. 점수에 얹매여 배움의 기회를 제한하거나, 학습 자체가 경쟁의 수단이 되는 것을 영국 교육은 경계한다. 이는 학습의 본질이 단순한 암기나 성취가 아니라, 개인의 성장과 가능성 확장에 있음을 강조하는 것이다.

이러한 교육 철학은 평가 방식에서도 잘 드러난다.

영국의 평가 체계는 국가 차원의 공식 시험과 학교 차원의 교사 평가가 병행되는 이원적 구조를 갖는다. 이는 **학습 성취를 객관적으로 검증**하는 동시에, 학생 **개개인의 학습 과정을 정밀하게 진단하고 지원**할 수 있도록 설계된 체계이다. 국가 주도의 평가, 예를 들어 Key Stage 2의 SATs(Standard Assessment Tests, 국가 표준 시험)나 Key Stage 4의 GCSE(General Certificate of Secondary Education) 시험은 정해진 기준에 따라 전국적으로 동일하게 시행되며, **학생의 성취 수준을 외부적으로 인증**하는 기능을 한다. 반면, 각 학교에서는 **교사의 판단에 따라 형성평가와 총괄평가를 실시하며, 수업 중 질문이나 과제, 단원별 테스트 등을 통해 학생의 이해도를 지속적으로 점검하고 피드백을 제공**한다. 이처럼 외부 표준화 평가와 내부 맞춤형 평가가 함께 작동함으로써, 영국의 이원적 평가 체계는 **교육의 공정성과 개별화**라는 두 가지 목표를 동시에 추구하고 있다.

### 2. 교육 단계별 구조 및 평가 체계 개요

[표1] 영국 교육 Key Stage별 평가 제도 요약

연령 (Age)	학년 (Year)	Key Stage	평가 내용 (Assessment)
3~4	Nursery	Early years	평가 없음
4~5	Reception	Early years	Baseline Assessment (입학 기초평가) 초기 언어, 수리 능력 등 교사 평가 기반 진단
5~6	Year 1	KS1	Phonics screening check (음운 인식 검사)
6~7	Year 2		국가 시험(SATs): 영어 독해 및 수학 *2023/24년부터 KS1 국가 시험은 선택사항(선택적 평가) 교사 평가(TA): 수학, 과학, 영어 읽기·쓰기
7~8	Year 3	KS2	평가 없음
8~9	Year 4		Multiplication tables check (곱셈표 검사)
9~10	Year 5		평가 없음
10~11	Year 6		국가 시험(SATs): 영어 독해, 수학, 문법·구두점·철자 교사 평가(TA): 영어 작문, 과학
11~14	Year 7~9	KS3	국가 차원의 평가 없음 (대부분 학교 자체 평가)
14~15	Year 10	KS4	일부 학생은 GCSE 일부 선택과목 조기 응시
15~16	Year 11		대부분의 학생이 GCSE 응시
16~17	Year 12	KS5 (six form)	A-level/IB 시작, 과목 선택 및 심화 학습
17~18	Year 13		A-level/IB/BTEC 최종 평가, 대학 입시/진로 결정

### 3. 국가 표준 시험(SATs: Standard Assessment Tests)

영국의 국가 표준 시험은 학생이 국가 교육과정에서 요구하는 학습 기준에 도달했는지를 점검하기 위해 시행된다. 이 시험은 단순히 성적을 산출하는 것을 넘어, 학습 부진을 조기에 발견하고 이에 대한 교육적 지원을 제공하는 데 목적이 있다. 또한 학교의 교육 성과를 측정하여 교육청과 정부가 정책을 수립하거나 예산을 배분하는 데에도 활용된다. 특히 GCSE와 같은 시험은 학생의 진로 결정, 대학 입시, 직업 교육 진입에 직접적인 영향을 미치며, 개인에게도 실질적인 의미를 지닌다.

국가 표준 시험의 결과는 학생 개인에게만 의미 있는 것이 아니라, 학교 평가와 국가 교육 행정 전반에도 영향을 미친다. 이 시험 성과는 Ofsted(the Office for Standards in Education)가 학교의 교육 품질을 판단할 때 핵심 지표로 사용된다. Ofsted는 교육의 질, 학생 성취도, 리더십 등을 종합

적으로 점검하는 기관으로, Key Stage 2의 SATs나 GCSE 성적은 학생 성취도의 객관적 근거로 활용된다. 시험 성과가 지속적으로 낮거나 개선이 없는 학교는 집중 점검 또는 개선 명령의 대상이 되며, 반대로 높은 성취도를 보이는 학교는 ‘우수(Outstanding)’나 ‘양호(Good)’ 평가를 받을 가능성이 높아진다. 이러한 Ofsted의 평가 결과는 일반에 공개되며, 학부모와 지역사회가 학교의 교육 수준을 비교하거나 학교 선택의 기준으로 활용할 수 있다.

이러한 맥락에서, 국가 표준 시험과 Ofsted는 영국 교육의 질 관리를 함께 견인하는 상호 보완적 핵심 기제로 작용하고 있다. 국가 차원의 시험은 학습 진단을 통해 학생을 지원하는 동시에, 학교의 책무성과 교육 행정의 공정한 판단을 가능하게 하는 주요 수단이 된다.

국가 표준 시험은 Key Stage별로 연령과 교육 목적에 맞게 체계적으로 구성되어 있다. Early Years에서는 초등 입학 전 아동의 발달 상태를 확인하는 Baseline Assessment가 실시되며, KS1과 KS2에서는 문자 해독 능력과 수리 능력 등을 점검하는 Phonics Screening, Multiplication Check, SATs가 운영된다. 중등 과정의 KS3는 국가 시험 없이 학교 자율의 내적 평가에 초점을 맞추며, 이후 KS4의 GCSE와 KS5의 A-level·BTEC·IB는 진로 및 고등교육 진입을 위한 핵심 자격시험으로 작용한다. 각 단계의 시험은 **학생 개인의 성취 수준을 표준화된 기준에 따라 객관적으로 검증**하는 동시에, 학교 교육의 질을 종합적으로 점검하고, 향후 진로 준비와 교육 정책 수립에 필요한 신뢰 가능한 자료를 제공하는 중요한 역할을 수행한다.

## 4. 교사의 평가(TA, Teacher Assessment)

### 가. Key Stage 2

영국의 초등 교육과정에서 Key Stage 2는 학생의 학습 성장을 본격적으로 평가하는 중요한 시기이다. 이 시기의 학생 평가 중 핵심적인 측면을 이루는 것은 바로 교사 평가(Teacher Assessment, TA)다. TA는 단지 시험 점수를 넘어서, 학생의 수업 참여와 수행 전반을 기반으로 한 교사의 전문적 판단에 의해 이루어진다.

TA는 일상적인 교수-학습 과정 속에서 자연스럽게 이루어지며, 학생의 글쓰기, 실기 과제, 구술 표현 등 다양한 수업 활동을 종합적으로 반영한다. 교사는 **단편적인 결과가 아닌, 시간의 흐름에 따른 학생의 이해와 수행 능력의 일관성을 중심으로 평가한다**. 특히 TA는 정해진 기준에 도달했는지를 판단하는 데 초점을 두되, 학생의 전인적 발달과 개별적인 학습 경로를 존중하며 이루어진다.

Key Stage 2에서 **법적으로 요구되는 교사 평가**는 영어 작문(English writing)과 과학(Science) 과목에 적용된다. 영어 작문은 ▲기대 기준 도달 전 단계(working towards), ▲기대 기준 도달(working at expected standard), ▲심화 수준(greater depth) 세 단계로 평가되며, 과학은 ‘기대 기준 도달 여부’로 단순하게 구분된다. 학생이 기준에 미달하는 경우에는 Pre-Key Stage 평가 기준 또는 Engage

ment model을 활용하여 학생의 수준에 맞는 평가가 가능하다.

**Piece C: Explanation**

**Frankie: pupil scripts**

**Piece A: Short Story**

Opening the Fridge

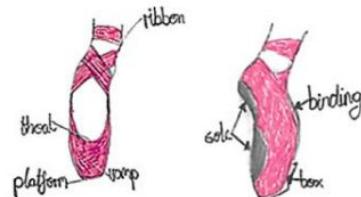
Slowly, Ewan peeped through the crack in his door. All was black. He took a step out. He could hear distant snoring as he crept across the landing. As his heart raced he stared into the darkness; he could hear the fridge urging him unwilling him to move. Now the stairs, the tricky bit. Suddenly a THUD!... He raced down the creaking stairs—even the seventh one that makes an earsplitting creak noise. He could see the white rectangle straight ahead of him. Then he opened it. He took a quick glance and saw the chocolate digestives. Then she. He could feel pair of eyes watching her in the darkness. Who was it? Had he been seen? His eyes darted around the room, his heart in his mouth. He grabbed the biscuits and ran for it.

"Ewan!" echoed a voice.

How Pointe Shoes Came To Be

Have you ever wondered why ballerinas look so beautiful and graceful on stage? Keep on reading to find out about what makes the Nutcracker you saw at Christmas the magical story that it is.

Pointe shoes are what make dancers different and beautiful. With their pink satin and silky ribbons, these shoes have been around since 1795. They were invented to make ballerinas look weightless when dancing, so they started spinning, balancing and jumping en pointe (on the tips of their toes). They are traditionally worn by women for a beautiful pad de deux (a solo dance with one man and one woman) but in some ballets men go en pointe too. There is an all male ballet company called Les Ballet Trockadero that had a very famous production of Swan Lake featuring men dancing en pointe as the female swans.



A diagram showing all the technical parts of a modern pointe shoe.

End-of-Key stage 2 statutory assessment – working at greater depth within the expected standard						
Name: Frankie	A	B	C	D	E	Collection
The pupil can:	Narrative	Description	Explanation	Newspaper report	Diary	Letter
• write effectively for a range of purposes and audiences, selecting the appropriate form and drawing independently on what they have read as models for their own writing (e.g. literary language, characterisation, structure)	✓	✓	✓	✓	✓	✓
• distinguish between the language of speech and writing and choose the appropriate register	✓	n/a	n/a	✓	✓	✓
• exercise an assured and conscious control over levels of formality, particularly through manipulating grammar and vocabulary to achieve this	✓	✓	✓	✓	✓	✓
• use the range of punctuation taught at key stage 2 accurately (e.g. semi-colons, dashes, colons and hyphens) and, when necessary, use such punctuation precisely to enhance meaning and avoid ambiguity.	✓	✓	✓	✓	✓	✓
(No additional statements for spelling or handwriting)						

[그림1] Key Stage 2 영어 작문 TA 프레임워크에 제시된 학생 작품과 평가 예시

TA에서 활용되는 'Pupil can' 진술문은 학생이 해당 수준에서 수행할 수 있어야 하는 행동이나 능력을 명확히 제시한다. 진술문에는 'most', 'many', 'some'과 같은 수식어가 포함되어 학생의 성취 빈도나 일관성을 구체적으로 드러낸다. 이러한 기준은 교사에게 명확한 평가 방향을 제공하면서도, 수업 속 다양한 활동을 자유롭게 반영할 수 있는 여지를 남긴다.

평가에 필요한 증거는 학생의 워크북, 프로젝트, 평가 메모, 시험 결과 등 다양한 수업 자료에서 자연스럽게 수집된다. 중요한 것은 양보다 질이며, 과도한 증거 수집은 지양된다. 예를 들어, 영어 작문에서 하나의 글이 여러 기준을 충족할 수 있고, 과학에서는 단 하나의 시험 문항이 중요한 판단

근거가 될 수 있다. 또한 6학년 교사는 이전 학년(3~5학년)의 평가 결과를 참고할 수 있어, 모든 내용을 다시 확인할 필요는 없다.

결과적으로, Key Stage 2의 교사 평가는 학생의 전인적 발달과 개별적인 학습 경로를 존중하며 이루어진다. 이는 시험 중심의 평가와는 다른, 교사의 전문성과 수업 현장의 관찰을 바탕으로 한 유연하고 통합적인 평가 체계라 할 수 있다. 학생의 성취를 단순한 수치로 환산하는 것이 아닌, 배움의 과정 자체를 존중하는 평가 문화가 Key Stage 2 교사 평가의 중심에 자리잡고 있다.

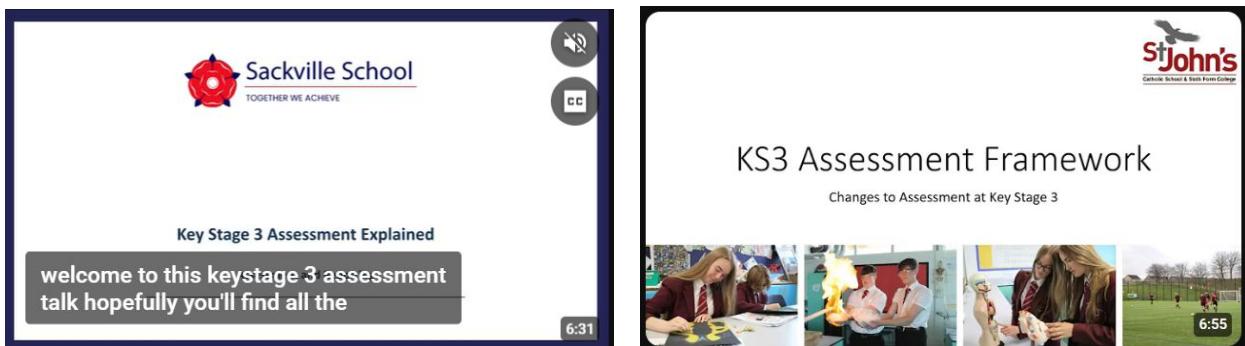
#### 나. Key Stage 3

Key Stage 2에서 실시된 국가 표준 시험 결과는 중등 과정 공식자격 시험인 GCSE 목표 성취 수준(Target Grade)을 설정하는 기초 자료로 활용된다. 교사들은 학생이 이 목표에 도달할 수 있도록 체계적으로 지도하며, 중등 5년간의 학업 성장을 추적·관리하고, GCSE 시험에 안정적으로 대비할 수 있도록 지원한다. 과거 Key Stage 3에서도 국가 단위 시험이 시행되었으나, 탐색 중심 학습과 자기 주도성 강화를 위해 2008년에 전면 폐지되었다.

국가 시험이 폐지된 이후에도, 학생의 성취 수준을 명확히 파악하고 지원하기 위한 학교 차원의 평가 체계는 지속적으로 발전해왔다. 하지만 일부 학교는 여전히 Ofsted에 성과를 입증해야 한다는 부담 속에서 GCSE 평가 기준을 Key Stage 3에 조기에 도입하고 있으며, 이로 인해 학생의 학습이 지나치게 시험 대비에 치우치는 부작용이 나타날 수 있다는 비판도 제기되고 있다.

과거에는 Key Stage 3에서 학생의 학업 성취를 판단하기 위해 National Curriculum Levels라는 국가 공통의 등급 체계가 사용되었다. 그러나 2016년, 영국 정부는 이러한 획일적인 평가 방식을 폐지하고 각 학교가 자율적으로 평가 체계를 수립하도록 권고하였다. 이 조치는 학생 개개인의 이해도와 학습 과정을 보다 정밀하게 반영하고, 교사 주도의 평가와 피드백을 활성화하기 위한 평가 체계의 전환이었다.

이에 따라 각 학교는 자체적인 평가 프레임워크를 설정하여, 교사들이 수업 중 진단 평가와 관찰을 통해 학생의 학습 상태를 면밀히 파악하고 개별 피드백을 제공하고 있다. 이러한 과정에서 평가는 전통적인 수치 등급 대신, **학생이 핵심 개념을 얼마나 깊이 이해하고 있는지를 서술적으로 표현하는 방식으로 이루어진다.** 이러한 질적 평가는 단순한 성취 결과가 아니라, 학생의 이해도, 사고력, 학습 지속성을 모두 반영하는 방식으로, Key Stage 3에서의 평가가 ‘성장 중심’ 교육을 실현하는 중요한 수단임을 보여준다.



[그림2] 학교별 Key Stage 3 Assessment 설명 영상

#### 4. 맷음말

영국의 평가 체계는 평가의 본질이 선발이나 서열화가 아니라, 학생의 성장을 지원하고 가능성을 실현하도록 돋는 데 있음을 분명히 보여준다. 이를 위해 국가 표준 시험과 교사 평가가 병행되는 이원적 구조를 채택하여, 교육의 객관성과 개별화를 동시에 추구한다. 국가 표준 시험은 모든 학생이 교육과정에서 요구하는 최소한의 성취 수준에 도달했는지를 점검하는 공통 기준으로 작용하며, 교육의 형평성과 책무성을 확보하는 기능을 한다. 한편, 학교 현장에서 이루어지는 교사의 평가는 학생 각자의 학습 여정과 이해 수준을 면밀히 진단하고, 그 잠재력을 최대한 발휘할 수 있도록 이끄는 역할을 맡는다. 이러한 이원적 평가 체계는 단순한 점수 산출이나 비교를 넘어, 학생 개인의 배움과 성장 그 자체에 집중하는 교육 철학을 담고 있다.

이제 한국 교육도, 평가를 결과 측정의 수단이 아니라 성장의 촉진 도구로 되돌려야 할 시점이다. 교사의 전문성과 학교 현장의 평가를 신뢰하고, 교육의 중심을 ‘사람’과 ‘성장’에 두는 구조로의 전환 이야기로 미래 교육이 나아가야 할 핵심 과제라 할 수 있다.

#### 【참고 자료】

- ▶ Department for Education (DfE). (2014). National curriculum in England: complete framework for key stages 1 to 4 – for teaching from September 2016
- ▶ Department for Education (DfE). (2014). Statutory guidance National curriculum in England: secondary curriculum
- ▶ Department for Education (DfE). (2024). Key stage 2 teacher assessment guidance 2025
- ▶ Standards and Testing Agency. (2015). National curriculum assessments: test frameworks
- ▶ Ofsted. (2022). Education inspection framework
- ▶ John McIntosh CBE. (2016). Commission on Assessment Without Levels: final report



## 프랑스의 학생 평가

발간위원 : 이한길 (포항제철자곡초등학교 교사)

### 1. 프랑스 교육 평가 시스템의 진화와 핵심 과제

#### 가. 프랑스 교육 평가의 중요성 및 최근 10년간(2015–2025) 변화의 흐름

평가는 프랑스 교육 시스템에서 학생 학습 성과를 측정하고 개별 학습 과정을 안내하며, 모든 교육 이해관계자에게 필수 정보를 제공하는 핵심 요소다. 지난 10년(2015–2025)간 평가 정책은 변화하는 사회적 요구와 교육학적 이해에 맞춰 시스템을 개선하려는 교육부의 지속적 노력을 보여준다. 이 기간에 통합 학습 기록부(Livret Scolaire Unique, LSU, 2015), 바칼로레아 개혁(2018), 학교 평가 시스템 도입(2020), ‘지식 충격(Choc des savoirs)’ 개혁(2024~) 등 주요 개혁이 시행됐다. 이런 흐름은 프랑스 교육이 정체되지 않고 역동적으로 발전해 왔음을 시사한다.

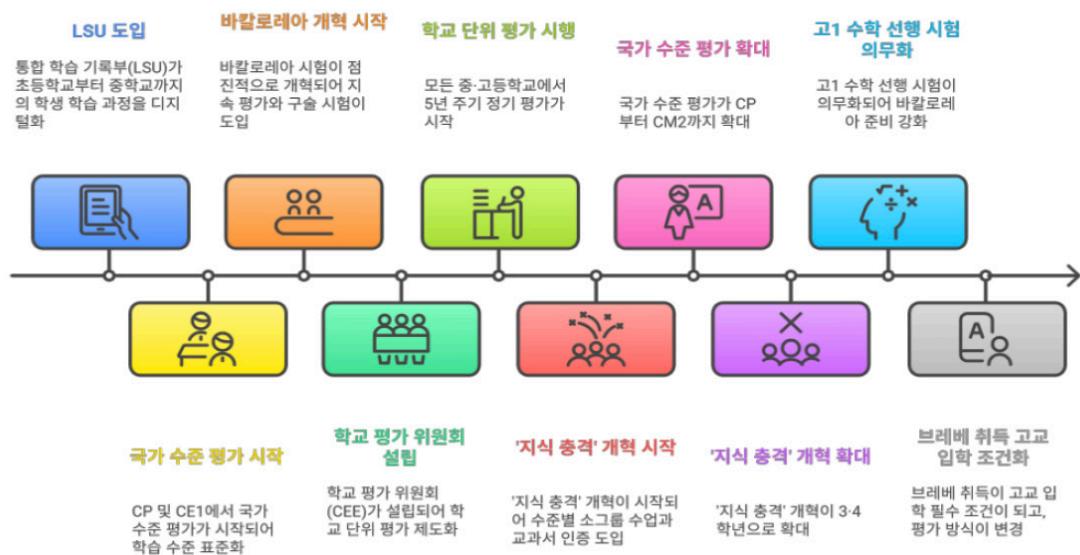
#### 나. PISA 2022 결과와 '지식 충격' 개혁의 배경

국제 학업 성취도 평가(PISA) 2022 결과는 프랑스 교육 시스템의 중대한 약점을 드러내며 ‘지식 충격(Choc des savoirs)’ 개혁을 촉발했다. 프랑스는 2018~2022년 수학에서 21점, 읽기 이해에서 19점 하락했으며, OECD 국가 중 사회경제적 배경에 따른 학력 격차가 큰 편이었다. 이 충격적 성과는 교육부로 하여금 학교 교육 및 평가 수준을 제고하는 대대적 개혁을 추진하게 했다. ‘지식 충격’ 개혁은 교육 철학이나 학술 논의보다 국제 비교 평가 같은 외부 지표가 정책에 미치는 영향이 커졌음을 보여준다. 이는 장기적으로 교육과정 설계, 교수법, 그리고 교육적 성공의 정의까지 변화시킬 여지를 만든다.

#### 다. 평가의 이중적 기능(교육학적 vs. 제도적)과 내재된 긴장

프랑스 교육 평가는 개별 학생 학습 진도를 지원하고 교수법을 개선하려는 교육학적 목표와, 국가 차원에서 시스템 성과를 측정해 정책 정보를 제공하려는 제도적 목표 사이의 긴장 관계에 있다. 교원 단체들은 대규모 국가 평가가 ‘불필요한 일반화’를 낳고 개별화 지원을 희생시키며 평가 업무와 학생 스트레스를 가중해 결국 통계로만 귀결된다고 비판한다. 이는 행정적 문제를 넘어 교육 목적에 대한 이념적 충돌을 반영한다. 교육부는 책무성과 성과 향상을 강조하는 반면, 교원 단체들은 교육학적 가치와 학생 중심 교육을 우선시하며, 각종 개혁이 이런 갈등을 심화시킨다.

## 프랑스 교육계의 평가 방식 개혁의 주요 이정표



[그림 1] 프랑스 교육계의 평가 방식 개혁의 주요 이정표 (관련 내용을 Napkin AI로 작성)

## 2. 주요 평가 제도 및 개혁 분석

### 가. 통합 학습 기록부(LSU) 및 국가 표준화 평가의 확대

#### 1) 통합 학습 기록부 (Livret Scolaire Unique, LSU)

2015년 법령으로 도입된 LSU는 초등학교 1학년(CP)부터 중학교 3학년(troisième)까지 학생의 학습 과정과 성취를 통합 기록·관리하는 국가 디지털 시스템이다. 다양한 학교 기록부를 대체해 의무 교육 전체 학업 이력을 일관되게 추적하고, 역량 중심 평가와 전통 점수 부여 간 불일치를 해소하며, 학부모에게 명확한 학습 정보를 제공한다. LSU는 학습 주기별 보고서, 이전 주기 최종 보고서, 공식 인증서를 포함하며, 주기 말에는 8개 공통 핵심 역량 속달도를 4단계로 평가한 종합 보

고서를 작성한다. 이 기록은 브레베(Diplôme national du brevet, DNB) 수여와 상급 학교 진학 결정에 핵심 자료가 된다. 또한 표준화된 평가 언어와 국가적 디지털 데이터베이스 구축을 통해 ‘지식 충격’ 개혁 이후 중앙 집중 데이터 분석과 목표 지향 개입의 기반을 마련한 전략적 단계였다.

## 2) 국가 표준화 평가(Repères)

2018년 CP(1학년)과 CE1(2학년)에서 시작된 국가 수준 평가는 2024년 9월 CP부터 CM2(5학년)까지 확대되었다. 이 평가는 교사에게 학생 학습 수준을 정확히 알려 맞춤형 교육을 지원하고, 학부모에게 자녀의 강·약점을 파악하게 한다. 국가적으로 기초학력 진단과 교육 불평등 해소를 목표로 한다. 확대는 ‘지식 충격’ 개혁 아래 저학년부터 체계적 표준화 데이터 수집을 지향하는 움직임으로 볼 수 있다. 교육부는 이를 진단 도구라 설명하지만, 중앙 당국이 학교·지역 성과를 비교하고 정책 효과를 측정하는 수단이 될 수 있다. 이는 교사 자율성을 제약하고 순응 압력을 높인다는 우려를 낳으며, 일부 교원 단체는 ‘실천 통제 강화 시도’라고 비판한다.

## 나. 바칼로레아 개혁

2018년부터 점진적으로 시행된 바칼로레아 개혁은 기존의 복잡한 최종 시험 부담을 줄이고 (8~12개에서 선행 시험 1개·최종 시험 4개로 축소), 지속된 학업 노력을 평가에 반영하며(지속 평가 40%+학교생활기록부 10%+과목별 지속 평가(E3C) 30%에 최종 시험 60% 합산), 구술 능력 평가 ‘그랑 오랄(Grand Oral)’을 도입해 고등 교육 준비와 진로 지도를 강화했다. 그러나 개혁은 학생들의 스트레스를 가중하고 새로운 불평등 문제를 야기했다. 지속 평가는 모든 점수가 중요해지며 강한 압박을 주었고, E3C는 시행 방식과 채점 기준 차이로 공정성 논란이 일었다. 또한 학교 재정·인력 부족으로 선택 과목 제공이 어려워지면서 도시 외곽·농촌 학교 학생들의 과목 선택권이 제한되어 교육 불균형이 심화되었다. 이는 준비 부족과 지역 다양성 미고려 시 의도치 않은 부작용이 생길 수 있음을 보여준다.

## 다. 지식 충격(Choc des savoirs) 개혁

지식 충격 개혁(2024)은 PISA 2022 결과에 따른 학력 저하·불평등 심화에 대응해 학력 향상, 불평등 감소, 기초 학습 강화를 목표로 한다. 첫째, 2024학년도부터 중학교 5·6학년 프랑스어·수학을 수준별 소그룹(약 15명) 수업으로 편성하고, 2025학년엔 3·4학년으로 확대한다. 교육부는 개인화 교육이라 하지만, 단일 중학교 원칙을 훼손해 분리·낙인을 초래한다는 비판이 있다. 둘째, 2027년부터 브레베(DNB) 취득을 고교 입학 필수 조건으로 하고, 미취득자는 준비반을 이수하도록 한다. 평가 방식도 지속 평가 40%, 최종 시험 60%로 변경해 선별 장벽을 도입했다는 우려가 제기된다. 셋째, 2026년부터 고1 수학 선행 시험을 의무화해 결과를 2027년 바칼로레아에 반영한다. 넷째,

초등 교과서에 교육부 인증 ‘라벨’을 부여하고, 우선 구매를 지원하지만, 특정 교수법 강요와 중앙 통제 강화라는 비판을 받는다. 이 일련의 조치는 유연·평등 지향에서 벗어나 중앙 통제를 강화하면 서도 학력 하락에서 상승으로 되돌리려는 정치적 의지를 보여준다.

#### 라. 학교 단위 평가

2019년 ‘신뢰의 학교를 위한 법률’로 학교 평가 위원회(CEE)가 설립되었으며, 2020년부터 모든 중·고교에서 5년 주기 정기 평가가 시행된다. 이 평가는 학교 구성원이 참여하는 자기 평가와 외부 평가팀의 외부 평가로 구성되며, 학습·진로, 학교생활·복지, 운영·제도 환경을 종합 점검한다. 목표는 학교 역동성 증진, 교육 질 향상, 학생 성공 지원, 교육 공동체 근무 조건·복지 개선이다. 학교 단위 평가는 기관 차원의 책무성을 강화하고 지속적 개선을 유도하는 시스템으로, 학생 성과뿐 아니라 학교 운영 정보를 제공해 데이터 기반 교육 거버넌스 모델을 공고히 한다.

### 3. 디지털 전환과 데이터 기반 평가의 명암

#### 가. LSU, LSL, LPI 등 디지털 도구의 역할과 상호연결성

프랑스 교육 시스템은 통합 학습 기록부(Livret Scolaire Unique, LSU), 고등학교 학습 기록부 (Livret Scolaire du Lycée, LSL), 포괄적 학습 과정 기록부(Livret de Parcours Inclusif, LPI)와 같은 디지털 도구를 통해 학생의 학업 이력과 지원 필요 사항을 체계적으로 관리한다. LSU는 초등학교부터 중학교, LSL은 고등학교 과정의 학업 성취를 기록하며, LPI는 특별한 교육적 요구가 있는 학생들을 위한 맞춤형 지원 계획(PPRE, PAP, GEVA-Sco, PPS 등)을 통합적으로 관리한다. 이들 디지털 기록부는 교육 지원 체계(Scolarité Services)라는 단일 국가 온라인 플랫폼과 에듀코넥트 (Educonnect)라는 국가 인증 시스템을 통해 학부모, 학생, 교직원이 접근할 수 있도록 상호 연결되어 있다. 이러한 포괄적인 디지털 생태계는 유아기부터 고등 교육 지원(Parcoursup 연계)까지 학생의 전체 교육 생애 주기를 관리하며, 행정 효율성을 높이고 교육 수준과 지원 서비스 간의 정보 단절을 해소한다. 이는 중앙 교육 당국이 전례 없이 상세하고 종단적인 학생 데이터를 확보하여 국가 수준의 분석과 정책 개입을 정교화할 수 있는 강력한 기반을 제공한다. 동시에, 방대한 개인정보의 중앙집중화는 데이터 프라이버시 및 활용 목적에 대한 잠재적 우려를 수반한다.

#### 나. 평가 데이터의 다중적 활용, 투명성 강화 및 책임성 문제

프랑스에서 수집된 평가 데이터는 교사의 수업 개선 및 학생 지원, 학부모의 자녀 학습 상황 이해, 교육청 및 감독관의 교육 활동 관리, 그리고 평가, 예측 및 수행국(Direction de l'évaluation, de la prospective et de la performance, DEPP)의 국가 교육 시스템 성과 측정 및 정책 결정 등

다양한 수준에서 활용된다. 이러한 다층적 데이터 활용은 학부모와 학생에게 교육 과정의 투명성을 높이는 긍정적 측면이 있지만, 동시에 평가를 교육학적 도구에서 책무성 관리 및 시스템 통제 체제로 전환하는 결과를 낳는다. 개별 학생 데이터가 교사와 학부모에게 교육적 정보를 제공하는 것을 목표로 하는 반면, 지역 및 국가 수준에서 데이터가 집계되고 분석되면서 학교와 교사에게 국가적 목표와 성과 지표에 부응해야 한다는 상향식 압력으로 작용할 수 있다. 이는 교육 현장의 자율성과 다양성을 위축시킬 위험을 내포하며, 평가의 본래 교육학적 목적이 제도적 관리 목표에 의해 잠식될 수 있다는 비판으로 이어진다.

### 3. 주요 쟁점과 교육계의 반응

#### 가. 교육 평등, 교사의 교육학적 자율성, 학생 및 교사의 부담 문제

지난 10년간 프랑스 교육 평가 개혁 과정에서 교육 평등, 교사의 교육학적 자율성, 그리고 학생 및 교사의 부담 문제는 지속적으로 핵심 쟁점으로 부각되었다. 특히 '지식 충격' 개혁의 수준별 그룹 운영과 고교 입학 시 브레베 필수화 조치는 사회경제적 배경에 따른 교육 불평등을 심화시키고 학생들에게 부정적 낙인을 찍을 수 있다는 심각한 우려를 낳고 있다. 또한, 국가 표준화 평가의 전면 확대와 교과서 라벨링 정책은 교사의 전문성과 교육 내용 및 방법 선택의 자율성을 침해하고 교육 실제에 대한 중앙 통제를 강화하려는 시도라는 비판에 직면해 있다. 여러 개혁에 걸쳐 교사들의 업무량 증가와 학생들의 과도한 학업 스트레스 문제 또한 반복적으로 지적되며, 이는 교육 개혁의 효과성과 수용성을 저해하는 주요 요인으로 작용한다. 이러한 문제들이 특정 개혁에 국한되지 않고 반복적으로 제기된다는 점은 프랑스 교육 시스템 내에 구조적인 긴장이 존재함을 시사한다. 이는 중앙 정부의 표준화, 성과 측정, 효율성 중심의 개혁 목표와 교육 현장의 자율성, 개별화 지원, 적정 노동 조건에 대한 요구가 충돌하는 근본적인 인식 차이를 반영한다.

#### 나. 주요 개혁에 대한 교원 단체의 비판적 시각과 우려

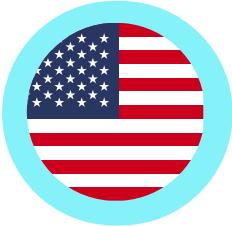
프랑스 주요 교원 단체들(Fep-CFDT, FSU-SNUipp, SNES-FSU, SE-UNSA 등)은 국가 평가 확대와 '지식 충격' 개혁에 강하게 반발한다. 이들은 전면 확대가 불필요·불공정하며 업무 부담과 학생 스트레스만 가중한다고 비판한다. 교과서 라벨링은 특정 학습 방법을 강요하고 전문성을 통제하는 조치로 본다. 개혁 전반은 교사 통제 강화와 사회적 선별 심화를 목표로, 교육과정 재편·교육학 규범 강화가 과학적 합의를 무시한다고 지적한다. 이들의 일관된 반대는 하향식 개혁이 전문성을 훼손하고 평등 이념을 약화할 수 있다는 깊은 이념적 갈등을 드러낸다.

#### 4. 맷음말

지난 10년간 프랑스 교육 평가 시스템은 통합 학습 기록부(Livret Scolaire Unique, LSU) 도입, 국가 표준화 평가(Repères) 확대, 바칼로레아 개혁, 학교 단위 평가, ‘지식 충격(Choc des savoirs)’ 등 주요 변화를 겪었다. 이들 개혁은 중앙집중화, 데이터 기반 관리, 결과 중심 교육 전환을 보여 준다. 그러나 교육학적 목표와 제도적 목표 간 긴장, 평등·공정성 확보의 어려움, 교사 자율성 대 표준화 간 갈등, 정책의 현장 수용 및 실행 문제 등 도전 과제는 여전하다. 특히 ‘지식 충격’ 개혁은 공화주의 이상과 충돌하며 심각한 논쟁을 야기했다. 이러한 과제가 반복되는 것은 단일 정책을 넘어 시스템 내부의 구조적·문화적 문제를 시사한다고 하겠다.

#### 【참고 자료】

- ▶ **Cadre général de l'évaluation 2022**, [https://www.vd.ch/fileadmin/user\\_upload/organisation/dfj/dgeo/fichiers\\_pdf/CGE/CGE.pdf](https://www.vd.ch/fileadmin/user_upload/organisation/dfj/dgeo/fichiers_pdf/CGE/CGE.pdf)
- ▶ **Message de la Ministre Anne Genetet à la communauté éducative...**, <https://www.education.gouv.fr/message-de-la-ministre-anne-genetet-la-communaute-educative-pour-le-futur-de-notre-ecole-415767>
- ▶ **PISA 2022: des résultats en forte baisse en France et dans la plupart des pays développés**, <https://www.apprentis-a-uteuil.org/actualites/education-et-scolarite/pisa-2022-des-resultats-en-forte-baisse-en-france-et-dans-la-plupart-des-pays-developpes>
- ▶ **Évaluations nationales en primaire : une généralisation ... – Fep-CFDT**, <https://fep.cfdt.fr/sinformer/vous-defendre/evaluations-nationales-en-primaire-une-generalisation-inutile-et-injuste>
- ▶ **Éducation nationale -Groupes de besoins, brevet, épreuve anticipée de mathématiques: les nouvelles mesures pour collèges et lycées | Service-Public.fr**, <https://www.service-public.fr/particuliers/actualites/A17850>



## 미국의 학습 평가

발간위원 : 이호연 (예천여자고등학교 교사)

### 1. 학생의 가능성에 주목하는 평가 혁신, 미국의 방향성

최근 미국 교육계는 단순히 점수로 학생을 평가하던 방식을 넘어서, 학생의 잠재력, 비판적 사고, 사회적 역량을 종합적으로 살펴보려는 방향으로 평가 체계를 전환하고 있다. 이러한 흐름은 코로나19 팬데믹 이후 더욱 가속화되었으며, 특히 학습 결손과 교육 격차가 부각되면서, 기존의 표준화 시험만으로는 학생의 실질적인 학습 상태와 성장을 판단하기 어렵다는 공감대가 확산됐다. 팬데믹 직후 실시된 전국 학업성취도 평가(NAEP) 결과, 미국 학생들의 수학과 읽기 성취도는 20년 만에 최저 수준으로 떨어졌고(NAEP, 2022), 학자와 교육자들은 기존 평가 방식이 “학생의 성장보다 선별에 집중해왔다”고 지적했다. 이에 따라 진단평가, 수행평가, AI 기반 평가(adaptive & automated assessment) 등 다양한 평가 방식이 주목받고 있다. 미국 교육계는 단순한 성적 산출 도구가 아니라, 학생의 개별적 성장, 정의로운 피드백, 맞춤형 지원을 이끌어내는 평가 패러다임으로 전환 중이다. 연방 교육부와 주 교육청, 교육연구기관 및 비영리 단체들이 이 평가 혁신에 함께 참여하고 있으며, 이는 정책 변화, 교육 플랫폼 기술 개발, 교사 전문성 강화 등 다양한 방식으로 실현되고 있다.

### 2. 진단평가(Diagnostic Assessment): 개별화 학습의 출발점

텍사스 교육청(TEA)은 2024년부터 STAAR(State of Texas Assessments of Academic Readiness) 평가의 일부 주관식 문항에 AI 기반 자동 채점 시스템을 도입하였다. 이 시스템은 자연어 처리(NLP) 기술을 활용하여 학생의 서술형 답안을 초기 채점하고, 이후 일부는 인간 채점자가 검토하는 방식으로 운영된다. 이를 통해 연간 약 1,500만 달러 이상의 비용 절감 효과가 기대되며, 채점의 일관성과 효율성을 높이는 데 기여하고 있다. 이러한 AI 기반 채점 시스템은 학생의 실시간 반응에 따라 문제의 난이도를 조절하는 적응형 평가(adaptive assessment)와 결합되어, 보다 정밀한 학습 진단이 가능하게 한

다. 이를 통해 교사는 학생 개개인의 학습 수준과 필요에 맞춘 맞춤형 지도를 제공할 수 있으며, 학생은 자신의 강점과 약점을 명확히 파악하여 학습 동기를 높일 수 있다.

### STAAR Test Development Process



[그림 1] STAAR Test Development & Review Process



#### STAAR U. S. History Blueprint Effective as of Academic Year 2022–23

Reporting Category	Number of Standards	Number of Questions	Number of Points
1: History	Readiness: 22 Supporting: 29	28–30	33–38
2: Geography and Culture	Readiness: 7 Supporting: 5	10–12	10–14
3: Government and Citizenship	Readiness: 5 Supporting: 10	8–10	9–13
4: Economics, Science, Technology, and Society	Readiness: 9 Supporting: 10	14–16	16–21
Item Types by Point	1-point questions (multiple-choice and non-multiple choice) 2-point questions (non-multiple choice)	50 14	50 28
<b>Total</b>		<b>64</b>	<b>78</b>

All TEKS, whether identified as readiness or supporting, are required to be taught in their entirety for a grade level or course.

**Readiness standards** are essential for success in the current grade level and important for preparedness for the next grade level or course. They address broad and deep ideas and require in-depth instruction. These standards make up approximately 55–70% of the total points on the base test.

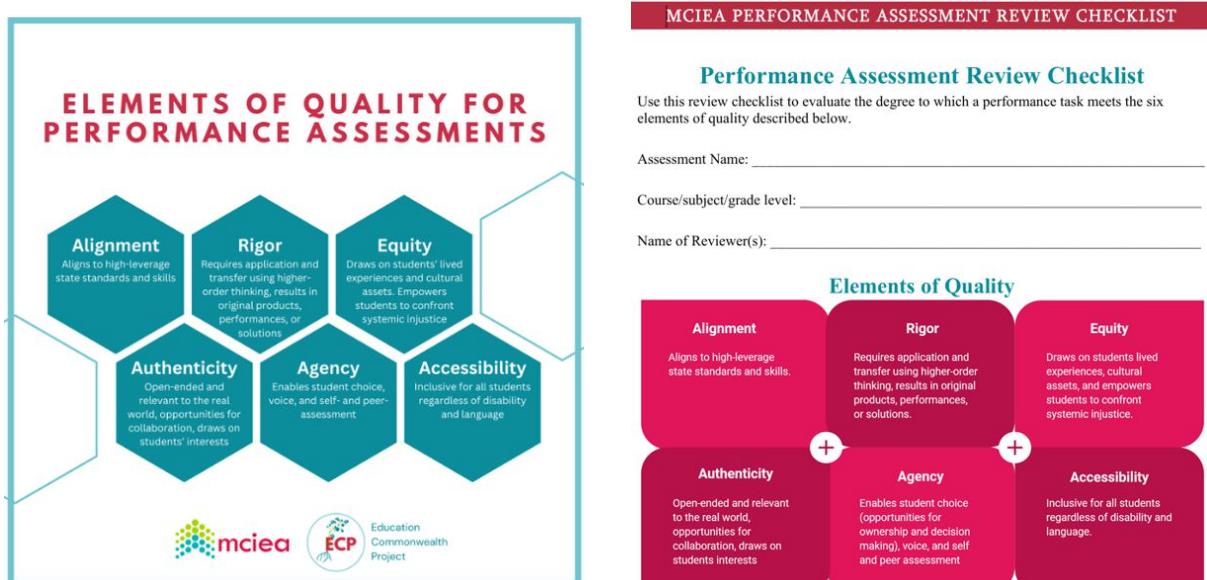
**Supporting standards** play a role in preparing students for the next grade or course but not one that is central. They address more narrowly defined ideas or concepts that are emphasized in grade levels below or above the current grade level or course. Supporting standards make up approximately 30–45% of the total points on the base test.

[그림 2] STAAR U.S. History Blueprint (2022–2023)

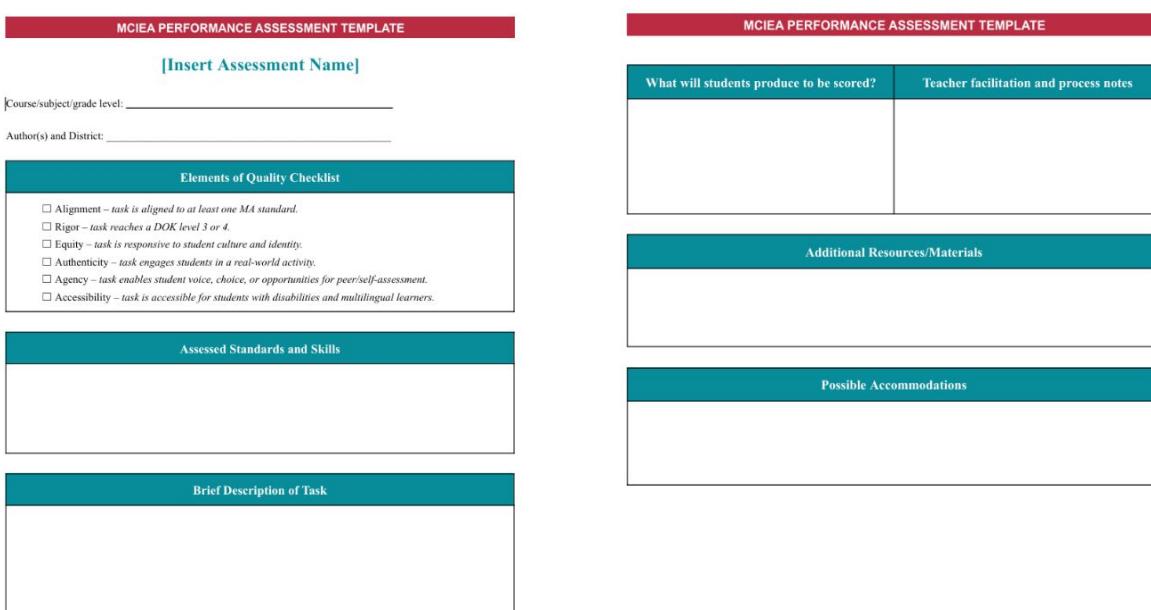
### 3. 수행평가(Performance-based Assessment): 비판적 사고와 창의성의 무대 가. 매사추세츠 MCIEA: 프로젝트와 포트폴리오 기반 졸업 평가

Massachusetts Consortium for Innovative Education Assessment (MCIEA)는 시험 중심의 졸업 요건에서 벗어나, 학생들이 직접 수행한 장기 프로젝트, 조사 보고서, 발표, 포트폴리오를 중심으로 졸업 요건을

평가하는 혁신 모델을 운영하고 있다. MCIEA는 매사추세츠 내 8개 학군과 교사 노조가 협력해 구성된 컨소시엄으로, 정성적 평가, 교사 자율성, 학생 주도 학습을 핵심 원칙으로 삼는다. 이들은 프로젝트 수행 과정에서 실제 사회와 연결되는 문제 해결, 자기 성찰, 협동 작업 등을 촉진하며, 평가 기준으로는 루브릭 기반의 채점 도구와 성장 포트폴리오가 사용된다. 또한 수행평가의 질을 담보하기 위해 정렬성, 엄정성, 형평성, 진정성, 주체성, 접근성의 6대 기준을 제시하고 있다. MCIEA 공식 보고서(2024)는 다음과 같이 설명 한다: “우리는 학생이 단순히 정답을 아는지를 보는 것이 아니라, 자신의 사고 과정을 설계하고, 결과를 전달하는 능력을 평가한다.” 이러한 방식은 기존 표준화 시험으로는 포착하기 어려운 학습의 깊이, 성장 경로, 사회적 역량을 드러낸다고 평가된다.



[그림 3, 4] MCIEA Performance Assessment Review Checklist and Quality Elements



**MCIEA PERFORMANCE ASSESSMENT TEMPLATE**

**[Insert Assessment Name]**

Course/subject/grade level: \_\_\_\_\_

Author(s) and District: \_\_\_\_\_

**Elements of Quality Checklist**

<input type="checkbox"/> Alignment – task is aligned to at least one MA standard.
<input type="checkbox"/> Rigor – task reaches a DOK level 3 or 4.
<input type="checkbox"/> Equity – task is responsive to student culture and identity.
<input type="checkbox"/> Authenticity – task engages students in a real-world activity.
<input type="checkbox"/> Agency – task enables student voice, choice, or opportunities for peer/self-assessment.
<input type="checkbox"/> Accessibility – task is accessible for students with disabilities and multilingual learners.

**Assessed Standards and Skills**

**Brief Description of Task**

**MCIEA PERFORMANCE ASSESSMENT TEMPLATE**

What will students produce to be scored?	Teacher facilitation and process notes
_____	_____

**Additional Resources/Materials**

**Possible Accommodations**

[그림 5, 6] MCIEA Performance Assessment Template



### Portfolios of Performance Pilot Planning Sheet

<b>District</b>	
<b>School</b>	
<b>Grade</b>	
<b>Subject</b>	
<b>Participating Educators</b>	

	Oct	Nov	Dec	Jan	Feb	Mar	Apr	May
<b>Power Standards</b> <i>(indicate when each will be assessed)</i>								
<b>Name of Performance Assessment</b> <i>(include link to materials: template, student directions, rubric)</i>								

Checklist:

- The full list of assessed standards represents the major areas of work and big ideas in the course.
- At least 3 performance assessments will be administered throughout the academic year.

[그림 7] MCIEA Portfolios of Performance Pilot Planning Sheet

#### 나. 캘리포니아 STEM 고등학교: 실제 문제 중심 설계 과제 도입

캘리포니아의 다수 STEM 특화 고등학교에서는 팀 프로젝트, 과학 실험 발표, 디자인 챌린지 등을 통해 수행평가를 실시하고 있다. 교사들은 단순 암기 테스트보다, 학생이 현실 문제를 정의하고 해결안을 제시하는 과정을 중시한다. EdWeek와의 인터뷰에서 한 교사는 이렇게 말했다. “지식 전달을 넘어, 학생들이 실제 문제에 창의적으로 접근하는 방식을 보게 되었습니다. 수행평가는 아이들의 사고 깊이를 이끌어내는 유일한 방법일 수 있어요.” 예를 들어, 한 공학 수업에서는 지역 사회의 교통 문제를 해결하기 위한 스마트 교차로 설계를 주제로, 학생들이 조사설계·3D 시뮬레이션 발표까지 수행하고, 동료 평가와 자기 평가까지 반영한 결과물을 제출했다.

#### 4. 형성평가(Formative Assessment): 교실 속 실시간 피드백

형성평가는 학생의 학습 과정을 실시간으로 파악하고 피드백을 제공하여 학습을 지원하는 평가 방식이다. 교육학자 Paul Black과 Dylan Wiliam은 『Inside the Black Box』(1998)에서 형성평가가 학생의 학습 동기와 성취도를 높이는 핵심 전략이며, 교사-학생 간 상호작용을 촉진하는 도구라고 분석한 바 있다. 최근에는 Nearpod, Kahoot!, GoFormative/Formative) 등 다양한 디지털 플랫폼을 통해 형성평가가 일상적으로 활용되고 있다. 이들 플랫폼은 실시간 피드백, 자동 채점, 반응 분석 기능을 통해 교사들이 학생의 이해도를 빠르게 진단하고 수업을 조정할 수 있도록 돋는다.

### 가. Nearpod: 실시간 참여와 시각적 피드백

샌프란시스코 통합교육구(SFUSD)는 Salesforce 재단과 협력해 Nearpod를 도입, 중학교 수학 수업에 커먼코어(Common Core) 기반 형성평가 활동을 적용하고 있다. 교사들은 'Draw It' 기능으로 학생의 풀이 과정을 실시간 확인하고, 'Time to Climb' 게임을 통해 학습 내용을 복습하며 참여도를 높인다. 또한 플로리다주의 핸드리 카운티 교육구는 영어 학습자(ELL) 학생을 위한 맞춤형 진단과 피드백을 위해 Nearpod를 적극 활용 중이다.

### 나. Kahoot!: 게임 기반 형성평가

Kahoot!는 수업 중간에 짧은 퀴즈를 통해 학생들의 이해도를 확인하고 피드백을 제공할 수 있는 게임 기반 학습 도구다. 미국 고등학교 교사들은 보통 3~5문항의 Kahoot! 퀴즈를 수업 중간에 실시하여 학생들의 오개념을 점검한 후, 결과에 따라 설명을 보강하거나 학습 전략을 수정한다. 한 고등학생은 “문제를 틀려도 바로 피드백이 와서 다시 시도할 수 있어 부담이 적다”고 말했다.

### 다. GoFormative (Formative): 실시간 개별 피드백 강화

Formative는 교사들이 다양한 질문 유형(객관식, 서술형, 수식 입력 등)을 통해 학생의 학습 상태를 실시간으로 파악하고 개별화된 피드백을 제공할 수 있는 플랫폼이다. AI 기반 질문 자동 생성 기능도 제공되며, 수학·과학·언어 예술 등 다양한 과목에서 활용되고 있다. 미국 전역 수천 개 학교가 사용 중이며, 교사들은 “Formative 덕분에 수업 중에도 학생 하나하나의 사고 과정을 볼 수 있어 지도 질이 달라졌다”고 평가한다.

## 5. AI 기반 평가: 자동화된 맞춤형 피드백과 학습 지원

AI 기술의 발전은 교육 평가 방식에도 혁신을 가져오고 있다. 특히 에세이 채점, 실시간 피드백, 감정 분석 등 다양한 영역에서 AI 기반 평가 도구가 활발히 활용되고 있으며, 이는 학습자 맞춤형 지원과 평가의 자동화를 가능하게 하고 있다.

### 가. ETS의 e-rater®

ETS(Educational Testing Service)는 AI 기반 자동 채점 시스템인 e-rater®를 개발하여, 문법, 문제, 구성 논리 등을 분석하고 점수를 제공하고 있다. 해당 시스템은 토플(TOEFL)과 같은 시험에 실제로 사용되며, 수험자에게 구체적인 글쓰기 개선 방향을 제시하고 있다.

### 나. Khan Academy의 Khanmigo

Khan Academy는 ‘Khanmigo’라는 AI 기반 글쓰기 코치를 통해, 학생들에게 실시간으로 글의 구조, 논리, 표현 등을 분석한 피드백을 제공하고 있다. 학생들은 에세이 초안에 대한 구체적인 조언을 받아 반복 학습을 할 수 있으며, 교사들은 이를 통해 글쓰기 지도에 들어가는 부담을 덜 수 있다.

#### 다. DreamBox Learning

DreamBox는 초·중등 수학 교육에 특화된 AI 기반 개인화 학습 플랫폼이다. 이 플랫폼은 학생의 문제 해결 패턴과 반응 속도를 실시간으로 분석하고, 이해도에 따라 학습 경로와 문제 유형을 자동 조절해 제공하고 있다. 이에 따라 학생 개개인의 학습 수준과 속도에 맞춘 피드백이 가능하다.

#### 라. Sown to Grow

Sown to Grow는 학생의 학습 감정과 동기를 데이터화하여 교사에게 제공함으로써, 정서적 상태에 따른 학습 지원을 가능하게 하는 플랫폼이다. 학생들은 수업 전후의 감정을 기록하고 목표를 설정하며, 교사는 이를 기반으로 맞춤형 피드백과 정서적 개입을 실천할 수 있다.

AI 기반 평가에 대해 교육계 일각에서는 우려의 목소리도 존재한다. RAND Corporation의 2023년 연구에 따르면, AI 알고리즘은 훈련 데이터에 따라 편향된 판단을 내릴 수 있으며, 이로 인해 공정성과 신뢰성에 대한 문제가 제기될 수 있다. 또한, 교육평가연구소(Institute for Assessment Research)의 레베카 민디치(Rebecca Mindich)는 Education Week와의 인터뷰에서 “AI 평가는 개인화 학습을 강화 할 수 있는 기회지만, 동시에 공정성과 프라이버시 문제가 우려된다”고 밝혔다. 한편, 캘리포니아의 한 고등학생은 EdSurge 인터뷰에서 “AI가 내 글을 평가해주니까 처음엔 놀랐는데, 지금은 실수를 지적해 주니까 큰 도움이 된다”고 말해, 학생들이 AI 피드백에 긍정적인 반응을 보이고 있음을 확인할 수 있다.

### 6. 맷음말

미국의 학생 평가는 점수 중심에서 학습 성장, 비판적 사고, 문제 해결, 협업 능력 중심으로 진화하고 있다. 전 미국 교육부 장관 존 킹(John King)은 National Public Radio(NPR) 인터뷰에서 “우리는 표준화 평가에 의존하던 관행에서 벗어나야 합니다. 학생의 다면적 역량을 볼 필요가 있습니다”라고 강조했다. 이러한 변화는 한국에도 중요한 시사점을 준다. 공정성, 신뢰성, 교사 역량 강화를 병행하며, 학생 개별 성장에 초점을 맞춘 평가 혁신이 필요한 시점이다.

#### 【참고 자료】

- ▶ <https://tea.texas.gov/student-assessment/staar>
- ▶ <https://www.mciea.org/performance-assessment>
- ▶ <https://www.ets.org/erater.html>
- ▶ <https://www.khanmigo.ai/writingcoach>
- ▶ <https://vorecol.com/blogs/blog-the-role-of-ai-in-adaptive-learning-systems-personalized-student-progress-tracking-through-automation-192564>



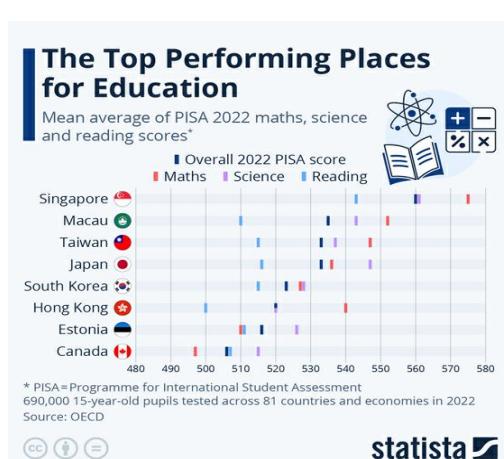
## 캐나다의 학생 평가

발간위원 : 안예린(포항송곡초등학교 교사)

학생 평가는 학습의 질 향상과 교육 형평성 확보를 위한 핵심 요소로, 평가와 학습은 서로 불가분의 관계에 있다. 캐나다 교육계는 일찌감치 평가의 확장적인 역할을 인식하고 교육과정 내에서 평가의 중요성을 강조하고 있다.

주와 준주로 이루어진 연방국가인 캐나다의 정부 구조 특성상, 캐나다의 교육 시스템은 국가 수준 교육부가 아닌 주 정부의 교육부 관할하에 작동한다. 각 학교급별 교육과정 또한 국가 차원의 단일 교육과정이 아닌 주/준주별 교육부마다의 독자적 교육과정 형태로 제시되며, 이는 평가에도 마찬가지로 적용된다.

한국, 일본과 같은 아시아 국가와 달리 캐나다 교육계는 표준화된 국가 수준의 일괄 평가보다는 주로 학교와 지역, 학생 특성에 맞는 개별적 평가 형식을 채택하고 있다. 그리고 이런 개별적 평가의 실효성을 확보하기 위해 교사의 전문성과 역량을 강조하고 있다.



[그림 1] 2022 PISA 결과 상위 8개 국가

전 국가적 수준의 표준화된 학업 성취도 평가가 시행되지 않음에도, 캐나다는 80여 개국의 만 15세 학생들을 대상으로 시행된 2022 PISA 검사 결과<sup>4)</sup> 읽기 507 점 (OECD 평균 476점), 과학 515점 (OECD 평균 485점), 수학 497점 (OECD 평균 472점)으로 5위 한국에 이어 서구권 교육 선진국 최고 수준인 8위를 차지했다. 특히 약 80%의 응시자들이 PISA 검사의 주 평가 항목인 수학과의 기초 성취 수준을 만족하였다는 점이 주목할 만하다.

4) Measuring Up: Canadian Results of the OECD PISA 2022 Study, The Performance of Canadian 15-Year-Olds in Mathematics, Reading, and Science: Highlights

이에 본 기사는 캐나다 K-12 교육 현장에서의 학생 평가의 원리와 각 유형, 유형에 따른 목적과 방식을 탐색하고 주별 최근 동향을 분석하여 한국 교육계의 학생 평가에 대한 시사점을 도출한다.

## 1. 캐나다의 학생 평가 철학과 목적

### 가. 캐나다 학생 평가의 철학

캐나다의 학생 평가 철학은 크게 학생 중심, 역량 중심, 형평성 및 포용성으로 요약할 수 있다. 표준화된 검사를 실시하기보다는 교사의 전문적 판단을 신뢰하고, 협력적 평가 문화를 구축하기 위해 학부모·지역사회와의 연계를 강조한다. 평가 기준 확보를 위해서는 각 주/준주별로 교육과정에 기초한 성취 기준 기반의 평가 루브릭과 관찰 기록 활용을 장려하고 있다.

### 나. 캐나다 학생 평가의 목적

캐나다에서 실시하는 학생 평가의 목적은 평가 유형에 따라 크게 세 가지로 구분하는데, 이중에서도 특히 **형성평가**(Assessment for Learning, AfL)와 **자기평가**(Assessment as Learning, AaL)을 강조하고 있다.

〈표 1〉 평가 목적에 따른 평가 분류

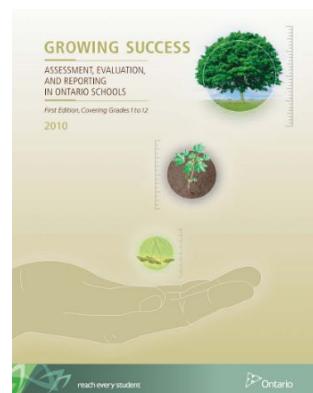
평가명	목적
<b>AoL: 총괄평가</b> Assessment <u>of</u> Learning	학습 종료 시점의 학습성취도 확인, 결과 중심 : 판단 및 성적 보고 이용, Summative assessment
<b>AfL: 형성평가</b> Assessment <u>for</u> Learning	학습 중 학습 진행 상황 파악 및 피드백 제공·학습 방향 조정, 과정 중심 : 학습 조정, Formative assessment
<b>AaL: 자기평가</b> Assessment <u>as</u> Learning	학습자 스스로 학습 과정 성찰 및 점검, 메타인지 중심 : 학습자의 자기주도 학습 촉진, Self assessment

## 2. 캐나다의 주별 평가 정책과 실행 방식

캐나다 각 주/준주별 교육부는 학생 중심 형성평가라는 공통된 철학에서 평가를 이해하고 활용하고 있지만, 주/준주의 사정에 맞는 다양한 방식으로 평가를 실행하고 있다. 주 수준의 K-12 대상 표준화된 학생 평가는 스스캐처월 주를 제외한 각 주/준주마다 이뤄지고 있으나 참가 학생 집단, 실시 연령 등 세부 내용은 모두 다르다. 2장에서는 몇몇 주들의 사례를 통하여 각 주의 평가 특징, 시기, 유형, 대상, 결과 통보 방법 등 다양한 학생 평가 관련 요소들을 분석한다.

## 가. 온타리오주의 CCRP

온타리오주에서는 Growing Success 정책을 통해 형성평가를 강화하는 방향으로 학생 평가를 실시하고 있다. 또한 주 교육부 홈페이지에서 평가와 감정(assessment and evaluation) 항목을 운영, 온타리오 주 학생 평가의 기본 원칙과 학생 성적 통보 방법, 지식·기능 영역, 평가 기준과 결과 서술 방식 등 교육 주체를 위한 다양한 평가 관련 정보를 제시하고 있다. 특히 온타리오주의 경우, Culturally Responsive and Relevant Assessment and Evaluation(이하 CRRP, 학생의 문화적인 배경을 고려한 평가) 운영을 강조하는 점이 두드러진다.

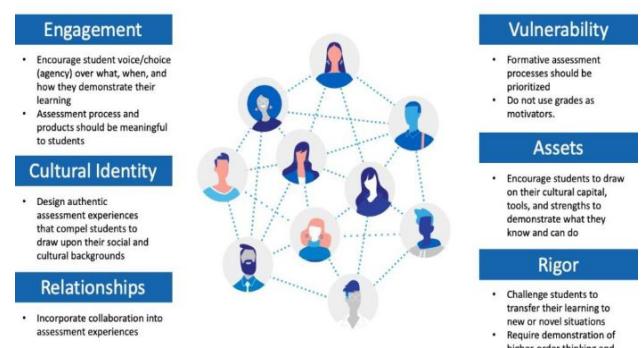


[그림 2] CRRT 작동 원리

CRRP의 실현을 위해서는 CRRT(Culturally Responsive and Teaching)이 우선되어야 한다. CRRT 과정에서 교사는 먼저 교사의 정체성과 편견이 학생 학습과 평가에 미치는 영향을 이해해야 한다. 이후 학생의 문화를 포용할 수 있는 수업을 통해 학생의 경험을 인간화하고, 학습 결과를 강화할 수 있다. 또한 교사 자신, 혹은 교수학습 과정에서의 문화적 편견을 해결하기 위해 지속적으로 자신을 성찰하고 다양한 성적 데이터를 비판적으로 분석해야 한다.

CCRP 평가의 주요 목적은 단순히 학생의 학습 수준을 확인하는 것이 아닌, 학생 학습을 개선하는 것에 있다. CRRP 평가는 교사가 각 학생과 그들의 사회문화적, 언어적 배경에 대해 학습할 수 있는 기회를 제공해 각 학생의 강점, 경험, 관심사 및 문화적 지식과 학습 방식을 학습 및 평가에 반영하게 한다.

또한 학습 및 평가 과정에서 교사는 학생이 보여주는 문화적 요소에 반응하는 방식으로 평가의 근거가 되는 학습 관련 증거를 다양하게 수집할 수 있다. 지속적인 서술적 피드백과 반응적 코칭은 학습의 과정일 뿐 아니라 CRRP 평가의 일환으로, 학생 학습 개선에 필수적인 역할을 수행한다. CRRP 과정에서 AI 및 디지털 에듀테크 또한 적극적으로 사용할 수 있다.



[그림 3] CRRP 구성요소<sup>5)</sup>

CRRP의 실현을 위해서는 CRRT(Culturally Responsive and Teaching)이 우선되어야 한다. CRRT 과정에서 교사는 먼저 교사의 정체성과 편견이 학생 학습과 평가에 미치는 영향을 이해해야 한다. 이후 학생의 문화를 포용할 수 있는 수업을 통해 학생의 경험을 인간화하고, 학습 결과를 강화할 수 있다. 또한 교사 자신, 혹은 교수학습 과정에서의 문화적 편견을 해결하기 위해 지속적으로 자신을 성찰하고 다양한 성적 데이터를 비판적으로 분석해야 한다.

CCRP 평가의 주요 목적은 단순히 학생의 학습 수준을 확인하는 것이 아닌, 학생 학습을 개선하는 것에 있다. CRRP 평가는 교사가 각 학생과 그들의 사회문화적, 언어적 배경에 대해 학습할 수 있는 기회를 제공해 각 학생의 강점, 경험, 관심사 및 문화적 지식과 학습 방식을 학습 및 평가에 반영하게 한다.

또한 학습 및 평가 과정에서 교사는 학생이 보여주는 문화적 요소에 반응하는 방식으로 평가의 근거가 되는 학습 관련 증거를 다양하게 수집할 수 있다. 지속적인 서술적 피드백과 반응적 코칭은 학습의 과정일 뿐 아니라 CRRP 평가의 일환으로, 학생 학습 개선에 필수적인 역할을 수행한다. CRRP 과정에서 AI 및 디지털 에듀테크 또한 적극적으로 사용할 수 있다.

5) Stembridge's (2020) Six Themes of Culturally Responsive Education Applied to Classroom Assessment

## 나. 브리티시컬럼비아(BC)주의 성적 전달 지침



브리티시컬럼비아주에서는 학생 평가 정책뿐 아니라 성적 통보 방식에 대한 가이드라인을 주 정부 홈페이지에 게시하고, 학부모와 교사를 위한 가이드라인을 배포하고 있다. 평가를 단순히 학습의 결과가 아닌 과정으로 인식하고 평가 결과 공유를 통해 차후 학습의 환경과 가정-학교 간의 협조적인 학습 환경 조성을 기대하는 주 교육부의 지침이 엿보인다.

해당 페이지에서는 특히 학생의 학습 결과를 가정으로 전달하기 위한 필수 수록 사항, 즉 성적표에 관한 지침을 포함하고 있다. 성적표에는 학생들과 학생의 가족들이 이해할 수 있는 언어로, 학생이 학습 과정에서 어떤 위치에 있는지와 차후 발전시킬 수 있는 영역에 대해 구체적이고 간결한 피드백이 포함되어야 한다. 또한 평가 결과 보고 과정의 일환으로 핵심 역량과 학생 목표 설정에 대한 **학생 자기평가**가 필수적으로 수록되어야 한다.

성적표에 성적을 표기할 때는 유치원에서 9학년(K-9)까지의 학생의 경우 성적표는 브리티시컬럼비아 학습 능력 척도를 사용해 표시하지만, 10~12학년은 문자 등급과 백분율(예시 B, 13%)을 명시하고 연말 학습 결과 요약서에는 졸업 상태에 대한 표기를 추가하여 졸업 요건 충족을 위해 필요한 정보를 확보하게 한다. 객관적인 성적 지표와 함께 K-12 전 학년에서 제공되는 서술형 피드백은 학생과 학생의 보호자에게 학생의 현재 능력과 발전 가능성에 대해 명확히 설명하는 역할을 한다.

또한 브리티시컬럼비아에서는 **학생 학습에 대해 자유롭고 유연한 소통**을 위해 최소 연 5회의 성적 통지와 같은 학생 학습 관련 소통 과정을 보장한다. 그중 3회는 서면 학습 결과 통지, 2회는 비공식적인 상담 등의 방식으로 진행된다. 의사소통 과정에서 교사는 학생과 보호자가 성적 관련 전달 내용을 명확하게 이해 할 수 있는지 그들의 문화와 언어 수준을 고려 해야만 한다.

이러한 학생 평가 정책은 브리티시컬럼비아 내 일반 공·사립 학교 학생들뿐 아니라, 캐나다의 또 다른 공용어인 프랑스어를 사용하는 브리티시컬럼비아 내 프랑스어 학교 학습자, 특수 교육 대상자, 성인 학습자와 온라인 학습 프로그램을 포함해 BC 교육시스템의 전 학습자에게 적용된다.



[그림 4] 브리티시컬럼비아 K-12 성적 통지 시기 및 필수 요소

## 다. 앨버타주

앨버타주 교육부에서는 저연령 초기 학습자 대상의 표준화된 평가를 확장하려 하고 있다. 2024년 12월, Demetrios Nicolaides 앨버타 교육부 장관은 기존의 초등학생 대상 의무 문해력, 수리력 평가의 대상을 확대해 매년 1월마다 유치원생 대상의 문해력, 수리력 평가를 실시하는 계획을 발표했다. 유치원생에게 표준화된 평가를 실시하는 일은 주 교육부 역사상 처음이다.

초등 1~3학년 학생들 또한 매 9월과 1월마다 의무적인 성취도 평가에 응시하게 되었으며, 해당 평가에서 추가적인 지원이 필요하다고 분류된 학생 대상으로는 6월에 추가적인 평가를 실시한다. 오는 2026년부터는 초등 4학년과 5학년 학생들 대상으로도 문해력, 수리력 평가가 실시될 예정이다. 초등학교 졸업까지 응시하는 시험이 기존 10회에서 최대 32회로 증가하는 것이다.

그러나 평가 계획 발표 이후 앨버타 교사 협회(이하 ATA)에서는 “학생들에게 필요한 것은 시험이 아닌 지원이다”며 우려를 표했다. ATA는 부진 학생 선별보다 소규모 학급 개설과 학생 대상 추가 지원이 더 중요하다고 주장했다. 이렇듯 전통적으로 학생 개별 맞춤형 평가를 중시했던 캐나다 내부에서도 표준화된 평가에 대한 논의가 활발히 진행 중이다.

## 3. 캐나다 학생 평가의 특징: 형평성과 포용성

### 가. ESL, ELL<sup>6)</sup> 학생

다언어 국가 특성상, 많은 캐나다 학생이 학업 과정에서 언어 장벽에 부딪히고는 한다. 학생 능력 평가 시 이런 언어 장벽 요인을 고려하기 위해 캐나다는 ESL, ELL 환경의 학생들을 대상으로 차별화된 평가를 제공한다.



[그림 5] 브리티시컬럼비아 ELL 평가 기준

### 나. 원주민(Indigenous) 학생

캐나다에서는 원주민 학습자의 문화를 존중하고, 원주민 학생들의 학업성취도와 학습 효능감을 향상할 수 있는 인정 기반 평가(Recognition based assessment)를 확산시키려 노력하고 있다.



[그림 6] 캘거리주 원주민 학생 평생 교육 프레임

### 다. 특수교육 학생

한국과 같이 특수교육 대상 학생을 위한 개별화 교육 계획(IEP, Individual Education Plan)을 제공하며 특수 교육 대상자의 평가도 일반 교육과정이 아닌 개별화 교육 계획을 따른다.

6) ELL: 능숙하게 영어를 구사할 수 있도록 교과수업 외적인 지원을 받는 학생. 캐나다의 경우 공용어인 영어와 프랑스어 지원을 받는 학생 모두를 지칭한다.

#### 라. 이주 배경 학생

캐나다에서는 CCRP와 같이 평가 도구의 개발 과정에서 이주 배경 학생의 문화적 어려움을 고려하여, 평가 내용과 방식의 문화적 편향을 배제하려 노력하고 있다.

### 4. 최근 캐나다 학생 평가 동향

최근 캐나다에서는 표준화된 평가에 대한 논의뿐 아니라 이전의 평가 방식을 개편하고, 평가 항목을 확장하려는 시도가 이뤄지고 있다. 역량 중심 교육과정에 맞는 역량 평가 방식 개편, AI 및 디지털 기술을 활용한 평가 시스템의 도입, 사회정서학습(SEI) 및 학생 주도 학습(Student-led learning) 등 다양한 분야의 평가에 대해 교육계 전반에서 활발한 논의가 이어지고 있다.

이러한 변화는 단순한 학업 성취도 측정을 넘어, 학생의 비판적 사고력, 문제 해결 능력 등 핵심 역량을 포괄적으로 평가하려는 방향으로 이뤄지고 있다. AI 기반 개별화 평가 시스템은 학생 맞춤형 피드백 제공 및 학습 성장 지원 도구로 주목받고 있으며, 사회정서 평가의 경우 학생의 정서적 안정 및 관계 형성 능력을 파악하여 전인적 발달을 도모하려는 시도의 일환이다.

이처럼 캐나다는 평가의 목적과 방법을 다각화하여 학생의 성장 가능성과 다양한 학습 경험을 반영할 수 있는 평가 체계를 구축하려는 움직임을 보인다.

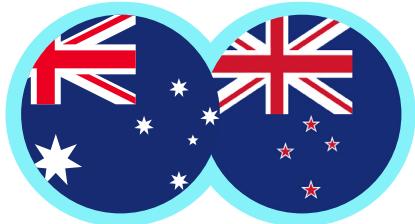
### 5. 맷음말 및 시사점

캐나다는 평가 과정에서의 학생 참여를 중시하고, 학생의 개별적 다양성을 적극 고려하여 평가 방식과 환경을 유연하게 운영하는 방향으로 학생 평가를 운영한다. 특히 평가 과정과 결과의 환류를 위해 교사의 전문성과 역량을 강조한다는 점도 한국 교육계에 좋은 시사점을 제시한다.

그러나 캐나다에서도 학생 평가 방식과 형태 개선에 대한 요구가 지속적으로 제기되고 있다. 캐나다 내부 주별 학습 불균형과 차이를 해소하고, 표준화된 평가를 통해 평가의 형평성과 신뢰성을 제고하면서도 개별화된 평가의 장점을 유지하기 위한 고민이 다양하게 이루어지고 있다.

#### 【참고 자료】

- ▶ <https://www.edcan.ca/articles/assessment-for-learning-across-canada/>
- ▶ <https://www.dcp.edu.on.ca/en/assessment-evaluation>
- ▶ <https://www.everylearnereverywhere.org/blog/culturally-responsive-teaching-and-humanizing-the-student-experience-in-the-age-of-ai/>
- ▶ <https://www2.gov.bc.ca/gov/content/education-training/k-12/administration/legislation-policy/public-schools/student-reporting>
- ▶ <https://globalnews.ca/news/10617855/alberta-teachers-association-elementary-school-student-assessments/>



## 호주와 뉴질랜드의 학생 평가

발간위원 : 최지원(상대초등학교 교사)

### 1. 호주의 학생 평가: NAPLAN

호주에서는 학생들의 학습을 개선하고 학업성취를 증진하고자 하는 데 목적을 두고 호주 교육과정평가보고위원회(Australian Curriculum, Assessment and Reporting Authority, ACARA)가 운영하는 국가 수준 학업성취도 평가(National Assessment Program – Literacy And Numeracy, NA PLAN)를 실시하고 있다. 이를 좀 더 구체적으로 살펴보면 다음과 같다.

호주의 학업성취도 평가(이하 NAPLAN)는 매년 3, 5, 7, 9학년을 대상으로 문해력과 수리력을 측정하기 위해 실시하는 것으로 우리나라의 기초학력 평가와 유사하다. 2008년부터 2022년까지는 매년 5월에 치러졌던 NAPLAN은 2023년부터 매년 3월에 치러지고 있으며, 2025년에는 3월 12일부터 24일까지의 기간 동안 실시되었다. 평가 영역은 읽기, 쓰기, 언어 규칙(문법, 철자, 구두점), 산수로 총 네 가지이며, 평가 문항은 철저히 전년도 영어과 및 수학과 교육과정의 학습 영역 및 분야와 일치하도록 출제한다. **호주 교육과정에서는 NAPLAN을 통해 측정하고자 하는 문해력과 수리력을 ‘일반 역량(general capabilities)’으로 칭하고, 이를 학교 학습 맥락을 통해 개발되고 교육과정 전반에 걸쳐 적용되는 일련의 지식, 기술, 행동 및 성향으로 정의하고 있다.<sup>7)</sup>**

평가 문항을 살펴보면 우리나라의 기초학력 평가와 몇 가지 차이점을 발견할 수 있다.<sup>8)</sup> 첫째, 읽기 영역에서 학생들에게 읽을거리 모음집과 답안지를 별도로 제공하고 있다. 읽을거리 모음집에는 다양한 장르의 텍스트가 수록되어 있는데, **동일한 학년 내에서도 다양한 수준에서 읽기 능력을 평가**

7) ACARA(2025). The Australian NAPLAN assessment framework. pp. 4.

8) NAPLAN 평가지(<https://www.acara.edu.au/assessment/naplan/naplan-2012-2016-test-papers>)

가할 수 있도록 다양한 길이와 수준의 텍스트를 제공한다. 학생들은 읽을거리 모음집을 읽어가면서 답안지에서 정답을 찾거나 쓰는 방식으로 문제를 해결하게 된다. 둘째, **쓰기 영역의 문항이 단지 서술형에 그치는 것이 아니라 실제로 한 편의 글을 쓰는 작문 시험을 위한 문항이다.** 학생들은 쓰기 문항을 통해 ‘글쓰기 자극(writing stimulus)’을 제공받고 서사적 글쓰기를 포함하는 상상 글쓰기, 정보 제공적 글쓰기, 설득적 글쓰기의 텍스트 유형 중에서 특정 유형의 텍스트로 답변을 작성해야 한다. 이러한 작문 형태의 쓰기 평가에서 존재할 수 있는 채점 곤란도를 해결하기 위해 채점 가이드를 개발 및 제공하고 있으며, 호주 전역에서 동일한 채점 기준이 적용될 수 있도록 평가자에 대한 집중적인 교육을 실시하고 있다.

**YEAR 5 READING**

Read *Bamboozled!* on page 3 of the magazine and answer questions 5 to 11.

**5** Bamboo is a type of

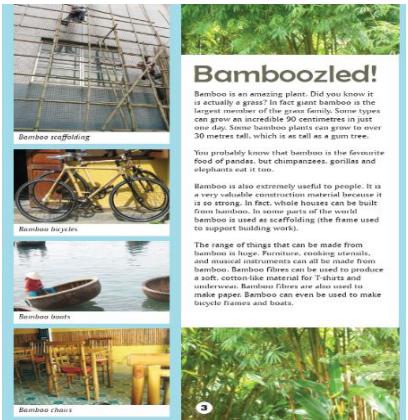
- tree.
- wood.
- bush.
- grass.

**6** According to the text, one way bamboo is amazing is that

- it can grow extremely fast.
- it grows near gum trees.
- it comes in many sizes.
- it grows anywhere.

**7** Which activity uses bamboo as scaffolding?

- plumbing
- building
- cooking
- sailing



**Bamboozled!**

Bamboo is an amazing plant. Did you know it is actually a grass? In fact giant bamboo is the largest member of the grass family. Some types of bamboo can grow up to 10 metres tall in just one day. Some bamboo plants can grow to over 30 metres tall, which is as tall as a gum tree.

You probably know that bamboo is the favourite food of pandas, but chimpanzees, gorillas and elephants eat it too!

Bamboo is also extremely useful to people. It is a very valuable construction material because it is very strong and flexible. Many houses are built from bamboo. In some parts of the world bamboo is used as scaffolding (the frame used to support a building).

The range of things that can be made from bamboo is huge. Furniture, cooking utensils, musical instruments, hats, clothes, paper, even medicine can be made from bamboo. Bamboo fibres can be used to produce a soft, cotton-like material for T-shirts and towels. Bamboo is also used to make filters to make paper. Bamboo can even be used to make bicycle frames and boats.

**YEAR 3 AND YEAR 5**

**Imagine**

Imagine if a character found an object that made something amazing happen.

Write a narrative (story) about the adventure.

You can use the characters and objects on this page OR you can have up your own.

Think about:

- the characters and where they are
- the complication or the problem to be solved
- how the story will end

Remember to:

- plan your story before you start
- choose your words carefully
- write in sentences
- pay attention to your spelling, punctuation and paragraphs
- check and edit your writing.

**Characters**

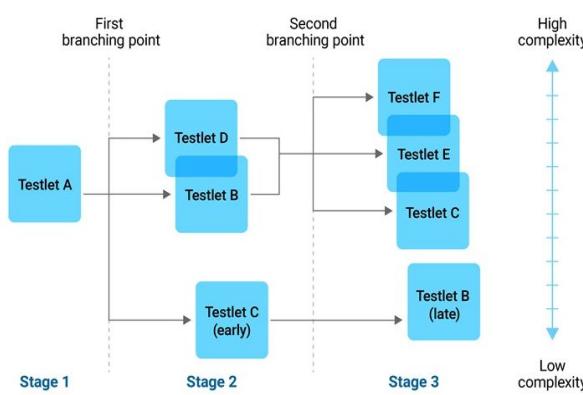


**Objects**



[그림 1] 5학년 읽기 평가지(좌), 읽기자료 모음집 일부(중), 글쓰기 자극 예시(우)

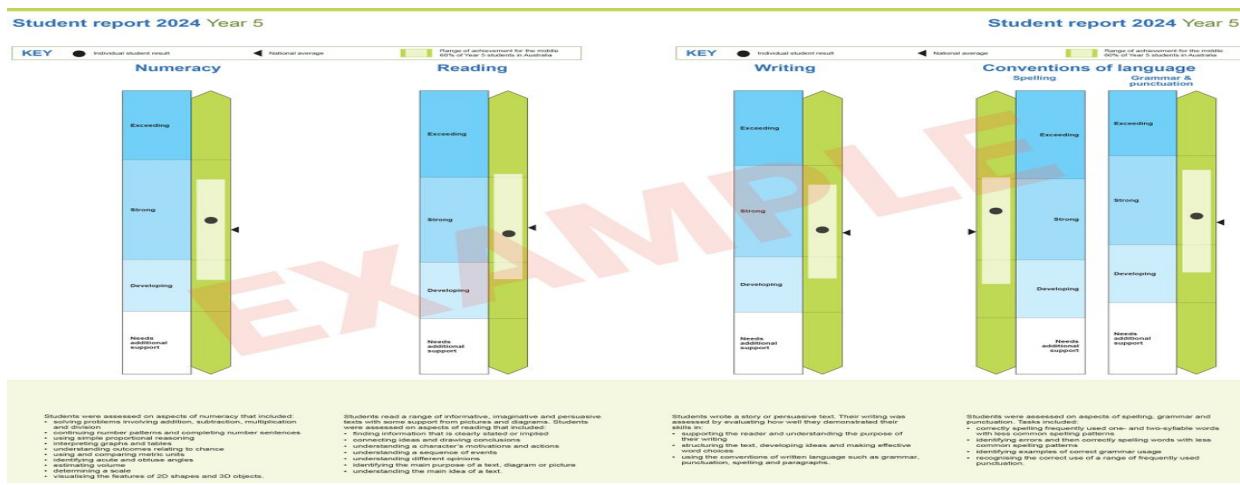
NAPLAN은 온라인으로 실시되며 다양한 문제 형식과 대화형 기능을 포함한다. 학생들은 클릭, 타이핑, 드래그를 통해 문제에 답할 수 있으며, 일부 문제에는 오디오 또는 대화형 도구가 포함되어 있다. 모든 문제는 복습 및 답안 작성, 플래그 지정, 나중에 다시 보기 등이 가능하다. 특히 NAPLAN 온라인 시험은 맞춤형 또는 적응형이므로, 학생들은 처음에 비슷한 문제들로 평가를 시작하지만 학생들 각각이 평가 초반에 어떠한 답변을 하는지에 따라 다음 문제들이 더 쉽거나 어렵게 제공된다. 이와 같은 맞춤형 평가는 각 학생의 시험 시간을 늘리지 않고 더 광범위한 질문을 함으로써 학생 성취도를 보다 정확하게 측정할 수 있으며, **학생들의 수준에 따라 차별화된 평가 도구를 제공할 수 있다는 강력한 이점이 있을 것으로 보인다.** 뿐만 아니라 시험 초반에 어려움을 겪는 학생들에게는 자신의 성적에 적합한 난이도의 문제가 주어지므로 학생들의 낙담 수준이 낮고, 평가가 학생의 능력에 맞춰 조정되기 때문에 학생들의 참여도가 높아지는 장점도 있는 것으로 나타났다.<sup>9)</sup>



- ▶ A 문제 세트로 평가를 시작하면 각 학생이 A 문제 세트에 답한 내용에 따라 두 번째로 나올 문제 세트가 결정됨(B or D)
- ▶ 두 번째 시험에서의 답변에 따라 가장 높은 복잡도(F), 중간 복잡도(E), 가장 낮은 복잡도(C) 등 최종 시험 문제의 난이도가 결정됨
- ▶ A 문제 세트에서 낮은 점수를 받은 학생은 바로 C 문제 세트(조기)로 이동한 후 B 문제 세트(후기)를 완료함

[그림 2] NAPLAN 맞춤형 시험 설계

평가 결과는 국가 결과, 학교 수준 보고서, 학생 개별 보고서 등 다양한 방식으로 보고된다. 국가 결과는 각 학년 및 영역별 주/지역별, 국가별, 성별, 영어 이외의 언어 배경별, 부모 직업 및 교육 수준별 결과가 포함된 대화형 온라인 보고서로 제공되는데, 2025년의 결과는 8월에 나올 것으로 예상한다.<sup>10)</sup> NAPLAN을 이수한 학교와 학생의 결과는 ‘My School’ 홈페이지에서 학교명을 검색 후 결과를 확인할 수 있으며, 학생 및 학교 요약 보고서(SSSR) 또는 주/준주 시험 관리 기관(TAA)이나 기타 교육기관에서 제공하는 보고서를 통해 학교에 제공된다. 학생 개별 보고서는 시험에 대한 일반 정보와 보고서를 읽는 방법을 포함해 각 평가 영역에서의 결과를 4가지 능숙도 수준(Exceeding, Strong, Developing, Needs additional support)으로 제공한다.



[그림 3] 5학년 학생 개별 보고서 예시

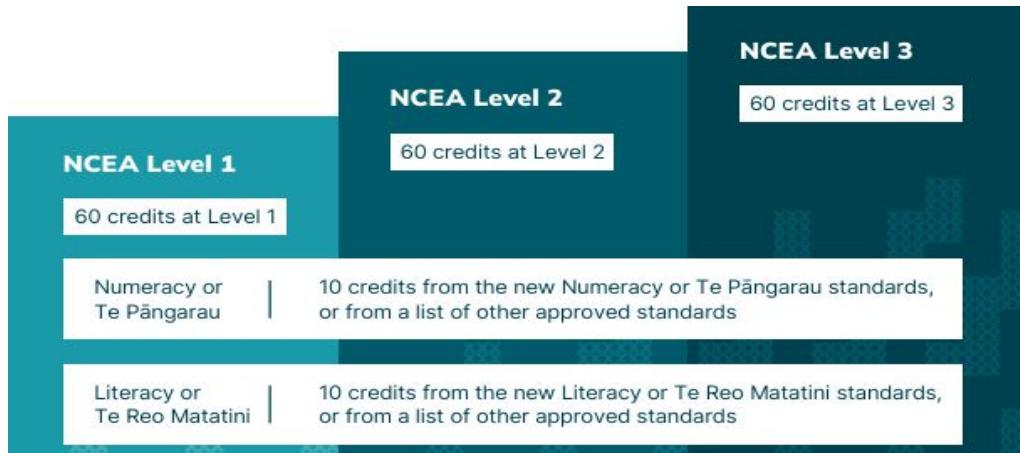
9) NAPLAN 온라인 맞춤형 테스트(<https://www.nap.edu.au/naplan/understanding-online-assessment/tailored-tests>)

10) NAPLAN 결과 및 보고서(<https://www.nap.edu.au/naplan/results-and-reports#isrs>)

## 2. 뉴질랜드의 학생 평가: NCEA

뉴질랜드에서는 11~13학년, 즉 고등학교 1~3학년 학생들을 대상으로 국가공인 학업성취 인증서 (National Certificate of Education Achievement, NCEA) 과정을 통해 평가하고 있다.

뉴질랜드의 국가공인 학업성취 인증서(이하 NCEA)는 레벨 1부터 레벨 3까지 총 3개 단계로 구성되어 있다. 보통 11학년 학생들은 레벨 1, 12학년은 레벨 2, 13학년은 레벨 3으로 진행하나, 일부 학교에서는 11학년과 12학년에 걸쳐 레벨 2 과정을 이수하기도 한다. **학생들은 순서와 상관없이 단계별 기준 학점(credits) 이상을 획득하고, 3단계 평가를 모두 거친 후에 총점을 받기 때문에 NCEA는 상당히 유연한 특징을 가지고 있다.**<sup>11)</sup> 단계별 기준 학점은 각 60학점이며, 여기에 문해력과 수리력 각 10학점을 합산한다.



[그림 4] NCEA Level 1~3 기준 학점

학생들은 NCEA를 준비하는 과정에서 학교에서 제공하는 40여 과목 중 자신의 능력과 관심에 따라 과목을 선택하여 공부할 수 있는데, 11학년은 영어, 수학, 과학과 더불어 두세 개의 과목을 추가 선택하고, 12학년과 13학년은 대학 진학 혹은 취업 방향을 고려하여 과목을 선택한다. 뉴질랜드에는 별도의 고등학교 졸업장이 없기 때문에 NCEA 학점 이수를 통해 대학 입학 자격인 UC (University Entrance)를 충족해야 하며, UE를 충족하지 못했을 경우에는 추가 학점을 이수해야 한다. NCEA 평과 결과는 뉴질랜드 고등학생들의 학업성취도를 파악하는 것뿐만 아니라, 국내 및 해외대학 진학, 장학금 수여, 취업 등에도 활용되고 있다.

11) NZQA(2024). Guide to NCEA, pp. 9.



[그림 5] NCEA Level 1~3

학생들이 선택하여 학습하고 평가에 응시하는 모든 과목은 표준(standards)에 의해 평가되며, 표준은 다양한 지식과 기술 분야를 설명한다. 예를 들어 레벨 1의 영어 표준은 [그림 6]과 같다. 각 학교에서는 어떤 과목을 제공할지, 어떤 기준으로 평가할지 결정하는데, 학생들이 이미 승인된 과제를 해결하거나 프로그램을 학습하는 것은 물론, **NCEA 가이드에 포함되지 않는 활동을 통해서도 표준을 달성하도록 할 수 있다.** 예를 들어, 운전면허증을 취득하는 것, 지역 또는 전국 카파 하카(kapa haka, 마오리 공연 예술) 대회나 폴리페스트(polyfest, 중등학교 폴리네시아 축제)에서 공연하는 것 등의 방법으로도 표준을 달성할 수 있다. 표준을 달성하면 해당 표준만큼의 가치가 있는 총 학점을 얻게 된다.

NCEA Level 1 English standards			
Demonstrate understanding of how context shapes verbal language use	Demonstrate understanding of specific aspects of studied text	Develop ideas in writing using stylistic and written conventions	Demonstrate understanding of significant aspects of unfamiliar texts
91924	91925	91926	91927
5 credits	5 credits	5 credits	5 credits
Internally assessed	Externally assessed	Internally assessed	Externally assessed

[그림 6] NCEA Level 1 영어 표준

NCEA 평가 방법은 내부 평가(internal assessment)와 외부 평가(external assessment)로 나누어 진다. 내부 평가는 교사들이 학교 내에서 학생들의 학업 성취를 자체적으로 평가하는 방식이며, 외

부 평가는 뉴질랜드 국가자격위원회 New Zealand Qualifications Authority, NZQA) 주관으로 실시하는 국가 단위의 평가로 그 결과는 NZQA에서 채점한다.

NCEA 표준을 달성했음을 알 수 있는 결과는 성취 표준(achievement standards), 단위 표준(unit standards), 기술 표준(skill standards)의 세 가지 유형으로 나누어 제시된다. 성취 표준은 내부 또는 외부에서 평가할 수 있고, 단위 표준과 기술 표준은 일반적으로 내부적으로 평가한다. 성취 표준을 달성한 정도는 달성하지 못함(Not achieved, N), 달성함(Achieved, A), 충족(Merit, M), 우수(Excellence, E)의 4단계로 제시되며, 단위 표준과 기술 표준 중 일부는 성취 표준과 같이 N, A, M, E 등급을 제공하기도 하나, N과 A 두 가지 등급만으로 제시되기도 한다.

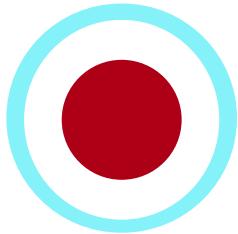
### 3. 맷음말

이러한 호주와 뉴질랜드의 학생 평가 제도는 우리나라의 기초학력 평가와 대학수학능력시험에 몇 가지 중요한 시사점을 제공한다. 먼저, 호주의 NAPLAN처럼 적응형 온라인 평가를 도입하는 것은 학생 개개인의 역량을 보다 정확히 진단하고, 학습 흥미를 유지하면서 평가의 스트레스를 줄이는 데 기여할 수 있다. 특히 평가 문항이 학생 개인의 수준에 맞춰 조정되는 점은 기초학력 미달 학생에게도 맞춤형 지원의 기회를 제공할 수 있다는 점에서, 우리나라의 다소 획일적인 문항 체계와 비교해볼 때 유의미한 개선 방향을 제시한다.

한편, 뉴질랜드의 NCEA는 학생의 선택권과 평가의 유연성 측면에서 시사점이 크다. 과정 중심의 평가와 다양한 인증 경로를 통해 학업성취를 입증할 수 있도록 함으로써 학생 개인의 진로와 관심을 존중한다는 점이 특징적이다. 또한 학교와 지역 사회의 활동까지 학업 성취로 연결하는 방식은 형식 교육을 넘어선 학습의 사회적 확장이라는 점에서 주목할 만하다.

#### 【참고 자료】

- ▶ NAP Australia, <https://www.nap.edu.au/home>
- ▶ ACARA Australia, <https://www.acara.edu.au/>
- ▶ My School by ACARA, <https://myschool.edu.au/>
- ▶ NSW Department of Education, “The assessment triangle to track learning”  
<https://education.nsw.gov.au/news/latest-news/the-assessment-triangle-to-track-learning>
- ▶ NZQA New Zealand, <https://www2.nzqa.govt.nz/ncea/about-ncea/>



## 일본의 학생 평가

발간위원 : 최광현(복주초등학교 교사)

### 1. 학생 평가의 변화 배경

#### 가. 변화의 배경

일본은 오랫동안 정기적인 시험을 통한 성적 산출식 평가에 의존해 왔지만 최근에는 학생 중심의 성장 촉진형 평가로의 전환이 활발히 이루어지고 있다. 이러한 변화의 배경에는 학생의 자율성과 사고력, 표현력 등 ‘비인지적’ 요소에 대한 관심 증가와 더불어 디지털 전환 가속화, 학습 다양성 확대, 문부과학성에서 제시한 새로운 학습지도요령(이하 교육과정) 등이 자리하고 있다. 이 외에도 등교 거부(不登校) 학생의 증가, 학력 격차 해소 필요 등 다양한 요인 또한 평가 방식의 다변화를 촉진하는 요인이다.

#### 나. 평가의 시기 및 방식

일본의 초·중·고등학교 평가는 크게 시험을 통한 ‘성적 평가’와 더불어 ‘수행 평가’, ‘출석 및 태도 평가’ 등으로 이루어진다.

이중 성적 평가는 학교급에 따라 평가 방식이나 평가의 주안점 등에서 다소 차이가 있는데, 먼저 초등학교는 기본적으로 ‘정기 평가’를 보지 않고 단원이 끝날 때마다 ‘단원 평가’를 실시하는 등 점수보다는 이해도를 중심으로 평가하는 경향이 있다.

반면 중·고등학교는 중간 평가와 기말 평가 등이 정기적으로 실시되는 등 지필 시험에 의한 평가 비중이 높다. 특히 고등학교는 대학 입시에 ‘내신 성적’이 중요한 역할을 함에 따라 성적 평가가 엄격히 진행된다는 특징이 있다.

일본의 학기는 한국과 달리 3학기제<sup>12)</sup>로 운영되며, 1학기는 4월~7월, 2학기는 9월~12월, 3학기

는 1월~3월로 구분된다. 따라서 중·고등학교에서 치러지는 정기 평가 또한 학기와 연동되어 그 시기가 결정되는데, 통상적인 중학교의 정기 평가 일정을 살펴보면 다음과 같이 정리할 수 있다.

〈표 1〉 정기 평가 일정

학기	정기 평가	일정
1학기	중간 평가	5월 중순~하순
	기말 평가	6월 하순~7월 상순
2학기	중간 평가	10월 상순~중순
	기말 평가	11월 하순
3학기	학년말 평가	2월 중순~하순 (중학교 3학년은 1월 하순에 실시)

## 2. 정기 평가과 단원 평가

### 가. 정기 평가 폐지 및 축소



[그림 1] 코지마치 중학교 교장 인터뷰 영상 (출처: Tokyo MX News)

2018년 도쿄도 치요다구립 코지마치 중학교는 일본에서 최초로 정기 평가 폐지를 결정했다. 정기 평가를 단원 평가로 바꾸고 재도전을 허용함으로써 학생들로 하여금 스스로 학습하는 자세를 기를 수 있도록 한 것인데, 이러한 결정은 전국적으로 주목을 받게 되었고 이와 비슷한 움직임이 전국 중학교를 중심으로 점차 확대되기 시작했다. 이후 사립 고등학교는 물론 공립 고등학교에서도 정기 평가를 폐지하는 학교가 늘어나고 있는 상황이다.

일본에는 정기 평가 실시와 관련된 법적 규정이 없다. 때문에 정기 평가를 언제, 몇 번 실시할지는 전적으로 각 학교의 판단에 달려있으며, 정기 평가가 아닌 다른 방식의 평가도 불가능한 것이 아니다. 물론 그럼에도 불구하고 지금까지는 성적 평가를 위해 빼놓을 수 없는 방법으로써 정기 평가가 강조되어 왔다.

12) 2022년 9월 기사(일본의 신학기 교육 준비) 참고

하지만 위에서 언급된 코지마치 중학교의 사례와 더불어 2022년부터 시행된 새로운 교육과정은 정기 평가를 중심으로 한 학습 평가 방식에 변화를 촉발했다. 기존에는 정기 평가의 결과를 중심으로 학생 성적표에 기록되는 점수를 산출했는데, 새로운 교육과정에서는 ‘자식·기능’(학습한 지식의 이해와 기술 숙달), ‘사고·판단·표현’(문제해결능력, 논리성, 창의적 사고), ‘주체적으로 학습에 임하는 태도·인간성’(자발성, 협동성, 성실성)이라는 세 가지의 관점에서 평가하는 ‘관점별 평가’를 강조함에 따라 각 학교에서는 정기 평가를 중심으로 한 학습 평가 방식의 개선 필요성을 느끼게 된 것이다.

[초등학교 성적표 양식(예)]

학습의记录 (教科評定)				
教 科	관점	목표	평가	
			よ う で き る	ち も 少 し
国 語	国語への関心・豊富・態度	国語に対する 관심をもち、豊富な表現力を持った。	○	
	豊す・聞く能力	友達の発表を聞いたり、自分の意見を述べたりする。	○	
	書く能力	目的や意図に応じて手紙を書いた。	○	
	読む能力	作品の雰囲気を味わったり、物語の構成要素を理解したりする。	○	
社会	言語についての知識・理解・技能	送りがなや仮名遣いに注意しながら文章を読み取る。	○	
	社会的行動の関心・意欲・態度	我が国の農業や水産業に関する知識を身につける。	○	
	農作物・資源活用の技術	地図や統計などの資料を効率的に読み取る。	○	
	社会的行動についての知識・理解	我が国の農業や水産業の特徴を理解する。	○	
算 数	数量の関心・意識・態度	数量や图形について既習の知識を活用する。	○	
	数学的な考え方	既習の内容をもとにして、問題を解決する。	○	
	数量や图形についての思考	数の仕組みがわからず、10倍、100倍などと計算する。	○	

[중학교 성적표 양식(예)]

教 科	학습목표(평가의 관점)	학습의樣子	1학기			2학기			3학기		
			蹟	蹟	蹟	蹟	蹟	蹟	蹟	蹟	蹟
国 語	国語に対する关心を深め、国語を身につけて話したり、的確に聞く	A									
	筋道を立てて、適切に文章を書く	B									
	様々な文章を的確に読み取る。	B	4								
	言語事項 (文法、漢字等)について	A									
社会	広く社会に対する关心をもち、意図的に社会に関する課題に対し、いろいろな情報を収集・分析する	A									
	社会的行動の関心・意欲・態度	B									
	各種資料に親しみ、それを選択・活用する	B									
	社会的な出来事の意義や特色を理解する	B									
数学	身の回りの事象を数学的にとらえる	A									
	事象を数学的にとらえられる見方	A									
	事象を数量・图形などで表現し、数量を計算する	A	5								
	自然の事物・現象に興味をもち、興味をもつて観察する	A									

[그림 2] ‘관점별 평가’ 반영 성적표 예시 (출처: 사이타마현 교육국)

참고로 관점별 평가는 단지 시험 점수뿐 아니라, 수업 중의 활동, 발표, 토론, 조사 보고서, 자기 평가 등 다양한 방식을 활용하여 복합적으로 학생을 바라보고자 하는 시도이다. 특히 루브릭을 통해 평가 기준을 사전에 제시하고, 평가 결과를 학생 피드백 및 학습 개선의 기회로 적극 활용한다 는 것이 특징적인 부분이라 할 수 있겠다.

문부과학성에서는 이러한 관점별 평가 기준의 일관성을 유지하기 위해 지속적인 교사 연수를 실시하고 있으며, 성취기준 도달 여부를 기록하는 방식으로 평가의 질을 높이기 위해 노력을 기울이고 있다.

## 나. 단원 평가

위에서 언급한 도쿄도 치요다구립 코지마치 중학교의 사례처럼 정기 평가를 폐지한 학교들 가운데 다수는 ‘단원 평가’를 도입하고 있다. 단원 평가란 말 그대로 특정 단원의 학습이 완료된 후, 해당 단원의 학습 내용이 학생들에게 잘 정착되었는지를 확인하기 위해 시행하는 평가 방식이다. 즉, 정기 평가처럼 성적이 한 번에 결정되는 것이 아니라 각 과목의 단원을 평가 기준으로 삼아 매일

의 학습이 성적에 연결되는 방식인 셈이다.

단원 평가의 장점을 살펴보면, 먼저 단원 평가는 정기 평가에 비해 출제 범위가 좁기 때문에 정기 평가와 달리 여러 과목을 벼락치기로 공부하는 학생들이 단원 평가 방식에서는 크게 줄어드는 경향이 있다. 게다가 교사들은 단원 평가를 통해 학생들의 학습 내용 이해도와 정착도를 보다 세밀히 파악할 수 있으며, 필요한 지도 방식을 조정하기에도 용이하다. 더불어 시험지 외의 다양한 방법으로 단원 평가를 시행할 경우 학생들을 더욱 다양적으로 평가할 수 있는 기회 또한 늘어날 수 있다.



[그림 3] 단원 평가 재도전 모습 (출처: VIEW next)

학생들 또한 단원이나 주제별로 자신의 이해도를 확인할 수 있어 이해도가 낮은 부분이 누적되기 전에 보충 학습을 할 수 있는 기회를 얻을 수 있으며, 일부 학교에서는 학생이 원하는 경우 여러 번 단원 평가에 재도전 할 수 있는 기회를 부여해 주는 ‘단원 평가 재도전’ 방식을 시행하고 있다.

실제로 이러한 장점을 근거로 2023년 가고시마현의 이쥬인 고등학교와 킨코완 고등학교는 정기 평가를 폐지하였는데, 해당 학교에서는 단원 평가와 루브릭 기반 수행 평가를 결합함으로써 종합적 성취도를 평가하고 있다. 이처럼 단원 평가의 도입은 ‘평가를 어떻게 학력 향상으로 연결할 것인가?’라는 교육의 본질적인 목표에 도달하기 위한 하나의 시도라 할 수 있겠다.

#### 다. 단원 평가의 한계

하지만 단원 평가 또한 분명한 한계점이 존재한다. 단원 평가는 각 학급의 수업 진행 속도에 맞추어 시행되기 때문에 평가 시기가 학급 간에 서로 다를 수 있다는 점이 가장 큰 단점으로 뽑힌다. 즉 단원 평가를 앞서 시행한 학급에서 아직 평가를 보지 않은 다른 학급으로 평가 문제를 유출할 가능성이 있기 때문에 공정성에 문제가 제기될 수 있는 것이다. 따라서 교사와 학부모는 단원 평가가 자신의 학습 도달도를 확인하고 이후 학습 개선에 활용되는 것임을 학생들에게 지속적으로 이해시키는 노력이 필요하다.

또한 자신의 상대적인 학력을 알고 싶어하는 경우에도 단원 평가는 정기 평가와 달리 전체적으로 비교가능한 점수를 확인하기 어렵다는 단점이 있다. 더불어 단원 평가는 단원의 수만큼 시행되기 때문에 정기 평가보다 평가 횟수가 많은 만큼 평가 문제의 출제, 채점, 성적 집계 등 교사의 업무 부담이 증가한다는 과제 또한 존재한다.

### 3. 다양한 방식의 평가 도입

#### 가. 전국 학력·학습 실태조사와 CBT 도입

일본은 문부과학성을 중심으로 매년 ‘전국 학력·학습 실태조사’<sup>13)</sup>를 시행하고 있다. 이 조사는 초등학교 6학년, 중학교 3학년 학생들을 대상으로 국어, 수학(또는 영어 등)의 학습 상황과 생활 실태를 진단하는 전국 단위의 진단 평가이다.

그런데 2023년부터 전국 학력·학습 실태조사에서 CBT(Computer-Based Testing) 도입이 추진됨에 따라 학생들은 개별 기기를 이용해 온라인으로 시험에 응시하게 되었는데, 이러한 CBT 도입을 통한 평가 실시에 따라 다음과 같은 변화 및 장점이 발생했다.

먼저 반응형 문항 구성으로 학생의 응답에 따라 난이도가 조절되는 적응형 구조의 진단 평가가 이루어지게 되었으며, 자동 채점 및 데이터 수집이 가능해 학생의 응답에 대한 빠른 피드백과 빅데이터 기반 학력 분석이 가능해졌다. 실제로 전국 학력·학습 실태조사 중 서술형 문항에 대한 AI 활용 자동 채점 방식이 현재 시험 적용 중인 상황이다. 이 외에도 장애학생들을 위한 확대 문자 제공이 가능하며 CBT를 이용한 학생들의 디지털 활용 능력 평가 또한 이루어지는 등 CBT 평가 방식의 다양한 장점이 보고되고 있는 상황이다.

참고로 전국 학력·학습 실태조사에서의 CBT 평가 방식은 2024년 129개 학교가 참여하였으며, 2027년까지 전국 확대를 목표로 하고 있다.

#### 나. 등교 거부(不登校) 학생을 위한 평가

문부과학성에 의하면 일본의 등교 거부<sup>14)</sup> 학생 수는 2024년 기준 약 34만 6천명으로 이러한 현상은 심각한 사회 문제로 대두되고 있다. 이에 문부과학성은 그 누구도 배움에서 소외되지 않도록 ‘COCOLO 플랜’을 수립하여 다양한 정책을 추진하고 있다.

특히 등교 거부 학생이 출석하지 않는 동안에도 학습이 지속될 수 있도록 다양한 평가 방안을 제시하고 있는데 몇 가지 주요 내용을 살펴보면 다음과 같다. 먼저 문부과학성은 등교 거부 학생이 비대면으로 실시간 수업 참여 또는 과제 제출 등을 할 수 있도록 ‘온라인 수업 출석 인정’ 방안

13) 2023년 11월 기사(일본의 기초학력 보장방안) 참고

14) 2024년 10월 기사(일본의 대안교육) 참고

을 마련함으로써 학교 외의 공간에서도 출석 및 성취도 평가가 이루어질 수 있는 여건을 마련했다. 또한 대면 수업이 어려운 학생들을 위한 다양한 학습 공간 제공을 위해 ‘교육지원센터 및 특례교 설립’ 방안을 제시하기도 했으며, 디지털 과제와 학습 로그를 바탕으로 개별 평가가 가능하도록 ‘1인 1기기 정책과 연계한 학습 추적’ 방안을 제시하는 한편, 일정 기간 내 평가 면제 또는 대체 평가를 실시할 수 있도록 ‘정서적 요인을 고려한 평가 유예 제도’를 실시하고 있다.

이렇듯 등교 거부 학생을 위한 포용적 평가 정책은 학생의 심리적 안정과 학습 지속성을 동시에 확보함은 물론 ‘출석이 평가의 전제가 되지 않는’ 시스템을 적절히 구현하고 있다는 평가를 받고 있다. 게다가 일부 학교에서는 학생의 학습기록을 포트폴리오로 축적하여 평가하는 사례도 나타나고 있는데, 이는 평가의 목적을 위해 유연한 평가 방식을 취하고 있는 일본의 정책을 잘 보여주고 있다.

#### 4. 맷음말

일본은 ‘평가를 통한 학생 성장’이라는 기조 아래 정기 평가 중심에서 벗어나 개별화되고 유연한 평가 방식을 실현하기 위해 다양한 노력을 이어가고 있다. 그런 의미에서 ‘관점별 평가’, ‘단원 평가와 수행 평가의 결합’, ‘디지털 기반 진단 평가의 확산’, ‘등교 거부 학생을 위한 포용적 평가 정책’ 등은 모두 학습자 중심의 교육을 실현하기 위한 노력의 일환이자, 결과 중심이 아닌 성장 중심의 평가 문화를 조성하고자 하는 시도로 볼 수 있겠다.

이처럼 일본의 학생 평가는 단기적인 정책이 아닌 교육의 질적 전환을 위한 구조적 개혁으로 작용할 것으로 보이며, 이를 위해 평가와 관련된 교사와 학교의 전문성 및 역량 향상을 위한 노력 또한 꾸준히 병행되고 있다.

#### 【참고 자료】

- ▶ 文部科学省, [https://www.mext.go.jp/content/20230418-mxt\\_jidou02-0\\_00028870-cc.pdf](https://www.mext.go.jp/content/20230418-mxt_jidou02-0_00028870-cc.pdf)
- ▶ 埼玉県教育局, <https://www.pref.saitama.lg.jp/documents/28045/gakkyuukeiei7-4.pdf>
- ▶ Tokyo MX News, <https://www.youtube.com/watch?v=YAB4QXr-Et4>
- ▶ VIEW next online, <https://view-next.benesse.jp/view/bkn-hs/article28211/>
- ▶ Yahoo Japan ニュース, <https://news.yahoo.co.jp/articles/26453b6472139e2440a21b14c55052bdc4744b28>