

2024 - 종 단 - 06

인천교육-2024-0273

# 인천교육종단연구 기반연구 교과역량 평가방안 제안

연구책임자 **최윤정** (이화여자대학교 부교수)

공동연구원 **곽예린** (이화여자대학교 강사)

연구보조원 **곽소망** (이화여자대학교 박사과정)

**우예진** (이화여자대학교 석사과정)



인천광역시교육청 정책기획조정관  
교육정책연구소

※ 이 연구는 인천광역시교육청 교육정책연구소의 연구지원비로 수행되었으며, 본 연구에서 제시된 정책 대안이나 의견 등은 인천광역시교육청의 공식 의견이 아니라 본 연구진(팀)의 견해를 밝힙니다.

※ 이 연구보고서 파일은 다음의 주소에서 다운받을 수 있습니다.

(<https://ice.go.kr> - 교육 - 교직원지원 - 교육정책연구소)

## 연구 요약

### 1. 서론

#### 가. 연구의 필요성 및 목적

본 연구는 2024 인천교육종단연구 기반연구의 일환으로, 교과역량 평가 방안을 제안하기 위해 진행되었다. **본 연구의 목적을 달성하기 위해 연구 문제를 다음과 같이 설정하였다.**

- 첫째, 국내 교육종단연구에는 어떤 사례가 있으며, 어떻게 시행되고 있는가?
- 둘째, 국외 교육종단연구에는 어떤 사례가 있으며, 어떻게 시행되고 있는가?
- 셋째, 기존 학업성취도 평가 결과의 활용 가능성은 어떠한가, 기존 종단연구에서는 어떻게 활용되고 있는가?
- 넷째, 교과역량 평가는 학교급에 따라 어떠한 차이를 지녀야 하는가?
- 다섯째, 교과역량 평가를 위한 시행 방식(온라인/오프라인)의 장단점과 실제 사례는 어떠한가?

#### 나. 연구 내용

본 연구는 1) 국내 교육종단연구 사례, 2) 국외 교육종단연구 사례, 3) 기존 학업성취도 평가 활용 가능성 4) 학교급별 교과역량 평가 방안, 5) 교과역량 평가 시행 방식에 대한 분석으로 내용을 구성하였다.

첫 번째 연구 내용은 기존에 시행된 국내 교육종단연구를 분석하였다. 국가 수준과 지역 수준으로 구분하여 본 연구에 적합한 사례를 선정하였으며, 연구 목적, 기간, 대상, 내용, 방법, 패널 유지 전략 등에 대해 제시하고 비교하였다. 또한 보다 실제적인 내용을 파악하고자 시도교육청 교육종단연구 담당자를 섭외하여 전문가 면담을 진행하였다.

두 번째 연구 내용은 국외 교육종단연구 사례로 텍사스 주 STARR 평가와 플로리다 주 FAST 평가를 분석하였다. 평가 개요, 변화, 실시 주기, 대상 학교급 및 학년, 교과목, 시행 방식 등을 소개하고 인천교육종단연구를 위한 시사점을 도출하였다.

세 번째 연구 내용은 기존 학업성취도 평가의 활용 가능성으로, 기존 학업성취도 평가의 대표적인 사례(국가수준 학업성취도 평가, 맞춤형 학업성취도 자율평가, 전국연합학력평가, 성취평가제, 책임교육학년제, 인천 기초학력진단평가)를 선정하여 조사하고, 기존 중단연구에서의 실제 활용 사례를 분석하였다.

네 번째 연구 내용은 학교급별 교과역량 평가 방안으로, 앞선 연구 내용에 기반하여 학교급에 따른 차이점을 중심으로 비교 분석하였다. 인천교육중단연구를 위해 권장될 수 있는 대상 학교급 및 학년, 그리고 대상에 따른 교과역량 평가 실시 방안에 대해 의견을 제시하였다.

다섯 번째 연구 내용은 교과역량 평가 시행 방식으로, 앞선 연구 내용에 기반하여 온라인 방식과 오프라인 방식의 장단점을 중심으로 시사점을 도출하였다. 또한 각 방식을 실시할 경우에 고려해야 할 부분에 대해 제안하였다.

## 다. 연구 방법

본 연구에서는 문헌 분석과 전문가 자문을 활용하였다.

문헌 분석은 기존 국내 중단연구 사례, 국외 중단연구 사례, 기존 학업성취도 평가 결과의 활용 가능성, 학교급별 평가 방안 및 평가 시행 방식 분석을 목적으로 하였다. 기존 교육중단연구 및 학업성취도 평가 주관 기관, 교육부, 시도교육청에서 발행한 보고서와 홈페이지 등에 게시된 자료를 대상으로 하였다.

전문가 자문은 교육중단연구의 시행과 관련된 현실적인 어려움 및 고려사항 등을 파악하는 데 목적이 있다. 시도교육청 교육중단연구 담당자를 섭외하였으며, 대상 학교급·학년·교과목 선정 이유, 교과역량 조사 및 평가 방식 선정 이유, 시행의 어려움이나 실제 현장의 의견 및 요구 사항 등을 다루었다.

## 2. 연구 결과

### 가. 국내 교육중단연구 사례 분석

국내 교육중단연구에서의 교과역량 평가 사례를 분석하기 위해 선행연구 분석 및 전문가 면담을 진행하였다. 연구 결과, 지역 수준 교육중단연구에서 평가 도구, 평가

방법, 평가 방향에 있어 공통점이 나타났다. 교과역량 조사 및 평가를 위한 도구는 설문조사 또는 별도의 검사를 실시하고 있었다. 검사의 경우, 선다형과 서술형 문항이 혼합된 검사 도구를 사용하고 있으며, 수직척도화를 위하여 공통문항 설계를 적용하였다. 평가 방향은 교과지식을 평가하기 위한 학업성취도 검사에서 역량 중심 평가로의 전환을 시도하였다. 주된 차이점은 평가 영역에서 나타났는데, 서울과 경기도는 국어, 영어, 수학을 대상으로 한 반면, 전남에서는 국어, 수학, 사회, 과학을 대상으로 하고 있었다.

#### 나. 국외 교육종단연구 사례 분석

국외 교육종단연구 사례로 텍사스 주의 STARR 검사와 플로리다 주의 FAST 검사는 실시 목적, 대상 학년, 시험 유형, 성취 수준에서 차이를 보였다. 가장 눈에 띄는 차이는 STARR 검사의 경우, FAST 검사에 비해 대상 교과목이 다양하였으며, 고등학교에서는 과목을 더욱 세분화하여 평가하고 있었다. 이는 STARR 검사의 주목적이 학년 진급 및 졸업 자격 평가인 반면, FAST 검사의 주목적은 학업성취도 추적 및 학습 진도 모니터링이기 때문으로 보인다. 두 검사의 유사성은 컴퓨터 기반 검사와 컴퓨터화 능력적응검사로 학업성취도를 평가하며, 특수한 환경의 학생들을 위해 원격 시험이나 지필평가(종이 기반 시험)를 제공하고 있다는 점과 교과역량보다는 교과별 내용 지식을 중심으로 평가한다는 점이다.

#### 다. 기존 학업성취도 평가의 활용 가능성

기존 학업성취도 평가는 실시 목적에 따라 대상 학교급과 학년이 달랐으며, 일반적으로 국어, 영어, 수학을 중심으로 하고 필요에 따라 사회, 과학 등의 교과를 함께 평가하고 있었다. 여러 평가 도구 중에서 국가수준 학업성취도 평가, 맞춤형 학업성취도 자율평가, 기초학력진단평가는 평가를 및 결과 보고 차원에서 역량 중심 교육과정을 반영하려는 노력이 보였으나, 실제로 평가하고 있는 내용은 교과지식 중심인 것으로 나타났다. 또한 기존 교육종단연구에서 초등학교와 중학교의 경우에는 별도의 검사를 시행하는 반면, 고등학교의 경우에는 기존 학업성취도 평가(전국연합학력평가 결과 등)를 활용하는 경향이 있었다. 이때, 일부 교육종단연구에서는 자기보고식 설문이 지니는 한계점인 응답의 신뢰성 문제를 극복하기 위해 교사 대상의 개별 학생 평가를 별도로 시행하여

학생의 응답과 교차 검증하는 방안을 적용하기도 하였다.

#### 라. 학교급별 교과역량 평가 방안

국내 교육종단연구 사례 분석을 통해 초등학교는 7개 연구, 중학교는 6개 연구, 고등학교는 3개 연구에서 별도의 평가를 개발 및 시행하고 있는 것을 확인하였다. 또한 국가 수준 교육종단연구에서는 기존 학업성취도 평가 결과를 활용하는 반면, 지역 수준 연구에서는 자체적인 평가를 개발하여 활용하고 있었다. 국외 교육종단연구 사례 분석 결과, STARR 검사의 경우, 초등학교, 중학교에서는 학년별 학업성취도 파악 및 학년 진급을 위한 표준화 검사로 활용되고 있었고, 고등학교에서는 과목별 EOC 시험으로 구성되어 과목별 심화 지식을 평가하고, 그 결과는 졸업 요건으로 활용되었다. FAST 검사의 경우, 초등학교와 중학교에서는 읽기 및 수학 성취도 추적을 위해 연 3회 시행되고 있었으며, 고등학교에서는 FAST 검사가 아닌 EOC 시험을 통해 주요 과목에 대한 심화 지식을 평가하고 졸업 요건으로 활용하고 있었다.

#### 마. 교과역량 평가 시행 방식

교과역량 평가 시행 방식과 관련하여 국내 교육종단연구 사례 분석 결과는 다음과 같다. 온라인 방식을 적용한 교과역량 평가 사례는 한국 초·중등 학교교육 성과 종단연구가 유일하였으며, 서울교육종단연구에서는 코로나19 상황으로 인해 학부모 설문에서 온라인 조사 방식을 활용하였다. 전문가 면담 결과, 온라인으로 교과역량 평가를 시행하는 것에 대해서는 다소 회의적인 의견이 있었다. 국외 교육종단연구 사례에서 텍사스 주 STARR 검사는 컴퓨터 기반 검사, 플로리다 주 FAST 검사 및 EOC 검사는 컴퓨터화 능력적응검사를 시행한다. 또한 컴퓨터를 이용한 검사와 태블릿 PC를 이용한 검사를 모두 허용하고 있었으며, 효율적인 시험 관리, 검사 보안 유지, 접근성 확대, 즉각적인 채점 및 피드백 제공이라는 장점을 적절히 활용하고 있는 것으로 나타났다.

### 3. 연구 제언

#### 가. 국내 교육종단연구 사례 분석

국내 교육종단연구 사례 분석을 통해 도출한 결론 및 제언은 다음과 같다. 평가 대상 교과는 시간과 비용, 평가 부담 절감의 측면에서 반드시 조사가 필요한 소수의 교과목으로 한정할 필요가 있다. 이에 본 연구에서는 주요 과목인 국어, 영어, 수학 중에서도 선행학습의 영향을 받기 쉬운 영어를 제외한 국어와 수학을 평가 대상 교과목으로 제안하고자 한다.

## 나. 국외 교육종단연구 사례 분석

국외 교육종단연구 사례 분석을 통해 다음과 같은 결론 및 제언을 도출하였다. 첫째, 텍사스 주의 STAAR 검사와 플로리다 주의 FAST 검사는 전수 조사 방식으로 시행되고 있으나, 이는 두 검사가 졸업 또는 진급 요건으로 반영되며, 참여가 의무화되어 있기에 가능한 것이었다. 따라서 전수 조사 방식이 갖는 장점에도 불구하고, 인천교육종단연구는 한국의 상황을 고려하여 표집 조사 방식으로 시행하여야 한다. 둘째, 역량 중심 교육과정과 긴밀하게 연계된 평가 도구를 개발하여야 한다. 앞선 두 검사에서는 전통적인 교과지식 중심의 평가를 진행하고 있었으나, 인천교육종단연구는 2022 개정 교육과정에서 추구하는 교육 목표를 반영하여 교과역량을 충실히 측정할 필요가 있다.

## 다. 기존 학업성취도 평가의 활용 가능성

기존 학업성취도 평가 결과의 활용 가능성에 대한 연구 결과로 도출된 결론과 제언은 다음과 같다. 기존 학업성취도 평가 결과를 활용할 경우의 장점은 시간과 비용, 평가 부담의 절감이다. 그러나 학생들의 실제 성적 자료를 수집하는 데에는 현실적인 한계가 있어, 기존 종단연구에서는 자기보고식 설문을 통해 학생이 성적에 대해 스스로 판단하고 응답하게끔 하였다. 그러나 자기보고식 설문을 통한 조사는 응답의 신뢰성 문제를 무시할 수 없으며, 수직척도화를 통한 학년 및 학교급 간 연계에 어려움이 있다. 따라서 본 연구에서 제언하고자 하는 바는 대상 학년이나 교과를 제한해서라도 종단연구의 목적에 맞는 별도의 검사를 개발 및 시행하는 것이다.

## 라. 학교급별 교과역량 평가 방안

학교급별 교과역량 평가 방안 연구에서 도출한 결론 및 시사점은 국내 교육종단연구, 국외 교육종단연구, 기존 학업성취도 평가 도구의 측면으로 구분해서 살펴볼 수 있다.

국내 교육중단연구 사례 분석 결과, 교과역량 조사 및 평가를 위해 학교급에 따른 특성을 적절히 반영하여야 한다. 고등학교의 경우에는 입시 등으로 인해 평가 부담이 클 수 있으므로 별도의 평가를 실시하는 대신 교과역량을 조사하는 방안을 고려하여야 한다. 국외 교육중단연구 사례 분석 결과에서는 미국과는 다른 한국의 상황(대학 입시 문제 등)을 고려하여 초등학교와 중학교에 한해 별도의 검사를 개발 및 시행할 것을 제안하였다. 또한 국외 사례에서도 기초 학습 능력으로 강조되고 있는 언어력, 수리력을 측정하기 위해 국어와 수학을 대상으로 해야 한다. 기존 학업성취도 평가 활용에 대한 연구에 따르면 초등학교와 중학교는 중단연구의 목적을 반영한 별도의 검사를 시행하는 것이 가장 바람직하나, 고등학교의 경우, 성취평가제에 따른 내신 성적이나 전국연합 학력평가 결과를 조사하는 등 학생과 교사의 부담을 줄이려는 노력이 필요하다.

#### 마. 교과역량 평가 시행 방식

교과역량 평가 시행 방안에 대해서는 국내 교육중단연구와 국외 교육중단연구 사례 분석을 통해 결론 및 시사점을 도출하였다. 국내 교육중단연구 사례 분석 결과에 따르면 교과역량을 제외한 내용들은 자료 수집의 용이성 등의 측면에서 온라인 조사 방식이 선호될 수 있다. 그러나 교과역량 평가는 온라인 환경에서의 응답률 감소와 불성실 응답 증가, 학교 간 인프라 격차, 검사 도구 개발 비용과 같은 문제가 일정 수준 이상 해결되기 전까지는 오프라인 평가 방식(지필평가)을 적용하고, 추후 온라인 평가 방식을 도입할 필요가 있다. 국외 교육중단연구 사례 분석 결과에 따르면 컴퓨터 기반 검사를 제안하는 바이나, 접근성 확대(장애 학생 지원), 문항 유형 다양화, 즉각적인 피드백 제공과 같이 지필평가와 차별화되는 온라인 평가 방식만의 장점을 충분히 살릴 수 있어야 한다. 그리고 온라인 평가를 시행하더라도 피치 못한 상황(기술적 제약 등)에 대비하여 오프라인 평가를 함께 준비하여야 한다. 또한 안정적인 온라인 평가 환경 구축을 위한 제도적 지원이 요구되며, 장기적인 관점에서 컴퓨터 기반 검사가 정착된 이후에는 컴퓨터화 능력적응검사까지 나아갈 수 있어야 한다.



# 목차 Contents

<b>I</b>	<b>서론</b>	<b>03</b>
	1. 연구의 필요성 및 목적 .....	3
	2. 연구 내용 및 방법 .....	4
<b>II</b>	<b>국내 교육종단연구 사례 분석</b>	<b>10</b>
	1. 국가 수준 교육종단연구 .....	10
	2. 지역 수준 교육종단연구 .....	24
	3. 요약 .....	41
<b>III</b>	<b>국외 교육종단연구 사례 분석</b>	<b>47</b>
	1. 텍사스 주 STAAR 시험 .....	47
	2. 플로리다 주 FAST 검사 .....	59
	3. 요약 .....	69
<b>IV</b>	<b>기존 학업성취도 평가의 활용 가능성</b>	<b>76</b>
	1. 학업성취도 평가 소개 .....	76
	2. 기존 학업성취도 평가의 종단연구 활용 사례 .....	88
	3. 요약 .....	91
<b>V</b>	<b>학교급별 교과역량 평가 방안</b>	<b>96</b>
	1. 국내 교육종단연구 학교급별 교과역량 평가 방안 .....	96
	2. 국외 교육종단연구 학교급별 교과역량 평가 방안 .....	103
	3. 기존 학업성취도 평가의 학교급별 활용 방안 .....	112
<b>VI</b>	<b>교과역량 평가 시행 방식</b>	<b>118</b>
	1. 국내 교육종단연구의 교과역량 평가 시행 방식 .....	118
	2. 국외 교육종단연구의 교과역량 평가 시행 방식 .....	119

<b>VII</b>	<b>결론</b>	<b>136</b>
1.	국내 교육종단연구 사례 분석 요약 및 결론 .....	136
2.	국외 교육종단연구 사례 분석 요약 및 결론 .....	138
3.	기존 학업성취도 평가 활용 방안 요약 및 결론 .....	139
4.	학교급별 교과역량 평가 방안 요약 및 결론 .....	140
5.	교과역량 평가 시행 방식 요약 및 결론 .....	144
	<b>참고문헌</b>	<b>148</b>

# 표 목 차 Contents

<표 I -1> 2022 개정 교육과정의 핵심역량 및 교과역량 .....	3
<표 II -1> KYPs 시기별 조사 내용 .....	11
<표 II -2> KCYPs2010 시기별 조사 내용 .....	12
<표 II -3> KCYPs2018 시기별 조사 내용 .....	13
<표 II -4> KEEP I 시기별 조사 내용 .....	15
<표 II -5> KEEP II 시기별 조사 내용 .....	16
<표 II -6> KELS2005 시기별 조사 내용 .....	17
<표 II -7> KELS2013 시기별 조사 내용 .....	20
<표 II -8> 서울학생역량 검사 성취 기준 및 문항 수 .....	29
<표 II -9> 전남교육종단연구 검사 출제 범위 및 문항 수 .....	36
<표 II -10> 시도교육청 면담 개요 .....	39
<표 II -11> 인지 영역 조사 및 평가도구와 영역 .....	42
<표 III-1> 학년 및 교과목별 STAAR 검사 부분점수 문항 유형 .....	50
<표 III-2> 2022-2023학년도 STAAR 시험 응시자 수 .....	53
<표 III-3> 고등학교 이전 STAAR 시험 교과목별 보고 범주 .....	58
<표 III-4> 고등학교 STAAR 시험 교과목별 보고 범주 .....	59
<표 III-5> FAST 읽기 검사 길이 .....	63
<표 III-6> FAST 수학 검사 길이 .....	63
<표 III-7> FAST 검사에 참여한 학생 수 (PM3) .....	63
<표 III-8> FAST 검사 소요 시간 .....	65
<표 III-9> FAST 읽기 검사 텍스트 유형 .....	68
<표 III-10> FAST 읽기 검사의 보고 범주별 평가 비율 .....	68
<표 III-11> FAST 수학 검사의 학년별 보고 범주 및 평가 비율 .....	69
<표 III-12> 텍사스 STAAR과 플로리다 FAST 검사 비교 .....	70
<표 III-13> 텍사스 STAAR 검사와 플로리다 FAST 검사의 교과목별 보고 범주 .....	70
<표 IV-1> 학업성취도 평가 평가틀 .....	76

<표 IV-2> 2016년~2019년 국가수준 학업성취도 평가 개요 .....	78
<표 IV-3> 2024년 국가수준 학업성취도 평가 개요 .....	79
<표 IV-4> 맞춤형 학업성취도 자율평가 개요 .....	82
<표 IV-5> 2024년 전국연합학력평가 개요 .....	85
<표 IV-6> 기존 학업성취도 평가 종류별 개요 .....	91
<표 IV-7> 기존 학업성취도 평가 결과 활용 사례 .....	92
<표 V-1> 국내 교육종단연구 주요 조사 대상 및 인지 영역 내용 .....	96
<표 V-2> KYPs 조사 문항 .....	98
<표 V-3> KYPs2010 조사 문항 .....	99
<표 V-4> KYPs2010 4차년도 조사 문항 .....	100
<표 V-5> KYPs2010 5차년도 조사 문항 .....	100
<표 V-6> KYPs2010 7차년도 조사 문항 .....	101
<표 V-7> KYPs2018 조사 문항 .....	101
<표 V-8> KEEP I 조사 문항 .....	102
<표 V-9> KEEP II 조사 문항 .....	102
<표 V-10> STAAR 수학 시험 학년별 특징 .....	104
<표 V-11> STAAR 읽기 시험 학년별 특징 .....	106
<표 V-12> STAAR 사회 시험 학년별 특징 .....	108
<표 V-13> STAAR 과학 시험 학년별 특징 .....	108
<표 V-14> 학교급별 텍사스 STAAR과 플로리다 FAST 검사 비교 .....	112
<표 V-15> 학교급별 기존 학업성취도 평가 종류 및 결과 활용 사례 .....	113
<표 V-16> 학교급별 기존 학업성취도 평가 결과 활용 사례 .....	114
<표 VI-1> 국내 교육종단연구 조사 방법 .....	118
<표 VI-2> FAST 검사 시스템별 사용 .....	126
<표 VI-3> 온라인 실시 방안 측면 텍사스 STAAR과 플로리다 FAST 검사 비교 .....	131

## 그림 목차 Contents

[그림 III-1] 준비 기준(Readiness Standards)과 지원 기준(Supporting Standards) .....	52
[그림 III-2] FAST 검사 성취 수준별 설명 .....	66
[그림 III-3] FAST 검사별 각 성취 수준의 점수 범주 .....	67
[그림 IV-1] 국가수준 학업성취도 평가 결과표 - 학교 단위 .....	80
[그림 IV-2] 국가수준 학업성취도 평가 결과표 - 학생 단위.....	81
[그림 IV-3] 맞춤형 학업성취도 자율평가 학생 대상 교과별 결과표(국어) .....	83
[그림 IV-4] 맞춤형 학업성취도 자율평가 학생 대상 종합결과표 .....	84
[그림 IV-5] 한국복지패널 아동부가조사 학교생활 학교 성적 문항 .....	89
[그림 IV-6] 한국교육고용패널조사 II 학생 내신 등급 관련 문항.....	90



| 인천교육종단연구 기반연구 - 교과역량 평가 방안 제안 |

# I

## 서론







## 서론

### 1. 연구의 필요성 및 목적

인천교육중단연구의 교과역량 평가는 2022 개정 교육과정의 흐름과 방향성을 충실히 반영하여야 한다. 2022 개정 교육과정 총론(교육부, 2022a)에 따른 교육 목표는 ‘깊이 있는 학습을 통한 핵심역량의 함양’이다. 또한 개별 학생의 진로와 적성, 학습 수준에 적합한 맞춤형 교육을 지향하며, 디지털 역량 강화를 위해 교수·학습 및 평가 체제의 변화를 추구한다. 교과역량이란 교육과정에서 제시하고 있는 핵심역량을 각 교과와 특성에 맞게 재구성한 것을 뜻한다. 2022 개정 교육과정의 핵심역량과 주요 교과목(국어, 수학, 영어)별 역량은 <표 I-1>와 같다.

<표 I-1> 2022 개정 교육과정의 핵심역량 및 교과역량

	교과목	역량
핵심역량		자기관리 역량, 지식정보처리 역량, 창의적사고 역량, 심미적감성 역량, 협력적 소통 역량, 공동체 역량
교과역량	국어	비판적·창의적 사고 역량, 디지털·미디어 역량, 의사소통 역량, 공동체·대인 관계 역량, 문화 향유 역량, 자기 성찰·계발 역량
	수학	문제해결 역량, 추론 역량, 의사소통 역량, 연결 역량, 정보처리 역량
	영어	영어 의사소통 역량

출처: 교육부(2022b, 2022c, 2022d)

교과역량에 대해 학교급에 따른 차이는 없었으며, 국어과는 ‘비판적·창의적 사고, 디지털·미디어, 의사소통, 공동체·대인 관계, 문화 향유, 자기 성찰·계발’ 역량, 수학과는 ‘문제해결, 추론, 의사소통, 연결, 정보처리’ 역량, 영어과는 ‘영어 의사소통’ 역량을 제시하고 있었다. 인천교육중단연구의 교과역량 평가가 단순히 교과지식을 평가하는

것이 아니라 ‘역량’을 평가하기 위해서는, 평가 도구 개발 단계에서 <표 I-1>에 제시된 교과목별 역량을 고려하여 문항에 충실히 반영할 수 있어야 한다.

본 연구의 목적은 인천교육종단연구에서 교과역량을 충실히 평가할 수 있는 방안을 제안하고자 하는 것이며, 연구 목적의 달성을 위해 다섯 가지의 연구 문제를 설정하였다.

첫째, 국내 교육종단연구에는 어떤 사례가 있으며, 어떻게 시행되고 있는가?

둘째, 국외 교육종단연구에는 어떤 사례가 있으며, 어떻게 시행되고 있는가?

셋째, 기존 학업성취도 평가 결과의 활용 가능성은 어떠하며, 기존 종단연구에서는 어떻게 활용되고 있는가?

넷째, 교과역량 평가는 학교급에 따라 어떠한 차이를 지녀야 하는가?

다섯째, 교과역량 평가를 위한 시행 방식(온라인/오프라인)의 장단점과 실제 사례는 어떠한가?

## 2. 연구 내용 및 방법

### 가. 연구 내용

#### 1) 국내 교육종단연구 사례 분석

첫 번째 연구 내용은 기존에 시행되었던 국내 교육종단연구의 사례를 다루었다. 우선 관련 기관 보고서를 포함한 연구 문헌을 탐색하였으며, 국가 수준에서 시행되는 종단연구와 지역 수준에서 시행되는 종단연구 사례를 구분하여 연구 목적, 기간, 대상, 내용, 방법, 패널 유지 전략을 조사하였다. 또한 시도교육청 교육종단연구 담당자를 대상으로 면담을 실시하여 실제 현장의 의견이나 시행의 어려움 등에 대한 실질적인 정보를 수집 및 분석하였다.

#### 2) 국외 교육종단연구 사례 분석

두 번째 연구 내용은 국외 교육종단연구 사례를 분석하여 시사점을 도출하는 것이다. 인천교육종단연구가 지역 단위로 시행되는 연구임을 고려하여, 미국의 주 단위 교육종단연구 중에서 텍사스 주 STARR 평가와 플로리다 주 FAST 평가를 조사하였다. 평가 개요, 변화, 실시 주기, 대상 학교급 및 학년, 교과목, 시행 방식 등에 대해 분석하였으며, 국내 교육 현장과의 차이점에 기반하여 결론을 유도하였다.

### 3) 기존 학업성취도 평가의 활용 가능성

세 번째 연구 내용은 교육종단연구를 위한 교과역량 평가를 별도로 개발 및 시행하는 대신, 기존 학업성취도 평가 결과를 활용할 수 있는지에 대해 분석하였다. 국가수준 학업성취도 평가, 맞춤형 학업성취도 자율평가, 전국연합학력평가, 성취평가제, 책임교육학년제, 인천 기초학력진단평가의 대상 학교급 및 학년, 교과목, 평가 방식 등을 조사하였으며, 기존 교육종단연구에서의 평가 결과 활용 사례를 분석하였다.

### 4) 학교급별 교과역량 평가 방안

네 번째 연구 내용은 학교급별 교과역량 평가 방안을 실제 사례에 기반하여 확인하고 인천교육종단연구를 위한 평가 방안을 제안하였다. 앞의 국내 교육종단연구 사례, 국외 교육종단연구 사례, 기존 학업성취도 평가 활용 사례에 대한 분석 결과를 토대로 학교급에 따라 어떤 차이점을 두고 있으며, 인천교육종단연구에서는 어떤 방식이 권장될 수 있는지 의견을 제시하였다.

### 5) 교과역량 평가 시행 방식

다섯 번째 연구 내용은 국내 교육종단연구 사례, 국외 교육종단연구 사례, 기존 학업성취도 평가 활용 사례에 대한 분석 결과에 근거하여 교과역량 평가 시행 방식으로 온라인 평가와 오프라인 평가를 비교하였다. 앞의 문헌연구 결과와 전문가 면담 결과를 종합하여 온라인 평가 방식과 오프라인 평가 방식의 장단점을 분석하였으며, 각 방식을 시행할 경우에 고려해야 할 점을 구체적으로 제시하였다.

## 나. 연구 방법

### 1) 문헌 분석

문헌 분석은 기존에 시행되고 있는 국내 및 해외 교육중단연구 사례, 기존 학업성취도 평가 결과 활용 가능성, 학교급별 평가 방안, 평가 시행 방식을 파악하고자 실시하였다. 교육중단연구 및 학업성취도 평가를 주관하는 기관에서 발행한 보고서와 교육부 및 시도교육청 홈페이지 등을 주된 분석 대상으로 하였다.

### 2) 전문가 자문

전문가 자문은 교육중단연구에 대한 실질적인 시사점을 도출하는 데 주목적이 있으며, 연구진 논의를 통해 서울, 경기, 전남 교육청의 교육중단연구 담당자를 섭외하였다. 원활한 면담 진행을 위하여 사전에 답변을 준비할 수 있도록 질문지를 배포하였으며, 질문지 내용은 최대한 문헌을 통해서는 파악하기 힘든 내용을 중심으로 구성하였다. 구체적으로는 대상 학교급·학년·교과목 선정 이유, 교과역량 조사 및 평가 방식 선정 이유, 평가 부담 등 시행에 있어서 어려움이나 실제 현장의 의견 및 요구 등을 포함하였다.

| 인천교육종단연구 기반연구 - 교과역량 평가 방안 제안 |

## Ⅱ

# 국내 교육종단연구 사례 분석





## 국내 교육중단연구 사례 분석

### 1. 국가 수준 교육중단연구

본 장에서는 국가 수준에서 진행되고 있는 교육중단연구 중에서도 한국청소년정책연구원(NYPI), 한국직업능력연구원(KRIVET), 한국교육개발원(KEDI), 한국교육과정평가원(KICE)에서 수행하고 있는 사례를 기관별로 연구(조사)의 목적, 기간, 대상, 내용, 방법, 패널 유지 전략 등을 중심으로 살펴보았다. 특히 인지 영역에서는 어떤 방법을 통해 어떤 내용을 수집 및 측정하고 있는지 검토하였다.

#### 가. 한국청소년정책연구원

##### 1) 한국청소년패널조사

한국청소년패널조사(Korean Youth Panel Survey, KYPS)는 청소년기의 진로 및 사회화 관련 항목을 조사하여 청소년기에 대한 이해를 증진하고 관련 정책 및 연구 발전에 기여하고자 하는 목적으로 실시되었다.

조사 기간은 2003년부터 2008년까지 6개년에 걸쳐 진행되었다. 조사 대상은 두 개의 패널로 구성되었는데, 중2 패널은 2003년 기준 중학교 2학년 3,449명과 그 학부모였고, 초4 패널은 2004년 기준 초등학교 4학년 2,844명과 그 학부모였다.

주요 조사 항목은 청소년과 학부모로 나누어 실시되었다. 청소년 대상으로는 인적사항, 직업선택 및 향후 진로설정, 진로준비(성적), 여가 활동, 일탈 행동, 생활영역별 시간배분 및 중요도, 자아관 등이 조사되었다. 학부모 대상으로는 가족구성 형태, 부모의 직업 및 근로형태, 월평균 가구 소득, 월평균 사교육비, 주거 형태 등이 조사되었다.

조사 방법은 청소년의 경우 대인면접조사 방식의 설문조사를 실시하였고, 학부모의 경우 전화면접조사를 실시하였다.

패널 유지를 위해 다양한 전략을 사용하였다. 상시관리 프로그램으로는 정기적 접촉 유지(우편물, 전화, 이메일), 패널관리용 홈페이지 운영 및 경품 행사, social network analysis를 활용한 거점 관리 등이 있었다. 또한 이탈위험패널을 위한 특별관리 프로그램

램으로 접촉 빈도 확대, 거점 및 친구를 통한 근황 파악, 추가 보상 제공 등을 실시하였다.

인지 영역은 학생 설문지를 통해 이루어졌으며, <표 II-1>과 같이 시기별 조사 내용을 구성되었다. 지난 학기 반 성적, 전교 성적, 전국 모의고사 성적을 직접 기입하게 하였다. 또한 과목별 지난 학기 반 성적을 5점 리커트 척도로 조사하였는데, 대상 과목은 국어, 영어, 수학, 사회영역, 과학영역, 음악, 미술, 체육이었다.

<표 II-1> KYPS 시기별 조사 내용

		2003	2004	2005	2006	2007	2008
중 2	지난 학기 반 성적 (반 전체 인원, 반 등수, 반 성적 비율)	○	○	○	○	○	○
	지난 학기 전교 성적 (전교 전체 인원, 전교 등수, 전교 성적 비율)	○	○	○	○	○	○
	가장 최근의 전국규모 모의고사 성적 비율	○	○	○	○	○	○
	과목별 성적 주관적 평가	○	○	○	○	○	○
초 4	지난 학기 반 성적 주관적 평가	-	○	○	○	○	○

출처: 박상현 외(2018)의 표 형식을 참고하여 재구성.

## 2) 한국아동·청소년패널조사2010

한국아동·청소년패널조사2010(Korean Children and Youth Panel Survey, KCYPS2010)의 주요 목적은 아동·청소년의 성장과 발달 과정을 종단적으로 추적하고 각 시기의 변화를 횡단적으로 관찰하여, 시간의 흐름 속에서 아동·청소년의 성장을 다각적으로 이해하고 정책 수립을 위한 기초자료를 제공하는 것이다.

조사 기간은 2010년부터 2016년까지 7년간 진행되었다. 조사 대상은 3개 중다패널로 구성되어 2010년 기준 초등학교 1학년 2,342명, 초등학교 4학년 2,378명, 중학교 1학년 2,351명을 유효표본으로 하였다.

주요 조사 항목은 개인발달(신체, 지적, 사회·정서 발달, 진로계획, 비행, 생활시간)과 발달환경(가정, 또래, 교육, 지역사회, 매체, 활동·문화)으로 나뉘었다.

조사 방법은 청소년 대상 개별면접조사와 보호자 대상 전화조사 또는 청소년(자녀)을



통한 유치조사를 병행하였다.

표본 유지를 위해 전화접촉, 경품행사, 뉴스레터 발송, 패널조사 홈페이지 구축, 이탈 위험패널 특별관리 등의 방법을 활용하였다.

인지 영역에서 지적 발달의 한 측면으로 성적을 중요한 요소로 보았으며, <표 II-2>와 같이 시기별 조사 내용을 구성되었다. 주요 과목뿐만 아니라 학교에서 공통적으로 운영되는 모든 과목을 포함하여 매년 조사되었다. 다만 성적은 응답자들이 민감하게 받아들일 수 있는 주제이므로, 2차년도부터 본격적으로 조사에 포함되었다. 서울 지역 초등학교에서 중간·기말고사가 폐지되고, 중학교에서도 성적 결과를 석차로 알리지 않게 되면서, 성적 질문 방식은 패널의 학년별로 달라졌다.

<표 II-2> KCYPS2010 시기별 조사 내용

		2010	2011	2012	2013	2014	2015	2016
초1	과목별 평가 여부	-	-	-	-	-	-	○
	과목별 점수	-	-	-	-	-	-	○
	과목별 주관적 평가	-	○	○	○	○	○	○
	전체 성적 만족도	-	-	-	○	-	○	○
	전체 성적 주관적 평가	-	-	-	-	-	-	○
초4	과목별 평가 여부	-	-	-	○	○	○	-
	과목별 점수	-	-	-	○	○	○	-
	과목별 주관적 평가	-	○	○	○	○	○	○
	전체 성적 만족도	-	○	○	○	○	○	○
	전체 성적 주관적 평가	-	-	-	-	-	-	○
중1	과목별 평가 여부	-	○	○	○	○	○	-
	과목별 점수(등급)	-	○	○	○	-	-	-
	과목별 주관적 평가	-	○	○	○	○	○	-
	전체 성적 만족도	-	○	○	○	○	○	-
	전체 성적 주관적 평가	-	-	-	-	○	○	-

출처: 박상현 외(2018)의 표 형식을 참고하여 재구성.

### 3) 한국아동·청소년패널조사2018

한국아동·청소년패널조사2018(Korean Children and Youth Panel Survey, KCYPS2018)의 주요 목적은 아동·청소년들의 성장과 발달의 변화를 종합적으로 파악할 수 있는 패널데이터를 구축하여 아동·청소년 관련 정책 수립 및 학술 연구를 위한 기초자료를 제공하는 것이다.

조사 기간은 2018년부터 시작되어 종료 시점을 정하지 않고 현재까지 지속되고 있다. 조사 대상은 2018년 기준 초등학교 4학년 2,607명과 중학교 1학년 2,590명의 두 개 코호트로, 보호자와 형제자매도 포함되었다.

주요 조사 항목은 아동·청소년의 개인발달(생활시간, 지적발달, 진로, 사회/정서/역량 발달, 비행, 건강, 신체발달 등)과 발달환경(가정, 학교, 친구, 매체, 활동/문화 환경 등), 그리고 보호자의 배경변인과 특성으로 구성되었다.

조사 방법은 주로 태블릿 PC를 활용한 개별 면접조사(Tablet Assisted Personal Interviewing, TAPI)를 사용했으며, 2020년에는 코로나19 상황에 따라 온라인 조사도 병행하였다.

표본 유지를 위해 패널 위촉장을 발급하여 패널로 참여하는 것에 대한 자긍심을 고취시키고 공신력 제고 효과를 얻고자 하였으며, 개인정보보호 안내문을 제작하여 조사 안내문과 함께 발송하였다. 아울러, 조사 참여 동의서(법정대리인 동의서), 비밀보장 각서, 조사 협조 공문, 조사원 신분증 및 명함, 답례품 영수증 등 본조사 시작 전 성공적인 실사를 위해 관련 자료들을 준비하고 예상 질문과 응답 예시, 조사 거절 사유별 대응 방안을 마련하여 조사를 진행하였다.

인지 영역에서 지적 발달의 한 측면으로 학업성취 항목과 관련하여 <표 II-3>과 같이 전 교과 성취도에 대한 주관적 평가와 학업성취 만족도를 5점 리커트 척도로 측정하였다.

<표 II-3> KCYPS2018 시기별 조사 내용

		2018	2019	2020	2021	2022	2023
초4, 중1	전 교과 성취도의 주관적 평가	○	○	○	○	○	○
	학업성취 만족도	○	○	○	○	○	○

출처: 박상현 외(2018)의 표 형식을 참고하여 재구성.

## 나. 한국직업능력연구원

## 1) 한국교육고용패널조사 I

한국교육고용패널조사 I (Korean Education and Employment Panel, KEEP I)은 국가 인적자원개발 정책 수립을 위한 기초자료 수집, 국가 교육정책 수립을 위한 기초자료 제공, 그리고 노동시장 정책 수립을 위한 기초자료 제공을 목적으로 하였다. 이 연구는 교육에서 노동시장으로의 이행 과정을 파악하고, 교육의 중장기적 성과를 분석하며, 인적자원개발 및 고용 정책의 효과를 검증하고자 하였다.

조사 기간은 2004년부터 2015년까지 12년간이었다. 여기서는 학교교육 기간에 한정하여 관련 내용을 정리하였다. 조사 대상은 기본적으로 2004년 1차 조사 당시의 중학교 3학년생 2,000명, 일반계 고등학교 3학년생 2,000명, 전문계 고등학교 3학년생 2,000명으로 총 6,000명의 학생이었다. 또한 학생의 보호자(가구) 6,000명, 학교 행정가 300명, 담임교사 1,112명도 포함되었다. 4차년도와 7차년도에 각각 신규 패널과 보정 패널이 추가되어, 최종적으로 12차년도까지 총 8,257명을 대상으로 조사가 진행되었다. 신규 패널의 표집은 일반계고 진학, 전문계고 진학 등 시간에 따른 표본의 특성 변화와 표본 탈락 등을 감안하여 설계되었는데, 전문계고 진학 비율이 30%로 향후 전문계고 진학자의 학교 특성별(공업계, 상업계, 기타 등) 진학을 분석하고 특목고(과학고와 외국어고) 표집이 필요하여 1,500명을 추가 표집하였다. 보정패널은 남성의 군입대 등으로 인해 노동시장의 성별 비교분석에 주는 어려움을 해소하고자 비교 코호트로 표집하였는데 전문대학을 졸업한 남성 337명과 4년제 대학을 졸업한 남성 420명을 표집하였다.

조사 방법은 한국직업능력개발원에서 자체 개발한 설문조사 솔루션(Krivet Information Survey System, KISS)을 기반으로 한 PDA와 노트북을 이용한 1:1 대면 조사였다. 1차년도에는 주로 PDA(Personal Digital Assistant)를 이용한 1:1 면접조사를 실시하였고, 2-3차년도에는 PDA 면접조사와 컴퓨터를 이용한 전화조사(Computer Assisted Telephone Interviewing, CATI)를 병행하였다. 4-12차년도에는 노트북을 이용한 면접조사, CATI, 온라인 웹조사 등 다양한 방법을 활용하였다.

조사 내용은 학생의 경우 학교생활, 가정생활, 학업성취, 진로계획, 자아관/직업관 등을 포함하였다. 가구 조사에서는 가족 구성원의 일반적 상황, 교육환경, 경제활동 상태, 가구 소득 및 지출 등을 조사하였다. 학교행정가 및 담임교사 조사에서는 학교의 일반현황과 학생평가 등을 다루었다.

인지 영역은 <표 II-4>와 같이 학생의 학교생활에서 주요 교과목에 대한 흥미도와

성취도를 주관적으로 평가하게 하였고, 학급 석차, 전교 석차, 성적 등급 등의 정보를 수집하였다. 또한 대학수학능력시험 성적 정보도 일부 패널에 대해 수집되었다.

<표 II-4> KEEP I 시기별 조사 내용

		2004	2005	2006	2007	2008 이후
중3, 고3	과목별 흥미도 주관적 평가	○	○	○	○	-
	과목별 성취도 주관적 평가	○	○	○	○	-
	과목별 선호도 주관적 평가	-	-	○	○	-
	학급 석차, 전교 석차 및 등급	-	○	-	-	-
	전체 성적 등급	-	-	○	○	-
	(교사) 전교 석차, 강점 과목, 약점 과목	○	-	-	-	-
	(동의자) 수능 점수	○	○	○	○	○(2008)

출처: 박상현 외(2018)의 표 형식을 참고하여 재구성.

## 2) 한국교육고용패널조사 II

한국교육고용패널 II(Korean Education and Employment Panel, KEEP II)는 국가 인적자원개발 정책, 교육정책, 노동시장 정책 수립을 위한 기초자료를 제공하기 위해 설계되었다. 구체적으로는 교육에서 노동시장으로의 이행 과정을 파악하고, 교육의 중장기적 성과를 분석하며, 인적자원개발 및 고용 정책의 효과를 검증하는 것이 주요 목적이었다.

조사는 2016년부터 시작되어 장기 추적조사로 계획되었다. 조사 대상은 국내 고등학교 2학년 학생 10,558명을 중심으로, 1차년도에는 학생을 둘러싸고 있는 가정 환경과 학교 환경, 담임교사의 특징 등에 대한 자료를 수집하기 위해 학생 설문과는 별도로 가정 설문, 학교 설문, 담임교사 설문을 실시하였으나 2차년도부터는 학생 응답자만 대상으로 하고 가구조사와 학교조사, 담임조사는 실시하지 않았다.

조사 방법은 1차년도에 학생들에 대해 조사원이 직접 학교를 방문하여 집단면접방식으로 전자기기를 활용한 대면 면접조사(Computer Assisted Personal Interview, CAPI)와 종이 설문지를 사용한 대면 면접조사(Paper and Pencil Interview, PAPI)를

실시하였고, 가구(보호자)에 대해서는 종이 설문을 유치하고 수거하였으며, 학교행정가와 담임교사에 대해서는 개별면접으로 PAPI 조사를 실시하였다. 2차년도부터는 노트북을 이용한 일대일 면접조사를 원칙으로 하되, 불가능한 경우 종이 설문지를 병행하였다. 4-6차년도(2020-2022)에는 코로나-19의 영향으로 온라인 조사도 일부 병행되었다.

조사 내용은 1차년도에 학생의 학교생활, 학습 및 사교육, 가정생활, 여가생활, 진로계획 및 직업의식 등을 포함하였고, 학교행정가와 담임교사를 통해 학교 현황과 학생평가 정보도 수집하였다. 2차년도부터는 고등학교 생활, 자격증, 대학생활, 교육훈련, 구직활동, 일자리 특성, 진로계획 등으로 조사 내용이 확장되었다.

인지 영역은 <표 II-5>와 같이 시기별 조사 내용을 구성하였다. 1차년도에 학생의 학교생활에서 학생들의 과목별 주관적 평가와 함께 담임교사가 평가한 전체 성적 등급 정보를 수집하였다. 2-3차년도에는 주요 과목별 내신등급 정보를 수집하여 학업성취도 변화를 추적할 수 있도록 하였다. 또한 학교알리미 데이터를 통해 국가학업성취도평가 결과를 수집하였다.

<표 II-5> KEEP II 시기별 조사 내용

		2016	2017	2018	2019 이후
고2	과목별 흥미도 주관적 평가	○	○	○	-
	과목별 성취도 주관적 평가	○	○	○	-
	과목별 선호도 주관적 평가	○	○	○	-
	(담임) 전체 성적 등급	○	-	-	-
	(학교알리미) 국어, 수학, 영어 국가학업성취도평가	○	○	○	○

출처: 박상현 외(2018)의 표 형식을 참고하여 재구성.

## 다. 한국교육개발원

### 1) 한국교육중단연구2005

한국교육중단연구2005(Korea Education Longitudinal Study, KELS2005)는 학생

들의 성장과 발달 과정을 장기적으로 추적하여 학교교육의 효과를 분석하고, 교육정책 수립을 위한 기초자료를 제공하며, 다양한 교육 관련 연구를 지원하는 것을 목적으로 실시되었다.

조사 기간은 2005년부터 2023년까지 총 3단계로 나누어 진행되었다. 1단계는 2005년부터 2010년까지 6년간 매년, 2단계는 2011년부터 2019년까지 8년간 2년마다, 3단계는 2023년에 1회 실시되었다. 여기서는 중등교육 기간에 해당하는 1단계 조사 내용만 포함하였다. 조사 대상은 2005년 당시 중학교 1학년 학생 6,908명이었다.

주요 조사 항목은 학생, 학부모, 교사, 학교로 나누어 실시되었다. 학생 대상으로는 학업성취도, 자기조절학습능력, 학교생활, 진로계획 등이 조사되었다. 학부모 대상으로는 가정배경, 교육지원, 자녀의 진로 등이 조사되었다. 교사 대상으로는 학생 평가, 교수활동 등이 조사되었다. 학교 대상으로는 학교 특성, 교육과정 운영 등이 조사되었다.

조사 방법은 학생들이 재학 중인 학교를 통한 집단 면접 방식과 개별 면접 조사, 우편 설문 및 전화 설문조사 방법이 병행되었다.

패널 유지 전략으로는 시도교육청 및 학교와의 협력을 통한 재학생 및 전학생 관리, 중도탈락자에 대한 개별 관리, 표본 학생들의 참여도 제고 방안 구안 등이 실시되었다. 또한 정기적인 협의회 개최, 조사 참여에 대한 보상 제공 등을 통해 지속적인 참여를 유도하였다.

인지 영역은 <표 II-6>과 같이 국어, 영어, 수학 영역의 학업성취도 검사를 통해 측정되었다. 다만, 고등학교 진학 이후에는 자체 시험에 대한 부담과 불성실 응답 가능성 등을 고려하여 국가 또는 시도교육청 수준의 성취도 평가 결과를 활용하는 방안을 모색하였다.

<표 II-6> KELS2005 시기별 조사 내용

	2005	2006	2007	2008	2009	2010	2011 이후
학업성취도 검사(국어, 영어, 수학)	○	○	○	-	-	-	-
(담임교사평가지) 전국 수준 점수 수집	-	-	-	-	○	○	-
(DB연계) 수능 성적 수집	-	-	-	-	-	○	○

출처: 박상현 외(2018)의 표 형식을 참고하여 재구성.

학생의 학업성취도 검사지는 국어, 수학, 영어 영역에 대하여 중학교 1, 2, 3학년을 대상으로 한 검사지이며, 학생들의 성취수준의 변화를 측정할 수 있도록 학년별로 동등화 과정을 거쳐서 개발하였다. 성취도 검사는 집단 또는 개별 시행이 가능한 지필 검사로 검사 시간은 정규 교과 시간인 45분이며 문항수로 나타낸 검사의 길이는 국어와 영어 35문항, 수학 25문항이다. 검사 채점 방법은 OMR 채점이다. 구체적인 문항 정보는 <부록 1>에서 <부록 5>에 제시하였다.

PARDUX(Burket, 1996)를 사용하여 교과별로 학년별 척도화작업을 수행하였다. 척도화 작업은 모든 능력 수준의 피검자 정보가 필요하고, 또한 문항반응이론(Item Response Theory, IRT)에 의한 추정 방법이 허용되는 검사 집단의 규모로 학년별로 약 5,000명의 학생을 대상으로 검사를 시행한다. 수집된 검사 문항 점수와 검사점수는 문항반응이론에 의거하여 3모수 IRT model로 문항 특성이 추정되며, PARDUX 프로그램을 적용하여 교과에 따른 학년별 척도화 작업을 수행한다. 척도화 작업이 수행되면, 검사 원점수와 생성된 척도점수와의 조건표를 제작한다.

자료 수집 설계에 따른 검사 자료가 수집되면, 모든 검사 문항들의 특성을 추정하는 작업을 수행한다. 검사 문항의 특성을 추정하기 위한 측정 모형은 2모수 및 3모수 IRT 모형에 의하며, 각 문항의 난이도, 변별도 및 추측도의 문항 특성 정보를 확보한다. 문항 특성 정보가 확보되면, 이에 기초하여 검사 특성곡선을 추정한다. 따라서 총 9종의 검사별로 검사의 난이도와 변별도를 반영하는 특성곡선이 추정된다. 문항과 검사의 특성이 추정되면, 이에 기초하여 피검자의 능력을 추정하는 작업을 수행한다. 문항 형식에 따라서 서답형 문항은 2모수 모형, 선다형은 3모수 모형을 적용하여 능력 모수를 추정한다.

척도화 작업을 위한 Software는 PARDUX computer program이다. 미국 중등학교 대상 검사 개발 기관인 CTB/McGraw Hill이 개발하였으며, 현재 미국 35개주 이상에서 검사개발 및 척도화 작업에 사용된다. 수직 척도화 작업과정에는 척도의 연속성을 확보하기 위한 SMOOTHING 작업과 척도의 최대 가능점수(highest obtainable scale score, HOSE)와 최소 가능 점수(lowest obtainable scale score, LOSE)를 결정하는 보완 작업이 필요하다. 이를 위하여 FLUX, PARMATE, TRUESCORE 등의 측정학적 분석용 컴퓨터 프로그램을 활용한다.

## 2) 한국교육중단연구2013

한국교육종단연구2013(Korea Education Longitudinal Study, KELS2013)은 한국 교육종단연구2005의 제한점을 보완하고, 초·중등 교육에 대한 새로운 종단연구를 구축하는 것을 목적으로 실시되었다. 학생들의 인지적·정의적 발달상황과 가정과 학교에서의 교육경험을 확인하고, 가정과 학교에서의 교육경험이 학생들의 인지적·정의적 발달에 어떠한 영향을 미치는지 파악하는 것을 주요 목적으로 한다. 또한 학교교육의 결과가 단기적으로는 상급학교의 진학에 어떤 영향을 미치며 장기적으로는 학생들의 직업 획득에 어떤 영향을 주는지 파악하는 것을 목적으로 한다.

조사 기간은 2013년부터 2030년까지 총 18년에 걸쳐 두 단계로 진행되었다. 1단계는 2013년부터 2020년까지 8년간 매년, 2단계는 2021년부터 2030년까지 10년간 진행되었다. 2단계에서는 대학교 1~2학년까지는 매년, 이후에는 2년마다 조사가 실시되었다. 여기서는 중등교육 기간에 해당하는 1단계 조사 내용만 포함하였다. 조사 대상은 2013년 당시 전국 242개 초등학교의 5학년 학생 7,312명이었다.

주요 조사 항목은 학생, 학부모, 교사, 학교장, 학교로 나누어 실시되었다. 학생 대상으로는 인지적·비인지적 성취, 자기주도학습태도, 학교생활, 진로 및 포부 등이 조사되었다. 학부모 대상으로는 가정배경, 교육지원, 자녀의 교육경험 등이 조사되었다. 교사 대상으로는 교사 배경, 전문성 신장, 학생 평가 등이 조사되었다. 학교장 대상으로는 학교경영, 교장업무 등이 조사되었다. 학교 대상으로는 학교배경, 교육과정, 교육지원 등이 조사되었다.

조사는 조사 대행기관을 선정하여 시행되었으며, 시도교육청의 협조하에 조사 대상 학교의 연구협력교사를 통해서 우편조사에 의해서 집단조사를 하고, 상담원이 학교를 방문하여 조사지를 검토한 후 송부하는 방법이었다. 도구는 주로 설문지를 통한 조사가 주로 이루어졌으며, 학생의 경우 기초능력검사지, 자기조절학습능력검사지, 학생 질문지 등이 활용되었다.

패널 유지 전략으로는 조사 위탁기관을 선정하여 표본 학생 관리를 맡기고 있지만, 조사 위탁기관이 표본 학생을 직접 관리하는 방식이 아니라 242개 학교의 연구협력교사를 통해 수행하고 있다.

인지 영역은 <표 II-7>과 같이 기초능력검사를 통해 측정되었다. 또한 교사가 작성하는 학생평가지를 통해 학생의 학업성취도가 평가되었다.



<표 II-7> KELS2013 시기별 조사 내용

	2013	2014	2015	2016	2017	2018	2019 이후
초등학교 기초능력검사 (국어, 영어, 수학)	○	○	-	-	-	-	-
중학교 학업성취도검사 (국어, 영어, 수학)	-	-	○	○	○	-	-
(학생설문지) 전국 수준(등급) 수집	-	-	-	-	-	○	○
(담임교사평가지) 전국 수준(등급) 수집	-	-	-	-	-	○	○

출처: 박상현 외(2018)의 표 형식을 참고하여 재구성.

초등학교 기초능력검사는 초등학교 4학년, 5학년, 6학년 학생들을 위한 국어, 영어, 수학교과와 기초능력검사로 검사의 종류는 각 학년별로 3개 교과 기초능력검사가 필요하므로 총 9종이며, 국어와 영어는 듣기검사가 포함된다. 검사 시행방법은 집단 또는 개별 시행이 가능한 지필검사이다. 학생들은 학급 단위로 검사를 치르지 않고, 소규모 단위로 성취도검사를 시행하거나, 개별 시행을 하게 된다. 즉, 검사의 시행조건이 표집된 학교에 따라 달라지므로 본 연구에서 기초능력검사 도구는 집단검사(예, 학급단위, 또는 소규모 집단)와 개별검사가 모두 가능한 표준화된 검사로 개발하였다. 검사의 시행시간은 초등학교의 정규 교과 시간인 40분이며 검사의 길이는 기초능력검사 도구의 문항수로 각 25문항이며 동등화 검사용으로 각 30문항이다. 구체적인 문항 정보는 부록에 제시하였다. 기초능력검사의 채점 방법으로 광학식 OMR 채점방식을 사용하였다.

학년별 척도화 작업을 위하여 다음의 두 가지 작업을 수행하였다. 첫째, 검사점수의 척도화 작업은 피검자의 능력이 계량화된 값이 검사점수이므로, 피검자의 능력수준을 심리적 단일 차원의 연속선(psychological continuum)상에 위치시키는 작업을 수행하였다. 즉, 검사점수를 일정한 규칙에 의하여 능력점수 척도로 전환하는 작업이 요청되었다. 둘째, 능력 점수 척도의 특성이 등간성을 갖도록 전환하는 작업을 수행하였다. 척도화 작업은 이 두 가지 목적을 성취하기 위한 것으로, 교과목에 따라서 학년별로 척도화 작업이 이루어지면, 학년 내에서 학생들 사이의 척도점수의 차이는 성취도의 차이로 해석할 수 있게 되며, 능력 점수의 차이 정보가 정밀하게 된다.

척도화 작업을 위한 문항모수의 추정을 위해 문항반응이론에 의한 추정방법을 사용하였고, 3개 학년의 3개 교과목 기초 능력 검사 총 9개 검사에 포함된 문항들의 문항모

수가 추정되었다. 이를 위해 문항반응이론에 의거하여 2모수 IRT 모형을 적용한 문항 특성을 추정하였으며, PARSCALE(Muraki 외, 1998) 프로그램을 이용하였다. 추정된 문항 모수의 특성을 기본적인 2모수 모형 척도상에서 제시하였고, 모형자료 적합도 지수 등 검사의 특성을 평가하는 자료를 제공하여 개발된 검사의 문항 적합성을 검증하였다.

중학교 학업성취도평가는 중학교 1학년~중학교 3학년 국어, 수학, 영어검사로 영어는 듣기검사가 포함된다. 검사 시행방법은 집단 또는 개별 시행이 가능한 지필 검사이다. 성취도 평가도구는 집단검사(예, 학급단위, 또는 소규모 집단)와 개별검사가 모두 가능한 표준화된 검사이다. 검사 시간은 중학교 정규교과 시간인 45분이며 검사의 길이(문항 수)는 국어 28문항, 영어 26문항, 수학 25문항이다. 검사채점 방법은 광학식 OMR 채점 방식을 사용하였다.

문항반응이론을 적용한 수직척도 개발을 위한 문항 특성 분석에는 2모수 문항반응모형을 적용한 문항모수를 추정하였으며, BILOG-MG(Zimowski 외, 1996) 프로그램이 이용되었다. 한국교육종단연구2005의 중학교 1, 2, 3학년 검사 척도의 수직 연계를 위해서 문항반응이론을 사용하였다(이규민 외, 2006; 2013). 기개발된 초등학교 4, 5, 6학년 수직척도와 연계성을 가정하여, 문항반응이론을 적용하여 중학교 1, 2, 3학년 수직척도를 개발하고, 나아가 학교급 간 연계화를 시도하였다.

2018년도부터 조사대상 학생이 고등학교에 진학함에 따라 별도의 시험 실시가 입시 및 학업에 부담을 줄 수 있다고 판단하고 전국연합학력평가 결과(등급)를 기재하는 문항을 신설하여 학생의 학업성취도를 파악하였다. 아울러, 데이터의 정확성을 높이기 위하여 교사를 대상으로 학생과 동일하게 조사대상 학생의 전국연합학력평가 결과를 조사하는 학생 평가지를 개발하여 실시하였다.

## 라. 한국교육과정평가원(KICE) 한국 초·중등 학교교육 성과 종단연구

한국교육과정평가원에서는 초·중·고등학교의 학교급별 교육성과를 추적할 수 있는 종단연구 기반을 마련하기 위해 종단조사 체제를 구축하고자 하였다. 이에 설계된 종단연구는 한국 초·중등 학교교육 성과 종단연구로 조사 기간은 초등학교 패널 2019년~2028년, 중학교 패널 2021년~2027년, 고등학교 패널 2023년~2026년이다.

조사 대상은 구체적으로 언급되지 않았으나 2019년 초등학교 4학년, 2021년 중학교

1학년, 2023년 고등학교 1학년을 대상으로 진행될 예정이다.

조사 방법은 웹 기반 중단조사 시스템을 활용하고자 하였다. 학생 검사와 배경변인 조사를 실시하여 학교장, 교사, 학생, 학부모를 대상으로 학교 특성, 교육과정, 교수학습 및 평가, 학교 풍토, 학생 활동, 가정 배경 등을 조사하고자 하였다.

인지 영역은 별도로 개발한 학생 검사를 사용하였는데, 학생 검사는 국어, 수학, 문제해결역량, 사회·정서적 역량으로 구성되며, 국어와 수학은 핵심교과지식과 교과역량으로, 문제해결역량은 문제해결역량 중심의 4개 교과의 교과지식을 연계하여 검사 내용을 구성하였다. 국어, 수학, 문제해결역량 영역은 서답형, 선다형 문항을 포함한 대문항(big item) 체제로 2종으로 개발하고 검사지 종류 간, 학년 간 공통문항을 포함하며 문제풀이 맥락의 메타인지, 인지 및 동기조절의 시험전략을 측정하는 자기보고식 문항을 포함하였다.

우선 학생 검사 평가틀의 내용 영역으로서 국어, 사회, 수학, 과학 교과 전문가들을 대상으로 4차례의 델파이 조사를 통해 영역별 성취기준과 국어, 수학 교과역량과 문제해결역량의 하위요소를 선정하였다. 또한 국어는 내용 영역, 국어 교과역량, 국어 자료의 세 차원으로, 수학은 내용 영역과 수학 교과역량의 두 차원으로, 문제해결역량은 핵심교과지식과 문제해결역량의 하위요소 2차원으로 평가틀을 구성하였다. 문항 개발 워크숍에서 충분한 논의를 거쳐 학생 검사 평가틀을 확정하였다.

평가틀에 기반하여 국어, 수학 교과에서는 4, 5학년 각각 20개씩 A, B 2종의 검사를 개발하였다. 문제해결역량 검사 문항은 대문항(big item) 형식으로 구성하여 4학년은 대문항 6개(하위 문항 17개), 5학년은 대문항 7개(하위 문항 수 22개, 4학년 공통 대문항 2개(하위 문항 6개)를 개발하였다. 문항 선제 워크숍에서 문항을 수정하고 선제하여 검사 문항의 완성도를 높이는 과정을 진행하였고 인지실험을 거쳐 다시 수정 및 보완한 후 예비검사 문항으로 최종 확정되었다.

국어 검사는 초등학교급의 학생 검사는 A, B형의 2종으로 제작하였으며, 전체 평균 정답률이 60% 정도가 되도록 난이도를 맞추려 하였다. 문항 개발 시 평가범위는 초등학교 4학년의 경우 4-1, 4-2, 3-1, 3-2학기, 5학년의 경우 4, 5학년 간 공통문항 및 5-1, 5-2학기의 내용을 다루었다.

수학 검사는 델파이 조사 결과에 근거하여 개발된 내용영역과 교과역량의 두 차원으로 구성된 평가틀에 따라 내용영역과 교과역량을 전반적으로 평가할 수 있는 문항을 출제하였다. 문항 유형은 선다형, 단답형, 서술형 3가지로 하여 전체적으로 서답형 문항이 40~60%를 차지하도록 문항을 개발하고 전체 평균 정답률이 60% 정도가 되도록

록 난이도를 맞추려 하였다. 또한 검사 시간 40분 안에 해결할 수 있도록 문항 수를 20~25개 정도 선정할 계획을 하였다. 그리고 교과서 없이 교육과정에 근거하여 문항을 출제했으므로, 예비검사의 문항 수를 충분히 확보하여 본검사에서 좋은 문항이 선정될 수 있도록 학년별로 A형과 B형 2종의 검사도구를 개발하도록 계획하였다. 검사도구당 문항 수는 20~25개를 개발하려 하였고, 델파이 조사 결과로 선정된 성취기준 각각에 대해 2개의 문항을 출제하여 초등학교 4학년은 총 54개, 초등학교 5학년은 총 62개의 문항이 개발되었다. 문항 개발 시 평가범위는 초등학교 4학년의 경우 3-2, 4-1, 4-2학기, 5학년의 경우 4-2, 5-1, 5-2학기 내용을 다루도록 하였다.

10년간 학생의 성취도 변화와 성장을 타당성 있게 분석하는 것을 목표로 삼기 때문에 전체 문항 대비 30%에 해당하는 총 6개의 4, 5학년 간 공통문항을 설정해 놓았다. 또한 양질의 문항을 최종적으로 선별하기 위해 A, B의 2종을 개발하였고, 이들 간에도 8개의 공통문항을 설정해 놓았다.

문제해결역량을 측정하는 문항은 학생들이 실생활에서 접할 수 있는 문제 상황 및 해당 문제를 해결하는 데 필요한 자료와 정보 등을 제시하고, 학생들이 문제를 인식, 파악하고 창의적이고 다양한 방법으로 문제를 해결해 가는 과정을 측정하는 것을 원칙으로 하고자 하였다. 따라서 문제해결역량 검사 문항은 자료와 2~4개의 하위 문항으로 구성된 대문항 형식으로 개발하였다. 문항은 국어, 사회, 수학, 과학 교과의 내용 영역과 문제해결역량을 아울러 평가할 수 있도록 하며 내용 영역은 각 교과의 성취기준을 중심으로 하며 내용 영역의 범위는 해당 학년의 교육내용을 중심으로 하되, 이전 학년의 교육내용도 포함하였다. 각 학년의 문제해결역량 검사 문항 수는 15~20개 정도가 되도록 개발하였고 문항 유형은 선택형 40% 내외, 서답형 60% 내외로 4, 5학년 검사의 공통문항을 개발하였다. 공통문항의 내용 영역은 가능한 한 사회, 과학 교과 교육과정에 제시된 핵심 개념 중 내용 요소가 초3-4학년군, 5-6학년군에 모두 포함되어 있는 핵심 개념에 해당하는 성취기준 중에서 선정하는 것을 원칙으로 하였다.

영역별 구체적인 문항 정보 및 예시문항은 <부록 6>에 제시하였다.

초등학교 4학년 160명, 5학년 133명 총 293명의 학생과 13명의 교사를 대상으로 학생 검사 인지 실험을 시행하였다. 각 학생은 국어, 수학, 문제해결역량의 한 영역을 응시하고 인지 실험 시행 학급 담당 교사는 시행과정에 참여하면서 모니터링 질문지를 바탕으로 점검하였다. 연구진도 시행 과정을 참관하였고 시행 시간 이후 학생 및 교사 대상 인터뷰를 진행하였다. 인지 실험 결과를 통해 문항을 수정 및 보완하였다. 문항을 새롭게 추가하거나 전면 수정한 경우는 없었고 예시답안을 추가하거나 일부 낮은 변별

도를 보인 문항에 한해서 오답지를 수정하였다. 아울러, 검사 시간이나 검사지 형식의 적정성 등에 대한 의견을 수렴하여 별도의 검사준비 시간을 마련하는 방안을 검토하여 예비검사에서 검사 준비 시간을 10분 배정하였다.

한국 초·중등 학교교육 성과 종단연구는 1차 기초연구에서 개발한 국어, 수학, 문제해결역량 학생검사의 평가틀을 웹 기반 검사로 구현할 수 있도록 수정 및 보완하였으며 1차 연도에 개발했던 지필식 문항을 웹 기반 문항으로 일부 전환하였고 새로 웹 기반 기술공학적 기능을 적용한 학생검사 문항을 개발하였다. 이를 바탕으로 1, 2차에 걸쳐 예비검사를 시행하여 학교 정보를 토대로 시행 매뉴얼을 제작하여 각 학교에 송부하고 연구진이 검사 시행 학교를 방문하여 시행의 전 과정을 점검하였다. 두 차례에 걸친 시행을 통해 웹 기반 검사에 대한 학생들의 긍정적 의견과 학교 현장에 대한 적용 가능성을 확인하였으며 학생검사에서 웹 기반 검사로의 전환 방안에 대한 정책적인 결정과 학교교육에서 컴퓨터 활용 능력을 높일 수 있는 교육을 강화할 필요가 있다는 시사점을 얻을 수 있었다.

## 2. 지역 수준 교육종단연구

본 장에서는 지역 수준에서 진행되고 있는 교육종단연구 중에서도 서울특별시교육연구정보원, 경기도교육연구원, 전라남도교육연구정보원에서 수행하고 있는 사례를 기관별로 연구(조사)의 목적, 기간, 대상, 내용, 방법 등을 살펴보았다. 특히 인지 영역에서는 어떤 방법을 통해 어떤 내용을 수집 및 측정하고 있는지 검토하고 각 기관에서 교육종단연구를 담당하고 있는 실무자 대상 면담을 진행하였다.

### 가. 서울특별시교육연구정보원

#### 1) 서울교육종단연구2010

서울교육종단연구(Seoul Education Longitudinal Study, SELS)는 서울특별시교육청 산하 학교에서의 학생들의 삶을 다각적이고 체계적으로 분석하고, 교육정책 및 재정 지원 사업의 성과를 평가하며 피드백을 제공하는 목적으로 실시되었다.

조사 기간은 2010년부터 2018년까지 9년간 진행되었다. 조사 대상은 2010년 당시 초등학교 4학년(5,059명), 중학교 1학년(4,544명), 일반계고 1학년(5,240명), 실업계고 1학년(1,216명) 학생들과 그들의 학부모, 교사, 학교 및 학교장을 포함하였다.

주요 조사 항목은 학생, 학부모, 교사, 학교장, 학교로 나누어 구성되었다. 학생 대상으로는 학교생활, 사회심리적 환경, 방과후학교 및 사교육 참여, 학습시간, 가정생활, 여가활동, ICT, 자기주도적 학습태도, 목표의식, 시민의식, 창의성, 건강 등이 조사되었다. 학부모 대상으로는 가정배경 특성, 학교교육에 대한 지각, 교육비, 자녀에 대한 교육 지원활동 및 기대 등이 조사되었다. 교사 대상으로는 배경 특성, 근무시간, 학교조직문화, 교수-학습 활동, 학생 평가 및 평가결과 활용, ICT 활용 등이 조사되었다. 학교장 대상으로는 배경 특성, 학교운영, 의사결정 권한 행사, 학교내외 구성원과의 관계 등이 조사되었다. 학교 대상으로는 교육과정 편성 운영, 학생평가 및 교과부진학생 지도, 교육시설 및 여건 등이 조사되었다.

조사 방법은 설문지를 통한 조사가 주로 이루어졌으며, 학교정보공시제에 의해 공시된 자료는 향후 연계될 수 있도록 설계되었다.

인지(지적) 발달 부분은 국어, 수학, 영어에 대한 학업성취도 검사를 통해 측정되었다. 이는 지필검사이며 검사 시간은 정규 수업시간과 동일하게 초등학교 4학년 40분, 중학교 1학년 45분, 고등학교 1학년 50분으로 설정되었다. 검사 범위는 해당 학교급의 학년에서 1학기(3월~6월 말) 교육과정에 포함된 내용으로 한정하였다. 문항 형식은 5지 선다형으로 문항 수는 초등학교 4학년 국어 30개, 영어 30개(듣기평가 문항 15개 포함), 수학 25개, 중학교 1학년 국어 35개, 영어 35개(듣기평가 문항 10개 포함), 수학 30개, 고등학교 1학년 국어 35개, 영어 35개(듣기평가 문항 9개 포함), 수학 30개였다. 공통문항의 경우, 각 과목의 검사 문항들 중 7~9개 문항(전체 문항의 약 20~30%)이 출제되며, 이들 문항은 전년도에 출제되었던 문항들로서 매년 실시되는 검사들을 동등화하여 학생들의 과목별 학습 수준의 변화를 파악하기 위해 활용된다.

2013년 성취도 검사는 중학교 1학년 45분, 고등학교 1학년 50분의 검사시간으로, 문항 수는 국어 35개, 수학 30개, 영어 35개(듣기평가 문항 9개)로 구성되었다. 공통문항은 1차년도와 3차년도에 출제되었던 문항들 중 일부를 선정하며, 각 과목마다 13~17개 문항(전체 문항의 약 45~50%)이 포함되었다. 공통문항은 검사동등화를 위해 반드시 필요하며, 이는 이후 학생들의 과목별 학습 수준의 변화를 파악하기 위해 활용된다. 구체적인 검사 세부 영역과 성취기준 및 성취수준은 <부록 7>에 제시하였다.

수직적도 개발을 위하여 인접(adjacent) 학년에 공통문항을 배치하여 분석하는 공통

문항 비동등집단 설계를 사용하였다. 본 연구는 낮은 학년의 문항을 공통문항으로 사용하여 높은 학년에 배치하였다. 공통문항을 배치하는 방법은 낮은 학년의 문항을 높은 학년에 배치하거나 높은 학년의 문항을 낮은 학년에 배치하는 두 가지 방법이 있다. 높은 학년의 문항을 낮은 학년에 배치하는 경우, 낮은 학년 학생들이 아직 배우지 않은 내용의 문항에 응답해야 하는 문제점이 발생하므로 본 연구에서는 낮은 학년의 문항을 높은 학년에 배치하였다.

수직척도 개발을 위한 절차는 6단계로 요약할 수 있다. 1) 모든 학년과 교과목에 대해 독립적으로 문항모수와 피험자의 능력모수를 추정하고, 공통문항의 난이도를 비교분석한다. 2) 공통문항의 적합도를 확인한다. Rasch모형의 내적합도가 1.5보다 높은 문항은 삭제한다. 스튜던트 잔차값을 통해 이상치를 확인하고, 절대값 2보다 큰 경우 해당 문항을 제외한다. 3) 삭제된 문항으로 인해 전체 수직척도 공통문항이 측정하고자 하는 교과내용에 변화가 없는지 확인한다. 4) 인접한 학년의 수직척도 지수를 계산한다. 이는 각 학년에서 공통문항의 난이도 평균 차이로 구한다. 5) 누적 수직척도 지수를 계산한다. 기준학년을 정하고, 학년이 올라갈수록 적합한 수직척도 지수를 적용한다. 6) 최종 결정된 수직척도 지수를 피험자 능력모수에 적용하여 척도점수를 계산하고, 수직척도의 안정성을 평가한다. 학년별 수직척도 지수 변화를 확인한다. 인접 학년 간 수직척도 공통문항의 상관관계를 확인한다. 수직척도화된 피험자 능력모수의 평균과 표준편차의 적절성을 확인한다. 이 절차를 통해 학생들의 학년 간 성취도 발달을 측정할 수 있는 안정적인 수직척도를 개발하였다.

수평척도 개발을 위하여 수직척도 개발과 마찬가지로 수평척도로 개발하고자 하는 동일한 학년에 공통문항을 배치하여 분석하는 공통문항 비동등집단 설계를 사용하였다. 2010년도 문항을 공통문항으로 사용하여 2013년도 검사에 배치하였다. 수평척도 개발을 위한 절차는 4단계로 요약할 수 있다. 1) 동일학년인 2010년과 2013년도 결과를 동일 척도상에 배치하기 위해 두 연도의 검사에 포함된 공통문항의 모수를 독립적으로 추정한다. 2) 공통문항의 안정성을 검토한다. 각 연도의 공통문항 평균을 계산하고, 두 평균의 차를 동등화 상수로 지정한다. 동등화 상수를 2013년도 문항모수에 적용하여 2010년도 문항과의 차이를 비교한다. 차이가 0.3을 넘는 문항은 순차적으로 공통문항에서 제외하고, 모든 문항이 0.3 이하가 될 때까지 이 과정을 반복한다. 3) 최종 결정된 공통문항의 모수를 2010년도 문항모수로 고정하여 2013년도 문항 모수와 피험자 모수를 재추정한다. 이를 통해 두 연도 학생들의 학업성취도를 비교 평가한다. 4) 동등화 과정에서 추정된 피험자의 모수와 척도점수의 최소값, 최대값을 결합하여

일차방정식을 구한 후 최종 수평척도점수를 계산한다. 이 절차를 통해 서로 다른 연도의 동일 학년 학생들의 학업성취도를 비교할 수 있는 수평척도를 개발하였다.

## 2) 서울교육종단연구2020

서울교육종단연구2020(Seoul Education Longitudinal Study, SELS2020)은 서울시 학생들의 학생 역량과 행복감에 영향을 미치는 요인을 장기간에 걸쳐 수집하고, 교육정책 및 학교의 교육활동이 학생에 미치는 효과를 분석하여 학교 교육활동을 개선하고 효과적인 교육정책을 수립하는 것을 목적으로 실시되었다.

조사 기간은 2020년부터 2026년까지 7년간이었다. 조사 대상은 초등학교 4학년, 중학교 1학년, 고등학교 1학년 학생과 학부모, 교사, 학교장이었다.

주요 조사 항목은 학생, 학부모, 교사, 학교장, 학교로 나누어 실시되었다. 학생 대상으로는 학생의 인구 통계학적 특성 및 배경, 학교 및 사회생활, 수업과 평가에 대한 인식, 방과후학교, 사교육, 가정생활, 학습태도, 진로, 역량 및 행복감 등이 조사되었다. 학부모 대상으로는 가정배경, 자녀의 교육경험, 학교 교육 참여, 부모-자녀 관계 등이 조사되었다. 교사 대상으로는 배경, 교직태도, 학교조직문화, 수업활동, 학생평가활동 등이 조사되었다. 학교장 대상으로는 배경, 학교 경영 목표, 의사결정 권한 등이 조사되었다. 학교 대상으로는 배경, 교육과정-수업-평가, 진로 및 취업지도 등이 조사되었다.

조사 방법은 오프라인 조사와 온라인 조사를 병행하였다. 학생 대상 조사에서는 온라인 조사 방법의 경우 오프라인 조사에 비해 응답률이 감소될 수 있고 무성의하게 답변할 가능성도 있어 학생을 대상으로 한 성취도 검사나 역량 검사를 시행하기에는 시스템적으로 매우 불안정하고 비용도 상당히 든다는 단점을 고려하여 온라인 조사 방법도 도입하기가 현실적으로 어렵다고 판단하였다.

조사 실행 방안으로 참여율을 제고하기 위해 4가지 방안을 실행하였다. 연구의 목적과 중요성을 적극적으로 홍보한다. 이를 위해 협력교사 워크숍을 개최하고, 교장 및 교감 연수에 종단연구 내용을 포함한다. 조사 대상자와 업무 협조자에 대한 물질적 보상을 확대한다. 학생들에게 제공하는 봉사시간과 간식비를 늘리고, 협력교사와 업무 협조 교사에 대한 보상을 증가시킨다. 또한 조사대상 교사(국영수 담당)에게도 혜택을 제공한다. 학생과 학부모에게 '피드백 시스템'을 도입한다. 조사 결과에 대한 피드백을 제공하여 학생들의 성장과 상대적 수준을 파악할 수 있게 한다. 지역 교육지원청이나 구청 수준의 결과 자료 제공을 고려한다. 고등학교 학생의 인지적 역량을 평가하는



방법을 개선한다. 기존 학력평가 결과 자료를 활용하는 방안을 모색하여 별도의 학업성취도 검사로 인한 부담을 줄인다. 이러한 방안들을 통해 조사의 중요성에 대한 인식을 높이고, 참여 동기를 강화하며, 실질적인 혜택을 제공함으로써 전반적인 참여율과 응답의 성실성을 향상시킨다.

SELS2020은 단순 지식 측정을 지양하고 교과역량 중심의 검사지 개발로 전환하여 급변하는 사회가 요구하는 인재 양성을 목표로 서울학생역량검사를 통해 인지 영역을 평가하였다.

### 3) 서울학생역량 검사

서울학생역량 검사는 국어, 수학, 문제해결역량 등 여러 역량을 평가하는데, 여기서는 인지 영역과 관련성이 높은 국어, 수학, 문제해결역량만 제시하였다. 국어 교과역량 평가들은 학습자들이 실제 삶 속에서 접할 수 있는 다양한 종류의 텍스트를 체계적으로 제시하였고 맥락을 세분화하여 제시하였다. 이에 내용 영역의 차원 외에 과정, 텍스트, 맥락 4가지 차원으로 구성되었다. 수학 교과역량 평가들은 내용, 과정, 맥락의 3가지 차원으로 구분되며 영어 교과역량 평가들은 내용, 과정, 텍스트, 맥락 4가지 차원으로 구성되었다.

2021년 초등학교 4학년, 중학교 1학년, 고등학교 1학년을 대상으로 검사를 개발하였으며 학교급별 부담, 학생의 성실도 등을 고려하여 학교급별 검사도구는 교과역량(국어, 수학, 영어)의 경우 초등학교 4학년, 중학교 1학년, 사회·정서적 역량은 초등학교 4학년, 중학교 1학년, 고등학교 1학년, 참여 자치 역량은 중학교 1학년, 고등학교 1학년을 대상으로 차별화하였다.

교과역량의 경우, 하나의 대문항(big item)에 2~3개의 하위문항이 포함되는 구조로 개발하였으며, 하위문항은 선다형, 단답형, 서술형이 골고루 포함되도록 하였고 서술형 문항이 전체 문항의 30%가 넘지 않도록 하였다. 검사도구는 개발한 평가들에 따라 가급적 모든 하위요소가 포함될 수 있도록 하였고 정규 수업 시간(초등학교 40분, 중학교 45분, 고등학교 50분) 내에 해결할 수 있는 문항을 개발하고자 하였다.

대상별 평가 내용과 문항 수는 <표 II-8>과 같이 편성하였다. 해당 학년의 1학기 이전 학년의 2학기 성취기준을 평가 내용으로 선정하였다.

&lt;표 II-8&gt; 서울학생역량 검사 성취 기준 및 문항 수

패널	대상	성취 기준	본 검사 문항 수(공통문항)		
			국어	수학	영어
초4	초4	초등학교 3-2학기 초등학교 4-1학기	20	21	20
	초5	초등학교 4-2학기 초등학교 5-1학기	25(5)	24(5)	20
	초6	초등학교 5-2학기 초등학교 6-1학기	22(5)	24(5)	20(2)
중1	중1	초등학교 6-2학기 중학교 1-1학기	24	19	24
	중2	중학교 1-2학기 중학교 2-1학기	25(5)	24(5)	24
	중3	중학교 2-2학기 중학교 3-1학기	23(5)	24(6)	29(5)

2019년 총 13개교에서 초등학교 4학년 269명, 중학교 1학년 226명, 고등학교 1학년 216명을 대상으로 예비검사를 시행하였다. 일관된 시행을 위해 시행 전 매뉴얼을 제작하여 제공하였고 시행 과정을 모니터링하고 문제점이나 긴급 상황에 대처할 수 있도록 초등학교와 중학교는 2명 이상, 고등학교는 1명 이상의 연구진이 직접 참관하였다.

예비검사 결과 정답률, 변별도, 선다형의 답지반응분포, 서답형 부분 점수 응답 분포를 분석하여 지나치게 어렵거나 쉬운 문항, 오답의 기능을 충분히 하지 못하는 문항, 변별력이 너무 낮거나 부적인 문항, 서답형의 경우 부분 점수 응답이 없는 문항인 경우 문항을 수정하였다. 구체적으로 정답률이 95% 이상인 문항이 있는지, 변별도가 0.1 이하이거나 부적 변별력인 문항이 있는지, 오답률이 25% 이상인 선택지가 있는지, 0점에 해당하는 서답형 부분 점수의 응답 분포가 50% 이상인 문항인지를 기준으로 개별 문항을 점검하고 수정하였다. 예비검사 분석 기준에 해당하는 문항은 개별 문항을 점검하고 수정하였다. 수학 교과와 경우 학생들의 서술형 답안 채점 결과를 바탕으로 채점기준을 정교화하고 유사 답안을 확보하였다.

2021년 초등학교 6개교, 중학교 5개교, 고등학교 5개교를 선정하여 학교별로 2개 학급을 대상으로 예비검사를 시행하였다. 시행의 효율성을 위해 새로 도입된 OMR 답안지 선다형 문항의 코딩과 서답형 문항 스캔 작업을 진행하였다. 채점 워크숍과

함께 서답형 채점을 진행하였다. 문항 분석 대상 학생 수는 초등학교 233명, 중학교 270명이었다. 교과역량은 고전검사이론(Classical Test Theory, CTT)에 의거한 분석을 실시하여 문항의 특성과 양호도를 검증하였으며, 수정된 문항에 맞추어 채점기준 또한 수정 보완하였다. CTT와 IRT에 따른 양호도를 검토하여 기준에 못미치는 문항에 대하여 문두나 답지(또는 오답지) 등의 수정을 고려하였다. CTT의 정답률과 변별도에 기반한 수정 기준은 정답률이 20%보다 작거나 같고 95%보다 크거나 같은 문항, 변별도가 0.2보다 큰 문항, 가급적 중간 수준의 난이도(정답률, 변별도) 문항이 아닌 경우, 음(-)의 값을 갖는 변별도 문항, 신뢰도에 부정적 영향을 주는 문항이었다. IRT 문항모수에 기반한 수정 기준은 문항적합도지수(S-X2) 기준 미달 문항( $p < .05$ ), 난이도가 -2보다 작거나 2보다 큰 문항, 변별도가 0.4보다 큰 문항, 가급적 중간 수준 난이도, 변별도 문항이 아닌 경우, 음(-)의 값을 갖는 변별도 문항이었다.

먼저 고전검사이론의 정답률과 변별도에 기반하여 정답률이 20%보다 낮거나 95%보다 높은 문항이 있는지, 변별도가 0.2보다 작은 문항이 있는지 살펴보았다. 가급적 중간 수준의 난이도(정답률, 변별도) 문항이 아닌 경우, 그리고 음의 값을 갖는 변별도, 신뢰도에 부정적 영향이 가는 문항을 수정하였으며 부가적으로 문항반응이론 문항모수에 기반하여 문항적합도를 검토하였다. 이상의 수정 기준에 해당하는 문항에 대해서는 개별 문항별로 점검하였다. 또한 이 과정에서 예비검사 시, 현장에서 제기된 25개 문항이 정해진 시간에 풀기 다소 부담스럽다는 의견을 함께 반영하여 문항 수를 조정하고자 하였다. 결과적으로 초등학교 5학년, 중학교 2학년 국어 역량 검사는 예비검사 결과에서 적합도가 낮은 문항 2개를 삭제하여 최종적으로 23개씩의 문항을 확정하였다. 수학 교과에서는 학생들의 서술형 답안 채점 과정에서 채점기준을 정교화하고 유사 답안을 확보하며 무응답, 이상 응답 등을 고려하여 문항을 수정 보완하였다. 결과적으로 초등학교 5학년 수학 역량 검사는 예비검사 결과에서 적합도가 낮은 문항 2개를 삭제하여, 최종적으로 KICE 종단연구의 문항 9개, 새롭게 개발한 문항 10개, 초등학교 4학년 검사지 문항 5개를 포함하여 24개의 문항을 확정하였다. 그리고 중학교 2학년 수학 역량 검사는 예비검사 결과에서 변별도가 낮은 문항 1개를 삭제하여 최종적으로 24개의 문항을 확정하였다. 교과별 역량 검사도구의 평가틀 하위요소별 비율과 문항정보표 및 예시문항은 <부록 8>~<부록 11>에 제시하였다.

2022년 초등학교 6학년, 중학교 3학년, 고등학교 3학년을 대상으로 초등학교 5개교, 중학교 4개교, 고등학교 2개교를 선정하여 예비검사를 실시하였다. OMR 답안지 선다형 문항 코딩과 서답형 문항 스캔 작업을 진행하고 채점 워크숍과 함께 서답형 문항

채점을 진행하였다. 분석 대상 학생 수는 초등학교 250명, 중학교 250명이었다. 교과역량은 고전검사이론에 따른 분석을 실시하여 문항의 특성과 양호도를 검증하였으며, 수정된 문항에 맞추어 채점기준 또한 수정 보완하였다. CTT에 따른 양호도를 검토하여 기준에 못미치는 문항에 대하여 문두나 답지(또는 오답지) 등의 수정을 고려하였다. CTT의 정답률과 변별도에 기반한 수정 기준은 정답률이 20%보다 작거나 같고 95%보다 크거나 같은 문항, 변별도가 0.2보다 큰 문항, 가급적 중간 수준의 난이도(정답률, 변별도) 문항이 아닌 경우, 음(-)의 값을 갖는 변별도 문항, 신뢰도에 부정적 영향을 주는 문항이었다.

먼저 고전검사이론의 정답률과 변별도에 기반하여 정답률이 20%보다 낮거나 95%보다 높은 문항이 있는지, 변별도가 0.2보다 작은 문항이 있는지 살폈다. 가급적 중간 수준의 난이도(정답률, 변별도) 문항이 아닌 경우, 그리고 음의 값을 갖는 변별도, 신뢰도에 부정적 영향이 가는 문항을 수정하였다. 수정 기준에 해당하는 문항에 대해서는 개별 문항별로 점검하였다. 또한 국어 교과에의 경우 이 과정에서 문항 수가 많아 정해진 시간에 풀기 다소 부담스럽다는 의견을 함께 반영하여 문항 수를 조정하고자 하였으며 위에서 문제가 된 초등학교 문항 2개와 중학교 문항 1개를 모두 삭제하였다. 수학교과는 예비검사 결과로 얻은 문항별 정답률, 변별도를 고려하여 문항을 수정하거나 삭제하고 학생들의 서술형 답안 채점 과정에서 채점기준을 정교화하고 유사 답안을 확보하며 무응답, 이상 응답 등을 고려하여 문항을 수정 및 보완하였으며 검사지 전체의 완성도를 위해 선다형 문항의 정답지가 유사한 분포를 이루도록 수정하거나 삽화를 수정하였다.

## 나. 경기도교육연구원

### 1) 경기교육종단연구

경기교육종단연구는 경기도 내 학생들의 학업성취, 정의적 발달, 민주시민의식 성장 등에 대한 정보를 체계적으로 수집하고, 이를 바탕으로 교사, 학교, 교육정책이 학생들의 성장과 변화에 미치는 영향을 분석하며, 나아가 학교 교육과정과 정책의 중장기적 효과를 평가할 수 있는 과학적이고 실증적인 기초자료를 구축하는 것을 목적으로 한다.

조사기간은 2012년부터 2021년까지로 2012년 당시 초등학교 4학년 3,541명, 중학교 1학년 4,051명, 고등학교 1학년 4,242명 학생과 학생의 부모(보호자), 담당교사를

대상으로 시작되었다. 초등학교에서는 담임교사와 교과전담교사, 중·고등학교에서는 조사대상 학생들의 담임교사와 국어, 영어, 수학, 과목 담당교사를 포함하였다. 2015년 4차년도 조사부터 중학교 1학년이 된 초등학생 패널만을 유지하는 것으로 조사 설계를 수정하여 2,773명을 추가 표집하였다. 2018년 7차년도 조사에도 고등학교 1학년 499명 추가 표집이 이루어졌고 패널 학생 수가 2명 이하인 학교는 학생과 학부모만 조사하였다.

주요 조사 항목은 학생, 학부모, 교사, 학교장, 학교로 나누어 실시되었다. 학생 대상으로는 학생 생활, 학교 적응, 진로 계획, 학업성취도 등이 조사되었다. 학부모 대상으로는 자녀와의 관계, 교육비, 학교 활동 참여 등이 조사되었다. 교사와 학교장 대상으로는 학교 운영, 교수활동, 학교 만족도 등이 조사되었다. 학교 대상으로는 설립유형, 교육과정 특성, 학생 수, 교직원 수, 시험 횟수 등이 조사되었다.

조사 방법은 설문조사와 국어, 영어, 수학 교과에 대한 학업성취도 검사를 통해 이루어졌다. 2018년 7차년도부터 국어, 영어, 수학 학업성취도 검사는 교육과정이 개편되어 기존의 검사문항이 적절하지 않아 실시하지 않았다.

국어, 영어, 수학 과목의 학업성취도 검사는 일반정규 수업시간과 동일하게 초등학교 40분, 중학교 45분, 고등학교 50분이었다. 교육과정을 충실히 반영하여 출제하였으며 출제 범위는 검사 대상 학년의 1학기 교육과정 내용을 포함하였다. 문항 수는 정규 수업시간 내에서 해당 과목의 특성과 검사에 참여하는 학생의 수준에 맞춰 국어 서술형 5문항, 수학 20문항, 영어의 경우 초등학교 25문항, 중학교 28문항, 고등학교 30문항으로 의사소통능력을 중시하는 교과 특성을 반영하여 듣기 평가 문항을 전체 문항의 초·중·고등학교는 50% 이상, 중·고등학교는 30% 이상 포함하였다.

1차년도(2012년)부터 3차년도(2014년)까지 각 패널의 성취도 검사지를 개발하였으며 여러 학년에 걸친 성취도 점수를 연계하고자 교과별, 학교급별 성취도점수의 수직척도를 구축하였다. 초등학교 패널의 경우 초등학교 4학년부터 중학교 3학년까지 교과별 학업성취도를 동일 척도상에서 비교할 수 있도록 학교급을 연결한 수직척도를 개발하였다. 초등학교 패널이 고등학교로 진학한 7차년도(2018년)에는 2015 개정 교육과정이 적용되는 시기로 고등학교 패널의 검사지를 활용한 성취도 검사가 적합하지 않아 성취도 검사를 시행하지 않았다.

표본의 이탈을 최소화하기 위하여 패널의 상세한 변동사항 파악, 표본학교 대상 설명회 개최 및 방문을 통한 협조 요청, 패널 학생의 개인정보 관리, 일관된 조사 문의처 유지 등의 방법을 통해 지속적인 조사 참여를 유도하고 표본을 체계적으로 관리하였다.

조사 참여로 패널 학생에게 제공한 것은 경기교육종단연구 참여 내용을 창의적체험활동으로 학교생활기록부에 기재하고 봉사시간을 부여하도록 하였다. 단, 학교 재량이기 때문에 학교마다 차이가 발생하였다. 협력교사에게는 위촉장을 수여하였다. 1-3차년도에는 경기도교육감의 위촉장을 수여하고 학교 관리자(교장, 교감)에게는 경기도교육연구원장의 위촉장을 수여하였으며 4-6차년도에는 위촉장을 수여하지 않았으며 7차년도부터는 경기도교육연구원장의 위촉장을 수여하였다. 패널 학생이 재학 중인 학교에 학교조사지원비를 제공하여 학생들의 간식비, 조사에 필요한 문구 등의 물품 구입, 설문 참여교사의 사례, 성취도 검사 감독비 등 학교 재량으로 활용하도록 하였다.

## 2) 경기기초학력종단연구

경기기초학력종단연구는 경기도 초등학교 저학년 학생들의 기초학력 성장과 관련된 다양한 요인을 종합적으로 분석하여, 데이터에 기반한 효과적인 기초학력 지원 정책과 조기 개입 방안을 수립하는 것을 목적으로 하며, 이를 통해 경기도 교육의 질 향상과 과학적인 교육정책 결정을 지원하고자 합니다.

조사 기간은 2021년부터 2023년까지로 2021년 초등학교 2학년 69개 학교 2,413명의 학생과 학부모, 125명의 교사를 대상으로 하고 2차년도에는 7개 학교를 추가표집하였으며 실제 참여 인원은 학생 4,286명, 학부모 3,272명, 교사 352명, 학교 72개이다. 조사 방법은 설문조사(학생용, 학부모용, 담임교사용, 학교 관리자용)와 기초 문식력(literacy) 및 기초 수리력 지필평가로 구성되었다.

조사 내용은 교육 여건 변수로 학생 대상 학생 특성, 가정환경 특성, 학부모 대상 가정 특성, 교사 대상 교사 인적 사항, 학급 특성, 학교 대상 학교 배경 특성을 포함하였으며, 교육 운영 변수로는 학생 설문에 학습 관련 정의적 특성, 학교풍토, 학교 교육활동, 학부모 설문에 자녀의 학교생활 및 교육 활동에 대한 인식, 부모의 교육적 지원 및 관심, 교사 설문에 교실 운영, 교사 및 학교풍토, 협력적 교육공동체, 수업 및 교육과정 운영, 기초학력 지도 관련 사항, 학교 설문에 기초학력 관련 사업 및 지원 시스템, 협력적 교육공동체, 교육활동, 역량 강화 지원 등이 포함되었다. 교육성과 변수로는 학부모 설문에 학교 및 교사에 대한 만족도와 기초학력 교육에 대한 만족도, 학생 대상 기초 문식력과 기초 수리력 수준이 포함되었다.

학생 기초학력은 기초 문식력(읽기, 쓰기), 기초 수리력, 학습에 대한 태도 및 사회·정서적 역량으로 구분된다. 문식력과 수리력 지필평가와 학생의 학습 태도 및 사회·정서적

역량에 대한 자기보고식 평가(초 3, 4)로 이루어진다. 기초 문식력과 기초 수리력은 각 20문항으로 구성되며 수정된 앵고프(Modified Angoff) 방법을 채택하여 기초학력 도달과 미도달을 구분하는 분할점수를 설정하고자 하였다. 이에 6명의 교직경력 10년 이상 전문가로 패널을 구성하였으며 3-4라운드에 걸친 반복 평정을 실시하고 각 라운드마다 토론과 영향자료 검토를 통해 합의점을 도출하였다. 문식력은 읽기 9문항 중 5점을 분할점수로 설정하였고 쓰기 11문항 중 7점을 분할점수로 설정하여 문식력 전체 20문항 중 12점을 분할점수로 설정하였으며 수리력 20문항 중 14점을 분할점수로 설정하였다. 구체적인 기초 문식력과 기초 수리력의 검사 구성 및 기초학력 최소능력 수행 수준 기술 내용은 <부록 12>와 <부록 13>에 제시되어 있다.

## 다. 전라남도교육연구정보원

### 1) 전남교육중단연구

전남교육중단연구는 전라남도 학생들의 인지적, 정의적, 사회적 변화 및 성장과정에 대한 종단적 자료를 수집하고, 이를 통해 학교교육 효과와 교육정책을 평가하며, 빅데이터에 근거한 교육정책 수립 기반을 마련하는 것을 목적으로 한다.

조사 기간은 2018년부터 2027년까지 10년간 진행될 예정으로 2023년 5차년도까지 기초분석이 완료되었다. 조사 대상은 2018년 초등학교 4학년 2,199명을 시작으로, 2019년 중학교 1학년 2,549명을 추가로 표집하여 두 개의 코호트를 대상으로 한다.

학업성취도 검사(인지 영역 검사) 2종과 학생, 학부모, 교사(학생평가지 포함), 학교장, 학교를 대상으로 하는 5종의 설문지를 통해 자료를 수집하며 주요 조사 항목은 학생, 학부모, 교사, 학교장, 학교로 나누어 구성되었다. 학생 대상으로는 인지 영역 검사 외에 비인지적 성취, 학습전략, 수업이해, 학교생활, 학교환경 인식, 시간 활용, 여가활동, 장애 포부, 고민 등이 조사되었다. 학부모 대상으로는 학생 및 가족 배경 특성, 교육경험 및 기대, 방과후학교·사교육 참여, 학교참여, 학교교육만족도 등이 조사되었다. 교사 대상으로는 교사 배경 특성, 교사직무, 자기 인식, 학생/학교에 대한 인식 등이 조사되었다. 학교장 대상으로는 배경 특성, 학교경영, 교장업무, 학교장 인식 등이 조사되었다. 학교 대상으로는 학교배경, 예산, 교육과정, 교육지원 등이 조사되었으며 교육행정정보시스템(National Education Information System, NEIS) 등 학교현황자료와 연계하여 조사되었다.

패널 관리를 위해 협력교원(교감, 패널학급 담임교사)을 조기에 구성하여 조사의 필요성과 운영 방법에 대해 철저히 안내하고 설명회를 실시하였다. 학생, 학부모, 교직원들의 이해와 협조를 얻기 위해 전남교육종단연구에 대한 홍보를 강화하고 학부모의 참여를 독려하기 위해 학생별 조사결과와 성장현황 보고서 제공을 고려하였다. 학교급 전환 시점에서 패널 이동상황을 파악하기 위한 별도 조사를 실시하고, 필요시 패널을 추가하고자 하며 장기적으로 지필 방식에서 웹기반 온라인 방식으로 전환하여 조사의 효율성을 높이하고자 하였다.

전남교육종단연구는 인지 영역 능력 수준의 종단적 비교를 중요 과제로 삼고 있으며 별도의 평가도구를 개발하여 사용하였다. 참고로 2019년 연구보고서에서는 학업성취도 검사 및 학업성취도라는 용어를 사용하였으나 이후 연구보고서에서는 인지 영역 검사 및 인지 영역 능력이라는 용어를 사용하고 있다.

## 2) 인지 영역 검사

2018년 초등학교 4학년을 시작으로 2019년 초등학교 5학년과 추가 패널인 중학교 1학년, 2020년 초등학교 6학년과 중학교 2학년, 2021년 중학교 1학년과 3학년, 2022년 중학교 2학년과 고등학교 1학년 대상 국어, 수학, 사회, 과학 4개 교과에 대한 검사 문항을 개발하였다. 전남교육청이 추구하는 미래핵심역량을 측정하기 위한 사고형 문항을 포함하여 개발하였다. 학년별 수준에 맞는 미래핵심역량 측정 문항 개발에는 시간과 예산에 한계가 있어 교과 영역의 지식을 측정하는 기존 문항도 함께 개발되, 사고형 문항을 추가하는 형태로 검사지를 구성하였다.

한국교육종단연구 및 타 시도교육청 교육종단연구의 경우 보통 국어, 영어, 수학 3개 교과를 검사 대상으로 하고 있으나 영어 교과의 경우 초등학교 4학년부터 시작되는 조사 대상 학년을 고려할 때 교육의 효과를 측정하기에는 교육 기간이 짧고, 선행학습 등 사교육의 영향으로 초등학교 단계에서는 적합하지 않다고 판단하여 영어 교과 대신 탐구 영역(사회, 과학 교과)을 추가하여 학생의 문제해결력이나 사고 능력을 측정하는 것이 더 적합하다고 판단하였다.

검사 문항은 국가수준 학업성취 및 평가 기준에 근거하여 개발하였으며 초등학교는 4지 선다형, 중학교는 5지 선다형을 기본으로 하되, 주관식 단답형과 서술형 문항을 함께 구성하였다. 학년별 수직척도화를 위해 공통문항 설계(common-item design)를 적용하였다.



교육중단조사연구에서 사용하는 검사는 학년별 척도화 작업과 학년 간 척도화 작업이 모두 요구되므로, 원칙적으로 초등학교 4학년 검사부터 고등학교 3학년 검사까지 전 과정의 검사를 자료수집 이전에 개발해야 하지만, 예산과 개발 일정 등 현실적인 상황을 고려하여 당해 사용할 검사를 우선 개발하되, 수직척도화를 위한 공통문항설계를 염두에 두어 개발하고자 하였다.

대상별 출제 범위와 본 검사 문항 수는 <표 II-9>와 같이 편성하되, 중학교는 자유학기(년)제를 고려하여 구성하였다.

<표 II-9> 전남교육중단연구 검사 출제 범위 및 문항 수

패널	대상	출제 범위	본 검사 문항 수(공통문항, 수평연계문항)			
			국어	수학	사회	과학
초4	초4	초등학교 4학년 1학기	15	15	5	5
	초5	초등학교 4학년 2학기 ~ 5학년 1학기	15(3)	15(3)	7(2)	7(2)
	초6	초등학교 5학년 2학기 ~ 6학년 1학기	15(3)	15(3)	8(2)	8(2)
	중1	초등학교 6학년 2학기 ~ 중학교 1학년 1학기	17(3, 3)	17(3, 3)	8(2, 2)	8(2, 2)
	중2	중학교 1학년 2학기 ~ 2학년 1학기	17(3, 3)	17(3, 3)	8(2, 2)	8(2, 2)
중1	중1	중학교 1학년 1학기	17	17	8	8
	중2	중학교 1학년 2학기 ~ 2학년 1학기	18(3)	18(3)	9(2)	9(2)
	중3	중학교 2학년 2학기 ~ 3학년 1학기	17(3)	17(3)	8(2)	8(2)
	고1	중학교 3학년 2학기 ~ 고등학교 1학년 1학기	17(3)	17(3)	8(2)	8(2)

예비 검사 문항 수는 본 검사 문항 수의 1.5배 이상으로 할당하여 예비조사 결과를 토대로 1배수 선정하였다. 문항 유형은 선다형(초등학교 사지선다형, 중학교 이상 오지선다형), 단답형, 서술형 등으로 구성하였으며 검사 시간은 정규 수업시간(초등학교 40분, 중학교 45분, 고등학교 50분)과 동일하게 설정하였다. 검사 취지에 맞는 문항 수를 적정화하고 문항 변별도 및 난이도를 조정하였으며 매년 진행될 검사의 수직척도화를 고려한 공통문항과 동일 학년 비교를 위한 수평연계문항을 포함하였다. 2020년의 경우 코로나19로 인한 학력 저하로 난이도 및 문항 개발 조절이 필요하였다.

검사는 크게 문항 개발과 예비조사 결과분석, 최종 문항 선정 단계로 구분하여 개발하였다. 문항 개발 단계에서는 교과별 내용 전문가인 현직 교사를 중심으로 문항을 개발하

였으며 1~4명의 교사를 문항개발팀으로 구성하였다. 각 교과에서 학교급별 교육과정 연계 측정 및 수직척도화를 위해 초등학교 교사와 중학교 교사가 교차 확인하도록 구성하였으며 자문과 운영지원 인력을 두었다.

개발 과정은 성취·평가기준 분석, 공통문항 선정, 예비조사 문항 개발, 예비조사 실시, 예비조사 결과분석, 본 검사 개발 순으로 구성된다. 예비조사에는 조사과정 및 학생 피드백을 기록할 수 있도록 각 학급에 담임교사 외 연구진 1인이 시험감독에 참여하여 학생들의 피드백을 기록하였다.

2018년 4개교 315명의 초등학교 4학년 학생을 대상으로 예비검사를 실시하고 예비검사 결과를 문항반응이론에 의거하여 분석하였다. 검사가 다분문항을 포함하고 있기 때문에 PARSCALE(v4.1) 프로그램을 사용하여 분석하였다. 2모수의 로지스틱 모형에 의거하여 분석하였으며, 탐구 영역의 경우 사회 문항(1~10번)과 과학 문항(11~20번)을 부분검사로 분할하여 분석을 실시하였다. 각 검사에 대한 전체문항정보곡선(Total Information Curve, TIC)을 보면, 검사별로 정도의 차이는 있지만 능력수준이 표준점수 -1에서 -2 사이의 학생들의 정보를 가장 많이 포함하고 있는 것으로 나타났다. 검사별 문항의 양호도를 측정하기 위해 고전문항이론에서의 난이도와 변별도, 그리고 문항반응이론에서의 2-모수 모형의 a(location: 난이도), b(slope: 변별도)를 제시하였다. 문항별로 문항특성곡선(ICC: Item Characteristic Curve)과 문항정보함수(IIC: Item Information Curve)를 비교하여 양호한 문항을 선정하는 기준으로 삼았다.

2019년 초등학교 1개교 85명, 중학교 1개교 109명의 학생을 대상으로 예비검사를 실시하여 예비검사 결과를 고전문항이론과 문항반응이론에 의거하여 분석하였다. 검사지에 따라서는 다분 점수를 포함하는 문항이 있었으나, 다분문항은 복수 개의 이분문항으로 분리하여 채점을 실시하여 모든 문항을 이분문항으로 변환한 뒤 BILOG-MG(v3.0) 프로그램을 사용하여 분석하였다. 2-모수 정규 오자이브 모형에 의거하여 분석하였으며, 탐구 영역의 경우 사회 문항과 과학 문항을 부분검사로 분할하여 분석을 실시하였다. 문항반응이론 기준 난이도  $\pm 2$  초과, 변별도 0.5 미만의 문항은 제외 문항으로 고려하였고 공통문항은 가급적 포함하였다.

2020년 초등학교 1개교 146명, 중학교 1개교 90명의 학생을 대상으로 예비검사를 실시하고 예비검사 결과를 고전문항이론에 의거하여 분석하였다. 검사지에 따라서는 다분 점수를 포함하는 문항이 있었으나, 다분문항은 복수 개의 이분문항으로 분리하여 채점을 실시하여 모든 문항을 이분문항으로 변환한 뒤 엑셀 및 IBM SPSS(v26)를 사용하여 분석하였다. 탐구 영역의 경우 사회 문항과 과학 문항을 부분검사로 분할하여

분석을 실시하였다. 예를 들어 0/1/2점 등 부분점수를 부여하는 2개의 부분문항으로 구성된 단답형 문항의 경우, 각 응답을 0/1점, 0/1점의 2개의 별개의 부분문항으로 나누어 분석하였다. 2개로 나누기 어려운 서술형 문항의 경우 0/1/2점을 0/1점으로 이분하여 채점하여 분석하였다. 문항 난이도가 지나치게 높거나 낮은 경우, 변별도가 상대적으로 낮은 문항은 제외 문항으로 고려하였고 공통문항은 가급적 포함하였다.

2021년 중학교 1개교 1학년 152명, 3학년 164명 학생을 대상으로 예비검사를 실시하였다. 결시자, 수검 의사가 전혀 없다고 판단되는 답안지(하나의 답안으로 표기, 그림 표기 등) 작성 학생 등은 분석 대상에서 제외하였다. 전체가 아닌 일부 문항(1/3 이상)에 의도적으로 하나의 답안으로 표기한 경우에는 문항추측도를 낮추기 위해 해당 문항을 무응답 처리하였다. 교과별 검사에 대한 문항내적일관성신뢰도(Cronbach's  $\alpha$ )를 산출하고 각 문항별로 해당 문항을 제거하였을 때의 신뢰도 변화를 산출하였다. 교과별 전체 신뢰도와 비교하여 이 수치가 큰 경우 제외 고려 대상 문항으로 판단하였다. 예비검사 결과를 고전검사이론과 문항반응이론에 의거하여 분석하고 교과별 검사 시간의 적정성을 검토하여 문항 유형별 문항 수를 조정하였다.

2022년 중학교 1개교 1학년 104명, 고등학교 1개교 1학년 164명을 대상으로 예비검사를 실시하였다. 답안지 중 선다형 문항을 하나의 답안으로만 표기한 경우가 일부 있었으나 그 빈도가 크지 않아 모두 유효한 답안지로 인정하였다. 교과별 검사에 대한 문항내적일관성신뢰도(Cronbach's  $\alpha$ )를 산출하고 문항별로 해당 문항을 제거하였을 때의 신뢰도 변화를 산출하였다. 교과별 전체 신뢰도와 비교하여 이 수치가 큰 경우 제외 고려 대상 문항으로 판단하였다. 예비검사 결과를 고전검사이론과 문항반응이론에 의거하여 분석하였다. 검사지에 따라서 다분점수를 포함하는 문항이 있었으나 이분 문항으로 처리하거나 부득이한 경우 복수의 이분문항으로 분리하여 채점을 실시하였고 모든 문항을 이분문항으로 변환하여 분석하였다. 고전문항이론에 근거하여 문항별 난이도와 변별도를 추출하여 분석하였다.

연도별 최종 문항의 문항정보표와 검사지는 <부록 14>, <부록 19>에 제시되어 있다.

## 라. 현장 면담

교육중단연구 내 학생의 인지역량 및 교과역량 영역을 조사하고 평가하는 과정에 대한 의견을 수렴하여 연구의 질적 향상을 도모하기 위해 <표 II-10>과 같이 서울교육

종단연구, 경기교육종단연구, 전남교육종단연구 담당자 면담을 진행하였다.

<표 II-10> 시도교육청 면담 개요

연구명	서울교육종단연구2010, 2020	경기교육종단연구2012	전남교육종단연구2018
일시	2024. 9. 9.	2024. 9. 12.	2024. 9. 12.
면담자	정○	김○○	신○○

면담 주제는 크게 인지역량 및 교과역량 영역 선택 과정, 인지역량 및 교과역량 영역 조사 및 평가 방식, 인지역량 및 교과역량 영역 조사 및 평가의 어려움과 개선 방안 등 3가지로 제시되었다.

세 시도교육청 모두 설문조사와 별개의 검사도구를 활용하여 인지 영역을 평가하였다. 서울교육종단연구의 경우 객관식과 서술형 문항을 혼합한 국어, 영어, 수학 교과 학업성취도 검사를 사용하였으며 경기교육종단연구는 국어, 영어, 수학 학업성취도 검사를 부분적으로 시행하였고, 전남교육종단연구는 영어 과목을 제외하고 국어/수학, 사회/과학 2교시로 구성된 학업성취도 검사(인지 영역 검사)를 실시하고 있었다.

서울교육종단연구에서는 1주기와 연계성, 주요 교과목으로서의 인식, 기본 학습 능력 측정에 적합, 문항 구성의 용이성 등을 고려하여 국어, 영어, 수학 교과목을 선택하였다. 사회·과학 과목의 경우 문항 구성이 어려워 제외하였으며 서울 지역 학생들의 영어 실력이 우수하여 교육과정 내 난이도 조절의 어려움을 보고하였다. 또한 시간이나 인력 등 학교 자원의 제약으로 여러 교과를 평가하기 어렵다고 하였다.

경기교육종단연구의 경우 평가 도구를 개편해야 하는 상황에서 2018년도부터 학업성취도평가를 수행하지 않았다. 기존 경기교육종단연구의 경우 연구설계 시점이 10여 년 전으로 최신 연구 동향과는 거리가 있어 2025년부터 진행될 예정인 2주기 경기교육종단연구에서는 국어, 수학, 사회, 과학 4개 교과로 변경하여 시행하고자 하였다. 기존에는 주요 교과이면서 조사의 용이성을 위해 국어, 영어, 수학 교과의 학업성취도에 초점을 두었으나 2주기에는 역량 중심 평가를 위해 국어, 수학, 사회, 과학 교과를 대상으로 하는 시나리오 기반의 단위 검사 형태로 개발할 예정이다. 측정하는 교과가 너무 많을 경우 협력 학교 지원, 예산 문제 등으로 현장 조사가 어려워질 수 있어 국어와 관련성이 깊은 교과인 영어를 배제하게 되었다고 한다. 특히 교과를 변경하는 데에는 영어 과목에서 학생들의 학업성취도와 변별력 문제, 교육과정과의 연계성, 학생

들의 참여와 동기 부여, 역량 중심 교육을 강조하는 교육정책의 변화가 영향을 미쳤다.

전남교육중단연구에서는 조사대상 학년이 초등학교 4학년부터 시작되어 교육의 효과를 측정하기에는 교육과정의 기간이 짧고, 선행학습 등 사교육 영향에 의해 성취도 차이가 크기 때문에 영어 교과를 초등학교 단계에서는 적합하지 않다고 판단하였다. 이에 따라 영어 교과 대신 탐구 영역을 추가하여 학생들의 문제해결력이나 사고 능력을 측정하는 것이 더욱 적합하다고 판단하였고 예비검사 결과에 따라 문항변별도가 낮은 사회, 과학 교과 문항을 국어, 수학 교과 문항과 합하여 국어/사회, 수학/과학 2교시로 운영하게 되었다.

평가방식과 관련하여 세 시도교육청 모두 지식 위주의 학업성취도 검사에서 역량 중심 평가로의 전환을 시도하고 있었다. 서울교육중단연구는 객관식 위주에서 서술형 및 역량 중심 평가로 전환하여 비판적 사고, 문제 해결 능력, 의사소통 능력 등 핵심역량 평가를 강화하며 시나리오 기반 문항 도입으로 실제 상황에서의 역량 평가를 시도하였다. 경기교육중단연구는 학업성취도에서 역량 중심 평가로 전환하기 위하여 2022 개정 교육과정 기반 ‘핵심역량’ 중심의 ‘깊이 있는 학습’ 평가 틀을 개발 중에 있으며 시나리오 기반 문항도 도입할 예정이다. 전남교육중단연구는 2025년부터 새로운 패션을 추가하여 역량 중심으로 문항을 개발할 계획이며 마찬가지로 시나리오 기반 단위 검사 형태로 개발할 예정이었다. 그러나 세 시도교육청에서 도입하고자 하는 시나리오 기반 문항은 교과와 별개의 역량을 대상으로 하는 검사에 해당하는 내용이었다.

서울교육중단연구는 교육정책 변화와 역량 중심 교육 필요성이 강조됨에 따라 학업성취도 평가에서 역량검사로의 전환을 모색하고 있는데, 객관식과 서술형으로 구성된 평가 방법에 사회적 역량을 평가하는 시나리오 기반 문항을 추가하였다. 오지선다형, 서술형, 단답형 등 다양한 평가 방식을 활용하여 종합적인 역량 평가가 가능하며 기본 학습 역량을 평가하는 데 효과적이지만, 서술형 문항의 시간 부담이 있으며 평가 방식의 일관성이 부족하고 학생 개인차 반영의 한계가 있어 일부 개선이 필요한 영역이 있다.

경기도교육청에서는 경기교육중단연구와 별개로 경기기초학력중단연구를 진행하고 있다. 경기기초학력중단연구는 코로나 이후 초등학교 저학년(2~4학년)의 기초학력 부진 문제가 심각해지면서 기초 문해력과 수리력을 측정하고 수직척도화를 통해 성장 발달을 추적하기 위해 설계되었다. 한편, 경기교육중단연구에서는 교과역량(인지적 역량)은 인지 검사 형태로 조사하고 있으나 2주기 교과역량 평가는 모두 시나리오 기반 문항으로 개발하였다. 다만 사회정서적 역량은 자기보고식 설문으로 진행될 예정이다. 시나리오 기반 문항 개발을 위해서는 교과지식을 활용하여 문제 상황을 해결할

수 있으면서 학생들이 실제로 경험했을 법한 실제적인 맥락에 기반한 주제를 찾을 필요가 있으며 교육적 맥락(학교), 개인적 맥락(학교 외 일상), 사회적 맥락(사회 문제) 등을 포괄한 주제 찾기에 어려움을 보고하였다. 역량 평가를 위한 서술형 문항의 필요성을 느꼈으나 문항 개발 및 채점 등에 필요한 시간과 비용의 한계로 2주기 평가에 도입하지는 못하였다. 최근 서술형 문항에 대한 AI 자동 채점 기술이 많이 발전됨에 따라 추후 도입할 전망이 있다.

주요 어려움과 개선방안에서 공통적으로 제시된 것은 평가의 객관성과 일관성 유지, 학생의 참여 독려, 교사의 업무 부담이 있었다. 이에 세 시도교육청 모두 온라인 평가 시스템 도입을 주요 개선방안으로 고려하고 있으며 객관식과 서술형의 균형, 프로젝트 기반 평가 등 다양한 평가 방식을 도입하려는 노력과 학생들의 참여를 독려하기 위한 문화상품권, 봉사활동 시간 인정 등의 보상을 제공하는 방안을 고려하고 있었다. 그러나 서울에서는 다른 시도에 비해 학생들의 동기 부여 문제를 더 강조하였으며 경기에서는 교육과정의 변화와 종단연구 방식의 괴리, 학교 현장의 인프라 문제를 지적하였으며, 전남은 문항 개발과 척도화의 어려움을 특별히 언급하였다. 이에 서울에서는 피드백 제공 강화와 학습 지원 확대를 구체적인 개선방안으로 언급하였으며 경기는 3년 단위의 교육종단연구를 제안하였고 전남에서는 채점자 교육 강화를 구체적인 개선방안으로 제시하였다.

패널 관리에 있어서는 공통적으로 협력 학교의 역할이 중요하다고 인식하고 있었으며 온라인 검사 도입, 조사 결과 공개 등에 있어서는 담당자마다 견해 차이가 있었다.

전체 면담에 대한 구체적인 내용은 <부록 20>에서 <부록 22>에 제시되어 있다.

### 3. 요약

국내 교육종단연구 사례 분석의 대상은 국가 수준 교육종단연구와 지역 수준 교육종단연구로 구분된다.

국가 수준 교육종단연구는 수행기관을 중심으로 한국청소년정책연구원(NYPDI)의 한국청소년패널조사(KYPS), 한국아동·청소년패널조사 2010(KCYPS 2010), 한국직업능력연구원(KRIVET)의 한국교육고용패널조사 I (KEEP I), 한국교육고용패널조사 II (KEEP II), 한국교육개발원(KEDI)의 한국교육종단연구 2005(KELS 2005), 한국교육종단연구 2013(KELS 2013), 한국교육과정평가원(KICE)의 한국 초·중등 학교교육

성과 종단연구가 있다.

한편 지역 수준 교육종단연구는 지역에 따라 서울특별시교육연구정보원의 서울교육종단연구 2010(SELS 2010), 서울교육종단연구 2020(SELS 2020), 경기도교육연구원의 경기교육종단연구(GEPS), 경기기초학력종단연구, 전라남도교육연구정보원의 전남교육종단연구가 있다. 지역 수준 교육종단연구의 경우 교육종단연구를 수행 중에 있는 시도교육청 가운데 시작 시기와 평가도구를 기준으로 세 개 시도교육청을 대상으로 선정하였다.

국내 교육종단연구는 <표 II-11>과 같이 인지적 영역과 정의적 영역, 환경적 영역을 조사 영역으로 하며 선다형, 서술형 문항을 혼용하는 지필평가와 자기보고식 설문조사, 수직척도화의 평가 방식을 특징으로 한다.

<표 II-11> 인지 영역 조사 및 평가도구와 영역

구분	운영기관	조사명	조사 및 평가 도구	조사 및 평가 영역
국가 수준 교육종단 연구	한국청소년정책 연구원	한국청소년패널조사	설문조사	전 교과
		한국아동·청소년패널조사2010	설문조사	전 교과
		한국아동·청소년패널조사2018	설문조사	전 교과
	한국직업능력 연구원	한국교육고용패널조사 I	설문조사	전 교과
		한국교육고용패널조사 II	설문조사	전 교과
	한국교육개발 원	한국교육종단연구2005	설문조사 학업성취도 검사	국어, 수학, 영어
		한국교육종단연구2013	설문조사 학업성취도평가	국어, 수학, 영어
	한국교육과정 평가원	한국 초·중등 학교교육 성과 종단연구	웹 기반 조사 및 검사	국어, 수학, 문제해결역량, 사회·정서적 역량
지역 수준 교육종단 연구	서울특별시교 육연구정보원	서울교육종단연구2010	설문조사 성취도 검사	국어, 영어, 수학
		서울교육종단연구2020	설문조사 서울학생역량 검사	교과역량(국어, 수학, 영어)

구분	운영기관	조사명	조사 및 평가 도구	조사 및 평가 영역
	경기도교육연구원	경기교육종단연구	설문조사 학업성취도 검사	국어, 영어, 수학
		경기기초학력종단연구	설문조사 지필평가	기초 문식력, 기초 수리력
	전라남도교육 연구정보원	전남교육종단연구	설문조사 인지 영역 검사	국어, 수학, 사회, 과학





| 인천교육종단연구 기반연구 - 교과역량 평가 방안 제안 |

# III

## 국외 교육종단연구 사례 분석





## 국외 교육종단연구 사례 분석

본 장에서는 국외 교육종단연구 사례를 분석함으로써 인천교육종단연구의 교과역량 조사 및 평가 방안에 대한 시사점을 도출하고자 한다. 이를 위해 텍사스 주에서 시행하는 STAAR 검사와 플로리다 주에서 시행하는 FAST 검사를 그 사례로 분석하였다. 텍사스 STAAR의 종단적 특징은 3학년부터 8학년까지 매년 동일 학생의 수학과 읽기/언어 능력을 추적하며, 고등학교에서는 EOC(End-of-Course) 시험과 연계된다. 이는 학년 진급 및 졸업 요건으로 활용되어 학생의 성장 경로를 효과적으로 파악할 수 있게 한다. 플로리다 FAST의 종단적 특징은 학년도 내에서 PM1(초기), PM2(중간), PM3(말기) 총 3회의 평가를 통해 3학년부터 10학년의 읽기, 3학년부터 8학년의 수학 능력 변화를 추적한다. 또한, 컴퓨터화 능력적응검사(CAT)를 활용하여 학생 개인의 성장 과정을 상세히 모니터링하는 특징을 가진다.

### 1. 텍사스 주 STARR 시험

#### 가. 평가틀(개요, 변화, 실시 주기, 대상, 교과, 방법 등)

##### 1) 개요

텍사스 학업 준비 평가(State of Texas Assessments of Academic Readiness, STAAR)는 학생이 각 학년, 과목, 또는 과정에서 정해진 텍사스 필수 지식과 기술(Texas Essential Knowledge and Skills, TEKS)을 어느 정도 학습하고 적용할 수 있는지를 측정하기 위해 설계된 표준 학업성취도 시험이다. 모든 STAAR 문항은 현재 시행 중인 해당 학년 및 과목 또는 과정의 텍사스 필수 지식과 기술(TEKS)에 직접적으로 맞추어져 있다.

STAAR는 텍사스 학생들이 전국 및 국제적으로 다른 학생들과 경쟁할 수 있도록 돕고, 학교와 교사가 학생들을 학업적으로 얼마나 잘 준비시키고 있는지를 평가하는 중요한 기능을 수행한다. STAAR는 학생들이 학업 경력 동안 특정 학년과 과목에서

평가받아야 한다는 연방 ‘모든 학생이 성공하는 법(Every Student Succeeds Act)’의 요구 사항을 충족시킨다.

STAAR는 3학년부터 8학년까지의 학생들을 대상으로 수학, 읽기/언어(Reading Language Arts, RLA), 과학, 사회 과목에서 온라인 평가를 시행하며, 3학년부터 5학년까지의 학생들을 대상으로는 스페인어로 된 온라인 시험도 제공한다. 구체적으로 STAAR 평가는 수학(3-8학년), 읽기/언어(RLA, 3-8학년), 과학(5, 8학년), 사회(8학년), 고등학교 종합 성취 평가(EOC, End-Of-Course)인 대수 I, 영어 I, 영어 II, 생물, 미국 역사 과목을 포함한다. 이때 고등학교 종합 성취 평가(EOC)는 고등학교 졸업 요건에 해당한다.

3-8학년 STAAR 시험은 봄에 시행되고, EOC 평가는 연중 시행된다. 학생이 치르는 시험 수는 학년에 따라 다르며, 대개 연간 2-4일간 시험이 치러진다. 학생들은 자신의 학교에서 시험을 치르고, 필요에 따라 별도의 장소에서 시험을 볼 수 있다. 이러한 STAAR 시험은 학생들의 학업성취도와 학년 진급 준비도를 평가하기 위한 것이다. 성취수준에 미치지 못한다면 재시험을 볼 수 있는 기회가 주어지거나 학년 진급이 불가할 수 있다.

STAAR 중간 평가(Interim Assessments)는 텍사스 필수 지식과 기술(TEKS)과 연계된 선택형 온라인 평가 세트로, 학생의 학업 진행 상황을 모니터링하고 STAAR 종합 평가에서의 성과를 예측하기 위해 고안되었다. 중간 평가는 학군에 무료로 제공되며, 학교나 교사, 학생의 성과와는 연계되지 않는다.

## 2) 개정된 STAAR 시험의 특징

2021년 제87회 텍사스 주의회에서 제정된 HB 3261 법안은 2022-2023학년도까지 주 평가를 온라인으로 시행할 것을 요구하였다. 온라인 시험 시행은 학생들이 교실에서 받는 것과 같은 편의를 제공하고, 빠른 시험 결과를 제공하며, 시험 운영을 개선하고, 비선다형 문항을 허용한다. 모든 학생은 최소한의 예외를 제외하고는 온라인으로 평가받아야 한다. HB 3906은 '선다형 제한'을 도입하여 STAAR 시험에서 선다형 문항이 전체 점수의 75%를 초과할 수 없도록 규정하였다.

다른 교과 영역(예: 사회, 과학, 수학, 예술 등)의 텍사스 필수 지식과 기술(TEKS)과 일치하는 내용을 참조하는 교과 간 정보 지문이 평가에서 증가하였다. 읽기/언어(RLA) 평가의 교과 간 지문은 다른 교과 영역의 주제를 포함하지만, 질문은 RLA 텍사스

필수 지식과 기술(TEKS)만을 평가하며, 학생들은 다른 교과 영역의 텍사스 필수 지식과 기술(TEKS)에 대한 이해로 점수를 받지 않는다.

STAAR 읽기/언어(RLA) 평가(3-8학년 RLA, 3-5학년 RLA 스페인어, 영어 I 및 II)는 읽기와 쓰기를 모두 평가하며, 모든 평가에는 확장된 구성형 응답, 즉 에세이가 포함된다. 에세이 구성 요소는 독립된 질문에 대한 응답을 작성하는 것에서 읽기 지문을 기반으로 응답을 작성하는 것으로 변경되었다. 학생들은 정보적(informational), 논증적(argumentative), 서신(correspondence) 중 하나의 형식으로 글을 작성하며, 5점 만점의 채점 기준을 사용해 평가된다. 채점 기준은 아이디어 개발과 언어 규칙의 두 가지 주요 요소로 구성된다.

### 3) 문항의 특성 (객관식, 주관식 등)

#### 가) 문항 유형 (교과별 다양한 유형의 문항)

수학, 읽기/언어(RLA), 과학, 사회, 그리고 고등학교 생물학 및 대수 I(Algebra 1)과 같은 과목에서 사용되는 문항 유형은 아래와 같이 다양하게 구분된다.

- 수학: 다지선다형(Multiple choice), 선택형(Inline choice), 다중 선택(Multiple select), 드래그 앤 드롭(Drag and Drop), 수식(Equation), 핫스팟(Hot Spot), 숫자 선(Number line), 그래프(Graph) 등의 유형이 있다.
- 읽기/언어(RLA): 다지선다형(Multiple choice), 다중 선택(Multiselect), 텍스트 강조(Highlight Text), 단문 구성형 응답(Short Constructed Response), 장문 구성형 응답(Extended Constructed Response), 선택형(Inline Choice) 등이 포함된다.
- 과학: 다지선다형(Multiple choice), 다중 부분(Multi part), 핫스팟(Hot Spot), 다중 선택(Multiselect), 드래그 앤 드롭(Drag and Drop), 단문 구성형 응답(Short Constructed Response) 등의 유형이 있다.
- 사회: 다지선다형(Multiple choice), 드래그 앤 드롭(Drag and Drop), 단문 구성형 응답(Short Constructed Response) 등이 있다.
- 고등학교 생물: 다지선다형(Multiple choice), 핫스팟(Hot Spot), 다중 부분(Multi part), 단문 구성형 응답(Short Constructed Response), 드래그 앤 드롭(Drag and Drop), 텍스트 입력(Text Entry) 등의 유형이 포함된다.

- 대수 I (Algebra I): 다지선다형(Multiple choice), 드래그 앤 드롭(Drag and Drop), 다중 선택(Multiple select), 그래프 작성(Graphing), 선택형(Inline Choice), 수식 편집기(Equation Editor) 등이 있으며, 총 50문항으로 구성된다.
- 역사(History): 다지선다형(Multiple choice), 다중 선택(Multiselect), 핫 텍스트(Hot Text), 다중 부분(Multi part), 단문 구성형 응답(Short Constructed Response), 매칭 테이블 그리드(Match Table Grid), 선택형(Inline Choice), 드래그 앤 드롭(Drag and Drop) 등의 유형이 있다.
- 영어 I (English I): 다지선다형(Multiple choice), 텍스트 입력(Text Entry), 단문 구성형 응답(Short Constructed Response), 다중 선택(Multiple select), 장문 구성형 응답(Extended Constructed Response), 선택형(Inline Choice) 등이 사용된다.
- 영어 II (English II): 다지선다형(Multiple choice), 단문 구성형 응답(Short Constructed Response), 매칭 테이블 그리드(Match Table Grid), 다중 선택(Multiselect), 장문 구성형 응답(Extended Constructed Response), 선택형(Inline Choice) 등의 유형이 포함된다.

#### 나) 부분점수 문항

수학, 읽기/언어(RLA), 과학, 사회, 고등학교 과목별 1점, 2점, 확장 구성형 응답 개수는 <표 III-1>과 같다. 여기서 1점 문항은 다지선다형 및 비다지선다형 모두가 포함되고, 2점 문항은 모두 비다지선다형이다.

<표 III-1> 학년 및 교과목별 STAAR 검사 부분점수 문항 유형

학년	과목	1점 문항	2점 문항	확장 구성형 응답	총계
초3	수학	23	7	-	30
	읽기/언어	38	2	1	41
초4	수학	24	8	-	32
	읽기/언어	38	2	1	41
초5	수학	26	8	-	34
	읽기/언어	38	2	1	41
	과학	25	7	-	32
초6	수학	29	7	-	36
	읽기/언어	42	2	1	45
중1	수학	30	8	-	38

학년	과목	1점 문항	2점 문항	확장 구성형 응답	총계
	읽기/언어	42	2	1	45
중2	수학	32	8	-	40
	읽기/언어	42	2	1	45
	과학	30	8	-	38
	사회	31	9	-	40
	대수 I	41	9	-	50
고등 학교	생물	37	8	-	45
	영어 I	48	3	1	52
	영어 II	48	3	1	52
	역사	50	14	-	64

출처: Texas Education Agency(2022a), Texas Education Agency(2022b), Texas Education Agency(2022c), Texas Education Agency(2022d), Texas Education Agency(2022e), Texas Education Agency(2022f), Texas Education Agency(2022g), Texas Education Agency(2022h), Texas Education Agency(2022i), Texas Education Agency(2022j), Texas Education Agency(2022k), Texas Education Agency(2022l), Texas Education Agency(2022m), Texas Education Agency(2022n), Texas Education Agency(2022o), Texas Education Agency(2022p), Texas Education Agency(2022q), Texas Education Agency(2022r)

#### 다) 문항 구성

검사는 준비 기준(Readiness standards)과 지원 기준(Supporting standards)에 해당하는 문항으로 분류되어 구성된다. 이때 준비 기준과 지원 기준에 대한 설명은 아래와 같고, 더 자세한 설명은 [그림 III-1]에 제시되어 있다.

준비 기준(Readiness standards)은 현재 학년에서 성공을 위해 필수적이며, 다음 학년 또는 과정에 대비하는 데 중요한 역할을 한다. 이 기준은 넓고 깊은 개념을 다루며, 심층적인 지도가 필요하다. 이러한 기준은 기본 시험의 총점수 중 약 55-70%를 차지한다. 또한 지원 기준(Supporting standards)은 다음 학년 또는 과정에 대비하는 데 기여하지만, 중심적인 역할을 하지는 않는다. 이 기준은 현재 학년 또는 과정보다 하위 또는 상위 학년에서 강조되는 좁게 정의된 개념을 다룬다. 지원 기준은 기본 시험의 총점수 중 약 30-45%를 차지한다.



Readiness Standards	Supporting Standards
<ul style="list-style-type: none"> <li>• are essential for success in the current grade or course</li> <li>• are important for preparedness for the next grade or course</li> <li>• support college and career readiness</li> <li>• necessitate in-depth instruction</li> <li>• address broad and deep ideas</li> </ul>	<ul style="list-style-type: none"> <li>• may be introduced in the current grade or course and emphasized in a subsequent year</li> <li>• may be reinforced in the current grade or course and emphasized in a previous year</li> <li>• play a role in preparing students for the next grade or course, but not a central role</li> <li>• address more narrowly defined ideas</li> </ul>

[그림 III-1] 준비 기준(Readiness Standards)과 지원 기준(Supporting Standards)

출처: Texas Education Agency(2024w)

#### 4) 검사 개발 (Test Development)

정교한 학생 평가 프로그램을 유지하는 것은 복잡하고 세밀한 시험 개발 과정을 수반하며, 텍사스 교육청(TEA)은 텍사스 공립학교 학생들의 학습 평가가 공정하고 정확하도록 교육자들의 의견을 반영하고 있다. STAAR 시험 문항(STAAR 스페인어 버전 포함)은 매년 개발되며, 교육자 위원회의 검토, 현장 테스트, 데이터 분석을 거친 후, 승인을 받으면 STAAR 문항 데이터베이스에 추가된다. 대부분의 경우 새로 개발된 문항은 매년 봄 STAAR 본 평가에 포함되지만, 독립적인 현장 테스트는 주기적으로 필요하며 2011년, 2015년, 2019년, 2022년에 시행되었다.

#### 5) 검사 관리 및 시행 (Test Administration)

모든 STAAR 평가-3-8학년 수학, 3-8학년 읽기/언어(RLA), 5학년과 8학년 과학, 8학년 사회 과목, 3-5학년 스페인어 수학, 3-5학년 스페인어 RLA, 5학년 스페인어 과학, 대수 I, 영어 I, 영어 II, 생물, 미국 역사-는 봄에 (1년에 한 번) 온라인으로 시행된다. STAAR 종합 성취 평가(EOC)는 6월과 12월에도 온라인으로 시행된다. 특수한 상황이 있는 학생들을 위해 STAAR의 종이 버전도 제공된다.

##### 가) 대상

미국 교육부(USDE)는 3학년부터 8학년까지 모든 학생에게 읽기와 수학 시험을 매년

실시할 것을 요구하며, 10학년부터 12학년 사이에는 최소한 한 번씩 시험을 치르도록 규정하고 있다. 모든 학생은 과학 시험도 3-5학년, 6-9학년, 10-12학년 사이에서 최소 한 번씩 치러야 한다. 샘플링 방식은 STAAR 평가의 목적을 충족시키지 못하기 때문에, STAAR 평가에서는 샘플링이 사용되지 않는다. 하지만 텍사스 평가 프로그램의 다른 부분에서는 샘플링이 사용된다. 예를 들어, 예비시험(Field Test)을 위해 샘플링을 활용해 모든 학생이 시험용 문항(Field Test Item)을 풀 필요 없이 문항의 품질에 대한 정보를 수집한다.

2023 STAAR technical digest에 제공되어 있는 STAAR 응시자 수는 아래 <표 III-2>에 제시되어 있다. STAAR EOC의 경우에는 12월, 봄, 6월 응시자 총계가 기입되어 있다.

<표 III-2> 2022-2023학년도 STAAR 시험 응시자 수

STAAR 시험		응시자 수
3학년	수학	370,006
	읽기/언어	356,558
	수학 (스페인어 버전)	16,454
	읽기/언어 (스페인어 버전)	30,213
4학년	수학	373,988
	읽기/언어	365,035
	수학 (스페인어 버전)	11,497
	읽기/언어 (스페인어 버전)	21,694
5학년	수학	378,663
	읽기/언어	372,677
	과학	378,742
	수학 (스페인어 버전)	8,483
	읽기/언어 (스페인어 버전)	15,991
	과학 (스페인어 버전)	9,775
6학년	수학	384,766
	읽기/언어	391,376
7학년	수학	331,698
	읽기/언어	400,416
8학년	수학	364,110
	읽기/언어	410,472
	과학	407,847
	사회	414,692
대수 I		608,559
영어 I		712,965
영어 II		612,128

STAAR 시험	응시자 수
생물	544,720
미국 역사	427,007

출처: Texas Education Agency(2024w)

## 나) STAAR 시험의 역할: 학년 진급, 졸업과의 연계

STAAR 시험은 학생들의 학년 진급, 졸업과의 직접적인 관련이 있기에 해당 시험을 치르는 학생들은 시험에 충실하게 참여할 수 있다. 학년 진급, 졸업과의 직접적인 연관성은 다음과 같다.

STAAR 시험을 통과해야 학년 진급이 가능한지 여부는 학생이 5학년 또는 8학년에 속한 경우에 해당된다. 이 경우, 학생은 읽기와 수학에서 STAAR 시험을 통과해야만 다음 학년으로 진급할 수 있다. 이 요건은 1월 1일과 첫 STAAR 시험 실시일 사이에 학군에 등록된 모든 학생에게 적용된다. 모든 학생은 STAAR 시험에 응시할 기회를 제공받아야 하며, 추가적으로 필요한 가속 학습 또한 제공되어야 한다(텍사스 교육법 §28.0211(a)(1)-(2); 19 텍사스 행정법 §§101.2001, 2003, 2006).

만약 5학년이나 8학년 학생이 STAAR 시험에서 불합격할 경우, 재시험 기회가 두 번 주어진다. 만약 첫 재시험에서도 불합격할 경우, 주에서 승인한 대체 평가를 두 번째 재시험으로 치르거나, STAAR 시험을 세 번째로 볼 수 있다(텍사스 교육법 §28.0211(b)).

5학년 또는 8학년 학생이 STAAR 시험에서 불합격하고 두 번의 재시험에서도 만족스러운 결과를 얻지 못한 경우, 그 학생은 진급하지 못할 수 있다. 그러나 학부모는 이 진급 보류에 대해 항소할 수 있으며, 학교의 교장 또는 그 대리인, 교사, 그리고 학생의 학부모로 구성된 성적 배정 위원회가 만장일치로 해당 학생이 진급 후에도 학업을 성공적으로 수행할 수 있다고 결정하면, 진급이 허용될 수 있다(텍사스 교육법 §28.0211(e)). 3학년부터 8학년까지의 학생이 STAAR 시험에서 불합격할 경우, 그 학생은 학년 진급 전 가속 학습을 완료해야 한다(텍사스 교육법 §§28.0211(a-2), 0217).

고등학교 학생들을 대상으로 이뤄지는 STAAR EOC 시험은 고등학교 졸업 요건에 해당한다. 고등학교 학생들은 대수 I, 생물, 영어 I, 영어 II, 미국 역사라는 5가지 과목의 EOC 시험을 통과해야 졸업할 수 있다. 11학년 또는 12학년 학생이 2과목

이하에서 STAAR 시험을 통과하지 못한 경우, 개별 졸업 위원회가 해당 학생이 졸업할 자격이 있다고 결정하면 졸업이 허용될 수 있다(텍사스 교육법 §§28.0258, 39.023(c); 19 텍사스 행정법 §101.3022). 또한, 고등학교 학생이 STAAR EOC 시험에서 불합격할 경우 학군은 그 학생에게 가속 학습을 제공해야 하며, 이 가속 학습은 종종 과외 형태로 이루어질 수 있다(텍사스 교육법 §§28.0211(a-2), 0217).

이와 같이 STAAR 시험이 학년 진급, 고등학교 졸업과 연계된다는 사실은 모든 학생들이 충실히 STAAR 시험에 응할 수 있도록 유도한다.

#### 다) 검사 전달 시스템(Test Delivery System, TDS)

STAAR 온라인 평가는 TDS(Test Delivery System)를 사용하여 진행된다. TDS는 시험 세션을 예약하는 데 사용되는 시험 관리자 인터페이스, 학생들이 시험에 참여할 수 있는 학생 인터페이스, 그리고 안전한 온라인 시험 환경을 제공하는 보안 브라우저 애플리케이션으로 구성되어 있다. TDS는 시험 데이터를 안전하게 전송하고 저장할 수 있으며, 학생 인구를 지원할 수 있도록 확장 가능하게 설계되어 있다. TDS의 아키텍처는 120만 명 이상의 동시 사용자를 원활하게 처리하여 안정성과 효율성을 입증하였다.

#### 라) 예외적인 시험 환경

STAAR 시험에 응시하는 학생들은 학교에서 제공하는 온라인 기기를 사용해 지정된 장소에서 시험을 치르게 된다. STAAR 시험은 해당 학년의 모든 학생들이 의무적으로 응시해야 하며, 시험 기간은 텍사스 주에서 공식적으로 지정된 일정에 따라 운영된다. 따라서 시험은 수업과 분리된 별도의 기간에 진행되며, 이 기간에 각 학교는 시험 응시를 위한 준비와 관리에 집중한다.

이와 달리 정해진 시험 기간에 학교에 출석할 수 없는 경우에는 대안적인 시험 환경을 제공한다. 당일 결석한 학생들을 위해 모든 학년, 과목, 과정에 대해 STAAR 시험 기간 동안 보충 시험 기회가 제공된다. 거주 지역에서 시험을 볼 수 없는 학생들은 타지역에서 시험을 치를 수 있다. 예를 들어, 여름 동안 휴스턴에서 댈러스로 간 학생은 댈러스에서 시험을 등록할 수 있다. 타지역 학생들은 해당 지역이 시험 준비를 할 수 있도록 정해진 기간 내에 등록을 완료해야 하며, 시험 당일에는 신분증을 지참해야

한다. STAAR EOC 평가를 통과하지 못했으나 나머지 졸업 요건을 충족하고 학교에 더 이상 재학하지 않는 응시자는, 시험 시행 기간 동안 시험을 시행하는 교육구에서 시험에 응시할 수 있다.

## 6) 검사 결과 보고

### 가) 성취 기준(Performance Standards)

성취 기준은 학생들이 주 전체 교육과정에서 학습하도록 기대되는 내용을 기준으로 시험 성적 수준을 직접적으로 연관시킨다. 표준 설정은 평가에서 성취 수준을 정의하는 컷 점수를 설정하는 과정이다. STAAR 표준 설정은 정책, 텍사스 필수 지식 및 기술 (TEKS) 내용 기준, 교사들이 학생들이 알아야 할 내용과 수행할 수 있어야 하는 기술에 대한 지식, 그리고 학생들의 주 평가 성과와 다른 평가와의 성과 정렬 정보를 고려했다. 표준 설정 위원회는 이러한 요소들의 상호작용을 각 STAAR 평가별로 신중히 검토한 다양한 이해 관계자들로 구성되었다.

STAAR 시험(스페인어 버전 포함)의 성취 수준 및 정의는 아래와 같다.

- 학년 수준 미달 (Did Not Meet Grade Level): 이 범주에 속하는 학생들은 지속적인 학업 지원이 없이는 다음 학년이나 과정에서 성공할 가능성이 낮다. 이 범주의 학생들은 평가된 지식과 기술에 대한 충분한 이해를 보여주지 못한다.
- 학년 수준 근접 (Approaches Grade Level): 이 범주의 학생들은 목표로 하는 학년이나 과정에서 맞춤형 학업 지원을 받으면 성공할 가능성이 있다. 일반적으로 평가된 지식과 기술을 익숙한 상황에서 적용할 수 있음을 보여준다.
- 학년 수준 달성 (Meets Grade Level): 이 범주의 학생들은 다음 학년이나 과정에서 성공할 가능성이 매우 높지만, 단기적인 학업 지원이 필요할 수 있다. 이들은 비판적으로 사고하고 평가된 지식과 기술을 익숙한 상황에서 적용할 수 있음을 보여준다.
- 학년 수준 초과 (Masters Grade Level): 이 범주의 학생들은 거의 학업 지원 없이 다음 학년이나 과정에서 성공할 것으로 기대된다. 이들은 익숙하거나 새로운 상황에서도 평가된 지식과 기술을 비판적으로 사고하고 적용할 수 있음을 보여준다.

### 나) 평가 보고서

텍사스 교육청(TEA)은 STAAR 및 STAAR 스페인어 평가에 대한 학생 성적 보고서

를 텍사스 공립학교 학군과 자유 등록형 차터스쿨에 제공한다. 각 STAAR 시험 시행 시 학생 성적표, 학생 라벨, 캠퍼스 명단, 요약 보고서, 데이터 파일 등이 제공된다. STAAR 및 STAAR 스페인어 시험의 봄 시험 결과는 연습 시험 사이트(Practice Test Site)를 통해 공개된다. 이 공개된 시험과 연계하여, 텍사스 교육청(TEA)은 학생 문항 분석 보고서 및 문항 분석 요약 보고서를 제공한다. 이 요약 보고서는 캠퍼스, 학군, 지역, 주 단위로 제공된다.

#### 다) 시험 결과 활용

시험 결과는 일정 기간 그룹의 성과를 평가하는 데 사용할 수 있다. 예를 들어, 4학년 읽기/언어 시험의 평균 변환 점수를 여러 해에 걸쳐 비교하여 학생 성과가 향상되고 있는지 파악할 수 있다.

또한, STAAR 성적은 서로 다른 인구 통계 그룹 또는 프로그램 그룹의 성과를 비교하는 데 사용할 수 있다. 예를 들어, 어떤 인구 통계 그룹이 가장 높은 평균 변환 점수를 받았는지, '학년 수준 근접' 성취 기준을 충족한 학생들의 비율이 가장 낮은 그룹은 어디인지 등을 분석할 수 있다. 이러한 분석을 통해 학군과 캠퍼스는 학생 그룹의 잠재적인 학업 약점을 식별하고, 성적 데이터를 바탕으로 교육 프로그램을 평가할 수 있다.

시험 점수는 학생이 추가적인 학습이나 지원이 필요한 과목을 식별하는 데도 활용될 수 있다. 예를 들어, 보고 범주 데이터는 학생이 어려움을 겪고 있는 영역을 파악하여 교사들이 효과적인 학습 개입을 계획할 수 있도록 돕는다. 시험 결과는 다른 성과 지표와 함께 학생 배치 결정을 내리는 데 참고되며, 이러한 결정을 내릴 때는 가능한 많은 학생 정보를 고려하는 것이 중요하다.

#### 나. 교과별 보고 범주 (reporting category)

텍사스 주 STAAR 시험은 텍사스 필수 지식 및 기술(TEKS)에 근거하여 고등학교 이전과 이후 교과목별 보고 범주(reporting category)를 정의하고 있다. 고등학교 이전과 고등학교 교과목별 보고 범주는 <표 III-3>, <표 III-4>에 제시되어 있다.

<표 III-3> 고등학교 이전 STAAR 시험 교과목별 보고 범주

과목	보고 범주(Reporting Category)			
수학	숫자 표현 및 관계	계산 및 대수적 관계	기하와 측정	데이터 분석 및 개인 금융 문해력
읽기/언어	읽기	쓰기		
과학	물질과 에너지	힘, 운동, 에너지	지구와 우주	생물과 환경
사회	역사	지리 및 문화	정부 및 시민권	경제, 과학, 기술, 사회

출처: Texas Education Agency(2022a), Texas Education Agency(2022b), Texas Education Agency(2022c), Texas Education Agency(2022d), Texas Education Agency(2022e), Texas Education Agency(2022f), Texas Education Agency(2022g), Texas Education Agency(2022h), Texas Education Agency(2022i), Texas Education Agency(2022j), Texas Education Agency(2022k), Texas Education Agency(2022l), Texas Education Agency(2022m), Texas Education Agency(2022n), Texas Education Agency(2022o), Texas Education Agency(2022p), Texas Education Agency(2022q), Texas Education Agency, 2022r)

<표 III-4> 고등학교 STAAR 시험 교과목별 보고 범주

과목	보고 범주(Reporting Category)				
대수 I	수와 대수적 방법	선형 함수, 방정식 및 부등식을 설명하고 그래프로 표현	선형 함수, 방정식 및 부등식 작성 및 해결	이차 함수와 방정식	지수 함수와 방정식
생물	세포 구조와 기능	유전 메커니즘	생물학적 진화와 분류	생물학적 과정과 시스템	환경 시스템 내 상호 의존성
영어 I, 영어 II	읽기	쓰기			
역사	역사	지리 및 문화	정부 및 시민권	경제, 과학, 기술, 사회	

출처: Texas Education Agency(2022a), Texas Education Agency(2022b), Texas Education Agency(2022c), Texas Education Agency(2022d), Texas Education Agency(2022e), Texas Education Agency(2022f), Texas Education Agency(2022g), Texas Education Agency(2022h), Texas Education Agency(2022i), Texas Education Agency(2022j), Texas Education Agency(2022k), Texas Education Agency(2022l), Texas Education Agency(2022m), Texas Education Agency(2022n), Texas Education Agency(2022o), Texas Education Agency(2022p), Texas Education Agency(2022q), Texas Education Agency, 2022r)

## 2. 플로리다 주 FAST 검사

### 가. 평가틀(개요, 변화, 실시 주기, 대상, 교과, 방법 등)

#### 1) 개요: 2024-2025 FAST 검사

2022년 3월, Ron DeSantis 주지사는 새로운 법안을 통해 기존의 학년별 플로리다 학업성취 기준 평가(Florida Standards Assessments, FSA)를 플로리다 학생 사고 평가(Florida Assessment of Student Thinking, FAST)라는 통합 검사 및 진도 모니



터링 시스템으로 대체하는 법안에 서명했다. 이 시스템은 플로리다 주 법령 제 1008.25(9)조에 의거하여 필수적으로 시행되며, B.E.S.T. 기준에 맞춰져 있다. FAST 검사를 통해 3~10학년을 대상으로 읽기(English Language Arts reading, ELA reading)을, 3~8학년을 대상으로 수학 교과목의 학업성취도를 평가한다. 모든 FAST 평가는 컴퓨터화 능력적응검사(CAT, Computer-Adaptive Test)로 진행되는데, 이는 학생이 문항에 올바르게 답할 경우 문항의 난이도가 점차 증가하고, 반대로 답을 틀릴 경우 난이도가 감소하는 방식이다. 그러나 이러한 난이도 조정에도 불구하고, 선택된 문항은 동일한 지식과 기술을 평가하며, 이는 시험 청사진에 의해 결정된다.

FAST 평가는 학기 중 세 번 시행되며, 각 진도 모니터링(Progress Monitoring, PM)은 학년 수준의 전체 시험 청사진을 다룬다. 진도 모니터링 PM은 다음과 같다.

- PM1: 학년 초에 시행되며, 이는 학생의 학습 시작점을 파악하고 교사들이 PM1에서 PM2까지의 학습 진도를 추적할 수 있도록 한다(8월 12일~9월 27일, 2024).
- PM2: 학기 중간에 시행되며, 이 시점까지 학생들은 학년 수준의 기준을 학습하게 되며, PM1의 점수와 비교할 수 있는 중간 점수를 제공한다(12월 2일, 2024~1월 24일, 2025).
- PM3: 학년 말에 시행되며, B.E.S.T. 기준에 대한 학생들의 숙련도를 측정하는 종합 평가 역할을 한다. PM1과 PM2는 정보 제공용이지만, PM3는 학교 평가에 사용된다(5월 1일~5월 30일, 2025).

위의 시험 기간은 각 학교가 최대한 유연하게 시험 일정을 조정할 수 있도록 설계된 것이며 특정 학교가 시험을 실제로 시행하는 데 소요되는 시간을 나타내지 않는다.

PM 1, 2, 3에 대해 구체적으로 설명하자면, PM1은 기준선 평가이고, PM2은 중간 평가의 역할을 한다. 이 두 평가에서 학생들이 아직 학년 수준에 도달하지 않았을 수 있지만, 이는 학생들이 학업 성공 경로에서 완전히 벗어났다는 것을 의미하지는 않는다. 이 결과는 정보 제공의 목적일 뿐이며, 추가적인 지도가 필요한 영역을 식별하는 데 사용된다. 따라서 PM1과 PM2의 결과는 학생 성취도를 나타내는 것으로 간주해서는 안 된다. 반면에, PM3는 학년 말에 종합 점수를 제공하는 최종 평가로, 학생이 해당 학년 수준의 교육 기준을 얼마나 숙달했는지를 측정한다.

## 2) 문항의 특성

### 가) 검사 문항 유형 (Item Types)

FAST 검사에서 읽기(English Language Arts reading, ELA reading) 영역의 문항 유형은 아래와 같이 설계되어 있다.

- 다지선다형(Multiple-Choice): 학생들은 네 개의 답안 중 하나의 정답을 선택한다.
- 선택 가능한 텍스트(Selectable Text): 텍스트에서 발췌된 문장이 제시되는 유형이다. 이 문항은 한 부분 또는 두 부분으로 구성될 수 있다. 두 부분으로 이루어진 문항에서 Part A는 학생에게 분석이나 추론을 요청하고, Part B는 Part A의 답을 뒷받침하는 텍스트를 선택하게 한다.
- 다중 선택(Multiselect): 학생은 제공된 옵션 중 정해진 수의 정답을 선택하도록 지시받는다. 이 유형은 하나의 정답만 선택하는 다지선다형과 다르다.
- 근거 기반 선택형 응답(Evidence-Based Selected Response, EBSR): 두 부분으로 이루어진 문항으로, Part A와 Part B에서 정답을 선택하도록 지시된다. 일반적으로 Part A는 다지선다형이고, Part B는 다지선다형 또는 다중 선택일 수 있다. Part A는 학생에게 분석이나 추론을 요청하고, Part B는 Part A의 답을 뒷받침하는 텍스트를 선택하게 한다.
- 테이블 매치(Table Match): 이 유형은 옵션이 열과 행에 제시된다. 옵션에는 단어, 구절, 문장, 인용문, 문단/지문 번호 또는 이미지가 포함될 수 있다. 학생은 열에서 올바른 옵션과 행에서 올바른 옵션을 매칭하여 상자를 선택하도록 지시받는다. 일반적으로 행 또는 열당 하나의 정답만 있으며, 정답의 수는 다를 수 있다.
- 멀티미디어(Multimedia): 기술을 활용한 콘텐츠는 이미지, 차트, 그래프, 타임라인과 같은 멀티미디어 요소를 포함할 수 있다. 오디오 요소는 3~10학년에서 포함되지 않는다. 멀티미디어 요소는 지문(자극 자료) 또는 평가 문항 내에 나타날 수 있다. 위에서 설명된 어떤 유형의 문항도 멀티미디어 콘텐츠를 평가하는 데 사용될 수 있다. 종이 기반의 대체 평가에서는 멀티미디어 콘텐츠가 수정되거나 동일한 보고 범주를 평가하는 종이 기반 문항 또는 자극 자료로 대체될 수 있다.
- 외부 복사 상호작용(External Copy Interaction): 이 유형은 학생들이 지문에서 텍스트를 직접 선택할 수 있는 버튼을 포함한다. 학생이 외부 복사 상호작용 버튼을 선택하면, 관련 자극 자료의 텍스트가 강조 표시되며 나타난다. 그 후 학생은 질문에 답하는 텍스트를 선택할 수 있다. 이러한 문항은 독립적으로 사용되거나 두 부분으로 이루어진 문항의 일부로 사용될 수 있다. 두 부분으로 이루어진 경우, Part A는 학생에게 분석이나 추론을 요청하고, Part B는 Part A의 답을 뒷받침하는 텍스트를

사용하게 한다. 다른 경우, 두 부분은 독립적으로 존재할 수 있다. 이 유형은 컴퓨터 기반 평가에서만 나타난다.

FAST 검사의 읽기 영역에 이어서 수학 영역의 문항 유형은 아래와 같이 설계되어 있다.

- 다지선다형(Multiple-Choice): 학생들은 네 개의 답안 중 하나의 정답을 선택한다.
- 편집 작업 선택(Editoring Task Choice): 학생은 드롭다운 메뉴에서 수식, 표현, 진술, 또는 다른 요소를 완성할 수 있는 옵션을 클릭한 후, 그 중에서 정답을 선택한다.
- 선택 가능한 핫 텍스트>Selectable Hot Text): 학생은 여러 옵션 중에서 하나 이상의 정답을 클릭하라는 지시를 받는다. 옵션 위에 마우스를 올리면 텍스트가 강조 표시되며, 이는 해당 텍스트가 선택 가능함을 나타낸다. 옵션은 목록, 본문 내에 포함되거나 표로 제시될 수 있으며, 학생은 이를 클릭해 선택한다.
- 다중 선택(Multiselect): 학생은 여러 옵션 중에서 모든 정답을 선택하라는 지시를 받는다. 이는 하나의 정답만 선택하는 다지선다형 문항과 다르다.
- 그래픽 응답 항목 표시(Grid): 학생은 점, 선, 또는 화살표 도구를 사용해 그래프에 응답을 만든다. 또한 숫자, 단어, 구절 또는 이미지를 선택하고 드래그 앤 드롭 기능을 사용해 이를 그래픽에 배치해야 할 수도 있다.
- 수식 편집기(Equation Editor): 학생은 문항에 맞게 숫자, 변수, 표현 또는 수식을 응답 상자에 입력한다. 학생은 응답을 만들기 위해 다양한 수학 기호를 포함하는 도구 모음을 제공받으며, 응답 상자는 문항의 텍스트와 분리되거나 텍스트 내에 포함될 수 있다.
- 매칭 항목(Matching Item): 학생은 열의 머리글이 행의 정보와 일치하는지 여부를 나타내기 위해 상자를 체크한다. 행 또는 열당 정답의 수는 다를 수 있다.

#### 나) 검사 문항 수

FAST 검사에서 읽기(ELA) 과목과 수학 과목의 문항 수는 <표 III-5>, <표 III-6>에 제시되어 있다.

<표 III-5> FAST 읽기 검사 길이

FAST 읽기(ELA) 검사 길이	
학년 군(범위)	문항 수
3-5	36-40 (+ PM3에서의 5개의 시험용 문항)
6-8	36-40 (+ PM3에서의 5개의 시험문항)
9-10	36-40 (+ PM3에서의 5개의 시험문항)

출처: Florida Department of Education(2024u)  
※ 참고: PM3에서 약 5개의 문항은 예비 시험을 위한 시험용 문항이다.

<표 III-6> FAST 수학 검사 길이

FAST 수학 검사 길이	
학년 군(범위)	문항 수
3-5	35-40
6-8	36-40

출처: Florida Department of Education(2024v)

3) 검사 관리 및 시행

가) 대상

법령에 따라, 플로리다의 모든 공립학교 학생들은 주 전역에서 시행되는 FAST 평가에 반드시 참여해야 한다. 학생들은 FAST 검사에서 수학, 읽기 과목의 시험을 치르고, 고등학교 학생들은 수학 과목의 경우 FAST 검사와 구분되는 EOC 시험을 치르게 된다. <표 III-7>은 2023년 봄에 시행된 온라인 시험에서 FAST와 EOC 검사(대수 I, 기하)에 참여한 학생 수를 학년 및 과목별로 제시하고 있다.

<표 III-7> FAST 검사에 참여한 학생 수 (PM3)

수학			읽기		
학년/검사	검사 인원	보고된 인원	학년	검사 인원	보고된 인원
3	221,144	221,088	3	221,662	221,585
4	197,888	197,851	4	201,164	201,109
5	206,364	206,300	5	207,502	207,452
6	211,163	210,954	6	216,412	216,236
7	151,498	151,197	7	209,599	209,397

수학			읽기		
학년/검사	검사 인원	보고된 인원	학년	검사 인원	보고된 인원
8	170,944	170,630	8	215,735	215,503
대수 I (EOC)	254,244	253,628	9	223,090	222,786
기하(EOC)	222,829	222,493	10	262,343	261,928

출처: Florida Department of Education, 2023a

### 나) 컴퓨터화 능력적응검사(CAT)

모든 FAST 검사는 컴퓨터화 능력적응검사로 시행된다. 컴퓨터화 능력적응검사(CAT)는 학생의 능력 수준에 맞춰 적응하는 컴퓨터 기반 평가 방식이다. CAT에서는 각 시험 문항이 컴퓨터 알고리즘에 의해 선택되며, 이는 제시된 시험 청사진을 기반으로 하여 내용 기준을 충족하고, 이전 문항에 대한 학생의 성과를 바탕으로 한다. CAT는 맞춤형 테스트라고도 불리며, 이전 문항이나 문항 세트에 대한 응답 결과에 따라 시험 문항의 난이도가 조정된다. 학생이 질문에 답할 때마다, 그 응답은 다음에 제시될 문항 또는 문항 세트를 결정하는 데 도움이 된다.

컴퓨터화 능력적응검사는 오랫동안 유치원부터 12학년까지의 학교, 고등 교육 기관, 다양한 자격증 및 인증 시험에서 전 세계적으로 사용되어 왔다. CAT 평가의 결과는 평가 대상자가 알고 있는 것과 할 수 있는 것을 유효하고 신뢰성 있게 측정하는 것으로 입증되었다.

### 다) 검사 소요 시간

PM3 평가는 PM1 및 PM2보다 약간 더 길며, 이는 평가에 예비 시험을 위한 시험용 문항(Field Test Items)이 포함되어 있기 때문이다. 시험용 문항(Field Test Items)은 시험 중에 포함된 새로운 문항들을 시험하기 위해 사용되는 문항들로, 학생들의 성적에는 반영되지 않는다. 이러한 문항들은 향후 정식 시험에서 사용할 수 있도록 문항의 질, 난이도, 신뢰성을 평가하기 위해 포함된다. 시험 주최 기관은 시험용 문항을 통해 문항의 성능을 분석하고, 실제 평가에서 사용할지를 결정한다.

모든 FAST 평가는 하루에 한 세션으로 진행되며, 학생들이 하루에 한 과목만 치르는 것이 권장된다. K-12 학생 평가국은 시험 관리 매뉴얼과 대본에서 일정 계획에 대한 지침을 제공한다. 일정 계획을 위해 세션 길이가 제공되지만, 배정된 시간이 끝난 후에

도 작업을 계속하는 학생들은 학교 시간이 끝날 때까지 시험을 완료할 수 있는 시간이 제공된다. 플로리다 주 전역 평가에서는 시간 제한이 시험의 일부가 아니기 때문에, 각 세션별 시간, 총 시험 시간, 그리고 표준 세션 시간 내에 시험을 완료하지 못한 학생들을 위한 연장 시간이 제공된다. 이로써 학생들이 압박 없이 성과를 충분히 발휘할 수 있는 환경을 조성한다. 학생들이 세션에 배정된 시간이 끝났음에도 여전히 문제를 푸는 중일 경우, 학생들은 일반적인 학교 시간의 길이까지 계속해서 시험을 볼 수 있다. PM별 상세한 검사 소요 시간은 아래 <표 III-8>에 제시되어 있다.

<표 III-8> FAST 검사 소요 시간

검사	학년 군(범위)	세션별 시간(분)
FAST ELA 읽기	3-10	90분 (PM1, PM2)
		120분 (PM3)
FAST 수학	3-5	80분 (PM1, PM2)
		100분 (PM3)
	6-8	100분 (PM1, PM2)
		120분 (PM3)

출처: Florida Department of Education(2024b)

#### 라) 예외적인 시험 환경

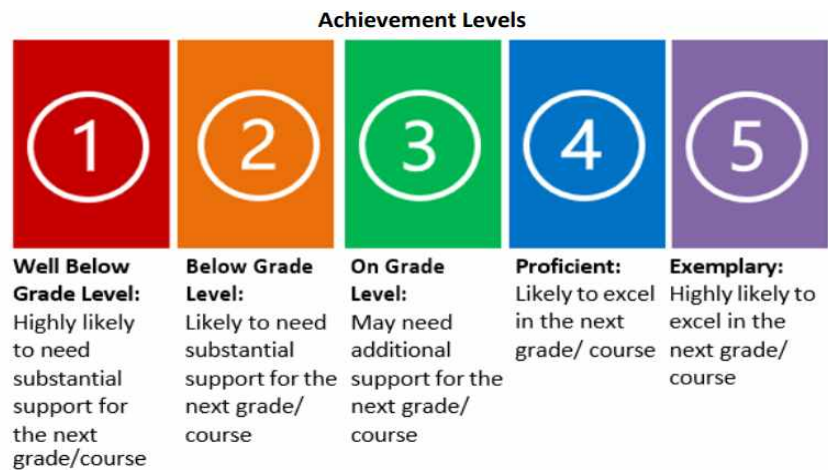
원격 시험은 PM1 및 PM2 기간 동안 3~10학년 FAST 평가에 대해 가상 학교 프로그램에 전일제로 등록된 학생들이나 병원/가정 교육 대상 학생들에게만 제공된다. 원격 시험 관리에 필요한 자료는 플로리다 주 전역 평가 포털에서 확인할 수 있다.

원격 시험은 PM3, B.E.S.T. 쓰기, 5학년 및 8학년 과학 시험, EOC(최종 종합 평가), K-12 FCLE, 또는 FAST ELA 읽기 재시험에서는 제공되지 않는다. 다만, 제한된 예외로 군 복무 중인 가족이 플로리다 외에 주둔 중인 FLVS(플로리다 가상학교) 학생들에게만 원격 시험이 허용된다.

#### 4) 검사 결과 보고 (Grades 3-10 FAST Reporting)

FAST 평가 결과는 학생이 시험을 완료한 후 24시간 이내에 플로리다 성적 보고 시스템(Florida Reporting System, FRS)에 보고된다. 점수는 B.E.S.T. 척도를 기준으

로 보고되며, 이는 주 교육 위원회(State Board of Education)에서 승인한 척도다. 학생들의 성취 수준은 1~5로 나뉘며, 레벨 1이 가장 낮고 레벨 5가 가장 높다. 성취 수준에 대한 설명은 각 학년과 과목에서 학생들이 무엇을 알고 할 수 있어야 하는지를 구체적으로 나타내며, 이는 해당 교육 기준에 따라 정의된다. 예를 들어, 성취 수준 3은 모든 평가에서 학년 수준 성과(On Grade Level)를 나타낸다. 이러한 성취 수준은 학생들의 평가 내용에 대한 성공 정도를 설명한다. FAST 검사 성취 수준별 설명은 [그림 III-2]에서, FAST 검사별 각 성취 수준의 점수 범주는 [그림 III-3]에서 확인할 수 있다.



[그림 III-2] FAST 검사 성취 수준별 설명

출처: Florida Department of Education(2024b)

Table 1. B.E.S.T. Scale Score Ranges for Each Achievement Level on FAST Assessments

Assessment	Level 1	Level 2	Level 3	Level 4	Level 5
Grade 3 ELA Reading	140–185	186–200	201–212	213–224	225–260
Grade 4 ELA Reading	154–198	199–212	213–223	224–236	237–270
Grade 5 ELA Reading	160–205	206–221	222–231	232–245	246–279
Grade 6 ELA Reading	161–208	209–224	225–236	237–249	250–284
Grade 7 ELA Reading	165–214	215–231	232–241	242–256	257–292
Grade 8 ELA Reading	169–219	220–237	238–250	251–261	262–300
Grade 9 ELA Reading	174–223	224–241	242–253	254–266	267–303
Grade 10 ELA Reading	179–229	230–246	247–257	258–270	271–308
ELA Reading Retake	179–229	230–246	247–257	258–270	271–308
Grade 3 Mathematics	140–182	183–197	198–208	209–224	225–260
Grade 4 Mathematics	155–199	200–210	211–220	221–237	238–273
Grade 5 Mathematics	158–206	207–221	222–233	234–245	246–285
Grade 6 Mathematics	168–212	213–228	229–238	239–253	254–287
Grade 7 Mathematics	175–222	223–234	235–246	247–257	258–288
Grade 8 Mathematics	183–226	227–243	244–253	254–262	263–291

[그림 III-3] FAST 검사별 각 성취 수준의 점수 범위

출처: Florida Department of Education(2024b)

## 나. 교과별 보고 범주 (Reporting category)

보고 범주(Reporting category)는 평가된 학생의 지식과 기술을 넓은 내용 영역으로 묶는 역할을 한다. 각 보고 범주는 유사한 기술 등을 기준으로 그룹화하며, 이는 각 학년과 과목에서 평가된다.

### 1) 읽기(English Language Arts, ELA)

FAST 읽기 평가는 B.E.S.T. (Benchmarks for Excellent Student Thinking) 기준에 따라 시행된다. 이 평가는 다양한 서면 자료를 활용하여 학생들의 읽기 이해 능력을 평가한다. 평가에 사용되는 지문은 소설적 지문(Fiction)과 정보적 지문(Informational)이라는 두 가지 유형으로 나뉜다. 정보적 지문(Informational)은 특정 주제에 대한 사실을 제공하며, 잡지나 신문 기사, 사설, 전기문 등이 포함된다. 이와 달리 소설적 지문(Fiction) 지문은 주로 독자의 즐거움을 위해 작성된 글로, 단편 소설,



시, 민담, 소설의 발췌문 등이 해당된다. 소설적 지문(Fiction)과 정보적 지문(Informational)의 비중은 <표 III-9>에 제시되어 있다.

일부 지문 세트에는 그림이나 삽화 등과 같은 멀티미디어 요소가 포함될 수 있다. 대부분의 지문은 출판된 자료에서 선택되지만, 일부는 플로리다 주 전역 평가를 위해 특별히 작성된 경우도 있다. 또한, 플로리다 법령 제 1008.22조에 따라 일부 읽기 지문은 사회 과학 관련 핵심 교과 내용을 포함해야 한다. 모든 학년 수준에서 FAST 읽기는 학생들이 <표 III-10>에 나열된 광범위한 보고 범주에서 무엇을 알고 있고, 무엇을 할 수 있는지를 평가한다.

<표 III-9> FAST 읽기 검사 텍스트 유형

학년 군(범위)	소설적(Fiction)	정보적(Informational)
3-10	50%	50%

출처: Florida Department of Education(2024d)

<표 III-10> FAST 읽기 검사의 보고 범주별 평가 비율

보고 범주 (Reporting Category)	평가 비율
산문과 시 읽기	25-35%
정보 텍스트 읽기	25-35%
다양한 장르 및 어휘 읽기	35-50%

출처: Florida Department of Education(2024d)

## 2) 수학

FAST 수학은 수학 B.E.S.T. (Benchmarks for Excellent Student Thinking) 기준에 기반한다. FAST 수학은 3학년부터 8학년까지의 학생들에게 시행되며 FAST와 구분되는 EOC 검사는 고등학교 학생들을 대상으로 시행된다. FAST 수학과 EOC 검사(대수 I, 기하)별 보고 범주는 <표 III-11>에 나타나 있다.

<표 III-11> FAST 수학 검사의 학년별 보고 범주 및 평가 비율

학년	보고 범주(Reporting Category)	평가 비율
3학년	수 감각과 덧셈 추론	23-29%
	수 감각과 곱셈 추론	
	분수 추론	
	기하학적 추론, 측정, 데이터 분석 및 확률	
4학년	정수에 대한 수 감각과 연산	31-37%
	분수 및 소수에 대한 수 감각과 연산	
	기하학적 추론, 측정, 데이터 분석 및 확률	
5학년	정수에 대한 수 감각과 연산	23-29%
	분수 및 소수에 대한 수 감각과 연산	
	대수적 추론	
	기하학적 추론, 측정, 데이터 분석 및 확률	
6학년	수 감각과 연산	33-42%
	대수적 추론	25-36%
	기하학적 추론, 데이터 분석 및 확률	25-36%
7학년	수 감각과 연산 및 대수적 추론	25-31%
	비례 추론과 관계	22-31%
	기하학적 추론	22-28%
	데이터 분석 및 확률	22-28%
8학년	수 감각, 연산 및 확률	22-28%
	대수적 추론	22-28%
	선형 관계, 데이터 분석 및 함수	25-31%
	기하학적 추론	22-28%
고등학교 (EOC) 대수 I	표현, 함수, 데이터 분석	31-38%
	선형 관계	31-38%
	비선형 관계	31-38%
고등학교 (EOC) 기하	논리, 관계, 정리	33-40%
	합동, 닮음, 작도	27-33%
	측정과 좌표 기하	33-40%

출처: Florida Department of Education(2024v)

3. 요약

텍사스 주에서 시행하는 STAAR 검사와 플로리다 주에서 시행하는 FAST 검사의 시험 목적, 대상 학년, 시험 유형, 시험 관리, 성취 수준, 특수 환경 시험 측면에서 비교 및 분석한 내용은 <표 III-12>에 제시되어 있다.

<표 III-12> 텍사스 STAAR과 플로리다 FAST 검사 비교

항목	텍사스 STAAR	플로리다 FAST
시험 목적	학년 진급 자격 및 졸업 자격 평가. EOC가 포함됨.	학업성취도 추적 및 학습 진도 모니터링. EOC가 포함되지 않음
대상 학년	3-8학년(수학, 읽기/언어, 과학, 사회), 고등학교 EOC 과목	3-10학년(읽기), 3-8학년(수학)
시험 유형	컴퓨터 기반 검사(CBT), 다지선다형, 서술형, 논술형 등	컴퓨터화 능력적응검사(CAT), 다지선다형, 텍스트 강조, 테이블 매치 등
시험 관리	연간 1회(봄), 고등학교 EOC는 연중 시행 가능	연간 3회(PM1, PM2, PM3)
성취 수준	4단계: 미달, 근접, 달성, 초과	5단계: 레벨 1-5
특수 환경 시험	IEP 및 가정학습 학생을 위한 원격 시험, 종이 시험 일부 제공	

이에 더해, 텍사스 주에서 시행하는 STAAR 검사와 플로리다 주에서 시행하는 FAST 검사의 교과목별 보고 범주(reporting category)에 대한 내용은 <표 III-13>을 통해 알 수 있다.

<표 III-13> 텍사스 STAAR 검사와 플로리다 FAST 검사의 교과목별 보고 범주

교과목	텍사스 STAAR	플로리다 FAST
수학	초등~중학교: • 숫자 표현 및 관계 • 계산 및 대수적 관계 • 기하와 측정 • 데이터 분석 및 개인 금융 문해력	초등~중학교: • 수 감각과 덧셈 추론 • 수 감각과 곱셈 추론 • 분수 추론 • 기하학적 추론 • 측정 • 데이터 분석 및 확률 • 정수에 대한 수 감각과 연산 • 분수 및 소수에 대한 수 감각과 연산 • 대수적 추론 • 비례 추론과 관계 등

교과목	텍사스 STAAR	플로리다 FAST
	<p>고등학교:</p> <ul style="list-style-type: none"> <li>• 수와 대수적 방법</li> <li>• 선형 함수</li> <li>• 방정식 및 부등식을 설명하고 그래프로 표현</li> <li>• 선형 함수</li> <li>• 방정식 및 부등식 작성 및 해결</li> <li>• 이차 함수와 방정식</li> <li>• 지수 함수와 방정식</li> </ul>	<p>고등학교: X</p>
읽기	<p>초등~중학교:</p> <ul style="list-style-type: none"> <li>• 읽기</li> <li>• 쓰기</li> </ul>	<p>초등~중학교:</p> <ul style="list-style-type: none"> <li>• 산문과 시 읽기</li> <li>• 정보 텍스트 읽기</li> <li>• 다양한 장르 및 어휘 읽기</li> </ul>
	<p>고등학교:</p> <ul style="list-style-type: none"> <li>• 읽기</li> <li>• 쓰기</li> </ul>	<p>고등학교:</p> <ul style="list-style-type: none"> <li>• 산문과 시 읽기</li> <li>• 정보 텍스트 읽기</li> <li>• 다양한 장르 및 어휘 읽기</li> </ul>
과학	<p>초등~중학교:</p> <ul style="list-style-type: none"> <li>• 물질과 에너지,</li> <li>• 힘, 운동, 에너지,</li> <li>• 지구와 우주,</li> <li>• 생물과 환경</li> </ul>	<p>초등~중학교: X</p>
	<p>고등학교:</p> <ul style="list-style-type: none"> <li>• 세포 구조와 기능,</li> <li>• 유전 메커니즘,</li> <li>• 생물학적 진화와 분류,</li> <li>• 생물학적 과정과 시스템,</li> <li>• 환경 시스템 내 상호 의존성</li> </ul>	<p>고등학교: X</p>
사회	<p>초등~중학교:</p> <ul style="list-style-type: none"> <li>• 역사</li> <li>• 지리 및 문화</li> <li>• 정부 및 시민권</li> <li>• 경제, 과학, 기술, 사회</li> </ul>	<p>초등~중학교: X</p>
	<p>고등학교:</p> <ul style="list-style-type: none"> <li>• 역사</li> <li>• 지리 및 문화</li> </ul>	<p>고등학교: X</p>

교과목	텍사스 STAAR	플로리다 FAST
	<ul style="list-style-type: none"><li>• 정부 및 시민권</li><li>• 경제, 과학, 기술, 사회</li></ul>	



# IV

## 기존 학업성취도 평가의 활용 가능성







## 기존 학업성취도 평가의 활용 가능성

교육중단연구에서 학생들의 교과역량을 파악하기 위해 새로운 교과역량 평가 도구를 개발 및 실시하는 것은 검사 제작의 어려움, 교사의 업무 부담 및 학생의 시험 부담 증가라는 측면에서 한계를 지닌다. 새로운 교과역량 평가 도구를 개발하는 대신 기존에 시행되고 있는 학업성취도 평가 결과를 교육중단연구에 활용할 수 있다면 비교적 효율적으로 학생들의 교과역량에 대한 정보를 수집할 수 있을 것이라 기대된다. 따라서 기존의 학업성취도 평가 결과를 교육중단연구에 활용할 수 있는지를 살펴보고, 실질적인 활용 방안이나 한계점 등에 대해 논의해 보고자 한다.

### 1. 기존 학업성취도 평가 소개

이 장에서는 현재 시행되고 있는 교과역량 관련 평가 중에서도 국가수준 학업성취도 평가, 맞춤형 학업성취도 자율평가, 전국연합학력평가, 성취평가제, 책임교육학년제, 그리고 인천광역시에서 시행하는 기초학력진단평가에 대해 소개하고자 한다.

#### 가. 학업성취도 평가

국가수준 학업성취도 평가와 맞춤형 학업성취도 자율평가는 교육과정에 기반한 학업성취도 평가와 비인지적 영역에 대한 설문 조사를 병행하여 개별 학생의 전인적인 성장 및 발달을 지원하기 위한 목적으로 시행되는 평가이다. <표 IV-1>과 같이 2015 개정 교육과정에 근거하여 영역, 역량, 맥락의 3차원 평가틀을 적용하고 있다.

<표 IV-1> 학업성취도 평가 평가틀

교과	영역	역량	맥락
국어	<ul style="list-style-type: none"> <li>듣기·말하기, 읽기, 쓰기, 문법, 문학</li> </ul>	<ul style="list-style-type: none"> <li>비판적·창의적 사고 역량, 자료·정보 활용 역량, 의사소통 역량,</li> <li>공동체·대인관계 역량, 문화 향유 역량, 자기 성</li> </ul>	<ul style="list-style-type: none"> <li>담화 텍스트, 문학 텍스트, 비문학 텍스트</li> </ul>

교과	영역	역량	맥락
		차별·개발 역량	
수학	<ul style="list-style-type: none"> <li>• (초) 수와 연산, 도형, 측정, 규칙성, 자료와 가능성</li> <li>• (중, 고) 수와 연산, 문자와 식, 함수, 기하, 확률과 통계</li> </ul>	<ul style="list-style-type: none"> <li>• 계산·이해, 추론, 문제해결, 의사소통, 정보처리</li> </ul>	<ul style="list-style-type: none"> <li>• 실생활 중심, 학문 중심</li> </ul>
영어	<ul style="list-style-type: none"> <li>• 듣기, 말하기, 읽기, 쓰기</li> </ul>	<ul style="list-style-type: none"> <li>• 영어 의사소통 역량, 지식정보처리 역량, 공통체 역량</li> </ul>	<ul style="list-style-type: none"> <li>• 개인적, 사회적</li> </ul>
사회	<ul style="list-style-type: none"> <li>• 지리, 역사, 일반사회</li> </ul>	<ul style="list-style-type: none"> <li>• 사실과 개념 이해력, 정보 활용 능력, 비판적 사고력, 문제 해결력 및 의사결정력</li> </ul>	<ul style="list-style-type: none"> <li>• 개인적, 지역·국가적, 전 세계적</li> </ul>
과학	<ul style="list-style-type: none"> <li>• 운동과 에너지, 물질, 생명, 지구와 우주</li> </ul>	<ul style="list-style-type: none"> <li>• 과학 원리의 이해 및 적용 능력, 과학적 탐구 및 문제 해결력, 과학적 의사소통 능력</li> </ul>	<ul style="list-style-type: none"> <li>• 실생활, 순수과학</li> </ul>

출처: 학업성취도 평가 지원 포털 - 학업성취도 평가 개관  
(<https://inaea.kice.re.kr/siteMain/overview.do>)

영역은 국어, 수학, 영어, 사회, 과학 교과에의 내용 영역을 의미하며, 역량은 교과에서 중점적으로 강화하고자 하는 능력으로 2015 개정 교육과정에서 제시한 역량을 재구성한 것이다. 맥락은 해결을 요구하는 문제 상황을 제시하고 있다.

학업성취도 평가는 준거참조평가로 실시되며, 국가 교육과정에 제시된 교과별 성취기준에 근거하여 4개 수준으로 결과를 산출한다. 평가대상 학년의 학생이 도달하기를 기대하는 교육과정 성취기준에 대하여 가장 높은 수준인 4수준은 ‘거의 모든 부분을 이해 및 수행하는’ 수준을, 3수준은 ‘상당 부분을 이해 및 수행하는’ 수준을, 2수준은 ‘부분적으로 이해 및 수행하는’ 수준을, 가장 낮은 1수준은 ‘이해 및 수행을 위해 많은 노력이 필요한’ 수준을 의미한다.

평가 실시 방식의 측면에서는 다양한 유형의 문항을 적용하고, 효율적인 평가 결과의 수집·처리·분석을 위해 국가 수준 학업성취도 평가와 맞춤형 학업성취도 자율평가 모두 2022년부터 지필평가(Paper-Based Test, PBT) 방식에서 컴퓨터 기반 평가(Computer-Based Test, CBT) 방식으로 전환되었다(교육부, 2021; 교육부, 2022).

### 1) 국가수준 학업성취도 평가(박인용 외, 2023; 서민희 외, 2020b)

국가수준 학업성취도 평가는 한국교육과정평가원에서 시행 및 관리하는 평가로서, 우리나라 교육의 질을 관리하고 책무성을 점검하고자 하는 목표 아래, 다음과 같은 세부적인 목적에서 시행되는 평가이다.

- 우리나라 학생들의 교육과정 교육목표에의 도달 정도 및 교과별 학업성취 추이 파악
- 학업성취도와 교육맥락변인(학생, 교사, 학교) 간 관련성 분석을 통해 학업성취에 영향을 미치는 요인 파악
- 학생들의 학업성취 수준에 따른 보정 교육으로의 연계
- 질 높은 평가 도구 개발로 현장의 교수·학습 방법 개선 및 평가 방법 선도

국가수준 교육성취도 평가 기본 계획이 수립된 1998년 이후로 현재까지 교육과정, 교육정책, 사회적 요구의 변화에 발맞춰 평가 시기, 영역, 대상 등의 측면에서 발전을 거듭하였다(박인용 외, 2023). 현재 보고된 2016년부터 2019년까지의 평가 대상, 영역, 범위, 문항 유형의 변화는 <표 IV-2>와 같다.

<표 IV-2> 2016년~2019년 국가수준 학업성취도 평가 개요

연도	평가대상	평가영역	평가범위	평가시간	시행시기	문항유형	
2016	중학교 3학년	국어, 수학, 영어 (설문)	중1~중2 전 과정, 3학년 1학기 과정	교과별 60분 설문지 20분	6월 21일	선다형, 서답형 ※국어, 영어 듣기평가 포함	
	고등학교 2학년		고1 전 과정				
2017	중학교 3학년		중1~중2 전 과정, 3학년 1학기 과정	교과별 60분 설문지 20분	6월 20일		
	고등학교 2학년		고1 전 과정				
2018	중학교 3학년		중1~중2 전 과정, 3학년 1학기 과정	교과별 60분 설문지 30분	6월 19일		
	고등학교 2학년		국어, 영어: 범교과 소재 수학: 수학I, 수학II				
2019	중학교 3학년		중1~중2 전 과정, 3학년 1학기 과정	교과별 60분 설문지 30분	6월 13일		선다형, 서답형, 복합형 ※국어, 영어 듣기평가 포함
	고등학교 2학년		국어, 영어: 범교과 소재 수학: 고등학교 ‘수학’	교과별 50분 설문지 30분			

연도	평가대상	평가영역	평가범위	평가시간	시행시기	문항유형
----	------	------	------	------	------	------

출처: 박인용 외(2023)에서 발췌

2016년에서 2019년까지 평가 영역(교과)은 국어, 수학, 영어로 유지되었다. 평가 범위의 경우, 2019년까지 중학교는 전 학년 전 과정이었으며, 고등학교의 경우, 2017년까지 3학년 1학기 과정, 2018년과 2019년에 국어, 영어는 범교과 소재, 2018년에 수학은 수학I, 수학II, 2019년에 고등학교 ‘수학’ 과정이었다. 전반적으로 국가 교육과정 변화의 흐름에 따라 평가 범위도 달라진 것으로 나타난다. 평가 소요 시간은 교과별로 60분에서 2019년에 50분으로 줄어들었으며, 문항 유형은 선다형, 서답형에서 2019년으로 오면서 복합형이 추가되었다. 2015 개정 교육과정에 의거하여 2019년부터 도입된 복합형 문항은 다양한 성취기준을 측정하기 위해 복합적인 자료를 제시하는 형태의 문항으로, 종합적인 분석 능력 등의 교과역량을 반영하고 있다(서민희 외, 2020a).

큰 틀에서 살펴보면 2000년부터 시행되어 초등학교 6학년, 중학교 3학년, 고등학교 1학년, 고등학교 2학년을 대상으로 표집 평가를 실시해 오다가 2009년부터 2016년까지 전수 평가를 실시하였으며, 2017년부터 현재까지는 중학교 3학년, 고등학교 2학년을 표집하여 평가하고 있다. 초등학교 학업성취도 평가 폐지 및 표집 평가로의 전환은 평가 부담 완화를 위한 결정이었으며, 전수 평가 시 발생할 수 있는 학교 간 서열화 문제를 방지하기 위해 표집 평가로 전환하였다. 표집 방법으로는 시도교육청, 지역 규모, 학교설립유형, 학교성별유형을 고려하여 우리나라 전체 학생의 3%를 2단계 층화군집 표집(two-stage stratified cluster sampling)하는 방법을 사용한다. 평가 대상 학교급, 학년, 교과목 등에 관련된 자세한 내용은 <표 IV-3>과 같다.

<표 IV-3> 2024년 국가수준 학업성취도 평가 개요

학교급	학년	교과	필수 여부	평가 범위	평가 시간	비고
중학교	3학년	국어	필수	중 1~2학년 전 과정, 3학년 1학기 과정	교과별 45분	‘수학’ 교과를 제외하고 음성 자료가 포함된 문항이 검사에 포함되어 있음
		수학	필수			
		영어	필수			
		사회	사회, 과학 중 한 과목 필수 시행			
		과학				
고등학교	2학년	국어	필수	국어, 영어: 범교과 소재 수학: 고등학교	교과별 50분	
		수학	필수			
		영어	필수			

학교급	학년	교과	필수 여부	평가 범위	평가 시간	비고
				'수학' 과목		

출처: 학업성취도 평가 지원 포털 - 국가수준 학업성취도 평가  
(<https://inaea.kice.re.kr/siteMain/nationStandExam.do>)

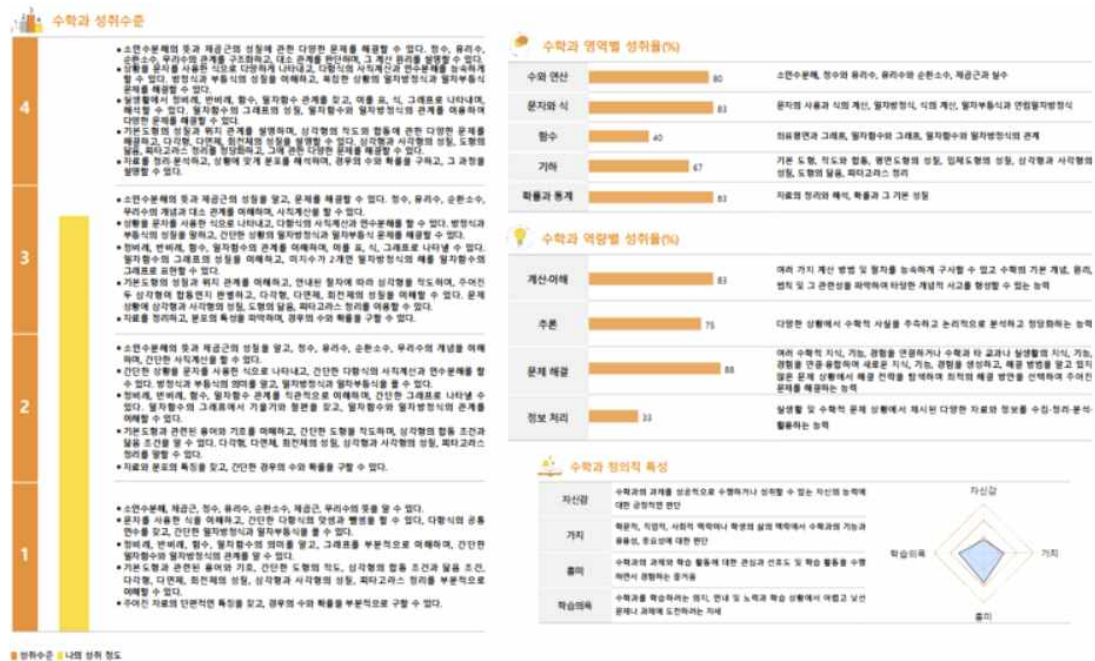
교과와 범위는 중학교 3학년의 경우 중학교 1~2학년 전 과정 및 3학년 1학기 과정을 평가하며, 국어, 수학, 영어, 사회, 과학 교과를 포함한다(사회와 과학 중 한 과목 선택 필수). 고등학교 2학년은 국어, 수학, 영어 교과를 평가하며, 국어와 영어의 경우 범교과 소재를, 수학의 경우 고등학교 '수학' 과목을 다룬다. 또한 학생들의 인지적 특성(교과역량)에 대한 평가와 함께 비인지적 특성(사회·정서적 역량 등)에 대한 조사도 함께 이루어지며, 학교 단위 평가 결과와 학생 단위 평가 결과를 모두 보고한다.

평가 결과는 학교 단위 결과와 학생 단위 결과로 나뉘고, 인지적 영역과 비인지적 영역의 결과를 모두 제공하며, [그림 IV-1], [그림 IV-2]와 같은 형태로 보고된다.



[그림 IV-1] 국가수준 학업성취도 평가 결과표 - 학교 단위

출처: 학업성취도 평가 지원 포털 - 국가수준 학업성취도 평가  
<https://inaea.kice.re.kr/siteMain/nationStandExam.do>



[그림 IV-2] 국가수준 학업성취도 평가 결과표 - 학생 단위

출처: 학업성취도 평가 지원 포털 - 국가수준 학업성취도 평가  
<https://inaea.kice.re.kr/siteMain/nationStandExam.do>

학교 단위 결과에서는 성취수준별 학생 비율과 각 교과 영역별 및 역량별 성취율을 제공한다. 학생 단위 결과에서는 교과별 성취 수준에 대한 설명과 학생의 현재 성취 수준을 표시하며, 각 교과 영역별 및 역량별 성취율을 보고한다. 또한 교과별 정의적 특성(자신감, 가치, 흥미, 학습 의욕)을 함께 제시한다.

국가수준 학업성취도 평가는 준거참조평가(criterion-referenced assessment)로서 성취기준에 따른 성취수준(준거) 설정을 위해 변형된 앙고프 방법(modified Angoff method)을 적용하고 있다. 또한 검사 동등화를 위해 추이 분석용 문항을 개발하여 시행하고 있다. 동등화를 통한 추이 분석을 위해 문항반응이론 기반의 연도별 척도 점수를 개발 및 연계하였으며, 성취수준에 따른 분할점수(cut score)를 연계하였다(서민희 외, 2020).

## 2) 맞춤형 학업성취도 자율평가

2022년 국가수준 학업성취도 평가가 컴퓨터 기반 검사 방식으로 전환됨에 따라, 교수·학습의 환류를 위해 개별 학생의 성취 특성에 대한 맞춤형 정보를 제공하려는 목적에서 맞춤형 학업성취도 자율평가가 도입되었다(박인용 외, 2023). 참여 희망 학교 또는 학급을 대상으로 지정된 평가 기간 중 원하는 시기에 자율적·상시적으로 참여할 수 있는 맞춤형 학업성취도 자율평가는 2022년 초6, 중3, 고2를 대상으로 시작되어 2023년 초5, 고1, 2024년 초3, 초4, 중1, 중2와 같이 연차적으로 대상 학년을 확대하였다. 평가의 시행 시기는 2024년을 기준으로 2월 중순에서 4월 중순, 즉 학기 말에서 다음 학기 초까지로 설정하고 있으며, 시행 시기에 따라 새로운 검사를 제공한다. 따라서 학교에서는 시행 일자를 자율적으로 선택하여 총합평가 또는 진단평가의 목적으로 활용할 수 있다. 2024년 현재 평가 대상 학교급과 학년, 교과, 범위, 소요 시간에 관련된 자세한 내용은 <표 IV-4>와 같다.

<표 IV-4> 맞춤형 학업성취도 자율평가 개요

학교급	학년	교과	평가 범위	시간
초등학교	초3	문해력, 수리력	초 1~2학년 범위의 문해력, 수리력 내용	교과별 40분
	초5	국어, 수학, 영어, 사회, 과학	초 4학년 전범위	
	초6		초 5학년 전범위	
중학교	중1	국어, 수학, 영어, 사회, 과학	초 6학년 전범위	교과별 45분
	중3		중 2학년 전범위	
고등학교	고1	국어, 수학, 영어, 사회, 과학	중 3학년 전범위	교과별 50분
	고2	국어, 수학, 영어	고 1학년 전범위	

출처: 학업성취도 평가 지원 포털 - 맞춤형 학업성취도 자율평가  
(<https://inaea.kice.re.kr/siteMain/atnmsExam.do>)

대상 학년은 초등학교의 경우 3, 5, 6학년, 중학교의 경우 1, 3학년, 고등학교의 경우 1, 2학년을 대상으로 하며, 평가 범위는 직전 학년에 속하는 전 범위에 해당한다. 초등학교 3학년은 문해력, 수리력에 대한 기초 지식을 평가하며, 초등학교 5학년부터 고등학교 1학년까지는 국어, 수학, 영어, 사회, 과학, 고등학교 2학년은 국어, 수학, 영어를 평가한다. 평가 소요 시간은 학교급별로 차이를 두고 있으며, 교과별로 초등학교 40분, 중학교 45분, 고등학교 50분으로 지정하고 있다.

맞춤형 학업성취도 자율평가에서는 학생의 전인적인 성장을 돕기 위해 교과 영역뿐만 아니라 비교과 영역의 평가도 함께 실시하고 있으며, 사회적 역량(공동체 의식, 협업, 갈등 해결), 정서적 역량(스트레스 대처, 회복 탄력성), 학교생활(수업 준비 및 참여도, 학교생활 행복도, 학생-교사 관계), 진로 성숙도(자기 이해, 진로 탐색, 진로 설계), 자기조절학습(학업적 자기효능감, 행동 통제, 학습 전략), 성취목표지향성(수행 접근 목표, 수행회피 목표, 숙달접근 목표)를 평가한다.

평가 결과로는 개별 학생의 인지적·비인지적 특성에 대해 종합적인 맞춤형 피드백을 제공한다. 맞춤형 학업성취도 자율평가에서 제공하는 학생 대상의 교과별 결과표 및 종합결과표 예시는 [그림 IV-3], [그림 IV-4]와 같다.

## 맞춤형 학업성취도 자율평가 결과 [국어]

· 학생의 지도 학업 성취수준은 3수준 이며, 학습결과상 3수준 학생이 많고, 할 수 있는 구체적인 내용은 3수준의 진술만을 할 수 있음.

· [지이표영역별 성취율] 각각 10개 항목의 성취율만 고려한 점, 듣기·말하기, 읽기·쓰기, 어휘·문법 영역을 말하는 것이며, 쓰기 영역은 좀 더 결합할 수 있도록 노력이 필요합니다.

**[지도작성·영양정보 작성권]** 영남도청 지도작성과 영양정보관리팀을 비롯하여, 양곡정책사업담당, 지방·영양협력담당, 인식조사담당, 식품안전담당, 저가영양, 계통영양 담당(이)가 담당하고 있으며, 농림축산식품부·농업·다민관계영양정보(별) 산정하기 위한 노력이 필요합니다.

· [국어의 정의와 특성] 국어 학습에 대한 자신감, 국어(語) 학습하고자 하는 학습의욕(意欲) 높고, 국어의 중요성에 대한 가치 인식(認識) 높음, 국어에 대한 흥미(興味) 높습니다.

[illegible]

[그림 IV-3] 맞춤형 학업성취도 자율평가 학생 대상 교과별 결과표(국어)

출처: 학업성취도 평가 지원 포털 - 맞춤형 학업성취도 자율평가  
<https://inaea.kice.re.kr/siteMain/atnmsExam.do>



맞춤형 학업성취도 자율평가 결과



[그림 IV-4] 맞춤형 학업성취도 자율평가 학생 대상 종합결과표

출처: 학업성취도 평가 지원 포털 - 맞춤형 학업성취도 자율평가  
<https://inaea.kice.re.kr/siteMain/atnmsExam.do>

교과별 평가 결과표에서는 교과의 4가지 성취수준과 개별 학생에게 해당하는 성취수준의 일반적인 성취 특성에 대한 정보를 제공하고 있다. 또한 개별 학생의 교과에 대한 영역별 성취율과 역량별 성취율, 정의적 특성에 대한 정보를 함께 제공한다. 종합 결과표는 문해력·수리력 검사만 보는 초등학교 3학년을 제외한 모든 학년에 제공되며, 교과별 성취수준과 교과 기반 정의적 특성 수준, 비교과 설문 결과를 막대그래프 형태로 보고한다.

맞춤형 학업성취도 자율평가는 컴퓨터 기반 검사 방식을 취하고 있으므로 기본적으로 PC 또는 노트북으로 참여할 수 있으며, 태블릿 PC 버전도 함께 제공하고 있다. 태블릿 PC 환경에서 검사에 참여하기 위해서는 한국교육과정평가원에서 개발한 'iNAEA'이라는 전용 애플리케이션을 기기에 설치해야 한다.

## 나. 전국연합학력평가

전국연합학력평가는 고등학교 재학생들의 학력 수준을 측정하고 대학수학능력시험

을 준비하기 위하여 전국의 시도교육청(2024년 기준 서울특별시교육청, 부산광역시교육청, 인천광역시교육청, 경기도교육청)에서 돌아가며 주관하는 평가로, 시험 일정과 응시 학년이 고정적이지 않다. 일반적으로 각 시도교육청에서 문항 출제 등의 업무를 담당하고, 채점 등의 업무는 한국교육과정평가원에서 수행한다. 2024년 학년별 전국연합학력평가의 시행 시기와 주관 시도교육청, 평가 영역은 <표 IV-5>와 같다.

<표 IV-5> 2024년 전국연합학력평가 개요

시행	주관	평가 영역		
		고1	고2	고3
3월	서울	국어, 수학, 영어, 한국사, 통합사회, 통합과학	국어, 수학, 영어, 한국사, 사회탐구, 과학탐구	국어, 수학, 영어, 한국사, 사회탐구, 과학탐구
5월	경기	-	-	국어, 수학, 영어, 한국사, 사회탐구, 과학탐구
6월	부산	국어, 수학, 영어, 한국사, 통합사회, 통합과학	국어, 수학, 영어, 한국사, 사회탐구, 과학탐구	국어, 수학, 영어, 한국사, 사회탐구, 과학탐구, 직업탐구, 제2외국어 및 한문
7월	인천	-	-	국어, 수학, 영어, 한국사, 사회탐구, 과학탐구
9월	인천	국어, 수학, 영어, 한국사, 통합사회, 통합과학	국어, 수학, 영어, 한국사, 사회탐구, 과학탐구	국어, 수학, 영어, 한국사, 사회탐구, 과학탐구, 직업탐구, 제2외국어 및 한문
10월	경기	국어, 수학, 영어, 한국사, 통합사회, 통합과학	국어, 수학, 영어, 한국사, 사회탐구, 과학탐구, 직업탐구, 제2외국어 및 한문	국어, 수학, 영어, 한국사, 사회탐구, 과학탐구, 직업탐구, 제2외국어 및 한문

출처: EBSi 홈페이지

(<https://www.ebsi.co.kr/ebs/ent/enta/retrieveExmSchedRng.ebs>)

평가 영역은 크게는 국어, 수학, 영어, 한국사, 탐구(사회/과학/직업탐구), 제2외국어 및 한문이다. 평가 문항은 국어, 영어, 한국사, 탐구 영역 모두 선다형으로 구성되어 있으며, 수학의 경우 선다형뿐만 아니라 단답형 문항을 포함한다. 또한 영어 영역에서는 듣기 평가가 시행된다.

주관 시도교육청 및 한국교육과정평가원에서 시험지 채점 및 성적표 인쇄 업무를

완료하면 학생들은 영역별 원점수, 표준점수, 학급 및 학교석차, 백분위, 등급, 정오표, 문항별 정답률, 보충학습이 필요한 문항 번호 등에 대해 상세한 정보를 제공받는다. 평가가 종료된 이후, 주관 시도교육청 홈페이지를 통해 영역별 문항, 정답 및 해설, 통계자료가 공개되며, 통계자료의 경우, 전국 단위의 기초자료 수준에서만 정보가 제공되므로 개별 학생 수준이나 문항 수준의 분석은 불가하다.

#### 다. 성취평가제

성취평가제란 교육부의 「중등학교 학사관리 선진화 방안(2011.12.13.)」에 따라 도입된 평가 제도로써, 상대 서열을 중시하는 규준참조평가 방식에서 교과별 성취기준에 근거한 준거참조평가 방식으로의 전환을 꾀하였다. 중학교의 경우, 2012학년도 1학년 입학생부터 전 교과에서 시행하였으며, 2014학년도에 전 학년에 적용되었다. 고등학교의 경우, 2012학년도 1학년 입학생부터 전문교과에 도입하였으며, 시범 운영기간이 끝난 후, 2014학년도 고등학교 1학년 입학생부터 보통교과에 적용하기 시작하여 2016학년도부터는 전체 고등학교에서 운영되었다.

중학교의 경우, 국어, 수학, 영어, 사회, 과학, 음악, 미술, 체육, 도덕, 기술·가정, 그 외 선택교과 등에 도입되었으며, 고등학교의 경우, 보통교과, 전문교과(마이스터고 등)에 도입되었다. 학생의 학업 성취 수준에 대해 교육과정에 근거하여 개발된 교과목별 성취기준에 비추어 평가하며, ‘A-B-C-D-E’ 또는 ‘A-B-C’ 또는 ‘P’와 같이 학생들의 성취기준 도달 정도를 준거 점수의 형태로 나타내게 된다(행정안전부 국가기록원, 2024.09.09.). 한국교육과정평가원의 학생평가지원포털 사이트에서는 성취평가제 시행을 위해 필요한 초·중·고의 교과별 성취기준과 성취수준, 예시 자료집 등에 대한 정보를 제공하고 있다. 또한 해당 사이트를 통해 성취평가 계획을 수립하고, 평가 결과를 입력, 산출, 관리할 수 있다. 그러나 예시 자료집이나 평가 계획 수립 및 결과 입력 시스템 등 전반적인 사이트 이용을 위한 접근 권한은 현직 교사에게만 부여되고 있다.

#### 라. 책임교육학년제(교육부, 2023.07.11.)

책임교육학년제란 학생들의 학업적 성장에 있어서 결정적인 시기인 초등학교 3학년, 중학교 1학년을 책임교육학년으로 지정하고, 교육격차를 막기 위해 집중적으로 지원하려는 목적에서 수립된 제도이다. 초등학교 3학년을 읽기·쓰기·셈하기에 기초하여 본격적으로 교과 학습을 시작하는 단계이고, 중학교 1학년을 초등교육을 기반으로 중등교

육이 시작되는 단계라는 점에서 책임교육학년으로 지정되었다. 책임교육학년제의 지원 내용은 학생들의 성취수준을 진단하기 위한 컴퓨터 기반 학업성취도 평가 시행, AI 맞춤형 학습 및 학습 관리 튜터링 연계 제공, 학습지원 대상의 확대 적용, 방학 중 기초 소양(언어·수리·디지털 능력) 강화를 위한 계절학기 운영이 있다. 또한 중학교 1학년의 경우, 자유학기제를 내실화하기 위해 진로·적성의 진단과 진로 체험, 인성 교육 프로그램 등을 제공하게 되며, 지원 대상을 기초학력 미달 학생(전체의 약 5%)에서 중·하위 수준 학생(전체의 약 30%)으로 확대하고자 한다.

## 마. 기초학력진단평가

기초학력진단평가란 코로나19로 인한 학습 결손 발생에 대응하고자 학생들의 교과별 기초학력 부진 영역을 조기에 진단 및 보정하기 위해 실시하는 검사이다. 교육부에서 발표한 ‘제1차 기초학력 보장 종합계획(2023~2027)’에 따르면 기초학력은 3R’s(읽기, 쓰기, 셈하기)와 직결되며, 학기 초 진단검사를 실시하고 교사의 보정 지도 및 향상도 진단이 학기 중 지속되는 구조로 이루어진다(교육부 공식 블로그, 2022.10.26.). 기초학력진단평가를 위해 전국의 시도교육청 차원에서 평가 자료 및 보정 학습을 위한 맞춤형 교육 자료 등을 지원하고 있다.

인천에서는 초등학교 3학년~고등학교 2학년 학생 중 학습 부진 및 경계 학생을 대상으로 하며, 대상 과목은 국어, 사회, 역사/한국사, 수학, 과학, 영어이다. 초등학교 3학년의 경우, 읽기, 쓰기, 셈하기를 평가하며, 초등학교 4학년부터는 교과목별 평가를 시행하는데, 교과 내에서도 읽기, 쓰기, 셈하기와 같은 영역을 구분한다. 평가 방식은 온라인 검사와 오프라인 검사를 모두 지원하고 있다(인천 교육청 기초학력 진단 보정 시스템 홈페이지). 또한 기초학력 진단 보정 시스템 홈페이지를 통해 기초학력 진단검사 도달 기준점수(정답한 문항의 수로 결정)를 공시하며, 학교급에 따른 영역별(읽기, 쓰기, 셈하기/수와 연산/수 이해와 활용) 기준점수 및 교과별 기준점수를 각각 제시한다. 기준점수에 도달하지 못한 학생은 학습 결손이 있는 것으로 판단되며, 미도달 학생을 대상으로 보정 학습 및 지도와 향상도 진단이 이루어진다.

기초학력진단평가의 경우, 검사 목적이 ‘기초학력 부진 학생 파악’이므로 최소한의 성취기준만을 반영하며, 검사 난이도가 매우 쉽다는 특징이 있다. 즉, 기초학력 도달 여부를 넘어서 학생들의 전반적인 학습 상태나 성장 및 발달을 면밀히 파악해야 하는 종단연구의 목적과는 다소 부합하지 않아 활용이 어려울 것으로 판단된다. 그러나

2022 개정 교육과정을 반영하여 학생들의 교과역량 측정을 위해 ‘읽기, 쓰기, 셈하기’라는 역량 기반 평가를 실시하고 있으며, 그에 따른 기준점수를 설정한다는 점은 주목할 만하다. 따라서 인천교육중단연구에서 학생들의 교과역량을 파악하기 위해 기초학력 진단평가 결과를 직접적으로 활용하는 것은 타당하지 않으나, 기초학력진단평가의 평가틀이나 기준점수 설정 방식 등 전반적인 구조를 참고할 수 있을 것으로 보인다.

## 2. 기존 학업성취도 평가의 중단연구 활용 사례

### 가. 한국교육중단연구(KELS2005, KELS2013)

한국교육중단연구는 한국교육개발원(KEDI)에서 2005년부터 수행하고 있는 국가 수준의 교육 조사 연구로서 학생의 학업적 경험과 성취에 대한 중단자료를 구축하고 지적, 정의적, 사회·문화적 성장에 영향을 미치는 학교에 대한 자료, 상급학교 진학과 교육 이동에 대한 자료를 수집함으로써 정성적 차원의 통계자료를 제공하고 교육정책의 수립 및 평가를 위한 기초자료를 제공하는 데 목적이 있다.

제16회 한국교육중단연구 학술대회 데이터 자료집(한국교육개발원, 2022.11.25.)에 따르면 한국교육중단연구2005에서는 학생의 인지적 발달 수준을 조사하기 위해 중학교 1학년부터 3학년까지 별도의 학년별 성취도 평가를 시행하였으나, 고등학생의 경우에는 기존의 학업성취도 평가 결과를 활용하였다. 5차년도에는 고등학교 1학년의 국가수준 학업성취도 평가(2008년 10월 시행) 점수를, 6차년도에는 고등학교 3학년의 대학수학능력시험 점수(2011학년도)를 조사하여 연계하였다.

한국중단연구2013에서는 초등학교 5학년부터 중학교 3학년까지는 학년별 학업적 기초능력검사를 시행하고 있으나, 고등학생의 경우에는 한국교육중단연구2005와 마찬가지로 기존의 학업성취도 평가 결과를 활용하고 있다. 학생 설문지(6차년도~8차년도)와 교사의 학생 평가지(6차년도, 7차년도)를 통해 학생의 전국연합학력평가 결과(국어, 수학, 영어 등급)를 조사하였다.

### 나. 한국복지패널(KWPS)

한국복지패널조사는 IMF 외환위기 이후 빈곤층, 근로빈곤층(working poor), 차상위

층(near poor)의 규모 및 생활 실태에 대한 변화 양상을 파악하고 복지 정책의 효과성을 제고하기 위한 목적으로 종단연구를 실시하고 있다. ‘2005년 인구주택총조사 자료에서 확률비례추출한 ‘2006년 국민생활실태조사’ 최종 조사 완료 가구의 소득 자료에 근거하여 일반 가구와 저소득층 가구를 구분하였으며, 두 계층으로부터 각각 3,500가구씩 총 7,000가구를 표본으로 선정하였다. 그 결과 최종 패널로 구축된 표본 가구는 총 7,072가구였으며, 표본 추출 과정에서 저소득층 가구는 예상되는 패널 소실 및 통계적 유의미성을 고려하여 과대표집하였다. 조사 방법은 조사원이 패널 가구를 직접 방문하여 응답자가 응답한 내용을 조사원이 기록하는 형태를 원칙으로 한다.

한국복지패널조사에서도 아동의 학업적 성장 및 발달에 대한 사항을 조사하고 있으며, ‘아동부가조사’ 내용 중 아동의 ‘학교생활’에 대한 하위 영역으로서 ‘학교 성적’이 이에 해당한다. ‘아동부가조사’는 3년 주기로 시행되며, 2006년 1차 조사 당시 초등학교 4, 5, 6학년 아동들을 표집하여 코호트 연구를 실시하였다. 자세한 문항 내용은 [그림 IV-5]와 같다.

문 3) (모든 응답자) 지난 1년 동안 여러분의 학교 성적은 어떻습니까? 다른 사람에게 알려지지 않으므로 솔직하게 응답하여 주시기 바랍니다.

문항	아주 못함	중간 이하	중간	중간 이상	아주 잘함
가 전과목 평균	①	②	③	④	⑤
나 국 어	①	②	③	④	⑤
다 수 학	①	②	③	④	⑤
라 영 어	①	②	③	④	⑤

[그림 IV-5] 한국복지패널 아동부가조사 학교생활 학교 성적 문항

출처: 한국복지패널조사(<https://www.koweps.re.kr:442/data/survey/list.do>)

아동의 학업적 능력으로 학교 성적(전 과목 평균, 국어 성적, 수학 성적, 영어 성적)에 대해 ‘아주 못함, 중간 이하, 중간, 중간 이상, 아주 잘함’의 5점 척도에 자기보고식으로 응답하도록 하고 있다.

#### 다. 한국교육고용패널조사 I, II(KEEP I, II)

한국교육고용패널조사 I, II는 인적자원개발 관련 정책의 기초자료로서 청년층의 교육 관련 정보를 조사하기 위한 목적에서 실시되었으며, 교육과 노동시장 간의 관계를 지속적으로 분석하고 데이터베이스를 구축하고자 한국직업능력개발원에서 진행한 연

구이다.

한국교육고용패널조사 I 은 2004년부터 2015년까지 중학교 3학년 코호트 2,000명과 고등학교 3학년 코호트 4,000명을 대상으로 조사를 진행하였으며, 학생, 학부모, 학교행정가, 담임교사 설문을 실시하였다. 학생의 학업성취도 관련 정보를 조사하기 위해 학생 설문지와 담임교사 설문지를 통해 학급 석차 또는 전교 석차, 등급 등의 정보를 입력하도록 하거나 학교 내신 성적을 9등급으로 분류하여 주관적으로 응답하게끔 하였다.

한국교육고용패널조사 II는 한국교육고용패널조사 I 의 후속 조사로서, 2016년부터 고등학교 2학년 패널 10,558명을 대상으로 하며, 학생, 학부모, 학교행정가, 담임교사 설문을 실시하고 있다. 학업성취도 정보를 수집하기 위해 담임교사 설문지와 학생 설문지에서 전체 성적 및 교과별 내신 성적을 9등급으로 분류하여 주관적으로 응답하게끔 하였다. 문항 예시는 [그림 IV-6]과 같다.

11. 본인의 고등학교 3학년 1학기 교과목 내신등급을 표시해 주십시오.

▶ 9) '전문교과' 는 특성화고 및 마이스터고 학생들만 응답해 주십시오.

교과목	상			중			하		
	1등급 4%이하	2등급 4%초과 ~11%이하	3등급 11%초과 ~23%이하	4등급 23%초과 ~40%이하	5등급 40%초과 ~60%이하	6등급 60%초과 ~77%이하	7등급 77%초과 ~89%이하	8등급 89%초과 ~96%이하	9등급 96%초과
1) 국어 [Y17SA01047]	①	②	③	④	⑤	⑥	⑦	⑧	⑨
2) 수학 [Y17SA01048]	①	②	③	④	⑤	⑥	⑦	⑧	⑨
3) 영어 [Y17SA01049]	①	②	③	④	⑤	⑥	⑦	⑧	⑨
4) 과학 영역 [Y17SA01050]	①	②	③	④	⑤	⑥	⑦	⑧	⑨
5) 사회 영역 [Y17SA01051]	①	②	③	④	⑤	⑥	⑦	⑧	⑨
6) 음악 [Y17SA01052]	①	②	③	④	⑤	⑥	⑦	⑧	⑨
7) 미술 [Y17SA01053]	①	②	③	④	⑤	⑥	⑦	⑧	⑨
8) 체육 [Y17SA01054]	①	②	③	④	⑤	⑥	⑦	⑧	⑨
9) 전문교과 [Y17SA01055]	①	②	③	④	⑤	⑥	⑦	⑧	⑨

[그림 IV-6] 한국교육고용패널조사 II 학생 내신 등급 관련 문항

출처: 한국교육고용패널조사(<https://www.krivet.re.kr/kor/sub.do?menuSn=20>)

### 3. 요약

기존 학업성취도 평가와 평가 결과를 활용한 사례에 대해 요약한 내용은 <표 IV-6>, <표 IV-7>과 같다.

<표 IV-6> 기존 학업성취도 평가 종류별 개요

분류	대상	영역	비고
2024 국가수준 학업성취도 평가	중3	국어, 수학, 영어, 사회, 과학	중 1~2학년, 3학년 1학기 범위
	고2	국어, 영어, 수학	범교과 소재, 고등학교 '수학'
맞춤형 학업성취도 자율평가	초3	문해력, 수리력	초 1~2학년 범위
	초5	국어, 수학, 영어, 사회, 과학	초 4학년 전범위
	초6		초 5학년 전범위
	중1	국어, 수학, 영어, 사회, 과학	초 6학년 전범위
	중3		중 2학년 전범위
	고1	국어, 수학, 영어, 사회, 과학	중 3학년 전범위
	고2	국어, 수학, 영어	고 1학년 전범위
전국연합 학력평가	고1	국어, 수학, 영어, 한국사, 통합사회, 통합과학	규준참조평가
	고2	국어, 수학, 영어, 한국사, 사회탐구, 과학탐구 (직업탐구, 제2외국어 및 한문)	규준참조평가
	고3	국어, 수학, 영어, 한국사, 사회탐구, 과학탐구 (직업탐구, 제2외국어 및 한문)	규준참조평가
성취평가제	중1~3	국어, 수학, 영어, 사회, 과학, 도덕, 기술·가정, 음악, 미술, 체육, 선택교과 등	준거참조평가 (성취기준-성취율)
	고1~3	보통교과(전체 고교), 전문교과(마이스터고 등)	준거참조평가 (성취기준-성취율)
책임교육 학년제	초3, 중1	언어, 수리, 디지털 능력 등	컴퓨터 기반 학업성취도 평가
기초학력 진단평가	초3~고2	읽기, 쓰기, 셈하기 국어, 사회, 역사/한국사, 수학,	준거참조평가 온라인/오프라인



분류	대상	영역	비고
		과학, 영어	

<표 IV-7> 기존 학업성취도 평가 결과 활용 사례

분류	대상	영역	비고
한국교육 종단연구2005	고1~3	고1 국가수준 학업성취도 평가, 고3 대학수학능력시험 점수	
한국교육 종단연구2013	고1~3	전국연합학력평가 등급 (국어, 영어, 수학)	
한국복지패널	초4~6 (2006 기준)	전과목 평균, 국어, 수학, 영어 (자기보고식 5점 척도)	코호트조사
한국교육고용 패널조사 I	중3, 고3 (2004 기준)	학급 및 전교 석차, 등급 (자기보고식 설문)	코호트조사
한국교육고용 패널조사 II	고2	전체 및 교과별 내신 등급 (자기보고식 설문)	패널조사



| 인천교육종단연구 기반연구 - 교과역량 평가 방안 제안 |

V

**학교급별  
교과역량  
평가 방안**





## 학교급별 교과역량 평가 방안

본 장에서는 앞서 분석한 국내외 교육중단연구 사례를 종합하여 인천교육중단연구의 학교급별 교과역량 조사 및 평가 방안에 대한 시사점을 도출하고자 한다. 먼저 국내 교육중단연구의 조사 대상과 인지 영역 조사 내용을 종합하여 학교급별 교과역량 조사 및 평가 방안을 정리하였고, 국외 교육중단연구 사례로 텍사스 주 STAAR 검사와 플로리다 주 FAST 검사를 비교하여 제시하였다. 또한 기존 학업성취도 평가 및 평가 결과 활용과 관련하여 학교급별 조사 방식의 차이점을 제시하고자 한다.

### 1. 국내 교육중단연구 학교급별 교과역량 평가 방안

본 장에서는 앞서 분석한 국내 교육중단연구의 조사 대상과 인지 영역 조사 내용을 종합하여 학교급별 교과역량 평가 방안을 정리하였다. 각 연구의 주요 조사 대상과 인지 영역 조사 내용은 <표 V-1>과 같다. 대부분의 연구가 초등학교 고학년부터 고등학교까지의 학생들을 대상으로 하고 있으며, 일부 연구는 학부모, 교사, 학교 행정가 등도 포함하고 있었다. 연구마다 조사 대상의 규모와 범위가 다양한데, 작게는 2,000명 정도부터 크게는 10,000명 이상까지 다양하였으며 특정 학년을 기준으로 시작하여 장기간 추적 조사를 하였다.

<표 V-1> 국내 교육중단연구 주요 조사 대상 및 인지 영역 내용

연구명	조사 대상			인지 영역 조사 내용
	초	중	고	
한국청소년패널조사(KYPS)	●	●		- 중2 패널: 지난 학기 반 성적, 전교 성적, 전국 모의고사 성적 직접 기입, 과목별로 지난 학기 반 성적 주관적 평가 - 초4 패널: 과목별로 지난 학기 반 성적 주관적 평가
한국아동·청소년 패널조사2010(KC YPS 2010)	●	●		- 초1 패널 : 과목별 주관적 평가, 전체 성적 만족도 - 초4 패널 : 과목별 주관적 평가, 전체 성적 만족도 - 중1 패널 : 과목별 점수(등급), 과목별 주관적 평가, 전체 성적 만족도 - 초등 패널의 중학교 진학 이후 : 과목별 점수, 주관적

연구명	조사 대상			인지 영역 조사 내용
	초	중	고	
				평가, 성적 만족도, 성적 주관적 평가 - 중등 패널의 고등학교 진학 이후 : 과목별 주관적 평가, 성적 만족도, 성적 주관적 평가
한국아동·청소년 패널조사2018(KC YPS 2018)	●	●		- 초4, 중1 패널 : 전 교과 성취도의 주관적 평가, 학업성취 만족도
한국교육고용패널 조사 I (KEEP I )		●	●	- 중3, 고3 패널 : 과목별 성취도, 선호도, 석차 및 등급, 수능 점수
한국교육고용패널 조사 II (KEEP II )			●	- 고2 패널 : 과목별 성취도, 선호도 주관적 평가, 국가학업 성취도평가
한국교육종단연구 2005(KELS2005)		●	●	- 중1 패널 : 국어, 영어, 수학 성취도 평가 - 고등학교 진학 이후 : 성취도 평가 실시 제외, 국가수준 평가 결과 연계
한국교육종단연구 2013(KELS2013)	●			- 초5 패널 : 국어, 영어, 수학 기초능력 검사 - 중학교 진학 이후 : 국어, 영어, 수학 학업성취도 검사 - 고등학교 진학 이후 : 전국 수준(등급) 수집
한국 초·중등 학교교육 성과 종단연구	●			초등학교 4, 5학년 대상 국어, 수학, 문제해결역량 조사 예정(예비검사만 실시)
서울교육종단연구 2010	●	●	●	초4, 중1, 고1 패널 : 국어, 수학, 영어 학업성취도검사
서울교육종단연구 2020	●	●	●	- 초4, 중1 패널 : 교과역량(국어, 수학, 영어) - 고1 패널 : 없음
경기교육종단연구	●	●	●	초4, 중1, 고1 패널 : 국어, 수학, 영어 학업성취도검사
경기기초학력종단 연구	●			초2 패널 : 기초 문식력, 기초 수리력 지필평가
전남교육종단연구	●	●		초4, 중1 패널 : 국어/사회, 수학/과학 인지영역 검사

국내 12개의 교육종단연구를 살펴보았을 때, 대부분의 연구가 초등학교 고학년부터 고등학교까지의 학생들을 대상으로 하고 있었으며, 일부 연구는 학부모, 교사, 학교 행정가 등도 포함하고 있었다. 연구마다 학교급별 또는 진학 이후 조사 내용이 크게 바뀌지는 않았으나 한국교육종단연구2005, 2013에서는 고등학교 진학 이후 성취도 평가를 시행하지 않고 국가수준 평가 결과를 연계하여 조사하였으며 서울교육종단연구

2020에서는 학교급별로 조사 역량에 차이를 두어 초등학교와 중학교에서는 국어, 수학, 영어 인지 역량을 측정하였고 고등학교에서는 인지 역량을 측정하지 않았다.

한국청소년패널조사는 <표 V-2>와 같이 학생 설문문을 통해 인지 영역을 조사하였는데, 중2 패널은 지난 학기 반 성적, 전교 성적, 전국 모의고사 성적 등을 직접 기입하게 하였다. 또한 과목별 지난 학기 반 성적에 대해 5점 리커트 척도(매우 못하는 수준, 못하는 수준, 중간, 잘하는 수준, 매우 잘하는 수준)로 평가한 내용을 조사하였는데, 대상 과목은 국어, 영어, 수학, 사회영역, 과학영역, 음악, 미술 체육이었다. 초4 패널은 과목별로 지난 학기 반 성적에 대한 주관적 평가만 조사하였다.

<표 V-2> KYPS 조사 문항

조사 대상	조사 내용	조사 방식
중2	지난 학기 반 성적 (반 전체 인원, 반 등수, 반 성적 비율)	직접 기입
	지난 학기 전교 성적 (전교 전체 인원, 전교 등수, 전교 성적 비율)	
	가장 최근의 전국규모 모의고사 성적 비율	
	과목별 성적 주관적 평가	5점 리커트 척도
초4	지난 학기 반 성적 주관적 평가	5점 리커트 척도

한국아동·청소년패널조사2010에서는 <표 V-3>과 같이 초1과 초4 패널은 국어, 수학, 바른 생활, 슬기로운 생활, 즐거운 생활 과목별 주관적 평가를 5점 리커트 척도(매우 잘 했다, 잘한 편이다, 보통이다, 못한 편이다, 매우 못했다)로 조사하였고 중1 패널은 국어, 수학, 영어, 과학, 도덕, 가정/기술, 음악, 체육 과목별 평가 여부(했다, 안 했다)와 성적(96점 이상, 95-90점, 89-85점, 84-80점, 79-75점, 74-70점, 69-65점, 64점 이하), 주관적 평가를 5점 리커트 척도(매우 잘했다, 잘한 편이다, 보통이다, 못한 편이다, 매우 못했다)로 응답하게 하였다. 마지막으로 자신의 성적에 대한 만족도를 4점 리커트 척도(매우 만족한다, 만족하는 편이다, 만족하지 않는 편이다, 전혀 만족하지 않는다)로 조사하였다. 다만 성적은 응답자가 민감하게 받아들일 수 있는 점을 고려하여 2차년도부터 조사에 포함하였다. 조사 당시 서울 지역 초등학교에서 중간 및 기말고

사가 폐지되었고, 중학교에서도 학생들에게 평가 결과를 석차로 알려주지 않게 되었기 때문에 주관적인 평가 위주로 문항을 구성하였다.

<표 V-3> KYPs2010 조사 문항

조사 대상	조사 내용	조사 방식
초1	과목별 주관적 평가	5점 리커트 척도
초4	과목별 주관적 평가	
중1	과목별 평가 여부	YES/NO
	과목별 점수(등급)	8개 구간 중 선택
	과목별 주관적 평가	5점 리커트 척도
	전체 성적 만족도	4점 리커트 척도

4차년도에는 <표 V-4>와 같이 조사 문항을 정교화하여 초1 패넬은 국어, 수학, 영어, 과학, 사회, 도덕, 실과, 음악, 미술, 체육 교과에 대한 주관적 평가를 5점 리커트 척도(매우 잘했다, 잘한 편이다, 보통이다, 못한 편이다, 매우 못했다)로 조사하였고 학교성적에 대한 만족도를 4점 리커트 척도(매우 만족한다, 만족하는 편이다, 만족하지 않는 편이다, 전혀 만족하지 않는다)로 조사하였다. 초4 패넬은 국어, 수학, 영어, 과학, 사회, 도덕, 가정/기술, 음악, 미술, 체육 과목별 성적 평가 여부(했다, 안 했다)와 성적(96점 이상, 95-90점, 89-85점, 84-80점, 79-75점, 74-70점, 69-65점, 64점 이하)을 조사하였고 전체 성적에 대한 만족도를 4점 리커트 척도(매우 만족한다, 만족하는 편이다, 만족하지 않는 편이다, 전혀 만족하지 않는다)로 응답하는 문항을 추가하였다. 중1 패넬은 국어, 수학, 영어, 제2외국어(한문 포함) 과목 성적 평가 여부(했다, 안 했다)와 성적(1등급, 2등급, 3등급, 4등급, 5등급, 6등급, 7등급, 8/9등급), 사회(국사, 지리, 정치, 경제 등)와 과학(화학, 물리, 생물, 지구과학 등) 과목 성적 평가 여부(했다, 안 했다)와 성적(A, B, C, D, E), 음악, 미술, 체육 교과 성적 평가 여부(했다, 안 했다)와 성적(우수, 보통, 미흡)을 조사하였고 과목별 주관적 평가를 5점 리커트 척도(매우 잘했다, 잘한 편이다, 보통이다, 못한 편이다, 매우 못했다)로 조사하였다.



<표 V-4> KYPs2010 4차년도 조사 문항

조사 대상	조사 내용	조사 방식
초1	과목별 주관적 평가	5점 리커트 척도
	전체 성적 만족도	4점 리커트 척도
초4	과목별 평가 여부	YES/NO
	과목별 점수	8개 구간 중 선택
	전체 성적 만족도	4점 리커트 척도
중1	과목별 평가 여부	YES/NO
	과목별 점수(등급)	8개 등급 중 선택
		5개 등급 중 선택
		3개 수준 중 선택
	과목별 주관적 평가	5점 리커트 척도

다만 5차년도에 <표 V-5>와 같이 중1 패널 성적 조사 문항을 다시 간략화하여 국어, 수학, 영어, 제2외국어(한문 포함), 사회(국사, 지리, 정치, 경제 등), 과학(화학, 물리, 생물, 지구과학 등), 음악, 미술, 체육 교과별 성적 평가 여부(있다, 없다)와 주관적 평가를 5점 리커트 척도(매우 잘했다, 잘한 편이다, 보통이다, 못한 편이다, 매우 못했다)로 조사하였고 전 과목 성적에 대한 주관적 평가를 7점 리커트 척도(매우 우수, 우수, 약간 우수, 보통수준, 약간 미흡, 미흡, 매우 미흡)로 조사하였고 학교성적에 대한 만족도를 4점 리커트 척도(매우 만족한다, 만족하는 편이다, 만족하지 않는 편이다, 전혀 만족하지 않는다)로 조사하였다.

<표 V-5> KYPs2010 5차년도 조사 문항

조사 대상	조사 내용	조사 방식
초1	과목별 주관적 평가	5점 리커트 척도
	전체 성적 만족도	4점 리커트 척도
초4	과목별 평가 여부	YES/NO
	과목별 점수	8개 구간 중 선택
	전체 성적 만족도	4점 리커트 척도
중1	과목별 평가 여부	YES/NO

조사 대상	조사 내용	조사 방식
	과목별 주관적 평가	5점 리커트 척도
	전체 성적 주관적 평가	7점 리커트 척도
	전체 성적 만족도	4점 리커트 척도

중1 패널이 제외된 7차년도에 경우에는 <표 V-6>과 같이 초1과 초4 패널을 대상으로 국어, 수학, 영어, 과학, 사회, 도덕, 가정/기술, 음악, 미술 체육 과목별 주관적 평가를 5점 리커트 척도(매우 잘 했다, 잘한 편이다, 보통이다, 못한 편이다, 매우 못했다)로 조사하였고 전체 성적에 대한 주관적 평가를 7점 리커트 척도(매우 우수, 우수, 약간 우수, 보통수준, 약간 미흡, 미흡, 매우 미흡)로 조사하였고 학교성적에 대한 만족도를 4점 리커트 척도(매우 만족한다, 만족하는 편이다, 만족하지 않는 편이다, 전혀 만족하지 않는다)로 조사하였다.

<표 V-6> KYPs2010 7차년도 조사 문항

조사 대상	조사 내용	조사 방식
초1, 초4	과목별 주관적 평가	5점 리커트 척도
	전체 성적 주관적 평가	7점 리커트 척도
	전체 성적 만족도	4점 리커트 척도

한편, 2018년부터 새로 시작된 초4와 중1 패널은 <표 V-7>과 같이 전 교과 성취도와 학업성취에 대한 만족도를 잘 모르겠음을 포함한 5점 리커트 척도로 측정하였다.

<표 V-7> KYPs2018 조사 문항

조사 대상	조사 내용	조사 방식
초4, 중1	교과 성취도	5점 리커트 척도+잘 모르겠음
	학업성취 만족도	

한국교육고용패널조사는 <표 V-8>과 같이 중3 패널과 고3 패널을 대상으로 1차년도에 국어, 수학, 영어, 과학 영역, 사회 영역, 음악, 미술, 체육, (전공) 과목별 성취도에

대한 주관적 평가를 5점 리커트 척도로 조사하였고 패널 학생의 담임 교사를 통해 해당 학생의 전교 석차, 상위 10% 이내에 드는 과목, 상위 11~30%에 드는 과목, 상위 70% 이상인 과목을 조사하였다. 2차년도에는 학생 패널을 대상으로 국어, 수학, 영어, 과학 영역, 사회 영역, 음악, 미술, 체육 과목별 성취도에 대한 주관적 평가를 5점 리커트 척도로 측정하였고, 학급 석차, 전교 석차와 등급을 조사하였다. 3차년도에는 국어, 수학, 영어, 과학영역, 사회영역, 음악, 미술, 체육, (전문교과) 과목별 성취도에 대한 주관적 평가를 5점 리커트 척도로 측정하였고 전체 성적 등급을 조사하였으며 4차년도에는 국어, 수학, 영어, 과학영역, 사회영역, 음악, 미술, 체육, 제2외국어/전문교과 과목별 성취도에 대한 주관적 평가와 전체 성적 등급을 조사하였다. 아울러, 2005학년도에서 2009학년도 사이에 대학 수능 시험을 본 패널 중, 성적 자료 제공에 동의한 학생만 성적 정보를 수집하였는데 고3 패널은 2005학년도부터 2009학년도까지, 중3 패널은 2008학년도와 2009학년도의 수능 점수를 조사하였다.

<표 V-8> KEEP I 조사 문항

조사 대상	조사 내용	조사 방식
중3, 고3	과목별 성취도 주관적 평가	5점 리커트 척도
	학급 석차, 전교 석차 및 등급	직접 기입
	전체 성적 등급	
	(교사) 전교 석차, 강점 과목, 약점 과목	
	(동의자) 수능 점수	통계DB 연계

2016년부터 새로 구성된 고2 패널을 대상으로 한 조사에서는 <표 V-9>와 같이 1차년도에 학생 설문에서는 국어, 수학, 영어, 과학 영역, 사회 영역, 음악, 미술, 체육, 전문교과 과목별 성취도에 대한 주관적 평가를 5점 리커트 척도로 측정하였고 담임교사 설문에서 해당 학생의 전체 성적 등급을 조사하였다. 2~3차년도에는 국어, 수학, 영어, 과학영역, 사회영역, 음악, 미술, 체육, 전문교과 과목별 내신등급을 조사하였다.

<표 V-9> KEEP II 조사 문항

조사 대상	조사 내용	조사 방식
-------	-------	-------

조사 대상	조사 내용	조사 방식
고2	과목별 성취도 주관적 평가	5점 리커트 척도
	(교사) 전체 성적 등급	직접 기입
	국가학업성취도평가 결과	통계DB 연계

## 2. 국외 교육종단연구 학교급별 교과역량 평가 방안

### 가. 텍사스 주 STARR 시험

#### 1) 시험의 성격

STAAR 검사는 3학년부터 고등학교 학생들을 대상으로 주기적으로 시행된다. 하지만 고등학교에 진학하면서부터는 STAAR 검사는 8학년(중3)까지는 수학, 읽기/언어(Reading Language Arts, RLA), 과학, 사회 과목으로 진행되지만, 고등학교에 진학한 뒤부터는 종합 성취 평가(End-Of-Course: EOC)로서 STAAR 검사를 치르게 된다. 8학년(중3)까지 이뤄지는 STAAR 검사와 고등학교부터 시행되는 STAAR 검사(EOC 검사)는 시험의 목적과 평가 대상이 상이하다. STAAR은 3학년에서 8학년까지 학생들의 학업성취도를 전반적으로 평가하고, 학년 진급 준비도를 파악하는 시험인 반면, STAAR EOC는 고등학교 과목별 평가로, 특정 과목을 이수한 후에 해당 과목의 학습 성과를 심화하여 평가하기 위해 설계되었다. 이때 과목은 대수 I, 영어 I, 영어 II, 생물, 미국사이고, 본 과목을 수강한 뒤에 시험에 통과하는 것이 고등학교 졸업 요건에 포함된다. 이러한 STAAR과 STAAR EOC 검사는 서로 다른 목적을 갖지만, 모두 STAAR 검사의 일종이다.

#### 2) 시험 과목

고등학교 이전에는 수학, 읽기/언어(Reading Language Arts, RLA), 과학, 사회 과목을 STAAR 과목으로 보았다면, 고등학교 이후에는 수학은 대수 I로, 읽기/언어(Reading Language Arts, RLA)는 영어 I, 영어 II로, 과학은 생물로, 사회는 역사

과목으로 보다 구체화된 것을 확인할 수 있다. 학교급에 따른 과목의 변화와 함께 과목별 보고범주는 <표 III-3>, <표 III-4>에 제시되어 있다.

### 3) 문항의 특징

검사는 준비 기준(Readiness standards)과 지원 기준(Supporting standards)에 해당하는 문항으로 분류되어 구성된다. 이때 준비 기준과 지원 기준에 대한 설명은 아래와 같고, 보다 구체적인 설명은 [그림 III-1]에 제시되어 있다.

준비 기준(Readiness standards)은 현재 학년에서 성공을 위해 필수적이며, 다음 학년 또는 과정에 대비하는 데 중요한 역할을 한다. 이 기준은 넓고 깊은 개념을 다루며, 심층적인 지도가 필요하다. 이러한 기준은 기본 시험의 총 점수 중 약 55-70%를 차지한다. 지원 기준(Supporting standards)은 다음 학년 또는 과정에 대비하는 데 기여하지만, 중심적인 역할을 하지는 않는다. 이 기준은 현재 학년 또는 과정보다 하위 또는 상위 학년에서 강조되는 좁게 정의된 개념을 다룬다. 지원 기준은 기본 시험의 총 점수 중 약 30-45%를 차지한다.

STAAR 시험의 경우 학년에 따라 기준 개수, 문항 개수, 배점이 달라지게 된다. 수학, 읽기/언어, 과학, 사회 과목에 따른 보고 범주별 기준 개수, 문항의 개수와 배점은 <표 V-10>, <표 V-11>, <표 V-12>, <표 V-13>에 제시되어 있다. 주목할 만한 점은, 수학 교과목의 경우 학년이 올라감에 따라 “숫자 표현 및 관계” 관련 내용은 감소하고, “계산 및 대수적 관계”, “기하와 측정”, “데이터 분석 및 개인 금융 문해력” 내용 범주 비중은 상대적으로 증가하고 있는 것을 확인할 수 있다. 또한 사회의 경우, “역사”, “경제, 과학, 기술, 사회” 내용 범주 비중이 학교급이 올라감에 따라 증가하는 것으로 보인다.

<표 V-10> STAAR 수학 시험 학년별 특징

보고 범주(Reporting Category)	기준 개수(Number of Standards)	문항 개수(Number of Questions)	배점(Number of Points)
<b>3학년</b>			
1. 숫자 표현 및 관계	준비(Readiness): 4 지원(Supporting): 10	7-9	10-12
2. 계산 및 대수적 관계	준비(Readiness): 4	11-13	13-17

보고 범주(Reporting Category)	기준 개수(Number of Standards)	문항 개수(Number of Questions)	배점(Number of Points)
	지원(Supporting): 10		
3. 기하와 측정	준비(Readiness): 4 지원(Supporting): 10	5-7	6-10
4. 데이터 분석 및 개인 금융 문해력	준비(Readiness): 4 지원(Supporting): 10	3-5	3-6
<b>4학년</b>			
1. 숫자 표현 및 관계	준비(Readiness): 3 지원(Supporting): 10	7-9	8-12
2. 계산 및 대수적 관계	준비(Readiness): 5 지원(Supporting): 7	10-12	12-16
3. 기하와 측정	준비(Readiness): 4 지원(Supporting): 7	8-10	9-13
4. 데이터 분석 및 개인 금융 문해력	준비(Readiness): 1 지원(Supporting): 4	3-5	3-6
<b>5학년</b>			
1. 숫자 표현 및 관계	준비(Readiness): 2 지원(Supporting): 4	5-7	5-9
2. 계산 및 대수적 관계	준비(Readiness): 6 지원(Supporting): 9	15-17	17-21
3. 기하와 측정	준비(Readiness): 3 지원(Supporting): 5	7-9	8-12
4. 데이터 분석 및 개인 금융 문해력	준비(Readiness): 1 지원(Supporting): 6	3-5	3-7
<b>6학년</b>			
1. 숫자 표현 및 관계	준비(Readiness): 4 지원(Supporting): 11	8-10	8-13
2. 계산 및 대수적 관계	준비(Readiness): 6 지원(Supporting): 11	13-15	14-19
3. 기하와 측정	준비(Readiness): 3 지원(Supporting): 3	5-7	5-9
4. 데이터 분석 및 개인 금융 문해력	준비(Readiness): 3 지원(Supporting): 10	6-8	6-10
<b>7학년</b>			
1. 숫자 표현 및 관계	준비(Readiness): 2 지원(Supporting): 5	4-6	4-8
2. 계산 및 대수적 관계	준비(Readiness): 5 지원(Supporting): 7	14-16	16-21
3. 기하와 측정	준비(Readiness): 4 지원(Supporting): 5	11-13	11-16
4. 데이터 분석 및 개인 금융 문해력	준비(Readiness): 2 지원(Supporting): 8	5-7	5-9

보고 범주(Reporting Category)	기준 개수(Number of Standards)	문항 개수(Number of Questions)	배점(Number of Points)
<b>8학년</b>			
1. 숫자 표현 및 관계	준비(Readiness): 1 지원(Supporting): 3	2-4	2-5
2. 계산 및 대수적 관계	준비(Readiness): 5 지원(Supporting): 9	15-17	17-22
3. 기하와 측정	준비(Readiness): 5 지원(Supporting): 9	14-16	15-20
4. 데이터 분석 및 개인 금융 문해력	준비(Readiness): 2 지원(Supporting): 6	5-7	5-9
<b>고등학교 대수 1</b>			
1. 수와 대수적 방법	준비(Readiness): 2 지원(Supporting): 11	9-11	9-14
2. 선형 함수, 방정식 및 부등식을 설명하고 그래프로 표현	준비(Readiness): 3 지원(Supporting): 8	10-12	10-16
3. 선형 함수, 방정식 및 부등식 작성 및 해결	준비(Readiness): 5 지원(Supporting): 7	12-14	12-18
4. 이차 함수와 방정식	준비(Readiness): 4 지원(Supporting): 4	9-11	9-14
5. 지수 함수와 방정식	준비(Readiness): 2 지원(Supporting): 3	5-7	5-9

출처: Texas Education Agency(2022a), Texas Education Agency(2022e), Texas Education Agency(2022f), Texas Education Agency(2022g), Texas Education Agency(2022h), Texas Education Agency(2022i), Texas Education Agency(2022j)

<표 V-11> STAAR 읽기 시험 학년별 특징

보고 범주(Reporting Category)	기준 개수(Number of Standards)	문항 개수(Number of Questions)	배점(Number of Points)
<b>3학년</b>			
1. 읽기	준비(Readiness): 12 지원(Supporting): 19	24-26	26-28
2. 쓰기	준비(Readiness): 13 지원(Supporting): 12	15-17	24-26
<b>4학년</b>			
1. 읽기	준비(Readiness): 13 지원(Supporting): 18	24-26	26-28
2. 쓰기	준비(Readiness): 12 지원(Supporting): 12	15-17	24-26
<b>5학년</b>			
1. 읽기	준비(Readiness): 13	24-26	26-28

보고 범주(Reporting Category)	기준 개수(Number of Standards)	문항 개수(Number of Questions)	배점(Number of Points)
	지원(Supporting): 18		
2. 쓰기	Readiness: 12 지원(Supporting): 12	15-17	24-26
<b>6학년</b>			
1. 읽기	준비(Readiness): 13 지원(Supporting): 17	26-28	28-30
2. 쓰기	준비(Readiness): 9 지원(Supporting): 8	17-19	26-28
<b>7학년</b>			
1. 읽기	준비(Readiness): 13 지원(Supporting): 17	26-28	28-30
2. 쓰기	준비(Readiness): 9 지원(Supporting): 8	17-19	26-28
<b>8학년</b>			
1. 읽기	준비(Readiness): 13 지원(Supporting): 17	26-28	28-30
2. 쓰기	준비(Readiness): 9 지원(Supporting): 6	17-19	26-28
<b>고등학교 English 1</b>			
1. 읽기	준비(Readiness): 14 지원(Supporting): 15	29-31	32-34
2. 쓰기	준비(Readiness): 7 지원(Supporting): 7	21-23	30-32
<b>고등학교 English 2</b>			
1. 읽기	준비(Readiness): 14 지원(Supporting): 15	29-31	32-34
2. 쓰기	준비(Readiness): 7 지원(Supporting): 7	21-23	30-32

출처: Texas Education Agency(2022k), Texas Education Agency(2022l), Texas Education Agency(2022m), Texas Education Agency(2022n), Texas Education Agency(2022o), Texas Education Agency(2022p)



<표 V-12> STAAR 사회 시험 학년별 특징

보고 범주(Reporting Category)	기준 개수(Number of Standards)	문항 개수(Number of Questions)	배점(Number of Points)
<b>8학년 사회</b>			
1. 역사	준비(Readiness): 16 지원(Supporting): 16	15-17	18-22
2. 지리 및 문화	준비(Readiness): 7 지원(Supporting): 10	8-10	9-13
3. 정부 및 시민권	준비(Readiness): 9 지원(Supporting): 13	8-10	9-13
4. 경제, 과학, 기술, 사회	준비(Readiness): 4 지원(Supporting): 8	5-7	6-9
<b>고등학교 미국 역사</b>			
1. 역사	준비(Readiness): 22 지원(Supporting): 29	28-30	33-38
2. 지리 및 문화	준비(Readiness): 7 지원(Supporting): 5	10-12	10-14
3. 정부 및 시민권	준비(Readiness): 5 지원(Supporting): 10	8-10	9-13
4. 경제, 과학, 기술, 사회	준비(Readiness): 9 지원(Supporting): 10	14-16	16-21

출처: Texas Education Agency(2022q), Texas Education Agency(2022r)

<표 V-13> STAAR 과학 시험 학년별 특징

보고 범주(Reporting Category)	기준 개수(Number of Standards)	문항 개수(Number of Questions)	배점(Number of Points)
<b>5학년 과학</b>			
1. 물질과 에너지	준비(Readiness): 1 지원(Supporting): 3	4-6	5-8
2. 힘, 운동, 에너지	준비(Readiness): 3 지원(Supporting): 2	6-8	7-10
3. 지구와 우주	준비(Readiness): 3 지원(Supporting): 10	8-10	9-13
4. 생물과 환경	준비(Readiness): 4 지원(Supporting): 4	10-12	11-15
<b>8학년 과학</b>			
1. 물질과 에너지	준비(Readiness): 5 지원(Supporting): 4	9-11	10-14
2. 힘, 운동, 에너지	준비(Readiness): 2 지원(Supporting): 5	7-9	8-12
3. 지구와 우주	준비(Readiness): 5	9-11	10-14

보고 범주(Reporting Category)	기준 개수(Number of Standards)	문항 개수(Number of Questions)	배점(Number of Points)
	지원(Supporting): 9		
4. 생물과 환경	준비(Readiness): 2 지원(Supporting): 11	9-11	10-14
<b>고등학교 생물</b>			
1. 세포 구조와 기능	준비(Readiness): 3 지원(Supporting): 3	8-10	8-13
2. 유전 메커니즘	준비(Readiness): 3 지원(Supporting): 4	8-10	8-13
3. 생물학적 진화와 분류	준비(Readiness): 3 지원(Supporting): 6	8-10	8-13
4. 생물학적 과정과 시스템	준비(Readiness): 3 지원(Supporting): 3	8-10	8-13
5. 환경 시스템 내 상호 의존성	준비(Readiness): 4 지원(Supporting): 3	8-10	8-13

출처: Texas Education Agency(2022b), Texas Education Agency(2024a), Texas Education Agency(2024b)

이어서 학년 및 교과목별 STAAR 시험 부분점수 문항 유형도 <표 III-1>에 제시되어 있다. 여기서 1점 문항은 다지선다형 및 비다지선다형 모두가 포함되고, 2점 문항은 모두 비다지선다형이다. 주목할 만한 점은, 학년이 올라갈수록 총 문항 수가 증가하고, 중학교에서 고등학교로 학교급이 바뀔 때, 특히 총 문항 수가 증가하는 양상을 확인할 수 있다. 또한 부분점수인 2점 문항도 중학교에서 고등학교로 학교급이 바뀔 때 증가하는 모습을 보인다.

## 나. 플로리다 주 FAST 검사

### 1) 시험의 성격

FAST(Florida Assessment of Student Thinking)와 EOCs(End-of-Course Assessments)는 각각 고유한 목적을 가지고 설계되었다. FAST는 학생들의 학업성취도 추적과 성장 평가를 목표로 한다. 이를 통해 학생들이 학업 중간에 어떤 부분에서

더 많은 지원이 필요한지 식별하는 데 도움을 주며, 연속적인 피드백을 제공하여 학습 개입의 근거를 마련한다. 반면, EOCs는 특정 과목에서 학생이 배운 내용을 평가하고, 과목별 성취도를 종합적으로 평가하는 데 중점을 둔다. EOCs는 수학, 과학, 역사와 같은 핵심 과목에서 학생들의 학업 성취를 종합적으로 측정하며, 일부 과목의 경우 졸업 자격을 결정하는 데 중요한 역할을 한다.

FAST는 연간 3회 시행되어 학생들의 학업 성취 변화를 추적하는 데 최적화된 평가 방식이다. 이러한 연속적인 평가를 통해 학년 동안의 학습 성장률을 모니터링하고, 필요한 경우 중간에 개입을 할 수 있는 기회를 제공한다. 반면, EOCs는 각 과목이 종료될 때 한 번 시행되며, 해당 과목에서 학생이 얼마나 성취했는지에 대한 종합적인 평가를 제공한다. FAST는 반복적인 평가를 통해 학생의 성장을 측정하는 반면, EOCs는 특정 시점에서의 학업성취도를 평가하는 점에서 차별된다.

FAST와 EOCs는 모두 컴퓨터화 능력적응검사(computer-adaptive test: CAT)로, 학생의 응답에 따라 문항의 난이도가 조정된다. FAST는 학생의 학업성취도에 맞춘 문제를 제공하여 개별화된 학습 평가를 가능하게 하며, 시험이 진행됨에 따라 학생의 능력에 맞춰 문제의 난이도가 실시간으로 조정된다. 이는 학생들의 학습 수준을 보다 정교하게 측정할 수 있는 장점을 제공한다. EOCs 역시 일부 과목(예: 대수 I, 기하)에서 컴퓨터화 능력적응검사를 채택하고 있으며, 특히 이러한 과목들은 중요한 수학적 기술을 평가하는 데 중점을 둔다. 그러나 EOCs는 특정 과목의 내용을 심층적으로 평가한다는 점에서 과목에 대한 집중적인 평가가 이루어지는 특징을 가진다.

FAST는 졸업 요건과는 직접적인 관련이 없다. 이는 주로 학생들의 학업성취도를 추적하고, 학습 과정에서 필요한 개입을 식별하기 위한 도구로 사용된다. 즉, FAST는 학업 성취의 연속적인 성장에 대한 피드백을 제공하는 평가 도구로 기능한다. 반면, EOCs는 일부 과목, 특히 대수 I EOC가 고등학교 졸업 요건과 직접적으로 연계되어 있다. 대수 I EOC에서 특정 점수 이상을 받는 것은 졸업 필수 요건 중 하나이며, 기하 EOC는 졸업 필수 요건에는 해당하지 않지만, 우등 졸업장 자격을 얻기 위해 활용될 수 있다. 또한 기하 EOC는 대수 I EOC의 대체 점수로도 사용되거나 우등 졸업장 자격을 얻기 위해 사용될 수 있어, 학생들에게 중요한 평가 요소로 작용한다. 생물학 I, 미국사, 시민 교육은 졸업 필수 과목은 아니지만, 선택적으로 졸업 요건을 충족하는 데 도움이 될 수 있다.

결론적으로, FAST 검사는 학년이 높아짐에 따라 성장 추적에서 졸업 자격 평가로 그 성격이 변화하며, 이 두 평가 방식은 상호 보완적으로 학생들의 학업성취도를 측정하

는 데 기여한다.

## 2) 시험 과목

플로리다 주의 학업 평가 시스템에서 FAST와 EOC는 서로 다른 학년과 과목을 평가하며, 두 평가의 가장 큰 차이점은 학생들이 시험에서 평가받는 교과목의 범위와 세부적인 과목 내용에 있다.

FAST 검사는 주로 초등학교와 중학교 학년을 대상으로, 읽기(ELA)와 수학을 평가한다. 읽기는 3학년부터 10학년(고2)까지, 수학은 3학년부터 8학년(중3)까지 평가가 이루어진다. 이 시험은 학업성취도와 학생들의 학습 성장 추적에 초점을 맞추며, 매년 3회 시행되어 학생들의 연속적인 학업 성취 변화를 관찰한다.

반면, EOC 시험은 고등학교 과목을 구체적으로 평가하는 방식으로 이루어진다. 수학 과목이 보다 심화되어 대수 I(Algebra 1)과 기하(Geometry)로 나뉘며, 추가적으로 생물학 I(Biology 1), 미국사(U.S. History), 시민 교육(Civics) 같은 과목이 포함된다. EOC는 각 과목이 종료될 때마다 한 번 시행되며, 연간 1회만 실시된다.

## 3) 문항의 특징

FAST 검사에서 시행하는 읽기 과목에 따른 보고 범주별 평가 비율은 학교급이 변화해도 일정하게 유지되는 반면, 수학 과목의 경우에는 학년에 따라 조금씩 차이가 있는 것으로 확인된다. 구체적인 변화는 <표 III-10>, <표 III-11>에 제시되어 있다.

FAST 검사에서 읽기, 수학 영역의 문항 수는 <표 III-5>, <표 III-6>과 같으며, 학년에 따라 큰 차이를 보이지 않는 것으로 파악된다.

## 4) 검사 소요 시간

FAST 검사에서 읽기, 수학 영역을 해결하는 데에 걸리는 검사 시간은 <표 III-8>과 같으며, 읽기 영역의 경우에는 같은 시간이 주어지지만, 수학 과목의 경우에는 과목의 특성상 학년에 따라 조금의 차이를 보이는 것으로 확인된다.

## 다. 텍사스 STAAR 검사와 플로리다 FAST 검사 비교

텍사스 주에서 시행하는 STAAR 검사와 플로리다 주에서 시행하는 FAST 검사의 시험 성격, 과목, 문항 특징, 성적 활용 측면에서 비교 및 분석한 내용은 <표 V-14>에 제시되어 있다.

<표 V-14> 학교급별 텍사스 STAAR과 플로리다 FAST 검사 비교

항목	텍사스 STAAR	플로리다 FAST
시험 성격	<ul style="list-style-type: none"> <li>- 초등~중학교: 주로 학년별 학업성취도를 평가하는 표준화된 시험 시행. 연간 1회 시행</li> <li>- 고등학교: STAAR 시험을 통해 평가. 과목별 EOC(End-of-Course)로 전환되어 특정 과목 학습 성취도를 심화 평가. 연간 1회 시행이지만 통과하지 못할 경우 추가 기회 제공</li> </ul>	<ul style="list-style-type: none"> <li>- 초등~중학교: 읽기(ELA)와 수학 성취도 추적에 중점. 연간 3회 시행</li> <li>- 고등학교(EOC): FAST 검사가 아닌 EOC 시험을 통해 주요 과목에서 심층 평가 (대수학, 기하학, 생물학, 역사 등). 과목 수강 후 1회 시행</li> </ul>
과목	<ul style="list-style-type: none"> <li>- 초등~중학교: 수학, 읽기/언어(RLA), 과학, 사회</li> <li>- 고등학교: 대수 I, 영어 I·II, 생물, 미국 역사와 같은 과목별 평가로 구체화</li> </ul>	<ul style="list-style-type: none"> <li>- 초등~중학교: 읽기, 수학</li> <li>- 고등학교(EOC): 대수 I, 기하학, 생물학, 미국 역사, 시민 교육 같은 과목으로 세분화</li> </ul>
문항 특징	세분화된 과목에 따라 보고 범주도 구체화되는 경향이 있음	
성적 활용	<ul style="list-style-type: none"> <li>- 초등~중학교: 학년 진급에 활용</li> <li>- 고등학교: 고등학교 졸업 요건에 포함</li> </ul>	<ul style="list-style-type: none"> <li>- 초등~중학교: 학업성취도 추적과 학습성장 평가</li> <li>- 고등학교: FAST 검사가 아닌 EOC 시험을 통과하는 것이 졸업 요건에 포함</li> </ul>

## 3. 기존 학업성취도 평가의 학교급별 활용 방안

본 장에서는 기존 학업성취도 평가와 교육종단연구에서의 평가 결과 활용과 관련하여 학교급별로 차이가 있는지 살펴보려고 하며, 자세한 내용은 <표 V-15>, <표 V-16>와 같다.

<표 V-15> 학교급별 기존 학업성취도 평가 종류 및 결과 활용 사례

분류	대상	영역	비고
2024 국가수준 학업성취도 평가	중3	국어, 수학, 영어, 사회, 과학	중 1~2학년, 3학년 1학기 범위
	고2	국어, 영어, 수학	범교과 소재, 고등학교 '수학'
맞춤형 학업성취도 자율평가	초3	문해력, 수리력	초 1~2학년 범위
	초5	국어, 수학, 영어, 사회, 과학	초 4학년 전범위
	초6		초 5학년 전범위
	중1	국어, 수학, 영어, 사회, 과학	초 6학년 전범위
	중3		중 2학년 전범위
	고1	국어, 수학, 영어, 사회, 과학	중 3학년 전범위
	고2	국어, 수학, 영어	고 1학년 전범위
전국연합 학력평가	고1	국어, 수학, 영어, 한국사, 통합사회, 통합과학	
	고2	국어, 수학, 영어, 한국사, 사회탐구, 과학탐구 (직업탐구, 제2외국어 및 한문)	
	고3	국어, 수학, 영어, 한국사, 사회탐구, 과학탐구 (직업탐구, 제2외국어 및 한문)	
성취평가제	중1~3	국어, 수학, 영어, 사회, 과학, 도덕, 기술·가정, 음악, 미술, 체육, 선택교과 등	준거참조평가 (성취기준-성취율)
	고1~3	보통교과(전체 고교), 전문교과(마이스터고 등)	준거참조평가 (성취기준-성취율)
책임교육 학년제	초3, 중1	언어, 수리, 디지털 능력 등	컴퓨터 기반 학업성취도 평가
기초학력 진단평가	초3~고2	읽기, 쓰기, 셈하기 국어, 사회, 역사/한국사, 수학, 과학, 영어	준거참조평가 온라인/오프라인

<표 V-16> 학교급별 기존 학업성취도 평가 결과 활용 사례

분류	대상	영역	비고
한국교육 종단연구2005	고1~3	고1 국가수준 학업성취도 평가, 고3 대학수학능력시험 점수	
한국교육 종단연구2013	고1~3	전국연합학력평가 등급 (국어, 영어, 수학)	
한국복지패널	초4~6 (2006 기준)	전과목 평균, 국어, 수학, 영어 (자기보고식 5점 척도)	코호트조사
한국교육고용 패널조사 I	중3, 고3 (2004 기준)	학급 및 전교 석차, 등급 (자기보고식 설문)	코호트조사
한국교육고용 패널조사 II	고2	전체 및 교과별 내신 등급 (자기보고식 설문)	패널조사





| 인천교육종단연구 기반연구 - 교과역량 평가 방안 제안 |

# VI

## 교과역량 평가 시행 방식





## 교과역량 평가 시행 방식

본 장에서는 앞서 분석한 국내외 교육종단연구 사례를 종합하여 인천교육종단연구의 교과역량 조사 및 평가 시행 방식에 대한 시사점을 도출하고자 한다. 먼저 국내 교육종단연구의 교과역량 조사 및 평가 방법을 종합하여 정리하였고, 국외 교육종단연구 사례로 텍사스 주 STAAR 검사와 플로리다 주 FAST 검사를 비교하여 제시하였다.

### 1. 국내 교육종단연구 교과역량 평가 시행 방식

국내 교육종단연구에서 별도의 평가도구를 개발하여 시행한 사례는 한국교육종단연구, 한국 초·중등 학교교육 성과 종단연구, 서울교육종단연구, 경기교육종단연구, 전남 교육종단연구가 있으며 국내 교육종단연구에서 교과역량 평가를 온라인으로 시도한 사례는 한국 초·중등 학교교육 성과 종단연구가 유일하다. 개별 평가도구에 대한 구체적인 내용은 1장에서 제시하였다.

각 연구의 조사 방법은 <표 VI-1>과 같다. 대부분의 연구에서 면접조사와 설문조사를 주요 방법으로 사용하고 있으며, 시간이 지남에 따라 PDA, 태블릿PC, 노트북 등의 전자기기를 활용한 조사 방식이 증가하는 추세를 볼 수 있었다. 또한 최근에는 코로나 19의 영향으로 온라인 조사 방식이 일부 되었음을 확인할 수 있다.

<표 VI-1> 국내 교육종단연구 조사 방법

연구명	조사 방법
한국청소년패널조사(KYPS)	청소년: 대인면접조사 방식의 설문조사 학부모: 전화면접조사
한국아동·청소년패널조사2010 (KCYPs 2010)	청소년: 개별면접조사 보호자: 전화조사 또는 청소년(자녀)을 통한 유치조사
한국아동·청소년패널조사2018	주로 TAPI(2020년 코로나19 상황에 따라 온라인 조사 병행)

연구명	조사 방법
(KCYPs 2018)	
한국교육고용패널조사 I (KEEP I)	1차년도 주로PDA를 이용한1:1 면접조사 → 2-3차년도 PDA 면접조사와 CATI 병행 → 4-12차년도 노트북을 이용한 면접 조사, CATI, 온라인 웹조사 등
한국교육고용패널조사 II (KEEP II)	1차년도 학생: 집단면접(CAPI, PAPI), 가구: 종이 설문(PAPI) 유치 후 수거, 학교행정가/담임교사: 개별면접 → 2차년도 이후 노트북을 이용한 일대일 면접조사 원칙, 불가능한 경우 종이 설문지 병행 → 4-6차년도 코로나-19의 영향으로 온라 인 조사 일부 병행
한국교육종단연구2005 (KELS2005)	학생: 재학 중인 학교를 통한 집단 면접 방식과 개별 면접 조사 우편 설문 및 전화 설문조사 방법 병행 오프라인 국어, 수학, 영어 학업성취도 평가
한국교육종단연구2013 (KELS2013)	주로 설문지를 통한 조사 오프라인 국어, 수학, 영어 학업성취도평가
한국 초·중등 학교교육 성과 종단연구	웹 기반 종단조사 온라인 국어, 수학, 문제해결역량(국어, 수학, 사회, 과학 교과) 학생 검사
서울교육종단연구2010	외부의 조사업체와 학교의 협력교사의 협조를 통해 우편으로 조사 오프라인 국어, 영어, 수학 성취도 검사
서울교육종단연구2020	오프라인 조사와 온라인 조사(학부모 설문)를 모두 활용 오프라인 국어, 수학, 영어 서울학생역량 검사
경기교육종단연구	협력학교 및 협력교사 선정을 통한 조사 오프라인 국어, 영어, 수학 학업성취도 검사
경기기초학력종단연구	오프라인 기초 문식력, 기초 수리력 지필평가
전남교육종단연구	협력학교 및 협력교사 선정을 통한 조사 오프라인 국어, 수학, 사회, 과학 인지영역 검사

## 2. 국외 교육종단연구 교과역량 평가 시행 방식

### 가. 텍사스 주 STAAR 시험

## 1) 온라인 STAAR 시험의 발전

STAAR 온라인 시험은 내장형 지원 기능과 함께 2017년 봄 시행부터 STAAR A(학습장애 또는 특별한 교육적 요구가 있는 학생들을 위해 조정된 시험)와 STAAR L(영어가 모국어가 아닌 학생들을 위해 조정된 시험)을 대체했다. 이 변경을 통해 학생들의 개별적인 필요에 따라 콘텐츠 지원, 언어 및 어휘 지원을 포함한 더 다양한 접근성 기능과 조정이 가능해졌다.

2018-2019 학년도에 STAAR 시험은 처음으로 미국 수어(ASL)와 리프레셔블 점자를 통해 온라인으로 시행되었다. 또한, 3~5학년 대상 STAAR 스페인어 시험도 처음으로 온라인으로 제공되었다. STAAR은 2022년 12월 시행부터 주로 온라인 평가 프로그램으로 전환되었다.

STAAR 중간 평가(Interim Assessments)는 TEKS(Texas Essential Knowledge and Skills)와 연계된 선택적 온라인 평가 세트로, 학생들의 학업 진행 상황을 모니터링하고 STAAR 종합 평가에서의 학생 성과를 예측하는 것이 목적이다. 중간 평가는 교육구에 무료로 제공되며, 책임성 평가와는 연결되지 않는다, 즉 성적이 공식적인 학교 평가에 반영되지 않는다.

2023년에 텍사스 주의 88회 입법부(2023년 정기 회기)에서 통과된 HB 1225 법안에 따르면, 학부모, 보호자, 또는 교사가 요청하는 경우, 학생에게 STAAR 시험의 종이 시험지를 제공할 수 있게 되었다. 이러한 요청은 가을 시험을 위해서는 9월 15일까지, 봄 시험을 위해서는 12월 1일까지 교육구에 제출해야 한다. 그러나, 종이 시험지를 제공받을 수 있는 학생의 수는 학군에 등록된 학생 수의 3%를 초과하지 않도록 제한된다. 또한, 이 3% 제한은 특별한 종이 시험을 받을 자격이 있는 학생들과는 별개의 규정이다. 이 법은 학부모와 교사들이 종이 시험을 요청할 수 있는 절차를 명확히 하고, 디지털 시험 대신 종이 시험을 선택할 수 있는 옵션을 확대한 중요한 변화이다.

## 2) 온라인 시험 환경

### 가) 시험 전달 시스템

STAAR 온라인 평가는 시험 전달 시스템(Test Delivery System, TDS)을 사용하여 시행된다. TDS에는 시험 세션을 일정에 맞춰 관리하는 시험 관리자 인터페이스, 학생

들이 시험에 참여할 수 있도록 하는 학생 인터페이스, 그리고 안전한 온라인 시험 환경을 제공하는 보안 브라우저 애플리케이션이 포함된다. TDS는 시험 데이터의 안전한 전송 및 저장을 가능하게 하며, 학생 시험 인구를 지원할 수 있도록 확장성을 유지한다. 또한, TDS 아키텍처는 120만 명 이상의 동시 사용자를 원활하게 처리하면서 안정성과 효율성을 입증했다.

### 나) 시험 시스템(ePEN 시스템)

STAAR 시험(스페인어 버전 포함)은 아이패드(iPad)와 같은 모바일 기기와, 크롬(Chrome), 엣지(Edge) 등에서도 시행 가능한 컴퓨터 기반 검사이다. 이때 사용되는 기기는 학교에서 제공된다. 이렇게 수집된 응답은 Pearson ePEN(electronic Performance Evaluation Network) 시스템을 통해 채점된다. 채점자는 TEA 승인 평가 기준표와 기준 답안에 접근하여 훈련, 자격 취득, 실제 채점 과정에서 이를 참고한다. 훈련과 자격 취득이 완료되면 학생들의 서술형 응답에 안전하게 접근할 수 있다. ePEN 응답 뷰어는 학생들의 서술형 응답을 스캔한 이미지를 표시하며, 채점자는 대비, 색상, 확대 기능을 조정하여 가독성을 높이고 피로를 줄일 수 있다.

특정 학생과 시험의 모든 서술형 응답은 고유 식별자를 통해 Pearson 채점 및 보고 과정에서 연결된다. 학생의 익명성을 보호하고 잠재적인 편견을 방지하기 위해 ePEN에서는 채점자가 학생 식별자나 기타 인구 통계 정보를 볼 수 없다.

응답은 학년, 과목 또는 코스별로 분류되어 ePEN 서버에 저장된다. 채점자가 응답을 채점하면, 더 많은 응답이 그들의 채점 대기열로 보내진다. 각 채점자는 응답을 독립적으로 읽고 컴퓨터 화면에서 점수를 선택한다. 채점 감독관, 채점 디렉터, 그리고 내용 전문가들은 각 응답을 읽은 채점자를 확인할 수 있다.

### 3) 온라인 시험 시행을 위한 조치

연구에 따르면, 어린 학생들도 온라인에서 타이핑할 수 있는 능력을 가지고 있다. 미국 교육 진흥 평가(NAEP)가 4학년 학생들을 대상으로 실시한 연구에서는 이들이 컴퓨터 프로그램을 사용하여 타이핑하고, 조직하고, 평가받을 만큼 충분히 잘 작성할 수 있음을 보여주었다. 최소한 20개 이상의 주에서는 초등학생들이 온라인에서 서술형 응답을 작성해야 한다. 또한, 텍사스 기술 TEKS는 3학년까지 학생들이 타이핑을 능숙

하게 할 수 있도록 요구하며, 여기에는 올바른 터치 키보딩 기술과 인체공학적 전략을 배우는 것이 포함된다. 타이핑에 어려움을 겪는 학생들은 시험 관리자가 학생 응답을 온라인 시스템에 입력할 수 있는 기존의 전사 지원을 받을 수 있다(Texas Education Agency, 2023).

## 가. 플로리다 주 FAST 검사

### 1) 온라인 FAST 검사의 발전

2014년, 플로리다주는 ELA와 수학에 대한 새로운 평가 시스템을 도입하기 위해 플로리다 표준(Florida Standards)에 맞춘 평가 개발을 추진하였다. 이에 따라 행정명령 13-276을 기반으로 한 제안 요청서가 발행되었으며, 이를 통해 Cambium Assessment, Inc.(CAI)와 계약이 체결되었다. CAI는 이전에 American Institutes for Research (AIR)로 알려져 있었으며, FSA(Fla. Standards Assessments)라는 새로운 평가 시스템을 개발하게 되었다.

이 새로운 평가는 플로리다 학업 표준에 맞춘 분석적 사고 측정에 중점을 두었으며, 심리측정학자들과 CAI, 플로리다 교육부(FDOE), 그리고 시험 개발 센터(TDC)의 전문가들이 모여 2015년 봄에 시행될 시험 양식을 개발하였다. 이때 유타주 학생 성장 및 우수성 평가(SAGE) 항목들이 플로리다 시험 양식에 사용되었으며, 통계적으로 적합한 문항만이 시험에 포함되었다. 이는 시험의 신뢰성과 타당성을 높이는 데 기여하였다.

2015년 봄에 FSA의 첫 운영 시험이 시행되었다. 3학년과 4학년 ELA 및 수학 평가는 종이 시험지로 진행되었고, 나머지 학년은 온라인 시험으로 시행되었다. 이때 대수 I의 EOC(End-of-Course) 성과는 졸업과 승급에 중요한 기준이 되었으며, 3학년 ELA와 10학년 ELA는 동백분위(Equipercentile) 동등화 연결 방식을 통해 평가 성과를 비교 분석하였다.

또한, 2015년 여름, 새로운 수직 척도가 설정되었으며, 이를 기반으로 310학년 ELA, 38학년 수학, 대수 I EOC, 기하 EOC 등에 대한 성과 기준이 정립되었다. 성과 기준 설정 과정에는 400명 이상의 교육자가 참여하여 교육 현장의 의견을 반영하였다.

2016년 봄, 4학년 ELA 읽기 부분이 온라인 시험으로 전환되었다. 그러나 IEP(개별 교육 계획)나 504 계획이 적용된 학생들은 여전히 종이 시험을 치를 수 있었다. 점수의

비교 가능성을 확보하기 위해, 2015년과 2016년의 시험 간 점수 동등화가 이루어졌다.

2017년 봄에는 3학년과 4학년 수학 평가가 온라인 방식으로 전환되었다. 이에 따라 대부분의 학생들은 온라인으로 시험을 보게 되었으며, 일부 학생들만 종이 시험을 선택할 수 있었다.

2018년 봄에는 대수 II 시험이 시행되지 않았다. 2019년에는 법안 7069에 따라 일부 학년과 과목이 다른 평가 방식으로 전환되었다. 4-6학년 읽기와 3-6학년 수학은 종이 시험으로 전환되었고, 7학년 쓰기 시험은 온라인 평가로 전환되었다.

## 2) 온라인 검사 관리 및 시행

FAST 시험 운영에 참여한 주요 인원은 지역 평가 코디네이터(DAC), 학교 관리자, 그리고 시험 감독관(TA)을 포함하였다. 이들은 시험을 감독하였다. TA를 위한 온라인 교육 과정이 제공되었다.

CAI(CAI Secure Browser)에서 개발한 보안 브라우저는 온라인 FAST 시험에 접근하기 위해 필요하다. 이 브라우저는 핫키, 복사 및 화면 캡처 기능을 비활성화하고, 인터넷과 이메일 같은 데스크탑 기능에 대한 접근을 차단함으로써 학생 시험 환경을 안전하게 제공했다.

학생들은 Windows, Chrome, Mac, iPad와 같은 여러 플랫폼을 통해 FAST 온라인 시험에 참여할 수 있었다. 시험 시행 전에, FAST 온라인 시험이 시행되는 모든 플랫폼에서 사용자 승인 테스트가 수행된다. 이 테스트는 시험 문항이 기대한 대로 표시되고, 장치 간 잠재적인 차이를 최소화하기 위해 시행된다. 교육 및 심리적 테스트에 대한 표준에 따라 CAI는 장치 간 플로리다 주의 시험 점수의 비교 가능성을 제공하기 위한 장치 비교 연구를 수행하였다.

시험 시행 전에, 사용자 승인 테스트는 문항이 의도한 대로 표시되고, 장치 간에 일관된 모양을 유지하는지 확인하기 위해 모든 승인된 플랫폼에서 수행된다. 엄격한 검토 과정이 이루어져 시험의 온라인 실행 방식과 일치하는 문항의 내용이 (예: 어휘, 그래픽, 단락 구분, 옵션 순서 등) 제대로 반영되었는지 확인한다.

## 가) 문항 개발 및 문제 은행

문항 개발 단계는 학업 표준과 청사진에 맞춰 학생들의 기술과 능력을 정확하게



측정하는 고품질 평가를 제작하는 다양한 활동을 포함한다. 매년 새로운 문항이 개발되어 필드 테스트를 거친 후 운영 문항 풀에 추가된다. 여러 요소가 새로운 문항의 개발을 결정하며, 문항 개발 팀은 문항 수, 문항 유형, 문항 난이도, 지식의 깊이(Depth of Knowledge, DOK) 수준, 각 스트랜드나 벤치마크별 문항 수와 같은 다양한 차원을 기준으로 분포 격차 분석을 수행한다. 총괄 온라인 평가는 운영 문항과 임베디드 필드 테스트(EFT) 문항을 포함하며, 이 문항들은 시험 전반에 걸쳐 무작위로 배치된다. 운영 문항은 학생의 점수를 계산하는 데 사용되며, EFT 문항은 채점되지 않으며 향후 운영을 위한 플로리다 학생 사고 평가(FAST) 시험 은행에 추가된다. 적응된 버전의 온라인 평가는 동등한 길이의 평가를 보장하기 위해 필드 테스트 슬롯에 필러 문항을 포함한다. 이러한 문항들은 필드 테스트 교정의 일부로 분석되지 않는다.

시험 문제 은행에는 플로리다의 각 과목과 학년 수준의 기준을 포괄하는 다양한 문항이 포함되어 있으며, 이 문항들은 여러 난이도를 가지고 있다. 이는 두 가지를 보장한다. 첫째, 각 평가 시도에서 시험 청사진을 충분히 충족할 수 있도록 충분한 다양한 문항이 제공된다. 둘째, 다양한 난이도 덕분에 각 시험 시도가 학생의 능력에 맞춰 맞춤형으로 제공된다. 시간이 지남에 따라 더 많은 문항이 개발되고 현장 시험을 거치면, 문항 은행은 점점 더 풍부해진다.

#### 나) 컴퓨터화 능력적응검사(CAT)의 알고리즘

CAT 알고리즘은 플로리다 학업 성취 기준에 맞추어 설계된 문항들이 포함된 공통 문항 은행에서 시험 문항을 선택하는 방식으로 작동한다. 문항 선택 과정은 시험 청사진과 학생의 성과에 의해 결정되며, 특히 시험 청사진은 학생들이 기준의 폭과 깊이를 반영한 시험을 보도록 보장하는 중요한 역할을 한다. 학생이 문항에 어떻게 응답하는지에 따라, 알고리즘은 다음 문항으로 더 어려운 문항을 선택할지, 혹은 덜 어려운 문항을 선택할지 동적으로 조정한다.

시험의 유형에 따라 알고리즘의 작동 방식은 약간 다르게 나타난다. 수학, 과학, 사회 과목에서는 개별 문항 수준에서 적응형으로 작동하며, 학생의 응답과 시험 청사진의 요구 사항을 기반으로 적합한 다음 문항을 선택한다. 5학년과 8학년의 과학 시험 및 생물학 I 시험의 경우, 동일한 자극과 관련된 여러 문항을 포함하는 문맥 의존(CD) 문항 세트가 포함될 수 있다.

ELA 읽기 시험에서는 지문 또는 지문 세트 수준에서 적응형으로 작동한다. ELA

읽기 시험의 문항은 특정 지문과 연결되어 있으며, 이를 지문 세트라고 한다. 알고리즘은 학생의 지문 세트에 대한 응답과 시험 청사진의 요구 사항을 바탕으로 다음 지문 또는 지문 세트를 선택하게 된다. 학생이 이전 지문 세트와 연관된 문항에 어떻게 응답했는지와 시험 청사진 요구 사항이 어떻게 충족되어야 하는지를 검토한 후, 적합한 지문 세트와 관련 문항이 선택된다. 지문 세트가 선택되면, 해당 세트는 고정되며 학생이 현재 세트 내에서 답을 변경하더라도 다음 문항으로 진행된다. 이를 통해 시험의 적응형 평가가 지문과 문항 수준에서 효율적으로 이루어진다.

컴퓨터화 능력적응검사(CAT)의 주요 장점은 시험 점수의 정확도와 평가 효율성 측면에서 두드러진다. 첫째, CAT는 학생의 능력 수준에 맞추어 문항을 조정함으로써, 모든 능력 수준에서 시험의 정확도를 향상시킨다. 이는 문항이 고정되어있는 시험에서 발생할 수 있는 특정 능력 수준에 대한 평가의 불균형을 해소하는 효과를 제공한다. 둘째, CAT는 학생의 능력에 동적으로 적응하는 특성 덕분에, 상대적으로 적은 수의 문항으로도 고정형 시험에 비해 높은 평가 정확성을 제공할 수 있다. 이러한 적응적 접근은 시험 시간이 단축되고, 학생의 피로도가 줄어드는 이점을 가진다. 셋째, CAT는 학생들이 시험 중에 서로 다른 문항을 접하게 함으로써 시험의 보안성을 강화한다. 이로 인해 문항 유출의 위험이 감소하고, 보다 신뢰성 있는 시험 환경이 조성된다. 마지막으로, CAT는 시험이 종료된 후 24시간 이내에 결과를 제공할 수 있어 신속한 피드백이 가능하며, 이는 학생들의 학습 및 성취 분석에 있어 즉각적인 정보를 제공하는 중요한 이점이 된다.

이러한 적응형 알고리즘의 구현에는 다양한 변수들이 적용된다. 예를 들어, 초기 학생 능력 추정 방식, 문항 재활용 알고리즘, 특정 내용 제약 등이 그 예이다. 이러한 알고리즘들은 시험이 진행되는 동안 학생의 능력 수준에 맞춘 문항을 제공하는 역할을 한다. 콘텐츠 요구 사항은 시험 전체 수준에서 최소 및 최대 문항 수로 정의되며, 각 범주별로 시험 청사진을 충족하도록 설계된다.

시험이 진행됨에 따라, 적응형 알고리즘은 문항 정보 함수에 기반해 특정 학생에게 가장 많은 정보를 제공할 수 있는 문항을 선택하게 된다. 이때 시험의 전반적인 흐름을 고려하여, 내용적으로 가치가 높은 문항들이 먼저 선택되며, 시험 후반부로 갈수록 최소 요구 사항을 충족하지 못한 범주의 문항들이 우선 선택된다. 즉, 시험 초반에 충족되지 못한 내용 범주를 채우기 위해 적응형 알고리즘이 작동하여 해당 범주에 맞는 문항이 출제될 가능성이 높아진다.

이후, 학생의 능력 수준에 맞춰 추가적으로 평가된 문항들이 최종 선택되며, 이 과정은

무작위로 이루어질 수 있다. 그러나 이러한 모든 과정에서 시험 청사진이 항상 우선되며, 문항의 콘텐츠와 능력에 맞춘 배치가 이루어진다. 또한, 초기 문항 은행 구성 시 적응형 알고리즘은 비활성화될 수 있으며, 이는 시험의 초기 보정 과정에서 충분한 데이터를 확보하기 위해 필요한 조치이다.

적응형 시험의 적절성을 평가하기 위해, 플로리다 교육부(FDOE)와 Cambium Assessment, Inc.(CAI)의 심리측정학자들이 다양한 시뮬레이션을 통해 알고리즘의 성능을 검토한다. 분석의 주요 지표로는 편향(bias), 평균 문항 난이도와 학생 능력 간의 상관관계(적응성을 측정하기 위한 지표), 문항 노출 빈도, 그리고 시험 청사진과의 일치도 등이 포함된다. 특히, 문항 은행의 깊이에 대한 주의가 요구된다. 학생이 여러 번 시험을 치를 경우, 시험 청사진을 충족하기 위해 문항이 재활용될 수 있으며, 이는 시험의 적응성에 영향을 미칠 수 있다. 따라서 문항이 적절하게 재활용되면서도 시험의 타당성과 신뢰성을 유지하기 위해서는 다중 기회 연구가 필수적이다. 이를 통해 문항의 재사용이 필요한 시점과 그 영향을 명확히 평가할 수 있다.

### 3) 온라인 시험 환경

3학년부터 10학년까지의 모든 평가와 종합 시험(EOC) 평가는 컴퓨터화 능력적응검사로 진행되며, 다음 시스템을 통해 시행 및 보고된다. 시험 계약자는 학생들에게 컴퓨터 기반 시험(Computer based Test: CBT)을 제공하기 위해 보안된 온라인 시험 플랫폼을 제공한다. 학생들은 고유한 로그인 정보가 포함된 시험 티켓을 사용해 로그인 하며, 다양한 도구와 기능을 사용해 시험을 탐색하고 상호작용한다. 이때 학생들은 개인용 디지털 기기가 아닌 학교에서 제공해주는 컴퓨터 또는 태블릿을 활용하여 FAST 검사에 응시할 수 있다. <표 VI-2>에서는 FAST 검사 시 사용되는 시스템에 대한 설명이 제시되어 있다.

<표 VI-2> FAST 검사 시스템별 사용

시스템	사용
검사 정보 배포 엔진(TIDE)	온라인 관리/학생 데이터
보안 브라우저를 사용하는 검사 배포 시스템(TDS)	온라인 시험 배포
플로리다 보고 시스템 (FRS)	결과 보고
가족 포털	학부모용 결과

시스템	사용
출처: Florida Department of Education, 2024b	

#### 4) 온라인 시험 시행을 위한 조치

##### 가) 인력별 역할 및 책임

플로리다 주의 K-12 학생 평가 시스템에서 시험 관리와 관련된 다양한 수준의 역할과 책임을 체계적으로 분담하여 시험이 원활하게 운영되도록 한다. 주 차원에서는 플로리다 교육부(FDOE)의 K-12 학생 평가국이 시험 관리 매뉴얼, 사용자 가이드, 교육 자료 등을 개발하고, 지침과 정책 정보를 제공하며, 시험 중에는 지역 직원들에게 지원을 제공한다. 또한, 시험 관리 팀을 통해 지역과 시험 계약자 간의 연락을 담당한다.

지역 차원에서는 지역 평가 코디네이터가 지정되어 플로리다 교육부(FDOE), 계약자, 그리고 지역 내 학교들 간의 연락을 담당한다. 대형 지역의 경우에는 여러 명의 직원들이 학교를 지원하며 시험 관리에 기여한다.

학교 차원에서는 학교 행정관이나 상담 교사가 학교 평가 코디네이터로 지정되어 지역과 학교 간의 소통을 담당하며, 기술 코디네이터는 시험에 필요한 기술적 환경을 설정하고 시험 중 기술적인 문제가 발생했을 때 해결하는 역할을 한다.

시험 세션에서 시험 관리자(TA)는 시험을 감독하며, 이들은 학교 구역의 교직원, 보통 교사들로 구성된다. TA는 시험 관리 전 교육을 받아야 하고, 시험 기간 동안 시험실에 상주해야 한다. 또한 시험 감독관(proctor)은 추가적인 모니터링 역할을 수행하며, 시험실 내 학생 수가 25명을 초과할 경우 필수로 배치된다. 시험 감독관은 학교 직원뿐만 아니라 학교 직원이 아니어도 될 수 있으며, 이들은 시험 전에 보안 절차에 대해 교육을 받는다. 비학교 직원 시험 감독은 시험 관리 절차에 직접 참여하지 않지만, TA를 보조할 수 있다. 시험 세션 중 시험 관리자는 가족 구성원에게 시험 관련 정보를 공유하면 안 되고, TA와 관련된 학생들은 다른 관리자로 배정된다. 또한, 학생의 부모나 보호자는 그 학생의 시험실에 있을 수 없다. 시험 관리에 관련된 모든 인력은 시험 보안 정책 및 절차를 충분히 숙지하고, 시험 관리 및 보안 동의서에 서명해야 하며, TA는 부정행위 방지 및 시험 무효화를 피하기 위한 금지 활동 동의서에도 서명해야 한다. 온라인 시험을 시행하는 데에 있어서 보다 구체적인 역할과 책임은 아래와 같다.

지역 평가 코디네이터(District Assessment Coordinator, DAC), 학교 평가 코디네이터(School Assessment Coordinator: SAC), 시험 관리자(Test Administrator:

TA) 각각은 온라인 시험 시스템에서 구체적인 역할과 책임을 지닌다.

- 지역 평가 코디네이터(District Assessment Coordinator, DAC): DAC는 지역 차원에서 시험을 조정할 책임이 있으며, 각 학교의 SAC가 적절한 교육을 받고 정책 및 절차를 인지하며, CAI(컴퓨터 보조 시험) 시스템을 사용할 수 있도록 훈련받았는지 확인한다.
- 학교 평가 코디네이터(School Assessment Coordinator, SAC): 각 시험 관리 전에 SAC와 DAC는 학생의 시험 자격이 TIDE에 올바르게 등록되었는지, 그리고 편의 제공 또는 시험 환경 설정이 정확한지 확인해야 한다. 컴퓨터 기반 시험에 참여하려면 학생이 TIDE에 해당 시험을 치를 자격이 있는 것으로 등재되어야 한다. SAC는 시험이 그들의 학교에서 플로리다 교육부(FDOE)에서 정한 시험 보안 및 기타 정책과 절차에 따라 진행되도록 책임을 진다. SAC는 시험 관리자(TA)를 식별하고 교육하는 데 주로 책임이 있다. SAC는 기술 코디네이터와 협력하여 컴퓨터 및 기기 시험을 위해 준비되었는지, 기술적 문제가 해결되었는지 확인한다. 시험 기간 동안 SAC는 시험 진행 상황을 모니터링하고, 모든 학생이 적절하게 시험에 참여했는지 확인하며, 필요한 경우 시험 문제를 처리한다.
- 시험 관리자(Test Administrator: TA): 시험 관리자(TA)는 FAST 평가를 감독하고, 각 SAC에서 제공하는 지침에 따라 시험 연습 세션을 수행한다. TA는 시험을 감독하기 전, 온라인 TA 인증 과정을 완료할 기회를 제공받는다. TA는 시험 환경을 준비하기 위해 필요한 매뉴얼과 사용자 가이드를 검토하고, 학생들이 책, 메모, 연습 용지(FAST 수학용으로 제공된 용지 외), 전자 기기(예: 휴대전화, 스마트폰, 스마트 워치 등)를 소지하지 않도록 해야 한다. TA는 매뉴얼에 명시된 지침에 따라 FAST 평가를 관리해야 하며, 시험 관리 중 발생한 어떠한 편차도 SAC에 보고해야 하며, 필요할 경우 SAC는 이를 DAC에 보고하고, 최종적으로 플로리다 교육부(FDOE)에 보고해야 한다. TA는 또한 특정 시험에 허용된 자료 외에 다른 자료가 사용되지 않도록 보장해야 하며, 시험에 필요한 추가 자원은 제공되지 않는다. 7학년 및 8학년 FAST 수학 평가에서는 시험 시스템(TDS)에 계산기가 포함되어 있다. 2022-2023 학년도에 6학년 FAST 수학 시험에서는 4기능 계산기가 도입되었으며, 해당 시험 세션 중에 학생들에게 휴대용 계산기를 제공할 수도 있다. 플로리다 교육부(FDOE)는 승인된 휴대용 계산기 목록과 금지된 기능 목록을 제공하며, 이 목록은 주 전역의 웹사이트에서 확인할 수 있다. 컴퓨터 기반의 FSA ELA 읽기 재시험 평가에서는 학생들이 시험 중에 오디오 지문을 듣기 위해 헤드폰 또는 이어

버드를 사용해야 한다. 플로리다 교육부(FDOE)는 학교에서 헤드폰이나 이어버드를 제공하지 않으며, 학생들은 자신의 것을 사용할 수 있다. 시험 첫날, TA는 학생들이 사용할 컴퓨터나 기기에서 헤드폰이 제대로 작동하는지 확인해야 하며, 이어폰이 제대로 작동하는지 확인하기 위해 신호음이 내장된 시스템을 사용한다. 시험이 진행되는 동안 학생들은 자신의 헤드폰 또는 이어버드를 사용하는지 확인해야 한다.

#### 나) 온라인 대체 시험(지필 시험)

주 전역 평가는 컴퓨터 기반 시험(CBT) 형식으로 진행되며, 자격이 있는 학생들을 위해 편의 제공으로 지필 시험(Paper based Test: PBT) 형식이 제공된다.

지필 시험을 보는 대상 학생들은 주로 장애가 있는 학생들로, 이들은 개별적인 필요에 맞춘 시험 편의를 제공받아야 한다. 장애 학생들은 장애인 교육법(IDEA)에 따라 개별 교육 계획(IEP) 또는 504 계획에 명시된 편의 제공을 통해 시험에 참여할 수 있다. 이러한 편의 제공에는 유연한 시험 방식(발표, 응답, 시간, 환경), 보조 기기 및 도구의 사용 등이 포함된다. 편의 제공은 학생의 교실 수업에서 정기적으로 사용되는 것이어야 하며, 시험의 타당성을 훼손하거나 보안을 위협해서는 안 된다.

장애 학생들이 지필 시험을 통해 시험에 참여할 수 있도록 필요한 편의를 제공하는 것은 학생의 능력이나 지식을 정확히 평가하기 위한 중요한 요소이다. 특히, 편의 제공 결정은 IEP 또는 504 팀에 의해 이루어지며, 그 내용은 해당 계획에 기록되어야 한다. 또한, 시험 전에 학생들이 시험 형식과 절차에 익숙해질 수 있도록 시험 준비 교육을 제공해야 한다.

공립학교에 재학 중이 아니거나 공립학교 서비스를 받지 않는 장애 학생들 또한, 플로리다 법령(F.S.) 1003.01(3)(a) 또는 플로리다 행정 규칙(F.A.C.) 6A-19.001(6)에 정의된 학생으로서 장애가 인정되고, 요청된 편의 제공이 정기적으로 교육에 사용된다는 증빙이 있을 경우 편의를 제공받을 수 있다.

#### 다) 온라인 검사를 위해 제공되는 자료

특정 컴퓨터 기반 시험(CBT)에는 시험의 효율성과 편의성을 높이기 위한 다양한 추가 자료가 제공된다. 예를 들어, 8학년 과학, 시민 EOC, 미국 역사 EOC 시험에서는 CBT 워크시트를 제공하여 학생들이 시험 중 메모를 작성할 수 있게 한다. FAST

수학 시험을 치르는 학생들에게는 연습 용지가 주어진다. 또한, 일부 평가에서는 온라인 계산기와 참조 시트가 제공되며, 이를 통해 학생들은 시험에서 필요한 자료를 신속하게 참조할 수 있다. 8학년 과학 및 생물학 I EOC 시험에서는 온라인 플랫폼을 통해 주기율표가 제공된다. 이러한 전자기기 및 자료는 모두 시험 기관에서 제공되며, 학생들에게 적절한 지원을 제공한다.

시험 관리 스크립트는 각 시험 세션에서 학생들에게 제공되는 시험 지침 및 정책을 표준화하는 데 기여한다. 모든 학생이 동일한 지침을 받을 수 있도록 설계된 이 스크립트는 시험실에 입장한 후 시험 감독관이 학생들에게 그대로 읽어야 한다. 스크립트에는 전자기기 사용 금지 정책, CBT 및 PBT 지침, 시험 규칙 확인서와 관련된 정보가 포함된다. 플로리다 교육부(FDOE)는 시험 보안을 강화하기 위해 학생 개인이 소지한 전자기기 사용을 엄격히 금지하며, 이를 위반할 경우 시험이 무효화될 수 있음을 명확히 경고한다. 반면, 시험에 필요한 컴퓨터와 계산기 같은 전자기기는 시험 기관에서 제공되며, 학생들은 이를 통해 시험을 치른다.

학생들은 시험 전 휴대전화나 다른 개인 전자기기를 사용할 수 없으며, 시험 중 다른 학생들과 대화하거나 시험 문제에 대해 도움을 주고받는 행위가 금지된다. 시험 내용은 보안이 유지되어야 하며, 시험 종료 후에도 문제나 내용을 외부에 공유하는 것은 엄격히 금지된다. 이러한 시험 규칙을 학생들에게 설명한 후, 학생들은 해당 규칙을 이해하고 준수할 것임을 확인하기 위해 확인서에 서명하거나 동의 체크박스를 선택해야 한다.

마지막으로, 시험 관리 스크립트는 각 시험 세션에 할당된 시간과 휴식 시간, 추가 시간이 필요한 경우의 스케줄링 지침을 포함하고 있다. 시험 감독관은 이를 기반으로 학생들이 시험을 효율적으로 진행할 수 있도록 관리해야 하며, 모든 시험 절차는 플로리다 교육부(FDOE)가 설정한 표준화 지침을 엄격히 준수하여 운영되어야 한다.

#### 라) 결과 보고 및 피드백

FAST 검사를 통한 결과 보고 및 피드백 제공은 플로리다 보고 시스템(FRS), 개별 학생 보고서, 성적 문의 시스템을 통해 이뤄진다.

먼저 플로리다 보고 시스템(Florida Reporting System, FRS)은 Cambium Assessment, Inc.에서 개발 및 제공한 동적 플랫폼으로, 주 전역 평가에 대한 학생 성적을 학교, 지역 및 주 차원의 사용자들이 접근할 수 있도록 한다. 시험이 컴퓨터화 능력적응검사 방식이기 때문에, 성적은 학생이 시험을 완료한 후 24시간 이내에 FRS에

보고된다. 교사는 자신이 담당하는 학생들의 평가 완료 데이터를 볼 수 있으며, 학교 차원 사용자(사립학교 관리자, 학교 평가 코디네이터, 학교 관리자, 학교 보고 접근자)는 학교의 모든 학생 성적을, 지역 차원 사용자(지역 평가 코디네이터, 지역 관리자, 지역 보고 접근자)는 지역 내 모든 학생 성적을 볼 수 있다.

개별 학생 보고서(Individual Student Reports)는 FRS와 가족 포털(Family Portal)에서 동적으로 생성되는 보고서로, 각 학생의 성적에 대한 일반적인 정보를 제공한다. 여기에는 점수, 성취 수준, 보고 범주별 성취 수준, 그리고 이전 성과(해당될 경우, ELA 읽기 및 수학에 한정)가 포함된다. 보고서에는 또한 해당 학생의 성과가 같은 시험을 치른 학교, 지역, 주의 다른 학생들과 어떻게 비교되는지도 명시되며, 유용한 학습 자원도 제공된다. FAST ELA 읽기 및 수학 평가의 경우, 개별 학생 보고서에는 학생이 표준을 밑도는지, 근접 또는 충족하는지, 초과하는지를 나타내는 추가 설명과 학부모가 학생의 학습 진전을 돕기 위한 방법이 제안된다. 이와 같은 확장된 설명은 2024-25 학년도에 과학과 사회과목 보고서에도 개발될 예정이다.

성적 문의 시스템(Score Inquiry System)은 만약 학생의 성적이 누락되었거나 보고되지 않았거나 이상하다고 판단되는 경우, 해당 학생의 학교 지역은 플로리다주 플로리다 플로리다 교육부(FDOE)에 성적 문의를 제출할 수 있다. 플로리다 교육부(FDOE)와 평가 계약자는 해당 사례를 조사하고, 그 결과를 학교 지역에 제공하며, 필요한 경우 성적 보고서를 수정할 수 있다. 문의 사례로는 누락된 성적이나 ELA 전문가를 통한 학생의 글쓰기 응답 점수 확인 등이 있다.

## 나. 텍사스 STAAR 검사와 플로리다 FAST 검사 비교

텍사스 주에서 시행하는 STAAR 검사와 플로리다 주에서 시행하는 FAST 검사의 시험 기기, 보안 시스템, 시험 관리 및 전달 시스템, 학생 지원 기능, 채점 및 피드백, 계산기 사용, 시험 준비 및 스크립트 측면에서 비교 및 분석한 내용은 <표 VI-3>에 제시되어 있다.

<표 VI-3> 온라인 실시 방안 측면 텍사스 STAAR과 플로리다 FAST 검사 비교

항목	텍사스 STAAR	플로리다 FAST
시험 기기	학교에서 제공되는 컴퓨터, 아이패드 등 다양한 기기에서 시행 가능.	학교에서 제공되는 컴퓨터 또는 태블릿 등 다양한 기기에서 시행 가능.



항목	텍사스 STAAR	플로리다 FAST
보안 시스템	보안 브라우저(TDS)를 사용하여 핫 키, 인터넷 접근 등 제한. Pearson의 ePEN 시스템을 통해 채점.	CAI 보안 브라우저 사용으로 핫키, 복사 및 인터넷 사용 제한. FAST 시험도구에 서 제공됨.
시험 관리 및 전달 시스템	TDS(Test Delivery System)를 통해 시험 관리. 여러 기기를 통한 확장 가능성 있음.	TIDE와 TDS를 통해 시험 배포 및 관리. 다수의 장치에서 사용 가능하며 장치 간 비교 연구 시행됨.
학생 지원 기능	특별한 교육적 요구가 있는 학생들을 위한 지원 기능 제공.	특별한 교육적 요구가 있는 학생들을 위한 지원 기능 제공.
채점 및 피드백	Pearson ePEN 시스템에서 문항 분석 및 응답자의 익명성 보장.	FAST는 시험 완료 후 24시간 내에 플로리다 보고 시스템(FRS)을 통해 성적 보고. 빠른 피드백 제공.
계산기 사용	수학 평가에서 온라인 계산기 제공. 필요시 휴대용 계산기 사용 가능하며, 관련된 제한 규정도 제공됨.	
시험 준비 및 스크립트	시험 관리 스크립트를 통해 시험 절차를 표준화. 전자기기 사용 금지와 관련된 엄격한 규정.	



| 인천교육종단연구 기반연구 - 교과역량 평가 방안 제언 |

# VII

## 결론





## 결론

본 연구에서는 인천교육중단연구 시행에 앞서 교과역량 평가 방안 마련을 위해 국내 교육중단연구 사례, 국외 교육중단연구 사례, 기존 학업성취도 평가 결과의 활용 가능성을 살펴보았으며, 학교급별 차이점 및 평가 시행 방식(온·오프라인)에 대해 분석하고 제안하였다. 각 장에서의 연구 결과에 대한 요약과 결론은 다음과 같다.

### 1. 국내 교육중단연구 사례 분석 요약 및 결론

먼저, 국내 교육중단연구 사례 분석에서는 인지 영역 조사 및 평가 도구, 평가 영역을 정리하고 지역 수준 교육중단연구 담당자를 면담한 내용을 바탕으로 세 가지 측면에서 개선 방안을 도출하여 제시하였다.

특히 지역 수준에서 인지적 영역 평가 방식을 구체적으로 살펴보면, 인지 영역 평가 방식은 평가 도구와 평가 방법, 평가 방향에 있어서 공통점이 나타났다. 평가 도구에서는 설문조사와 별도의 검사 도구를 활용하며, 정규 수업시간과 동일한 시간을 배정하여 선다형과 서술형 문항을 혼합한 검사 도구를 사용하고 있었다. 다음으로 평가 방법에 있어서도 수직척도화를 위한 공통문항을 설계하여 적용하였으며, 예비검사를 통한 문항 검증 및 수정 과정을 거치고 대개 문항반응이론에 기반한 분석을 실시하였다. 평가 방향과 관련해서 초기에는 교과지식을 바탕으로 한 학업성취도를 평가하였다가 교육과정의 변화에 발맞춰 역량 중심 평가로의 전환을 시도하고 있었다.

주된 차이점으로는 평가 영역과 문항 구성, 척도화 방법을 들 수 있는데 서울과 경기도는 국어, 영어, 수학 3개 교과를 대상으로 전남은 영어를 제외한 국어, 수학, 사회, 과학 4개 교과를 평가 영역으로 하였다. 문항을 구성할 때 서울은 5지 선다형 중심으로 하였으며 전남은 학교급별 선다형 문항 수를 차등하여 초등은 4지 선다형, 중등 이상은 5지 선다형 문항을 구성하였다. 척도화 방법에서 서울은 PARSCALE을 활용하였으며 전남은 부분검사 분할방식을 적용하였다.

국내 교육중단연구 사례를 분석한 결과, 크게 두 가지 측면에서 인천교육중단연구를 위한 시사점을 도출할 수 있다.

첫째, 교과역량을 조사 및 평가하고자 할 때에는 학교급에 따라 다르게 접근할 필요가 있다. 국내 교육종단연구 사례를 살펴보면, 대부분 초등학교와 중학교, 고등학교에 걸쳐 여러 개의 코호트를 운영하면서 학교급에 따라 조사 내용을 달리하고 있음을 확인할 수 있다. KYPS와 KCYPS 2010의 경우 초등학교에 비해 중학교에서 수집하는 조사 내용이 더 세부적이고 다양하며, 중학교나 고등학교 진학 이후 조사 내용을 추가하는 경향을 보였다. 한편, 초등학교나 중학교를 대상으로 조사를 시작했던 KELS 2005와 KELS 2013에서는 고등학교 진학 이후 별도의 평가 도구를 실시하지 않고 국가 수준 평가 결과를 연계하여 수집하는 데 그쳤다. 2주기 종단연구를 시작한 서울교육종단연구 2020에서는 평가 도구를 제작할 때부터 고등학교를 대상으로는 인지 영역 관련 내용을 제외하였다. 따라서 고등학교의 경우 별도의 시험 실시가 입시 및 학업에 부담을 줄 수 있으며 응답의 성실성 등을 고려할 때, 초등학교와 중학교에서는 교과역량을 평가하더라도 고등학교는 제외하는 방안을 고려할 필요가 있다.

둘째, 교과역량을 평가하고자 할 때 여러 교과 대신 필수 교과를 선별하여 실시할 필요가 있다. 국내 교육종단연구 사례를 살펴보면, 국가 수준 교육종단연구에서 전체 교과 성적을 조사 내용에 포함하는 경우가 있었다. 그러나 성적이라는 정보의 민감성을 고려하여 1차년도 조사 내용에서 제외하거나 중간고사와 기말고사 폐지 및 성적 공개와 관련된 교육 현장의 변화에 따라 조사 내용을 변경하는 등 일정한 데이터 수집에 어려움을 겪었다. 그 대안으로 별도의 평가도구를 개발하여 시행하는 사례가 특히 지역 수준 교육종단연구에서 일반적인 경향으로 나타났으나 학생 대상 설문조사 이외에 평가를 시행하는 데 시간적·물리적 제약을 고려하지 않을 수 없었다. 평가도구의 개발 및 실시 과정에서 막대한 예산과 인력, 그리고 시간이 소요되며 장시간 조사 및 평가에 참여하는 부담감 역시 상당하여 종단연구의 지속성을 해칠 수 있었다. 또한 지역 수준 교육종단연구에서는 단순히 여러 교과지식을 평가하기보다는 역량에 초점을 맞추며 다양한 교과를 접목하여 평가를 시행하려는 노력이 두드러졌다. 이에 전남교육종단연구에서는 초등학교 4학년이라는 조사 대상 학년을 고려하여 교육의 효과를 측정하기에는 교육 기간이 짧고, 선행학습 등 사교육의 영향이 클 것으로 예상되는 영어 교과 대신 국어, 수학 교과에 탐구 영역(사회, 과학 교과)을 접목하여 국어/사회, 수학/과학 등 2교시로 평가도구를 구성하여 시행하였다. 서울교육종단연구와 경기교육종단연구에서도 역량 중심 평가로 전환을 시도하면서 국어와 수학 교과 이외의 교과는 별도의 교과를 검사 대상으로 설정하기보다는 여러 교과를 통합하여 하나의 역량을 검사 대상으로 상정하고 있다. 따라서 교과역량을 평가하고자 할 때 구체적으로 측정하고자 하는 인지 영역을

고려하여 국어, 수학 등을 선별하여 필수 교과 위주의 교과역량을 평가하는 방안을 고려할 필요가 있다.

## 2. 국외 교육종단연구 사례 분석 요약 및 결론

텍사스 주의 STAAR 검사와 플로리다 주의 FAST 검사의 사례를 살펴보았을 때, 두 검사는 각각의 목적, 대상 학년, 시험 유형 및 성취 수준에서 차이를 보인다.

STAAR는 학년 진급과 졸업 자격 평가를 목적으로 하며, 3-8학년(수학, 읽기/언어, 과학, 사회)과 고등학교의 EOC 과목을 대상으로 연간 1회(봄) 시행된다. 시험 유형은 컴퓨터 기반 검사(CBT)로, 다지선다형, 서술형, 논술형을 포함하며, 성취 수준은 ‘미달, 근접, 달성, 초과’의 4단계로 나뉜다. 반면, 플로리다 FAST는 학업성취도 추적과 학습 진도 모니터링을 목적으로 하며, 3-10학년(읽기)과 3-8학년(수학)을 대상으로 연간 3회(PM1, PM2, PM3) 시행된다. 이 시험은 컴퓨터 적응형(CAT) 방식을 채택하여 다지선다형, 텍스트 강조, 테이블 매치 등의 문제 유형을 제공하며, 성취 수준은 레벨 1에서 5까지의 5단계로 구성된다.

위와 같이 차이를 보이지만, STAAR과 FAST는 유사한 특징도 가지고 있다. 구체적으로는 특수 환경에 처한 학생들을 위해 IEP 및 가정학습 학생 대상 원격 시험과 일부 종이 시험을 제공한다. 또한 교과역량보다는 각 교과의 내용 범주를 중심으로 평가를 진행하는 특징을 보인다.

위와 같이 텍사스 주의 STAAR 검사와 플로리다 주의 FAST 검사의 사례를 살펴보았을 때, 인천교육청에서 시행할 예정인 교육종단연구에 대한 제안은 다음과 같다.

첫째, 표집 조사 방식으로 종단연구를 진행할 필요가 있다. 텍사스 주의 STAAR과 플로리다 주의 FAST 검사는 졸업 요건과 연계되어 있어 전수조사가 가능하다. 이는 법적 요구와도 밀접한 관련이 있어, 모든 학생이 매년 시험에 응시하도록 의무화되어 있다. 특히, STAAR는 학년 진급 여부와 졸업 자격을 결정하는 주요 시험으로 학생들이 반드시 응시해야 하며, 이를 통해 전수조사가 자연스럽게 이루어진다. 하지만 한국의 경우 학년 진급과 졸업을 위한 시험이 의무적으로 시행되지 않는 상황이다. 현재 시행 중인 국가수준학업성취도 평가조차 전수조사에서 표집 조사로 전환된 바 있으며, 이는 전수조사가 현실적으로 어려운 한국의 상황을 보여준다. 따라서 표집 조사 방식으로 종단연구를 시행하는 것을 제안한다.

둘째, 역량 중심 교육과정과의 연계를 고려하여야 한다. 텍사스 주의 STAAR과 플로리다 주 FAST 검사의 평가 내용은 각각 과목별 보고 범주(Reporting category)를 기준으로 학생들의 성취도를 평가하지만, 이를 역량 중심 교육과정과 직접적으로 연결하기는 어렵다. STAAR 검사는 과목별로 지식과 기술을 평가하고 있지만, 역량에 중점을 두기보다는 전통적인 학업 성취에 중점을 두고 있다. 또한, FAST 검사는 컴퓨터화 능력적응검사(CAT)를 통해 학생들의 개별 학습 진도를 모니터링하는 데 초점을 맞추고 있지만, 역량을 직접적으로 측정하는 체계는 갖추지 않았다. 한국의 역량 중심 교육과정을 반영한 종단연구에서는, 단순한 학업성취도 평가에 그치지 않고, 학생들이 실생활에서 문제를 해결하고 비판적 사고를 활용하는 역량을 평가하는 방법을 고려해야 한다. 역량 중심 평가 항목을 포함하여 학생들의 문제 해결 능력, 비판적 사고, 협력 및 의사소통 능력을 평가할 수 있는 보고 범주를 개발하고, 종단적으로 역량의 발전을 추적하는 연구를 제안한다. 이는 미래 사회에서 요구되는 역량을 평가하고, 학생들의 성장 경로를 더욱 다각적으로 파악할 수 있도록 도울 것이다.

### 3. 기존 학업성취도 평가 활용 방안 요약 및 결론

기존 학업성취도 평가는 평가 시행 목적이나 결과 활용 방향에 따라 대상 학교급이 다르게 나타났다. 일반적으로 주요 교과인 국어, 영어, 수학을 중심으로 필요에 따라 사회, 과학 등의 교과에 대한 평가를 추가로 시행하고 있으며, 직전 학년이나 이전 학년의 전 범위에서 검사 문항을 출제하고 있었다. 역량 중심 교육과정 반영 측면에서 살펴보면, 국가수준 학업성취도 평가, 맞춤형 학업성취도 자율평가, 기초학력진단평가는 평가를 및 결과 보고 차원에서 영역, 성취기준, 역량과 같은 개념을 제시함으로써 교과지식이 아닌 역량을 측정하고자 하였다. 그러나 실질적으로는 교과지식을 묻는 검사에 가까운 성격을 띠고 있는 것으로 파악되었다.

기존 종단연구의 인지적 발달 조사 항목과 관련하여 초등학교, 중학교에서는 별도의 학업성취도 평가를 시행하는 편이었으나, 고등학교는 입시 문제와 관련된 평가 부담 등의 한계로 기존 학업성취도 평가 결과를 조사하고 있었다. 그러나 개별 학생의 국가수준 학업성취도 평가 결과 또는 전국연합학력평가 결과 자료를 직접적으로 수집하는 것은 어렵기 때문에, 학생 설문지를 통해 자기보고식 형태로 응답하도록 하고 있었다. 이러한 방안은 학생의 주관적인 응답이기 때문에 자료의 신뢰도가 떨어질 수 있다는



한계를 지니므로, 신뢰도를 높이기 위한 방안으로 교사 대상의 학생 평가지를 별도로 실시하여 개별 학생의 학업성취도에 대해 교차 검증하고 있었다.

기존 학업성취도 평가의 활용 가능성에 대해 분석한 결과로 도출된 결론 및 시사점은 다음과 같다. 기존 학업성취도 평가 결과를 활용할 경우, 문항 개발에 드는 비용과 시간을 절감할 수 있다는 이점이 있다. 또한 자체적으로 개발한 교과역량 평가를 실시하게 되면 교사와 학생에게 부담이 가중될 수 있다는 측면에서 별도의 교과역량 평가 도구를 시행하는 것보다 기존의 평가 결과를 활용하는 방향이 적합할 수 있다. 그러나 학생에게 자기보고식 설문으로 내신 또는 전국연합학력평가 등급을 묻는 방식은 실제 학생의 교과역량을 정확히 반영할 수 있는지에 대해 신뢰성 문제가 제기될 수 있다. 또한 교육종단연구의 경우, 학생의 인지적 성장 및 발달을 추적하려는 목적을 달성하기 위해 일정 비율의 가교문항(anchor item)을 포함하는 등 수직척도화를 위해 설계 단계에서의 노력이 요구된다. 그러나 기존 학업성취도 평가의 경우, 일반적으로 검사 설계 단계에 개입할 수 없으므로 해당 평가에 대한 성적 향상 정도가 학생의 성장 결과라는 결론을 내리기 어렵다. 특히 국가수준 학업성취도 평가의 경우, 전국의 중학교, 고등학교 학생의 3%를 표집하여 검사를 실시하므로, 교육종단연구의 패널 학생이 꾸준히 표집되지 않는다면 그 결과를 얻을 수 없다는 한계가 있다.

따라서 기존 학업성취도 평가 결과를 활용하고자 할 경우, 어떤 평가와 과목에 대한 정보를 수집할 것인지, 자기보고식 설문의 신뢰성 문제는 어떻게 해결할 것인지, 수직척도화와 관련하여 어떤 식으로 학생의 인지적 성장·발달을 볼 것인지에 대해 신중하게 결정해야 한다. 하지만 기존 학업성취도 평가 결과를 활용하는 것이 갖는 다양한 한계점을 고려한다면, 대상 학교급 및 교과를 제한하더라도 교육종단연구의 목적에 맞는 별도의 교과역량 평가 도구를 개발 및 시행하는 것이 가장 바람직한 방안으로 보인다.

## 4. 학교급별 교과역량 평가 방안 요약 및 결론

### 가. 국내 교육종단연구의 학교급별 실시 방안

본 장에서는 별도의 평가 도구를 개발하여 실시한 총 8개의 교육종단연구를 분석하였다. 분석대상은 한국교육종단연구2005, 한국교육종단연구2010, 한국 초·중등 학교교육 성과 종단조사, 서울교육종단연구2010, 서울교육종단연구2020, 경기교육종단연구, 경기기초학력종단연구, 전남교육종단연구이다.

초등학교에서는 7개 연구가 별도의 평가를 실시하였다. 한국교육종단연구2013에서는 국어, 영어, 수학 기초능력 검사를 개발하여 실시하였다. 한국 초·중등 학교교육 성과 종단조사에서는 국어, 수학의 핵심교과지식과 교과역량을 평가하고, 4개 교과(국어, 수학, 사회, 과학) 통합형 문제해결역량을 평가하였다. 서울교육종단연구2010에서는 국어, 수학, 영어 학업성취도 검사를 시행하였고 서울교육종단연구2020에서는 국어, 수학, 영어 교과역량 중심의 서울학생역량 검사를 시행하였다. 경기교육종단연구에서는 국어, 영어, 수학 학업성취도 검사를 시행하였고, 경기기초학력종단연구에서는 기초 문식력과 기초 수리력을 평가하였다. 전남교육종단연구에서는 국어/사회, 수학/과학을 2교시로 구성하여 기초능력 검사를 시행하였다.

중학교에서는 6개 연구가 별도의 평가를 시행하였다. 대부분 초등학교와 동일한 평가체제를 유지하였으나, 한국교육종단연구의 경우 다음과 같은 특징을 보였다. 한국교육종단연구2005의 경우 국어, 수학, 영어 학업성취도 검사를 시행하였고, 한국교육종단연구2013에서는 국어, 영어, 수학 기초능력 검사를 시행하였다.

고등학교에서는 평가 방식에서 뚜렷한 차이를 보였다. 자체 평가를 실시한 연구는 서울교육종단연구2010, 경기교육종단연구, 전남교육종단연구로 초·중학교와 동일한 평가체제를 유지하였다. 서울교육종단연구2020에서는 고등학교 대상 평가에서 교과역량 평가를 제외하였다. 반면 기존 평가 결과를 활용한 연구는 한국교육종단연구2005, 2013으로 별도 평가를 실시하지 않고 다음의 기존 평가 결과를 수집하였다. 국가수준 학업성취도평가 결과(국어, 사회, 수학, 과학, 영어), 전국연합학력평가 결과(언어, 수리가형, 수리나형, 영어), 대학수학능력시험모의평가 결과이다.

이상 살펴본 내용을 바탕으로 도출한 특징 및 시사점은 다음과 같다. 우선, 국가수준 연구는 고교 진학 후 기존 평가 결과를 활용하는 반면, 지역 수준 연구는 대체로 자체 평가 도구를 지속해서 활용하였다. 평가 과목은 대체로 국어, 수학, 영어가 기본이나 지역에 따라 사회, 과학을 포함하였다. 최근 교과지식 평가에서 교과 역량 중심 평가로 전환하는 추세를 보였고, 고교단계에서는 학생 부담 등을 고려해 기존 평가 결과를 활용하는 경향이 나타났다.

#### 나. 국외 교육종단연구의 학교급별 실시 방안

국외 교육종단연구 사례로서 텍사스 주의 STAAR 검사와 플로리다 주의 FAST 검사의 사례를 살펴보았을 때, 학교급별 시행 특징은 다음과 같다.

초등학교와 중학교에서는 STAAR 검사가 학년별 학업성취도를 평가하는 표준화된 시험으로 활용되며, 수학, 읽기/언어(RLA), 과학, 사회 과목을 대상으로 연 1회 시행한다. 시험 결과는 학년 진급에 중요한 자료로 사용된다. 고등학교에서는 STAAR가 과목별 EOC(End-of-Course) 시험으로 구성되어 있으며, 대수 I, 영어 I·II, 생물, 미국 역사 등의 과목을 심화 평가한다. 이 평가 결과는 졸업 요건의 일부로 포함되며, 연 1회 시행하되, 통과하지 못한 학생에게는 재시험 기회가 제공된다. 플로리다의 FAST 검사는 초등학교와 중학교에서 읽기(ELA)와 수학 성취도 추적을 중점으로 연간 3회 시행되며, 학업성취도와 학습 성장 평가의 자료로 활용된다. 고등학교에서는 FAST 검사가 아닌 EOC 시험을 통해 주요 과목(대수 I, 기하학, 생물학, 미국 역사, 시민 교육)을 심층적으로 평가하며, 각 과목 수강 후 1회 시행된다. 졸업 요건에는 FAST 검사가 아닌 EOC 시험의 통과 여부가 포함된다.

위와 같이 국외 교육중단연구 사례로서 텍사스 주의 STAAR 검사와 플로리다 주의 FAST 검사의 사례를 살펴보았을 때, 인천교육청에서 시행할 예정인 교육중단연구에 대한 제안은 다음과 같다.

첫째, 초등학교, 중학교에 한하여 별도의 성취도 검사를 개발 및 시행할 것을 제안한다. STAAR와 FAST 모두 고등학교 시험 부담을 줄이기 위해 다양한 방식으로 시험의 빈도와 형식을 조정했다. FAST는 초등학교와 중학교에서 연간 3회씩 치르지만, 고등학교에서는 연간 1회로 축소된다. 이는 고등학생들의 시험 부담을 줄이고, 학업 성취를 더 집중적으로 평가하기 위한 조치일 수 있다. STAAR 시험 역시 고등학교에서는 각 과목별로 졸업 요건을 충족하기 위한 종합 평가인 STAAR EOC로 전환된다. 또한 STAAR와 FAST 모두 고등학교에서는 졸업 시험과 밀접하게 연계되며, 학년 진급 및 졸업 요건을 충족하기 위한 평가로 활용된다. 이러한 이유로, 고등학교에서의 STAAR와 FAST는 과목별로 세분화하여 검사를 시행할 수 있다. 국내 고등학생들의 대입 준비 부담을 줄이기 위해 초등학교와 중학교에 한해 개별 성취도 검사를 시행할 것을 제안한다. 고등학생들의 대입 준비 부담을 완화하기 위해서는 이미 치르고 있는 수능 또는 모의고사 결과를 중단연구에 활용하는 것이 효율적이나, 고등학생 시험 결과 공유를 위한 관련 기관의 협조가 현실적으로 어렵다. 이에 따라 인천교육중단연구에서는 초등학교와 중학교 학생들을 대상으로 개별 성취도 검사를 시행할 것을 권장한다.

둘째, 교과목의 경우, 국어, 수학 교과를 대상으로 교과역량을 평가하는 것이 바람직하다. FAST 검사는 주로 초등학교와 중학교 학년을 대상으로, 읽기(ELA)와 수학을

평가한다. 읽기는 3학년부터 10학년(고2)까지, 수학은 3학년부터 8학년(중3)까지 평가가 이루어진다. STAAR 검사는 읽기(RLA)와 수학을 모두 3학년부터 8학년까지 평가하고, 과학은 5학년과 8학년에서, 사회는 8학년에서만 시행한다. 고등학교에서는 대수 I, 영어 I, 영어 II, 생물, 역사로 세분화된 과목을 평가한다. 이처럼 FAST와 STAAR 검사는 공통적으로 읽기와 수학은 3~8학년에 걸쳐 지속적으로 평가하여, 학생의 인지 역량을 종단적으로 추적하고 있다. 또한 읽기와 수학은 기초 학습 능력을 강화하고, 인지 발달과 사고 구조를 형성하는 데에 필수적인 학습 과정이라는 점에서 국내에서도 국어와 수학을 대상으로 교과역량 평가를 시행할 것을 제안한다.

#### 다. 기존 학업성취도 평가의 학교급별 활용 방안

기존 학업성취도 평가와 관련하여 학교급별 활용 방안에 대한 요약, 결론 및 시사점은 다음과 같다.

먼저, 기존 학업성취도 평가의 대상은 2024 국가수준 학업성취도 평가의 경우, 중학교와 고등학교 학생의 일정 비율을 표집하고 있었다. 맞춤형 학업성취도 자율평가는 초등학교, 중학교, 고등학교 중에서 참여를 희망하는 학교를 대상으로 하며, 전국연합학력평가는 대학수학능력시험을 준비하는 목적이 있으므로 전국의 고등학교 학생들을 대상으로 한다. 성취평가제는 중학교, 고등학교 학생들의 내신 평가를 위해 준거참조평가 방식으로 실시되며, 책임교육학년제는 초등학교 3학년, 중학교 1학년을 학습적으로 중요한 전환기로 지정하여 교육적 지원과 평가를 실시하는 제도이다. 기초학력진단평가는 기초학력 보장을 목적으로 학생들의 학습 부진 및 결손을 진단하는 검사이며, 인천에서는 초등학교 3학년부터 고등학교 2학년까지를 대상으로 한다.

국가수준 학업성취도 평가는 표집평가라는 점에서, 맞춤형 학업성취도 자율평가는 희망 학교만을 대상으로 한다는 점에서 교육종단연구의 교과역량 조사를 위해 활용하기에는 한계가 있다. 전국연합학력평가 결과는 고등학생 패널의 교과역량 조사에 활용할 수 있으며, 성취평가제에 따른 내신 점수는 중학생, 고등학생 패널의 교과역량을 조사하는 데 활용할 수 있을 것으로 보인다. 책임교육학년제를 통해 제공되는 평가 결과의 경우, 특정 학년만을 대상으로 하는 제도이므로 교육종단연구에는 활용이 어려울 것으로 보인다. 기초학력진단평가는 기초학력 부진 여부 파악을 목적으로 함으로써 검사 난이도가 매우 쉬워 학생들을 변별하기가 어려우며, 기초학력 도달/미도달에 대한 정보만을 제공하므로 학생의 종단적 변화를 보기 위해 기초학력진단평가를 활용

하는 것은 적합하지 않은 것으로 판단된다.

교육종단연구에서 기존 학업성취도 평가 결과를 학교급에 따라 어떤 식으로 활용하고 있는지에 대한 분석 결과, 한국교육종단연구2005에서는 학생들의 교과역량을 조사하기 위해 고등학교 1학년의 국가수준 학업성취도 평가 점수와 고등학교 3학년의 대학수학능력시험 점수를 수집하였다. 국가수준 학업성취도 평가가 전수평가 방식에서 표집평가 방식으로 전환되었으므로, 해당 평가 결과를 지속적으로 활용하기에는 어려움이 있었다. 따라서 한국교육종단연구2013에서는 고등학교 1~3학년 모두 전국연합학력평가 등급(국어, 영어, 수학)을 응답하도록 하였다. 한국복지패널은 초등학교 5학년을 대상으로 한 조사로서, 학교 성적 중 전과목 평균과 국어, 영어, 수학 점수에 대해 자기보고식 척도로 응답하도록 하였다. 한국교육고용패널조사 I, II에서는 학교 성적을 석차나 등급 수준에서 조사하였다. 또한 자기보고식 설문은 한계를 극복하고자 교사의 학생 평가지 응답 결과와 교차검증하거나 교육행정정보시스템(NEIS)과 연계하는 등의 방안을 활용하고 있었다.

따라서 학교급별 교과역량 조사를 위해 기존 학업성취도 평가 결과를 활용할 경우, 본 장에서 제안하고자 하는 바는 다음과 같다. 고등학교를 대상으로 하는 경우, 전국연합학력평가 결과나 성취평가제에 따른 성취율(A, B, C, D, E 수준)을 응답하도록 할 수 있다. 중학교는 성취평가제에 따른 성취율을 조사한다면 고등학교 자료와 연계하는데 가장 용이할 것으로 보인다. 초등학교, 중학교, 고등학교를 모두 연계하고자 하는 경우에는 학교 내신 성적에 대한 조사가 가장 적절할 것으로 보이나, 초등학교는 학교 성적에 대한 자기 평가와 같은 형태의 질문이 과연 적절할 것인지 고려해 볼 필요가 있다. 또한 직접적으로 시험 성적을 수집하는 데 현실적인 어려움이 있어 대부분의 교육종단연구에서는 자기보고식 설문을 활용하고 있는데, 응답의 신뢰도를 확보하기 위해서는 교사 설문을 통해 교차검증을 하는 등의 노력이 필요하다.

## 5. 교과역량 평가 시행 방식 요약 및 결론

### 가. 국내 교육종단연구의 평가 시행 방식

본 연구에서 다룬 국내 교육종단연구 사례 중에서 웹 기반 종단조사를 설계한 평가는 한국 초·중등 학교교육 성과 종단연구가 유일했다. 한국 초·중등 학교교육 성과 종단연구는 2회의 예비검사를 시행하였으며, 1차 예비검사에서는 참여자 일부가 자동 로그아웃

움이 되는 현상이 발생하기도 하였다. 그러나 1, 2차 시행에서 확인한 결과, 학생들이 문제 풀이 방법에 높은 흥미와 호기심을 보였으며, 컴퓨터 등 IT에 익숙한 학생들에게 저학년에도 적용할 수 있는 방식으로 판단되었다. 한편, 서울교육중단연구(SELS 2020)는 코로나19 상황을 고려하여 학부모 설문제 온라인 조사를 활용하여 실시하였는데, 온라인 조사 방법의 경우 오프라인 조사에 비해 응답률이 감소될 수 있고 무성의하게 답변할 가능성도 있어 학생을 대상으로 한 성취도 검사나 역량 검사를 시행하기에는 시스템적으로 매우 불안정하고 비용도 상당히 든다는 단점을 고려하여 학생 대상 조사는 온라인 조사 방법을 도입하기가 현실적으로 어렵다고 판단하였다.

경기교육중단연구 역시 학교급이 올라갈수록 학부모 설문지 회수율이 낮아지는 문제가 있어 대상별로 온라인 설문조사 방식의 적합성을 검토할 필요성이 제기되기도 하였으나 마찬가지로 학생을 대상으로 한 성취도 검사나 역량 검사를 온라인으로 실시하는 방안에 대해서는 언급하지 않았다. 시도교육청 교육중단연구 담당자 면담에서도 온라인 검사를 시행하기 위해서는 학교 현장의 인프라 문제가 해결되어야 한다는 의견이 있었다. 아울러, 온라인 검사를 개발하는 경우 태블릿PC, 데스크톱PC, 모바일 등 접속 기기 환경에 따라 별도 개발이 필요하며 AI 자동 채점 등 기능을 활용하기 위해서는 막대한 비용이 필요하다는 점이 한계점으로 지적되었다.

결론적으로 학교 소재지 등에 따른 인프라 차이와 비용 문제를 극복하지 않으면 교과역량 평가에 온라인 검사를 도입하기는 시기상조로 여겨진다. 다만, 조사의 용이성을 위해 온라인 설문을 도입하는 방안을 검토할 필요가 있다.

#### 나. 국외 교육중단연구의 평가 시행 방식

국외 교육중단연구 사례로서 텍사스 주의 STAAR 검사와 플로리다 주의 FAST 검사를 살펴보았을 때, 두 검사는 온라인 평가의 효율성과 보안성을 높이기 위해 다양한 기술적 접근 방식을 채택하고 있다.

텍사스 STAAR 검사와 플로리다 FAST 검사는 모두 다양한 기기에서 시행 가능하며, STAAR은 학교에서 제공되는 컴퓨터, 아이패드 등을, FAST는 컴퓨터 또는 태블릿을 이용해 평가를 진행한다. 두 시험 모두 보안성을 강화하기 위해 각각 보안 브라우저(TDS와 CAD)를 사용하여 찾기, 복사 및 인터넷 접근을 제한하고, STAAR은 Pearson의 ePEN 시스템을 통해 채점하며, FAST는 플로리다 보고 시스템(FRS)에서 24시간 내 성적 보고를 통해 빠른 피드백을 제공한다. 이와 함께 FAST는 CAT(Computerized

Adaptive Testing)을 적용하여 평가의 효율성과 정확성을 높이고 있다.

시험 관리 및 전달 시스템에서도 유사한 접근을 보이는데, STAAR는 TDS(Test Delivery System)를 통해 시험을 관리하고 여러 기기에서의 확장 가능성을 제공하는 반면, FAST는 TIDE와 TDS 시스템을 통해 시험을 배포 및 관리하며 기기 간 비교 연구를 시행한다. 또한, 두 시험 모두 특별한 교육적 요구를 가진 학생들을 위한 지원 기능을 제공하여 다양한 학습자에게 접근성을 보장한다.

수학 평가의 경우 STAAR과 FAST는 온라인 계산기를 제공하며, 필요에 따라 휴대용 계산기 사용을 허용하고 이에 대한 제한 규정을 명시하고 있다. 시험 준비와 관련하여서는 시험 관리 스크립트를 통해 절차를 표준화하며, 전자기기 사용 금지와 관련된 엄격한 규정을 적용하여 일관된 시험 환경을 유지한다.

위와 같이 국외 교육중단연구 사례로서 텍사스 주의 STAAR 검사와 플로리다 주의 FAST 검사를 살펴보았을 때, 인천교육청에서 시행할 교육중단연구에 대한 제안은 다음과 같다.

첫째, 컴퓨터 기반 검사(Computer based Test: CBT)를 실시할 경우, 그 장점을 충분히 이해하고 살릴 수 있어야 한다. 텍사스 주에서 시행하는 STAAR 시험은 최근 몇 년간 컴퓨터 기반 평가(CBT)로 전환되면서 다양한 장점을 보여주고 있다. STAAR의 CBT는 시험 응시자의 개별적인 요구를 충족시키기 위해 다양한 접근성 기능을 제공한다. 예를 들어, 학습 장애나 언어적 지원이 필요한 학생들을 위해 텍스트 확대, 텍스트 음성 변환, 언어 및 어휘 지원이 포함된 시험 환경을 제공한다. 또한 iPad와 같은 모바일 기기뿐만 아니라 Chromebook, Windows, Mac 등 다양한 기기에서 응시할 수 있다. 이를 통해 학교가 보유한 다양한 기술 자원에서도 시험이 원활하게 진행될 수 있으며, 학생들은 친숙한 기기에서 시험을 보게 되어 응시 편의성이 높다. 그리고 Pearson ePEN 시스템을 통해 안전하게 채점되고, 시험 관리 시스템(TDS)을 사용하여 응시 데이터를 보안 처리한다. 이에 더해 수동적인 시험지 배포 및 회수, 채점 절차가 생략되면서 시험 관리자의 부담이 줄어들고, 시스템을 통해 시험 세션을 예약하고 관리할 수 있어 전반적인 시험 관리 과정이 간소화된다. 인천교육중단연구에서도 STAAR의 컴퓨터 기반 검사(CBT) 방식을 참고하여 온라인 시험을 도입할 것을 제안한다. STAAR와 같이 CBT를 통해 빠른 피드백, 다양한 지원 기능, 높은 보안성을 제공할 수 있다. 당장 컴퓨터 기반 검사 형태로 시행하는 데에 무리가 있다면, 설문조사부터 컴퓨터 기반 검사 형태로 시행하는 것도 방법이 될 수 있다.

둘째, 컴퓨터화 검사를 실시하더라도 장애 학생 및 피치 못할 사정을 대비한 종이

기반 시험을 별도로 준비하여야 한다. 온라인 기반 평가가 매우 유용하더라도, 장애가 있는 학생이나 기술적 제약으로 인해 컴퓨터를 사용하기 어려운 학생들을 위한 종이 기반 시험이 준비되어야 한다. STAAR은 모두 장애 학생이나 기술적으로 온라인 시험에 참여할 수 없는 학생들을 위해 지필 시험을 제공한다. 인천교육중단연구에서 컴퓨터 기반 검사로 시행하더라도 장애 학생이나 특수한 상황에 처한 학생들을 위해 종이 시험을 적절히 준비해야 한다. 특히 인천교육중단연구에서 온라인 기반 시험을 도입할 때, 지필 시험과 같은 대체 방안을 마련하여 공정성을 확보할 필요가 있다.

셋째, 컴퓨터화 검사 실시를 위해 필요한 기기를 제공할 수 있도록 정책을 수립하고 다양한 기기를 지원하여야 한다. 온라인 시험을 원활하게 시행하기 위해서는 학생들이 다양한 디지털 기기를 사용할 수 있어야 하며, STAAR은 이 점에서 중요한 사례를 제공하고 있다. STAAR 시험은 아이패드, 컴퓨터, 크롬북 등 다양한 기기에서 응시할 수 있도록 하고 있으며, 각 학교에서 학생들에게 필요한 기기를 제공하고 있다. FAST 역시 다양한 기기에서 시험을 응시할 수 있도록 보안 브라우저를 통해 시스템을 운영하고 있으며, 시험 기기는 학교에서 제공된다. 이 시스템은 시험 응시 환경을 표준화하여, 학생들이 사용하는 기기와 관계없이 동일한 경험을 할 수 있도록 보장한다. 따라서 인천교육중단연구에서도 온라인 시험이 원활하게 진행될 수 있도록 아이패드, 컴퓨터, 또는 태블릿과 같은 다양한 기기를 제공하는 정책이 수립되어야 한다. 이를 통해 학생들은 기기에 구애받지 않고 시험에 참여할 수 있으며, 디지털 격차로 인한 불이익을 최소화할 수 있다. 특히, 학교에서 디지털 기기 사용에 대한 충분한 지원을 제공해야 하며, 학생들이 기기 사용에 어려움을 겪지 않도록 하는 사전 준비가 필요하다.

넷째, 국외 교육중단연구 사례 분석에서 소개된 바와 같이 컴퓨터화 능력적응검사(CAT)가 지니는 장점을 고려하여 적절히 활용하는 것이 바람직하다. 컴퓨터화 능력적응검사(CAT)는 학생의 응답에 따라 시험 난이도가 실시간으로 조정되는 평가 방식으로, STAAR 검사에 사용되고 있다. CAT는 학생의 실력에 맞춰 난이도를 조정해 학생 개인의 학업 성취를 더 정확하게 평가할 수 있다. 또한 고정형 시험보다 더 적은 문항으로도 학생의 성취 수준을 정확히 측정할 수 있어, 학생의 피로도를 줄이고 평가 시간을 단축할 수 있다. 마지막으로 모든 학생이 동일한 문항을 보지 않으므로, 시험 부정행위나 문항 유출의 위험이 감소한다. 인천교육중단연구에서 장기적으로 볼 때, CBT에서 더 나아가 CAT를 도입함으로써 학업성취도를 정확히 진단하고 학생들에게 맞춤형 학습 기회를 제공할 수 있을 것으로 기대된다.



## 참 고 문 헌



- 강호수 외(2017). 경기교육중단연구 5차년도 기초분석 보고서. 경기도교육연구원.
- 강홍수 외(2014). 경기교육중단연구 3차년도 문항개발 보고서. 경기도교육연구원.
- 교육부. (2020). 국가수준 학업성취도 평가 개선 방안. 교육부 교육기획보장과.
- 교육부. (2022). 2021년 국가수준 학업성취도 평가 결과 및 학습 지원 강화를 위한 대응 전략 발표. 교육부 보도자료. 2022.6.13.
- 교육부. (2022a). 초·중등학교 교육과정 총론. 교육부 고시 제2022-33호 [별책 1].
- 교육부. (2022b). 초등학교 교육과정. 교육부 고시 제2022-33호 [별책 2].
- 교육부. (2022c). 중학교 교육과정. 교육부 고시 제2022-33호 [별책 3].
- 교육부. (2022d). 고등학교 교육과정. 교육부 고시 제2022-33호 [별책 4].
- 김경희 외(2017). 한국 초·중등 학교교육 성과 중단조사 체제 구축을 위한 기초 연구 (Ⅰ). 한국교육과정평가원.
- 김경희 외(2018). 서울학생역량 검사지 개발을 위한 기초 연구. 서울특별시교육청교육연구정보원 교육정책연구소.
- 김경희 외(2018). 한국 초·중등 학교교육 성과 중단조사 체제 구축을 위한 기초 연구 (Ⅱ). 한국교육과정평가원.
- 김경희 외(2019). 서울교육중단연구 2020 학생역량 검사지 개발 연구. 서울특별시교육청교육연구정보원 교육정책연구소.
- 김경희 외(2021). 서울교육중단연구 2020 2차년도 학생역량 검사지 개발 연구. 서울특별시교육청교육연구정보원 교육정책연구소.
- 김경희 외(2022). 서울교육중단연구 2020 3차년도 학생역량 검사지 개발 연구. 서울특별시교육청교육연구정보원 교육정책연구소.
- 김성식 외(2007). 한국교육중단연구2005(Ⅲ):조사개요보고서. 한국교육개발원.
- 김성식 외(2008). 한국교육중단연구2005(Ⅳ). 한국교육개발원.
- 김성식 외(2016). ‘서울교육중단연구 7차년도’ 기초분석 보고서. 서울특별시교육연구정보원 교육정책연구소.
- 김성식 외(2017). ‘서울교육중단연구 8차년도’ 기초분석 보고서. 서울특별시교육연구

- 정보원 교육정책연구소.
- 김성식 외(2018). ‘서울교육중단연구 9차년도’ 기초분석 보고서. 서울특별시교육연구정보원 교육정책연구소.
- 김성훈 외(2015). 제6차년도 서울교육중단연구 학업성취도검사 척도점수 개발 연구. 서울특별시교육연구정보원.
- 김성훈 외(2016). 제7차년도 서울교육중단연구 학업성취도검사 척도점수 개발 연구. 서울특별시교육연구정보원 교육정책연구소.
- 김양분 외(2006). 한국교육중단연구 2005(Ⅱ). 한국교육개발원.
- 김양분 외(2010). 한국교육중단연구2005(VI). 한국교육개발원.
- 김양분 외(2013). 한국교육중단연구(KELS)2013(Ⅰ): 조사개요보고서. 한국교육개발원.
- 김양분 외(2014). 한국교육중단연구(KELS)2013(Ⅱ): 조사개요보고서. 한국교육개발원.
- 김양분 외(2015). 한국교육중단연구(KELS)2013(Ⅲ): 조사개요보고서. 한국교육개발원.
- 김양분 외(2016). 한국교육중단연구(KELS)2013(Ⅳ): 조사개요보고서. 한국교육개발원.
- 김위정 외(2021). 경기도교육중단연구 1주기 사업 평가 및 개선방안. 경기도교육연구원.
- 김위정 외(2021). 경기도교육중단연구 9차년도 및 진로진학조사 기초분석. 경기도교육연구원.
- 김자영 외(2018). 경기도교육중단연구 6차년도 기초분석 보고서. 경기도교육연구원.
- 남궁지영 외(2017). 한국교육중단연구(KELS)2013(V): 조사개요보고서. 한국교육개발원.
- 남궁지영 외(2018). 한국교육중단연구2013(VI): 조사개요보고서. 한국교육개발원.
- 류한구 외(2005). 한국교육중단연구 2005(1): 예비조사 보고서. 한국교육개발원.
- 박경호 외(2019). 한국교육중단연구2013(VII): 조사개요보고서. 한국교육개발원.
- 박경호 외(2020). 한국교육중단연구2013(VII): 조사개요보고서. 한국교육개발원.
- 박상현 외(2018). 서울교육중단연구 2020 학생패널 기본 설계 및 실행 방안 연구. 서울특별시교육청교육연구정보원 교육정책연구소.

- 박인용, 김미림, 박민호, 곽민호, 강혜진, 박도영, 김완수, 임황규. (2023). 국가수준 학업성취도 평가의 단계적 적응형 평가 체제 구축: 검사 설계 및 척도화. 연구보고 RRE 2023-8. 한국교육과정평가원.
- 박현정 외(2010). ‘서울교육종단연구’ 조사설계 및 기초분석 연구. 서울특별시교육연구정보원.
- 박현정 외(2011). ‘서울교육종단연구 2차년도’ 기초분석 보고서. 서울특별시교육연구정보원.
- 박현정 외(2011). 서울교육종단연구 2010 사용자 매뉴얼. 서울특별시교육연구정보원.
- 박현정 외(2012). ‘서울교육종단연구 3차년도’ 기초분석 보고서. 서울특별시교육연구정보원.
- 서동기 외(2021). 전남교육종단연구 3차년도 조사개요. 전라남도교육연구정보원 교육정책연구소.
- 서민희, 김미림, 김완수, 이소라, 한정아, 김성훈, 서영숙. (2020a). 컴퓨터 기반 국가수준 학업성취도 평가 도입을 위한 검사 설계 방안. 연구보고 RRE 2020-3. 한국교육과정평가원.
- 서민희, 김완수, 김미림, 한정아, 손윤희. (2020b). 2019년 국가수준 학업성취도 평가 결과: 고등학교. 한국교육과정평가원. 연구자료 ORM 2020-24-2.
- 성기선 외(2012). 경기교육종단연구 연구설계 및 문항개발 연구. 한국교육연구네트워크.
- 성기선 외(2013). 경기교육종단연구 2차년도 문항개발 연구. 경기도교육연구원.
- 송승훈 외(2019). 전남교육종단연구 조사개요 및 1차년도 문항설계. 전라남도교육연구정보원 전남교육정책연구소.
- 송승훈 외(2020). 전남교육종단연구 2차년도 조사 개요. 전라남도교육연구정보원 교육정책연구소.
- 송승훈 외(2023). 전남교육종단연구 1~5차년도 기초분석. 전라남도교육연구정보원 교육정책연구소.
- 신병건 외(2023). 전남교육종단연구 5차년도 조사개요. 전라남도교육연구정보원 교육정책연구소.
- 안영은 외(2021). 서울학생종단연구 2020 1차년도 결과 분석 보고서. 서울특별시교육

- 청교육연구정보원.
- 안영은 외(2021). 서울학생종단연구 2020 조사설계 및 시행 기반 구축 보고서. 서울특별시교육청교육연구정보원.
- 이경상 외(2003). 한국 청소년 패널조사(KYPS) I 조사개요 보고서. 한국청소년개발원.
- 이경상 외(2004). 한국 청소년 패널조사(KYPS) II 조사개요 보고서. 한국청소년개발원.
- 이경상 외(2005). 한국 청소년 패널조사(KYPS) III 조사개요 보고서. 한국청소년개발원.
- 이경상 외(2006). 한국 청소년 패널조사(KYPS) IV 조사개요 보고서. 한국청소년개발원.
- 이경상 외(2007). 한국 청소년 패널조사(KYPS) V 조사개요 보고서. 한국청소년정책연구원.
- 이경상 외(2008). 한국 청소년 패널조사(KYPS) VI 조사개요 보고서. 한국청소년정책연구원.
- 이규민 외(2013). 한국교육종단연구(KELS)2013(I) 초등학교 기초능력검사 도구 개발 및 수직 연계 연구. 한국교육개발원.
- 이규민 외(2015). 한국교육종단연구(KELS)2013(III) 중학교 학업성취도검사 개발 및 수직연계 연구. 한국교육개발원.
- 이윤선 외(2013). 서울교육종단연구 학업성취도검사 척도점수 개발 연구. 서울특별시교육연구정보원.
- 이윤선 외(2014). 서울교육종단연구 학업성취도검사 척도점수 개발 연구. 서울특별시교육연구정보원.
- 이정연 외(2015). 경기교육종단연구 1~3차년도 기초분석 연구. 경기도교육연구원.
- 이정연 외(2016). 경기교육종단연구 4차년도 기초분석 연구. 경기도교육연구원.
- 이정연 외(2019). 경기교육종단연구 7차년도 기초 분석. 경기도교육연구원.
- 이정연 외(2020). 경기교육종단연구 8차년도 기초 분석. 경기도교육연구원.
- 임현정 외(2009). 한국교육종단연구2005(V). 한국교육개발원.
- 지은림 외(2012). 서울교육종단연구 학업성취도검사 수직척도점수 개발 연구. 서울특

별시교육연구정보원.

최용환 외(2013). ‘서울교육종단연구 4차년도’ 기초분석 보고서. 서울특별시교육연구정보원.

하추자 외(2022). 2021년 전남교육정책 연구보고서 전남교육종단연구 4차년도 조사 개요. 전라남도교육연구정보원 교육정책연구소.

하형석 외(2018). 한국아동·청소년패널조사IX: 사업보고서. 한국청소년정책연구원.

하형석 외(2019). 한국아동·청소년패널조사X: 사업보고서. 한국청소년정책연구원.

하형석 외(2020). 2020 한국아동·청소년패널조사: 사업보고서. 한국청소년정책연구원.

하형석 외(2021). 2021 한국아동·청소년패널조사: 사업보고서. 한국청소년정책연구원.

한국교육고용패널 1차(2004)~12차(2015)년도 조사 사용자지침서. 한국직업능력개발원.

한국교육고용패널 I 1차(2004)~12차(2015)년도 설문지. 한국직업능력개발원.

한국교육고용패널 II 1차~6차년도 사용자지침서. 한국직업능력개발원.

한국교육고용패널 II 1차~6차년도 설문지. 한국직업능력개발원.

홍세희 외(2014). ‘서울교육종단연구 5차년도’ 기초분석 보고서. 서울특별시교육연구정보원.

홍세희 외(2015). ‘서울교육종단연구 6차년도’ 기초분석 보고서. 서울특별시교육연구정보원.

황진구 외(2022). 2022 한국아동·청소년패널조사: 사업보고서. 한국청소년정책연구원.

황진구 외(2023). 2023 한국아동·청소년패널조사: 사업보고서. 한국청소년정책연구원.

Florida Department of Education. (2022). Geometry Sample Test Materials Answer Key. Retrieved from [https://flfast.org/content/contentresources/en/2022\\_BEST\\_GEO\\_Answer-Key\\_CBT.pdf](https://flfast.org/content/contentresources/en/2022_BEST_GEO_Answer-Key_CBT.pdf)

Florida Department of Education. (2023a). State of Florida Benchmarks for Excellent Student Thinking (B.E.S.T.) 2022–2023 Volume 1 Annual Technical

- Report. Retrieved from [https://www.fldoe.org/core/fileparse.php/5663/urlt/BEST\\_22-23\\_Technical\\_Report\\_Volume1\\_AnnualTechnicalReport.pdf](https://www.fldoe.org/core/fileparse.php/5663/urlt/BEST_22-23_Technical_Report_Volume1_AnnualTechnicalReport.pdf)
- Florida Department of Education. (2023b). State of Florida Benchmarks for Excellent Student Thinking (B.E.S.T.) 2022–2023 Volume 2 Test Development. Retrieved from [https://www.fldoe.org/core/fileparse.php/5663/urlt/BEST\\_22-23\\_Technical\\_Report\\_Volume2\\_TestDevelopment.pdf](https://www.fldoe.org/core/fileparse.php/5663/urlt/BEST_22-23_Technical_Report_Volume2_TestDevelopment.pdf)
- Florida Department of Education. (2024a). 2024–25 B.E.S.T. Algebra 1 and Geometry End-of-Course Assessments Fact Sheet. Retrieved from <https://www.fldoe.org/core/fileparse.php/20102/urlt/2425BESTEOCFactSheet.pdf>
- Florida Department of Education. (2024b). 2024–25 Florida Assessment of Student Thinking (FAST) Fact Sheet Grades 3–10 ELA Reading and Grades 3–8 Mathematics. Retrieved from <https://www.fldoe.org/core/fileparse.php/20102/urlt/2425FASTGrd310FS.pdf>
- Florida Department of Education. (2024c). Algebra 1 Sample Test Materials Answer Key. Retrieved from [https://flfast.org/content/contentresources/en/2024\\_BEST\\_ALG1\\_Answer-Key\\_CBT.pdf](https://flfast.org/content/contentresources/en/2024_BEST_ALG1_Answer-Key_CBT.pdf)
- Florida Department of Education. (2024d). Florida Statewide Assessments Guide 2024–2025. Retrieved from <https://www.fldoe.org/core/fileparse.php/20102/urlt/K12SAG.pdf>
- Florida Department of Education. (2024e). Florida’s Statewide Grades 3–10 and End-of-Course (EOC) Computer-Adaptive Tests (CAT)
- Florida Department of Education. (2024f). Grade 3 FAST ELA Reading Sample Test Materials Answer Key. Retrieved from [https://flfast.org/content/contentresources/en/2024\\_FAST\\_3R\\_Answer-Key\\_CBT.pdf](https://flfast.org/content/contentresources/en/2024_FAST_3R_Answer-Key_CBT.pdf)
- Florida Department of Education. (2024g). Grade 4 FAST ELA Reading Sample Test Materials Answer Key. Retrieved from [https://flfast.org/content/contentresources/en/2024\\_FAST\\_4R\\_Answer-Key\\_CBT.pdf](https://flfast.org/content/contentresources/en/2024_FAST_4R_Answer-Key_CBT.pdf)
- Florida Department of Education. (2024h). Grade 5 FAST ELA Reading Sample Test Materials Answer Key. Retrieved from [https://flfast.org/content/contentresources/en/2024\\_FAST\\_5R\\_Answer-Key\\_CBT.pdf](https://flfast.org/content/contentresources/en/2024_FAST_5R_Answer-Key_CBT.pdf)

entresources/en/2024\_FAST\_5R\_Answer-Key\_CBT.pdf

Florida Department of Education. (2024i). Grade 6 FAST ELA Reading Sample Test Materials Answer Key. Retrieved from [https://flfast.org/content/entresources/en/2024\\_FAST\\_6R\\_Answer-Key\\_CBT.pdf](https://flfast.org/content/entresources/en/2024_FAST_6R_Answer-Key_CBT.pdf)

Florida Department of Education. (2024j). Grade 7 FAST ELA Reading Sample Test Materials Answer Key. Retrieved from [https://flfast.org/content/entresources/en/2024\\_FAST\\_7R\\_Answer-Key\\_CBT.pdf](https://flfast.org/content/entresources/en/2024_FAST_7R_Answer-Key_CBT.pdf)

Florida Department of Education. (2024k). Grade 8 FAST ELA Reading Sample Test Materials Answer Key. Retrieved from [https://flfast.org/content/entresources/en/2024\\_FAST\\_8R\\_Answer-Key\\_CBT.pdf](https://flfast.org/content/entresources/en/2024_FAST_8R_Answer-Key_CBT.pdf)

Florida Department of Education. (2024l). Grade 9 FAST ELA Reading Sample Test Materials Answer Key. Retrieved from [https://flfast.org/content/entresources/en/2024\\_FAST\\_9R\\_Answer-Key\\_CBT.pdf](https://flfast.org/content/entresources/en/2024_FAST_9R_Answer-Key_CBT.pdf)

Florida Department of Education. (2024m). Grade 10 FAST ELA Reading Sample Test Materials Answer Key. Retrieved from [https://flfast.org/content/entresources/en/2024\\_FAST\\_10R\\_Answer-Key\\_CBT.pdf](https://flfast.org/content/entresources/en/2024_FAST_10R_Answer-Key_CBT.pdf)

Florida Department of Education. (2024n). Grade 3 FAST Mathematics Sample Test Materials Answer Key. Retrieved from [https://flfast.org/content/entresources/en/2024\\_FAST\\_3M\\_Answer-Key\\_CBT.pdf](https://flfast.org/content/entresources/en/2024_FAST_3M_Answer-Key_CBT.pdf)

Florida Department of Education. (2024o). Grade 4 FAST Mathematics Sample Test Materials Answer Key. Retrieved from [https://flfast.org/content/entresources/en/2024\\_FAST\\_4M\\_Answer-Key\\_CBT.pdf](https://flfast.org/content/entresources/en/2024_FAST_4M_Answer-Key_CBT.pdf)

Florida Department of Education. (2024p). Grade 5 FAST Mathematics Sample Test Materials Answer Key. Retrieved from [https://flfast.org/content/entresources/en/2024\\_FAST\\_5M\\_Answer-Key\\_CBT.pdf](https://flfast.org/content/entresources/en/2024_FAST_5M_Answer-Key_CBT.pdf)

Florida Department of Education. (2024q). Grade 6 FAST Mathematics Sample Test Materials Answer Key. Retrieved from [https://flfast.org/content/entresources/en/2024\\_FAST\\_6M\\_Answer-Key\\_CBT.pdf](https://flfast.org/content/entresources/en/2024_FAST_6M_Answer-Key_CBT.pdf)

Florida Department of Education. (2024r). Grade 7 FAST Mathematics Sample

- Test Materials Answer Key. Retrieved from [https://flfast.org/content/contentresources/en/2024\\_FAST\\_7M\\_Answer-Key\\_CBT.pdf](https://flfast.org/content/contentresources/en/2024_FAST_7M_Answer-Key_CBT.pdf)
- Florida Department of Education. (2024s). Grade 8 FAST Mathematics Sample Test Materials Answer Key. Retrieved from [https://flfast.org/content/contentresources/en/2024\\_FAST\\_8M\\_Answer-Key\\_CBT.pdf](https://flfast.org/content/contentresources/en/2024_FAST_8M_Answer-Key_CBT.pdf)
- Florida Department of Education. (2024t). Graduation Requirements for Florida's Statewide Assessments. Retrieved from <https://www.fldoe.org/core/fileparse.php/7764/urlt/GradRequireFSA.pdf>
- Florida Department of Education. (2024u). Test Design Summary and Blueprint: FAST ELA Reading and B.E.S.T. Writing. Retrieved from <https://www.fldoe.org/core/fileparse.php/20102/urlt/TDS-FAST-ELA.pdf>
- Florida Department of Education. (2024v). Test Design Summary and Blueprint: FAST Mathematics and B.E.S.T. EOCs. Retrieved from <https://www.fldoe.org/core/fileparse.php/20102/urlt/TDS-FAST-Math.pdf>
- Frequently Asked Questions (FAQ). Retrieved from <https://www.fldoe.org/core/fileparse.php/20102/urlt/CompAdaptTestFAQ.pdf>
- Texas Education Agency. (2022a). STAAR Algebra I Blueprint Effective as of Academic Year 2022–23. Texas Education Agency. Retrieved September 12, 2024, from <https://tea.texas.gov/student-assessment/staar/staar-algebra-i-blueprint.pdf>
- Texas Education Agency. (2022b). STAAR Biology Blueprint Effective as of Academic Year 2022–23. Retrieved October 17, 2024, from <https://tea.texas.gov/student-assessment/staar/staar-biology-blueprint.pdf>
- Texas Education Agency. (2022c). STAAR English I Blueprint Effective as of Spring 2023. Texas Education Agency. Retrieved September 12, 2024, from <https://tea.texas.gov/student-assessment/staar/staar-english-i-blueprint.pdf>
- Texas Education Agency. (2022d). STAAR English II Blueprint Effective as of Spring 2023. Texas Education Agency. Retrieved September 12, 2024,



from <https://tea.texas.gov/student-assessment/staar/staar-english-ii-blueprint.pdf>

Texas Education Agency. (2022e). STAAR Grade 3 Math Blueprint Effective as of School Year 2022-23. Texas Education Agency. Retrieved September 12, 2024, from <https://tea.texas.gov/student-assessment/staar/staar-3-math-blueprint.pdf>

Texas Education Agency. (2022f). STAAR Grade 4 Math Blueprint Effective as of School Year 2022-23. Texas Education Agency. Retrieved September 12, 2024, from <https://tea.texas.gov/student-assessment/staar/staar-4-math-blueprint.pdf>

Texas Education Agency. (2022g). STAAR Grade 5 Math Blueprint Effective as of School Year 2022-23. Texas Education Agency. Retrieved September 12, 2024, from <https://tea.texas.gov/student-assessment/staar/staar-5-math-blueprint.pdf>

Texas Education Agency. (2022h). STAAR Grade 6 Math Blueprint Effective as of School Year 2022-23. Texas Education Agency. Retrieved September 12, 2024, from <https://tea.texas.gov/student-assessment/staar/staar-6-math-blueprint.pdf>

Texas Education Agency. (2022i). STAAR Grade 7 Math Blueprint Effective as of School Year 2022-23. Texas Education Agency. Retrieved September 12, 2024, from <https://tea.texas.gov/student-assessment/staar/staar-7-math-blueprint.pdf>

Texas Education Agency. (2022j). STAAR Grade 8 Math Blueprint Effective as of School Year 2022-23. Texas Education Agency. Retrieved September 12, 2024, from <https://tea.texas.gov/student-assessment/staar-grade-8-math-blueprint.pdf>

Texas Education Agency. (2022k). STAAR Grade 3 Reading Language Arts Blueprint Effective as of School Year 2022-23. Texas Education Agency. Retrieved September 12, 2024, from <https://tea.texas.gov/student-assessment/>

staar/staar-3-rla-blueprint.pdf

Texas Education Agency. (2022l). STAAR Grade 4 Reading Language Arts Blueprint Effective as of School Year 2022–23. Texas Education Agency. Retrieved September 12, 2024, from <https://tea.texas.gov/student-assessment/staar/staar-4-rla-blueprint.pdf>

Texas Education Agency. (2022m). STAAR Grade 5 Reading Language Arts Blueprint Effective as of School Year 2022–23. Texas Education Agency. Retrieved September 12, 2024, from <https://tea.texas.gov/student-assessment/staar/staar-5-rla-blueprint.pdf>

Texas Education Agency. (2022n). STAAR Grade 6 Reading Language Arts Blueprint Effective as of School Year 2022–23. Texas Education Agency. Retrieved September 12, 2024, from <https://tea.texas.gov/student-assessment/staar/staar-6-rla-blueprint.pdf>

Texas Education Agency. (2022o). STAAR Grade 7 Reading Language Arts Blueprint Effective as of School Year 2022–23. Texas Education Agency. Retrieved September 12, 2024, from <https://tea.texas.gov/student-assessment/staar/staar-7-rla-blueprint.pdf>

Texas Education Agency. (2022p). STAAR Grade 8 Reading Language Arts Blueprint Effective as of School Year 2022–23. Texas Education Agency. Retrieved September 12, 2024, from <https://tea.texas.gov/student-assessment/staar/staar-8-rla-blueprint.pdf>

Texas Education Agency. (2022q). STAAR Grade 8 Social Studies Blueprint Effective as of Academic Year 2022–23. Retrieved October 17, 2024, from <https://tea.texas.gov/student-assessment/staar/staar-8-social-studies-blueprint.pdf>

Texas Education Agency. (2022r). STAAR U. S. History Blueprint Effective as of Academic Year 2022–23. Retrieved October 17, 2024, from <https://tea.texas.gov/student-assessment/staar/staar-us-history-blueprint.pdf>

Texas Education Agency. (2023, November). Texas Assessment Program Frequ

ently Asked Questions. Student Assessment Division. Retrieved from <https://tea.texas.gov/student-assessment/student-assessment-overview/texas-assessment-program-faqs.pdf>

Texas Education Agency. (2024a). Spring May 2024 Grade 8 Science Answer Key. <https://tea.texas.gov/student-assessment/staar/released-test-questions/2024-staar-8-science-key.pdf>

Texas Education Agency. (2024b). STAAR Grade 5 Science Blueprint for the 2024-2025 School Year Assessing the Overlap Curriculum. Retrieved October 17, 2024, from <https://tea.texas.gov/student-assessment/staar/2024-2025-staar-5-science-blueprint.pdf>

Texas Education Agency. (2024c). STAAR Grade 8 Science Blueprint for the 2024-2025 School Year Assessing the Overlap Curriculum . Retrieved October 17, 2024, from <https://tea.texas.gov/student-assessment/staar/2024-2025-staar-8-science-blueprint.pdf>

Texas Education Agency. (2024d). STAAR Spring 2024 Algebra I Answer Key. <https://tea.texas.gov/student-assessment/staar/released-test-questions/2024-staar-algebra-i-key.pdf>

Texas Education Agency. (2024e). STAAR Spring 2024 Biology Answer Key. <https://tea.texas.gov/student-assessment/staar/released-test-questions/2024-staar-biology-key.pdf>

Texas Education Agency. (2024f). STAAR Spring 2024 English I RLA Answer Key. <https://tea.texas.gov/student-assessment/staar/released-test-questions/2024-staar-english-i-key.pdf>

Texas Education Agency. (2024g). STAAR Spring 2024 Grade 3 Math Answer Key. <https://tea.texas.gov/student-assessment/staar/released-test-questions/2024-staar-3-math-key.pdf>

Texas Education Agency. (2024h). STAAR Spring 2024 Grade 3 RLA Answer Key. <https://tea.texas.gov/student-assessment/staar/released-test-questions/2024-staar-3-rla-key.pdf>

- Texas Education Agency. (2024i). STAAR Spring 2024 Grade 4 Math Answer Key. <https://tea.texas.gov/student-assessment/staar/released-test-questions/2024-staar-4-math-key.pdf>
- Texas Education Agency. (2024j). STAAR Spring 2024 Grade 4 RLA Answer Key. <https://tea.texas.gov/student-assessment/staar/released-test-questions/2024-staar-4-rla-key.pdf>
- Texas Education Agency. (2024k). STAAR Spring 2024 Grade 5 Math Answer Key. <https://tea.texas.gov/student-assessment/staar/released-test-questions/2024-staar-5-math-key.pdf>
- Texas Education Agency. (2024l). STAAR Spring 2024 Grade 5 RLA Answer Key. <https://tea.texas.gov/student-assessment/staar/released-test-questions/2024-staar-5-rla-key.pdf>
- Texas Education Agency. (2024m). STAAR Spring 2024 Grade 5 Science Answer Key. <https://tea.texas.gov/student-assessment/staar/released-test-questions/2024-staar-5-science-key.pdf>
- Texas Education Agency. (2024n). STAAR Spring 2024 Grade 6 Math Answer Key. <https://tea.texas.gov/student-assessment/staar/released-test-questions/2024-staar-6-math-key.pdf>
- Texas Education Agency. (2024o). STAAR Spring 2024 Grade 6 RLA Answer Key. <https://tea.texas.gov/student-assessment/staar/released-test-questions/2024-staar-6-rla-key.pdf>
- Texas Education Agency. (2024p). STAAR Spring 2024 Grade 7 Math Answer Key. <https://tea.texas.gov/student-assessment/staar/released-test-questions/2024-staar-7-math-key.pdf>
- Texas Education Agency. (2024q). STAAR Spring 2024 Grade 7 RLA Answer Key. <https://tea.texas.gov/student-assessment/staar/released-test-questions/2024-staar-7-rla-key.pdf>
- Texas Education Agency. (2024r). STAAR Spring 2024 Grade 8 Math Answer Key. <https://tea.texas.gov/student-assessment/staar/released-test-questions/2024-staar-8-math-key.pdf>

tions/2024-staar-8-math-key.pdf

Texas Education Agency. (2024s). STAAR Spring 2024 Grade 8 RLA Answer Key. <https://tea.texas.gov/student-assessment/staar/released-test-questions/2024-staar-8-rla-key.pdf>

Texas Education Agency. (2024t). STAAR Spring 2024 Grade 8 Social Studies Answer Key. <https://tea.texas.gov/student-assessment/staar/released-test-questions/2024-staar-8-social-studies-key.pdf>

Texas Education Agency. (2024u). STAAR Spring 2024 RLA English II Answer Key. <https://tea.texas.gov/student-assessment/staar/released-test-questions/2024-staar-english-ii-key.pdf>

Texas Education Agency. (2024v). STAAR Spring 2024 U.S. History Answer Key. <https://tea.texas.gov/student-assessment/staar/released-test-questions/2024-staar-us-history-key.pdf>

Texas Education Agency. (2024w). Technical Digest 2022-2023. Texas Education Agency. Retrieved from <https://tea.texas.gov/student-assessment/reports-and-studies/2023-technical-digest.pdf>

Texas Education Agency. (n.d.). Sampling and the Texas Assessment Program. Retrieved from <https://tea.texas.gov/texas-schools/accountability/academic-accountability/performance-reporting/4-sampling-and-staar-program.pdf>

Texas Education Code §28.0211 (2023).

Texas Education Code §28.0211 (2023).

Texas Education Code §28.0217 (2023).

Texas Education Code §28.0258 (2023).

Texas Education Code §39.023 (2023).

19 Texas Administrative Code §§101.2001, 2003, 2006 (2023).

19 Texas Administrative Code §101.3022 (2023).

## 홈 페이지



교육부 홈페이지.

<https://www.moe.go.kr/boardCnts/viewRenew.do?boardID=340&lev=0&statusYN=W&s=moe&m=020201&opType=N&boardSeq=95642>에서  
2024년 10월 17일 검색.

교육부 공식 블로그. <https://if-blog.tistory.com/13725>에서 2024년 11월 7일 검색.  
인천 교육청 기초학력 진단 보정 시스템.

<https://basic.ice.go.kr/pt/intro/intro.do>에서 2024년 11월 7일 검색.

학업성취도 평가 지원 포털-국가수준 학업성취도 평가.

<https://inaea.kice.re.kr/siteMain/nationStandExam.do>에서 2024년 10월  
17일 검색.

학업성취도 평가 지원 포털-맞춤형 학업성취도 자율평가.

<https://inaea.kice.re.kr/siteMain/atnmsExam.do>에서 2024년 10월 17일  
검색.

학업성취도 평가 지원 포털-학업성취도 평가 개관.

<https://inaea.kice.re.kr/siteMain/overview.do>에서 2024년 10월 17일 검색.

한국교육개발원 홈페이지-제16회 한국교육중단연구 학술대회 데이터 자료집.

[https://www.kedi.re.kr/khome/main/announce/selectNoticeAnnounceForm.do?board\\_sq\\_no=1&selectTp=0&article\\_sq\\_no=34617](https://www.kedi.re.kr/khome/main/announce/selectNoticeAnnounceForm.do?board_sq_no=1&selectTp=0&article_sq_no=34617)에서 2024년 10월  
17일 검색.

한국복지패널조사 홈페이지.

<https://www.koweps.re.kr:442/data/survey/list.do>에서 2024년 10월 17일  
검색.

한국직업능력연구원 홈페이지. <https://www.krivet.re.kr>.

행정안전부 국가기록원.

<https://www.archives.go.kr/next/newsearch/listSubjectDescription.do?id=009347&pageFlag=&sitePage=>에서 2024년 10월 17일 검색.

EBSi 홈페이지.

<https://www.ebsi.co.kr/ebs/ent/enta/retrieveExmSchedRng.ebs>에서  
2024년 10월 17일 검색.

## 인천교육종단연구 기반연구

－ 교과역량 평가 방안 제안

---

발 행 일	2024년 12월
발 행 인	인천광역시교육감 도 성 훈
발 행 처	인천광역시교육청
주 소	인천광역시 남동구 정각로 9(구월동) (우)21554 <a href="http://www.ice.go.kr">http://www.ice.go.kr</a> 032-423-8232
I S B N	978-89-6960-111-7(93370) (비매품)

---

※ 이 책의 저작권은 인천광역시교육청에 있으며 일부 또는 전부의 무단 복제를 금합니다.