

『8차 교바사 함공모 월례토론회』 기초 참고 자료

-순서-

I. 교육정책포럼(한국교육개발원, 2012.02.10)

- 중등학교 성적 평가제도 개선의 이유와 의미(성태제, 2012)
- 중등학교 성적 평가제도 개선의 내용과 정책방향(박희동, 2012)
- 중등학교 성적 평가제도 개선에 대한 이해와 추진과제(김성숙, 2012)
- 미국의 학생평가: 성취기준 중심의 절대평가(송미영, 2012)
- 영국의 학생평가: 국가교육과정의 성취목표를 중심으로 한 준거지향평가(김주아, 2012)

II. 세계 주요국의 평가를 위한 노력

- 2010 교육과정·교육평가 국제동향 연구 사업 -미국- (한국교육과정평가원, 2010)
- 2010 교육과정·교육평가 국제동향 연구 사업 -핀란드- (한국교육과정평가원, 2010)

III. 교육과정과 교육평가의 연계방안(곽병선, 1998)

IV. 미국의 고부담 시험의 가능성과 한계(박세훈, 2010)

V. 고부담 시험 시대의 사회과교육: 학교 책무성의 신화, 학력, 그리고 교사 전문성의 관계(조대훈, 2011)

VI. 서울지역 학교교육 실태조사 분석 결과(전교조 서울지부, 2010.06)

VII. 일제고사 파행 전국 사례 요약(전교조, 2011.07)

VIII. 초등학교 중간·기말고사 폐지 찬반논란 및 그 이후(관련자료, 2010~2011)

□ 참고자료1. Wales의 사례

- What happens when high stakes testing stops?(Michael Reiss외 2명, 2010)
- Wales Eliminates National Exams for Many Students(Jeff Archer, 2006)

□ 참고자료2. 「OECD Education Working Papers No.58」(2011.04)

- Integrating Formative and Summative Assessment

□ 참고자료3. 「PARCC / SBAC」에 대한 브리프 (www.12center.org, 2011.07 update)

2012년 2월



I. 교육정책포럼(한국교육개발원, 2012.02.10.)

□ 중등학교 성적 평가제도 개선의 이유와 의미(성태제, 2012)

-중등학교의 미래를 밝힐 성취평가제 도입-

※ 교육에 대한 정의는 다양하지만 인간의 행동을 변화시키는 목적적 행위라고 하는 데는 이견이 없다. 그러기에 교육에는 목표가 있어야 하고 교육목표의 달성 여부를 확인하는 작업이 필요하다. 이를 고전적 의미의 교육평가라 한다.

● 평가의 구분 및 비교

- ✓ 평가의 관점에 따라 상대비교평가와 절대평가로 구분하며, 선진화된 평가방법으로는 능력참조평가와 성장참조평가가 있다.
- ✓ 우리나라 교육현장에서는 주로 상대비교평가와 절대평가를 적용하여 왔으며, 상대비교평가를 전반적으로 적용하고 있는 것이 현실이다. 2011년 12월13일 교육과학기술부가 발표한 '중등학교 학사관리 선진화 방안'에서 도입하는 성취평가제는 순위 경쟁에 찌든 중등교육의 문제점을 획기적으로 개선할 수 있다는 점에서 기대가 크다.

상대비교평가	절대평가
<ul style="list-style-type: none"> • <i>현행 중등학교 학사관리 방식</i>으로 교육계에서 오래 전부터 비판받아왔으며, 교육선진국에서는 일찍이 폐기된 방식임 • 핀란드에서는 상대평가를 지양하며, 교육과정에서 설정한 성취목표에 학생들이 도달했는지 여부만 평가하고 있음. • 미국은 1960년대 말부터 상대평가의 폐해를 지적하고 절대평가로 전환함 • 올바른 인성을 길러주기보다 오히려 학생 간의 배타적 경쟁심을 강화시켜 옴- 흡스가 말한 '만인에 의한 만인의 투쟁'을 교육의 당연한 원리로 여겨지게 함 • 상대비교평가의 학문적 용어는 규준참조평가로서 규준에 비추어 서열 정보를 제공하는 평가이다. 이는 점수 비교의 대상이 다른 사람이 되는 것으로, 학생이 받은 원점수의 위치가 비교대상 집단에서 어느 정도인지를 알려주는 것이다. 따라서 이 방법은 학생이 받은 점수의 상대적 위치는 쉽게 알 수 있지만 그 학생이 무엇을 얼마나 알고 있는지에 대해서는 알 수가 없으며, 학력 향상을 위해 어떤 내용을 얼마나 가르치고 배워야 하는지 파악할 수도 없다. • '높은 순위가 최고'라는 잘못된 인식을 확산시키면서, 성공적인 사회생활이나 학문 활동에 필요한 문제해결능력이나 탐구정신, 창의력 함양보다는 '순위 올리기'에만 집중하게 만든다. 즉 경쟁으로 인한 심리적 불안, 이기심 등을 조장하면서 올바른 인격체로서의 성장을 저해한다. 	<ul style="list-style-type: none"> • 인간은 각기 저마다 타고난 고유한 특성이 있기에 성취해야 할 목표 수준을 정해 놓고, 그것과 비교하여 학생의 능력을 향상시키도록 유도하는 것이 바람직하다. • 절대평가는 학생 개인들을 상호 비교하는 것이 아니라 학생의 성취수준을 어떤 준거에 비추어 판단하는 평가로 준거참조평가라 한다. 이는 학생 간 경쟁을 조장하지 않기에 탐구정신을 고양하고 상호협동을 유도하여 학생의 인성교육이나 정신 건강에도 바람직하다. 또한 준거에 비추어 제공되는 평가 결과는 학생들이 무엇을 얼마나 알고 있는지 알려주기에 교수-학습이론에 적합하며, 이에 따라 교육과정 개선에 도움이 되고 학습효과를 증진시키는 데 도움이 된다. • 1990년대 중반 이후에 고등학교 학생들의 성적을 절대평가로 전환하였지만, 대학진학은 상대비교평가에 의해 결정되어 학생, 학부모는 보다 후한 고교 내신성적을 요구하게 되어 '점수부풀리기', '고교등급제'의 논란을 야기하였다. 절대평가를 위한 여건 조성이 되지 못한 상태에서 등장한 이러한 문제점들은 고교내신을 다시금 상대비교평가로 바꾸게 하였으며, 이는 교육의 후진화라 하겠다. • 절대평가, 즉 준거참조평가에서 논란이 되는 것은 준거(cut score, criterion)설정이다. 우리나라 중·고등학교에서는 절대평가를 위한 준거설정의 경험이 거의 없다고 봐도 과언이 아니므로 준거 설정 방법에 대한 실질적이고 전문적인 연수가 절대적으로 필요하다. • 우리나라에서도 선진국처럼 중·고등학교에서 절대평가를 체계적으로 실행하게 된다. '12년 3월부터 전문교과를 중심으로 특성학교·마이스터고에서 적용하며, 보통교과에 대해서는 시범 적용 기간을 거쳐 2014년부터 본격 적용할 계획이다.

	<ul style="list-style-type: none"> • 그 간의 시행착오를 되풀이하지 않기 위해 학생평가의 절차와 방법, 그리고 결과보고 단계까지 절대평가 이론에 맞게 이루어져야 한다. 최대한의 준비와 새로운 접근 방식으로 학교현장에 성공적으로 적용시켜 보다 나은 교육환경을 만들어야 한다.
--	---

□ **중등학교 성적 평가제도 개선의 내용과 정책방향(박희동, 2012)**

● **평가제도 개선 추진배경**

글로벌 지식기반 사회가 요구하는 인재 양성을 목적으로 개정된 교육과정도 평가제도가 뒷받침되지 않으면 온전하게 성공하기 어렵다. 현행 석차 9등급제는 교육과정에서 제시한 '학습목표를 충실하게 달성했는지'를 평가하기 보다는 '누가 누구보다 잘 했는지'를 평가하기 때문에 교육의 본질을 살리는 데는 근본적으로 한계가 있다고 꾸준히 지적이 제기되었다. 학생측면에서는 다른 학생보다 잘 해야 한다는 부담 때문에 학습스트레스가 높고 급우 간 지나친 경쟁으로 학습 공동체 의식이 사라지는 문제, 교사측면에서는 변별력을 높이기 위해 함정문제를 출제하려는 경향, 학교에서는 학생의 적성과 소질에 따른 과목 개설이 소인수라는 이유로 개설되기 어려운 문제 등을 극복하고자 교과부는 지난해 말 성취평가제 도입을 골자로 한 '중등학교 선진화 개선방안'을 발표하였다.

● **평가방법의 질적 개선**

서술형 평가의 질 제고와 확대를 위해 교육과정의 성취기준을 토대로 창의성·인성 함양이 가능하도록 완전한 문장 형식으로 진술하는 예시문항이 다수 포함된 서술형 평가 매뉴얼을 개발하고, 시·도별 핵심교원 연수와 핵심교원의 모든 학교 전 교원 대상 전달연수를 실시하여, 2013학년도까지 지역 및 교과별 특성 등을 고려하여 서술형 평가 비율이 일정수준(20%~40%) 이상까지 확대되도록 유도한다.

● **성취평가제 도입**

아래의 성취수준 F의 도입 여부는 시범운영을 거쳐 2013학년도 하반기에 결정할 예정임

단계	현행	변경	적용대상	
			교과영역	과목
2단계	이수	P, (F)	교양교과(논리학 등)	일반과목
			기초교과(영·수)	기본과목
4단계	우수, 보통, 미흡	A, B, C, (F)	체육·예술교과	일반과목
			기초교과(국·영·수)	일반과목, 심화과목
			탐구교과(사·과)	일반과목, 심화과목
6단계	-	A, B, C, D, E, (F)	생활교과(기술·가정/제2 외국어/한문)	일반과목, 심화과목
			체육·예술교과	심화과목

- ✓ 성취평가제 도입에 따라 학교생활기록부 성적표기 방식은 중학교의 경우, '수-우-미-양-가'에서 'A-B-C-D-E-(F)'로 변경되고, 과목별 석차는 삭제되며, 성적 부풀리기 방지, 평가의 난이도, 점수 분포 등에 관한 정보 제공을 위해 원점수/과목평균(표준편차)는 병기된다. 고등학교의 경우 학교생활기록부의 석차등급은 삭제되고 6단계 성취도가 기재되며, 현행 원점수/과목평균(표준편차)는 병기된다.

- ✓ 교과부는 금년도(2012)부터 2년간 고교 100개교에 대한 시범운명을 실시하여 제기될 수 있는 모든 문제점을 점검하고, 가장 현장 적합성이 높은 방안으로 2014학년도 1학년에 입학하는 학생부터 성취평가제를 적용한다.

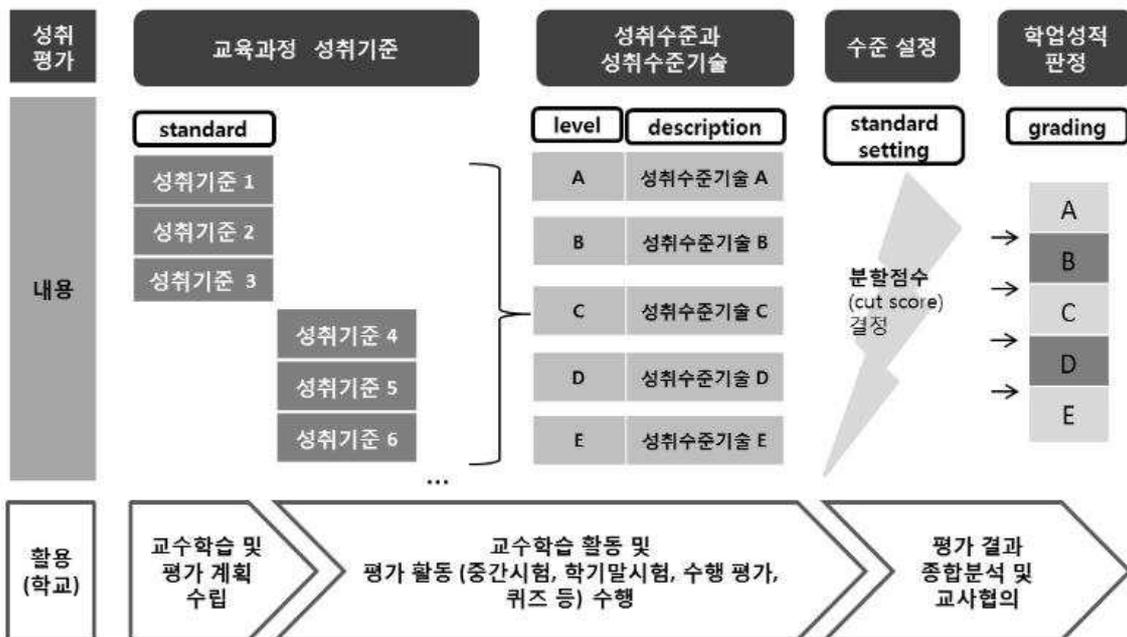
● **학사관리 선진화 요건 조성**

- ✓ 교과부는 한국교육과정평가원(이하 평가원)을 학사관리 선진화 추진 전담기관으로 지정하였으며, 평가원은 금년도 6월까지 2009개정 교육과정에 따른 교과별 성취기준·성취수준을 개발하고 교원 연수를 통해 학교현장을 적극 지원한다.
- ✓ 교과부는 교원의 평가 관련 연수를 확대하고, 성적 부풀리기 등 학사관리가 부실한 교원 및 학교에 대해서는 엄중 조치하여 공정한 평가를 위한 책무성을 강화한다. 교과목별 성취율을 학업성적관리규정에 반영하도록 하여 학업성적관리위원회 역할도 강화한다.
- ✓ 중등학교 평가제도는 입시제도와 밀접하게 관련되어 있다. 지난해부터 입학사정관제 지원사업 평가 시 '신입생 구성의 다양성' 지표를 반영하여 지역균형선발, 사회적 배려대상자 전형, 입학사정관제 등으로 농어촌·중소도시 일반고 학생들의 진학기회를 확대하고 있다. 교과부는 대교협과의 지속적인 협의를 통해 대학의 입학전형이 중등학교의 학사운영을 왜곡시키지 않도록 할 계획이다.

□ **중등학교 성적 평가제도 개선에 대한 이해와 추진과제(김성숙, 2012)**

● **성취평가제의 올바른 이해**

성취평가제는 준거참조평가(절대평가)를 포함하고 있다. 실제로 학교에서 학생의 성적산출을 위해 성취평가제를 적용하려면 아래의 그림처럼 개념화할 수 있다.



- ✓ 각 교과에서 일정기간 동안에 습득하여야 할 지식, 기능, 태도 등 교육과정에 근거한 성취기준은 다양한 교수학습 활동을 통해 가르치고 배우게 되며, 교사는 가르친 내용에 부합하는 평가를 실시하는 것이다.

✓ 성취평가제 운영절차

- i. 한 학기 평가계획에 따라 교수학습활동에 부합되는 퀴즈, 중간시험, 수행평가, 학기말 시험이 진행된다.
- ii. 학생에 대한 학기말 성적을 부여할 때 평가결과를 종합하여 각 성취수준을 결정하는 점수를 설정하는 과정 즉, 수준설정(standard setting)과정을 수행한다.
- iii. 수준설정에 따른 분할점수가 결정됨에 따라 개별 학생에 대한 최종성적이 부여되는 것이다.

● 성취평가제의 성공적 정착을 위한 과제

- ① 2009 개정 교육과정에 따라 개발하고 있는 성취기준과 성취수준이 성취평가제와 연계되도록 지원한다.
- ② 성취평가제의 올바른 이해를 위해 현장 교사들이 숙지하고 적용할 수 있는 평가전문성 함양 연수프로그램을 개발, 시행한다.
- ③ 금년부터 운영되는 성취평가제 시범학교를 통하여 철저하게 모니터링한 결과를 피드백할 수 있도록 준비한다.
- ④ 성취평가제 결과에 대한 신뢰성 확보의 문제와 대입전형자료로서의 내신 활용 문제를 점검하여야 한다.

● 교실 수업이 교사책임 지도아래 활성화되려면 학생을 가르치는 교사의 평가권을 인정해야 한다. 교사는 학생이 무엇을 알고 무엇을 할 수 있는가를 판정할 수 있어야 하며, 이는 학생평가가 교과 성취기준에 근거한 평가이어야 함을 의미한다. 그동안 상대평가를 적용해 오면서 교사는 학생평가를 위한 전문적 판단을 소홀히 해왔다. 교사가 가르친 지식과 내용을 얼마나 습득하였는가에 대한 평가라기보다 시험점수에 따라 학생들의 등급이 자동적으로 판정되어 왔기 때문이다. 따라서 중등학교의 평가제도 변화를 위해 무엇보다도 우선되어야 하는 것은 교사가 학생의 교과별 성취정도에 따라 절대적 판정을 부여할 수 있는 평가전문성이 제고되어야 함과 동시에 교사의 고유한 평가권을 신뢰할 수 있는 풍토가 보장되어야 할 것이다.

□ 미국의 학생평가: 성취기준 중심의 절대평가(송미영, 2012)

● 미국의 학생평가

- ✓ 미국 중등학교의 학생평가는 담당 교사가 제시한 학습 목표에 도달했는가를 평가하는 절대평가 방식으로 이루어진다. 학년 전체 학생들을 상호 비교하는 것이 아니라 학습내용을 충실히 이수했는지에 초점을 둔다는 것이 특징이다.
- ✓ 물론, 개별교사가 학생들을 평가한 정보가 상급학교 진학 자료로 활용되지만, 이는 이차적인 목적이다. 각 학교는 주(州)의 성취기준과 교육청의 평가 지침에 기초하여 평가계획을 수립, 실행한다.
- ✓ 일반적으로 평가는 과제, 수업시간 중 활동내용, 지필평가 등으로 이루어지며, 평가 결과는 지역과 학교 급에 따라 차이가 있지만 과목별로 학생의 성취 정도에 따라 A, B, C, D, Fail 등으로 수준을 구분하여 표기하고, 학생들의 상대적인 석차는 기록하지 않는다.
- ✓ 개별학생의 성취수준을 확인하여 한명 한명에게 취약한 부분을 보완해주는 것을 중요한 교육활동으로 삼고, 교사들은 학생들이 주에서 설정한 성취기준을 달성할 수 있도록 교

육하고, 성취기준에 근거하여 평가하는데 중점을 두고 있다.

● 학생에게 주에서 설정한 성취기준을 학습하도록 지도하고 학습한 성취기준의 깊이와 범위를 평가하게 된 배경

- ✓ 연방정부가 2002년에 학생들의 학업성취 향상을 목적으로 입법한 NCLB법안에 따라 각 주는 학년별로 학생들이 성취해야 할 목표, 성취기준을 정하고, 학생들이 성취기준을 달성했는지 여부를 판단하기 위해 주 단위의 평가를 주기적으로 실시하고, 그 결과를 보고하여야 한다. 그 영향으로 각 주는 성취기준을 설정하였고, 학생들은 해당 주에서 제시한 성취기준을 학습하여야 한다.
- ✓ 따라서 주 단위에서 시행되는 대규모 평가도 성취기준에의 도달 정도를 평가할 수 있도록 평가도구를 개발하여 시행할 뿐만 아니라, 학교 내 학생평가도 학생들이 성취기준을 성취하였는가를 평가하는 것이다.

● 미국은 각 주마다 설정한 성취 목표인 성취기준을 중심으로 교육활동이 구성됨

- ✓ 주에서 제시하는 성취기준을 달성하기 위해 학교 교육과정을 교사가 자율적으로 구성하여 실현하고, 학생평가는 성취기준에 근거해 절대평가로 수행된다.
- ✓ 절대평가라고해서 성적 분포의 형태를 가정하지 않고 성적별 학생비율이 설정되는 것은 아니다. 성취기준에 도달한 다양한 수준이 평가 결과에 반영될 수 있도록 과목별로 A~F 등의 한 문자로만 평가하는 것이 아니라 아래의 예처럼 과목의 성취기준과 그 성취수준을 보고하는 방식으로 변화하고 있는 주가 증가하고 있는 추세이다.
- ✓ 이러한 성적 보고 방식은 학생과 학부모에게 학생의 강약점에 대한 보다 상세한 정보를 제공함과 동시에 성취기준과 보다 직접적으로 관련지어 평가 결과를 해석할 수 있게 하여 그 결과를 보다 유용하고 종합적으로 활용할 수 있게 하였다.

Academic Performance Level –Subject Areas	
NAME	GRADE
Demonstrates a thorough understanding of the grade level content and consistently applies the benchmarks, and/or concepts, and/or processes/procedures in a variety of contexts	A
Demonstrates understanding of the grade level content and applies the benchmarks, and/or concepts, and/or processes/procedures in a variety of contexts	B
Demonstrates understanding and application of most of the benchmarks, and/or concepts, and/or processes/procedures of the grade level content, needs teacher support to demonstrate proficiency	C
Demonstrates limited understanding and application of the benchmarks, and/or concepts, and/or processes/procedures of the grade level content; needs more instruction and/or practice to demonstrate proficiency	D
Rarely demonstrates understanding and ineffectively applies benchmarks, and/or concepts, and/or processes/procedures of the grade level content	R
Incomplete	I
No Evidence	N
Satisfactory for Pass/Fail courses	S
Unsatisfactory for Pass/Fail courses	U

Subject Areas				
COURSE	Q1	Q2	Q3	Q4
HLA1012Q1 English 9 – Smith, A.	C			
HLA1012Q2 English 9 – Smith, A.		H		
STANDARDS/BENCHMARKS: 1 Read and understand a variety of materials 2 Write and speak for a variety of purposes and audiences 3 Use conventional grammar, usage, sentence structure, punctuation, capitalization, spelling 4 Apply thinking skills in many areas of literacy 5 Locate/select/use relevant information from a variety of sources 6 Read and recognize literature as a record of human experience				

- 미국의 학생평가는 '성취기준(standards)' 중심의 절대평가라는 점에서 우리나라의 학생평가 상황에 중요한 시사점을 준다. 교육과정을 성취기준 형태로 진술하고 이에 대한 달성 여부를 평가하고자 하는 것은 미국을 비롯한 여러 나라에서 1990년대부터 시작되어 현재까지 지속되고 있는 교육개혁 동향의 하나이다. 일정한 교육 기간이 끝났을 때 학생들이 학습의 결과로서 나타내 보일 수 있는 것들을 중심으로 교육목표를 세우고, 교육과정을 성취기준 형태로 진술한다. 교육과정에서 제시된 성취기준을 중심으로 교육활동이 진행되고, 학생평가는 성취기준에 비추어 학생이 얼마나 성취하였는가를 살피는 절대평가가 된다.

- 성취기준 중심의 학습이 강조되면 학생의 수행을 다른 학생들의 수행과 비교하는 것이 아니라 성취기준과 비교하는 방식으로 평가한다. 이러한 맥락에서는 교육목표-교육과정-교육평가가 서로 일관된 체계를 구성하게 된다. 이는 서열화 중심의 학생평가에서 학습자의 학습을 돕는 평가로 전환이 이루어지게 하기 위한 요건이라 할 수 있다. 이와 같이 우리나라의 성취기준을 중심으로 한 2009 개정 교육과정은 현재의 내신 상대 등급제에서 새로운 형태의 절대평가체제로 전환하는 구심점 역할을 할 수 있을 것으로 기대된다.

□ 영국의 학생평가: 국가교육과정의 성취목표를 중심으로 한 준거지향평가(김주아, 2012)

- 영국은 학생들의 낮은 학업성취도의 원인을 국가수준의 교육과정 질관리 체제의 부재로 진단하고, 1988년 교육개혁법에 근거하여 국가교육과정을 도입한 바 있다. 영국의 국가교육과정은 교과별, 핵심단계별로 교육내용과 성취목표(attainment targets)를 제시하고 있다.
 - ✓ 교육내용은 학생들이 배워야만 하는 내용을 의미하며, 성취목표는 "모든 학생들이 각 핵심단계별로 습득해야 할 지식, 기술, 이해"를 의미한다.
 - ✓ 국가교육과정에서는 핵심단계별로 해당 연령대에 성취하도록 기대되는 성취목표를 정해 놓고 있다. 예를 들면, 아래 표처럼 핵심단계 1에서는 수준2까지 도달할 것을 기대하며, 핵심단계 4에서는 학생들의 학업성취수준을 GCSE와 같은 국가자격시험을 통해 확인한다.

<국가교육과정 평가의 핵심단계별 성취목표 달성 수준>

핵심단계	연령	학년	국가교육과정 평가	국가교육과정 수준
1	5-6	1		
		2	교사의 평가(국어, 수학, 과학)	수준 2
2	7-8	3		
		4		
	9-10	5		
3	10-11	6	국가표준화시험, 교사의 평가(국어, 수학, 과학)	수준 4
	11-12	7		
4	12-13	8		
	13-14	9	교사의 평가(국어, 수학, 과학, 그 외 기초과목)	수준 5/6
4	14-15	10	GCSE 혹은 다른 국가자격시험	GCSE 평정

출처(2012. 2. 12 발췌)

http://www.direct.gov.uk/en/Parents/Schoolslearninganddevelopment/ExamsTestsAndTheCurriculum/DG_4016665

- 영국은 국가교육과정 평가의 양 축을 이루는 국가표준화시험과 교사의 평가 결과를 함께 공시하고 있다. 2010년 핵심단계2와 3의 국가교육과정 평가결과 비교표를 살펴보면, 국가표준화시험 결과와 교사의 평가결과가 거의 유사한 것을 볼 수 있다. 이러한 결과는 교사의

평가 결과의 객관성 확보 근거로 활용되며, 일각에서는 이를 근거로 교사의 평가가 표준화 검사를 모두 대체해도 좋을 것이라는 의견을 제시하고 있다.

우리나라도 성적 부풀리기 등의 우려를 불식시키기 위한 한 방안으로 교사의 평가 결과와 객관적인 평가 결과를 비교준거로 함께 제시하는 방안도 고려해 볼만하다.

<2010년 KS2 & 3 국가표준화시험 평가 결과와 교사의 학생평가 결과 비교>

구분	국가표준화시험	교사의 평가
KS2*	국어	81%(남 76%, 여 85%)
	읽기	84%(남 81%, 여 87%)
	쓰기	75%(남 64%, 여 79%)
	수학	80%(남 80%, 여 80%)
KS3**	국어	79%(남 73%, 여 86%)
	수학	80%(남 79%, 여 81%)

* 국가교육과정 성취목표 수준 4이상을 받은 학생 비율

** 국가교육과정 성취목표 수준 5이상을 받은 학생 비율

출처(2012. 2.12 발췌) : DfE: National Curriculum Assessments at Key Stage 2 & 3 in England,

<http://www.education.gov.uk/rsgateway/DB/SFR/s000949/index.shtml>

● 위에서 살펴본 공식적인 국가교육과정 평가 외에 단위학교 수준에서 교사들이 학기말마다 실시하는 학생평가도 국가교육과정 성취목표에 비추어 평가하는 준거지향평가(절대평가) 형태를 취하고 있다. 교사의 관찰, 상시 퀴즈, 과제물, 에세이, 발표 등 평상시 학업 과정에 기초하여 등급이 부여되며, 이와 함께 과제 수행 과정과 학습 태도 등 학업성취 상의 특징이 자세히 기록된다. 학생들의 상대적인 서열을 나타내는 정보는 제공되지 않으며 A, B, C, D, F 등의 종합평정 등급과 학생들의 학업성취 상의 장점과 단점이 종합적으로 기록된다.

● 학교현장에 성취목표에 기반한 평가가 정착되기 위해서는 교사들을 위한 전문적인 연수와 체계적인 안내가 필요하다. 성취목표를 교수-학습과 평가 상황에서 구체적으로 해석해 내는 교사 개인 차원의 훈련과 함께 교과 협의회 등 공동 협의를 통해 개인의 주관적인 판단을 객관화시켜 나가는 공동의 노력이 필요할 것으로 보인다. 또한 영국의 경우 성취목표는 학생의 성취정도를 판단하는 잣대, 기준으로서의 역할 뿐 아니라, 교수-학습의 지향점으로서의 역할을 하고 있다. 성취목표가 교수-학습과 평가에서 길잡이 역할을 할 때 교사는 교과서에 얽매어 가르칠 필요가 없게 되며, 성취목표를 중심으로 내용을 재구성하고, 교재를 채택하며, 다양한 교수-학습 방법을 활용할 수 있게 된다. 이러한 성취목표 중심의 교수-학습 접근은 교과서 중심 수업에서 학생들이 생각하고 참여하게 하는 수업으로 전환하는 데 있어 기폭제 역할을 할 수 있을 것으로 기대된다.

II. 세계 주요국의 평가를 위한 노력

□ 2010 교육과정·교육평가 국제동향 연구 사업 -미국- (한국교육과정평가원, 2010)

♣ 차세대 공통 평가제도

A Unique Moment in Time: Creating the Next Generation Comprehensive Assessment Systems To Serve All Our Children Well Dr. Forgione's Keynote Address(10.8.12)

● 공통 평가제도 소개

미국 교육부가 오바마 정부 교육정책인 Race to the Top(RTTT)에 대한 교육 예산 배분을 공통핵심 교과목표를 채택하는 주에 한정한다고 발표함으로써 '10 8월 현재 34개의 주들이

공통교과 목표를 채택한다고 발표하였다. 이와 함께 미 교육부는 공통교과 목표에 바탕을 두는 미국 전역에서 실시할 수 있는 공통 평가제도에 대한 계획을 발표하였다. RTTT평가제도 혹은 공통평가제도는 다음의 사항을 중점으로 한다.

- ✓ 개개 학생의 능력뿐 아니라 진보(성장)과정 측정
- ✓ 개개 학생의 대학 및 취업 준비를 위해 각 학년에서 고등학교까지 장기적 학업성취도를 측정하기 위한 체제 마련
- ✓ 평가결과가 다음 사항에 유용한 정보를 제공
 - 학습, 교습, 프로그램 개발 및 향상
 - 학교 효율성 결정
 - 교사, 교장의 자질 및 효율성 평가
 - 학생의 고등학교 졸업, 대학 입학 결정

● **공통 평가제도 컨소시엄**

- ✓ 미 교육부는 개개 주들이 그룹을 형성하여 어떠한 공통 평가제도 체제를 마련할 지에 대한 계획서를 제출하도록 요구하였으며, 계획서가 통과될 시에 총 3억5천만 달러(\$350 million)을 지원한다고 발표함
- ✓ 새로운 컨소시엄 평가제는 현 NCLB 평가 제도를 대체하게 되며, 2014-2015학년도에 첫 실시가 될 예정임
- ✓ 각 주들의 교육부들은 3개의 컨소시엄 중에 선택 참여하였으며, 그 중 PARCC와 SBAC는 초등학교-고등학교(K12) 평가 제도를 제안하고 있으며, SCOBES는 고등학교 평가제도에 대한 계획서를 제출하였음

④ **PARCC(Partnership for Assessment of Readiness for College and Careers) 컨소시엄 계획서 요약**

- 3천백만 명의 학생이 포함되는 총 25개의 주가 참여함
(알라바마, 애리조나, 아칸소, 캘리포니아, 콜로라도, 델라웨어, D.C., 조지아, 일리노이, 인디애나, 켄터키, 루이지애나, 메릴랜드, 메사추세츠, 미시시피, 뉴햄프셔, 뉴저지, 뉴욕, 노스다코타, 오하이오, 오클라호마, 펜실베이니아, 로드아일랜드, 사우스캐롤라이나, 테네시)
- 이들은 RTTT평가를 위한 미 교육부의 자금을 얻기 위해 협력하며, K-12(유치원-고등학교) 학년을 위한 영어/언어와 수학 평가 제도를 2014-2015년까지 마련하겠다고 제안함
- 또한 26개주에 걸쳐 200개의 2, 4년제 대학들이 참여하여 고등학교 평가제도 설립에 도움을 주겠다고 발표함

PARCC평가제도 계획서의 특징
1. 참여한 주들은 이 공통 평가제와 공통 수행 교과목표를 채택할 것이다.
2. 이 평가제도는 대학과 취업 준비 지식 및 능력과 연관되어 마련될 것이다. <ul style="list-style-type: none"> • 고등학교 영어, 수학 평가는 공통교과목표와 연관될 것이고, 대학 1학년 필수 과목을 성공적으로 이수할 수 있는 지식과 능력을 평가할 것이다. • 3-8학년(중3) 및 고등학교 1,2학년 평가들은 학생들이 고등학교 졸업에 필요한 지식, 능력을 갖추었는지를 체계적으로 평가할 수 있는 장기 연관 체제를 마련할 것이다.
3. 이 평가제도는 각 학년에서 일련의 수업 평가(Through-course)가 실시되며, 학생의 학업성취도를 종합적으로 평가할 학년말 시험을 포함한다. <ul style="list-style-type: none"> • 시험평가들은 학생들이 배운 영역을 평가하고, 그 평가 기간과 학생들이 배우는 기간이 근접하도록 하며, 즉각적인 피드백을 주어 교사들이 학생들이 필요한 점을 인식하고 수업을 조정할 수 있도록

한다.
4. PARCC 평가제도는 컴퓨터에 기초한 평가로 이루어진다. <ul style="list-style-type: none"> 평가 결과가 더욱 신속히 보고될 수 있으며, 평가 비용도 절약될 것이다. 교과목표를 효율적으로 측정할 수 있도록 테크놀로지를 이용한 문항을 포함한다. (상위의 사고를 요하는 수행평가 문항을 포함하여, 21세기에 필요한 지식 및 기술을 측정하고, 복잡한 사고를 유출할 수 있도록 한다.)
5. 교사들은 학생들의 평가 결과를 수시로 볼 수 있으며, 이를 수업 계획에 반영할 수 있다. <ul style="list-style-type: none"> PARCC는 교사들이 평가결과를 수업에 효율적으로 이용할 수 있도록 도움 자료 및 가이드라인을 개발할 것이다.

② SBAC(SMARTER Balanced Assessment Consortium) 컨소시엄 계획서 요약

- 2009년 12월에 31개주로 결성되어 미 교육부가 요구한 공통교과목표를 기준한 공통 평가 제도 마련을 목표로 결성되었다.
- 워싱턴 주가 이 컨소시엄을 현재 대표하고 있으며, 이 중 많은 주들이 PARCC 컨소시엄에도 포함되어 있다. (두 컨소시엄 모두 공통 평가제도의 계획안 마련 단계이므로, 많은 주들이 두 컨소시엄 모두에 관심을 갖고 있기 때문임)

SBAC평가제도 계획서의 특징
1. 선택 형성평가 혹은 벤치마크 평가 <ul style="list-style-type: none"> 학교, 교사의 결정에 따라 선택적으로 시행될 중간 평가로 computer adaptive형식을 취한다. 평가 범위는 작지만 깊이 있는 문항 및 태스크를 포함하며, 학생의 취약점을 측정하는 진단평가 결과를 제공하는데 그 목적을 둔다. 학생 평가 결과는 교사와 연계되어 제공되어 교사의 수업 계획에 도움을 주도록 한다.
2. 수행평가 과제 <ul style="list-style-type: none"> 컨소시엄에서 지정한 평가 기간 중 마지막 12주는 학생들에게 읽기, 쓰기에서 각각 1과제, 수학에서 2과제를 완수하도록 한다. 각 과제는 1-2수업시간이 소요될 것이며 학생들이 먼저 계획, 정보, 아이디어, 다른 자료 수집 등 프로젝트를 개발하고, 발표, 전시, 리포트로 표출하도록 한다. 교사와 컴퓨터 자동 채점이 함께 적용될 것이며, 과제 평가 결과는 2주안에 발표되도록 한다.
3. 교사의 수업준비에 도움이 될 수 있는 정규적인 평가제도 제공
4. 의무적 학년말 시험 <ul style="list-style-type: none"> 이 학년말 시험은 읽기, 쓰기, 수학에서 40-65문항으로 구성되며, 학년말 마지막 12주 동안에 computer adaptive형식으로 치러질 것이다.(사지선다형, 서술형, 테크놀로지를 이용한 문항들 포함) 컴퓨터를 이용한 자동 채점 및 교사의 채점이 함께 적용될 것이며, 과제 평가 결과는 2주안에 발표되도록 한다. 학교의 특성에 따라 다른 채점 체제가 개발될 수도 있다.
5. 모든 평가들은 학생들에게 대학 및 취업 준비를 목표로 하고, 이에 따른 학업성취도, 진보(성장) 과정도를 제공하는 것을 목표로 한다.

③ SCOBES(State Consortium on Board Examination Systems) 컨소시엄의 고등학교 평가제도 계획서

- 영어, 수학, 과학, 역사, 예술 과목에서 고등학교 평가제도 마련에 초점을 둔다.

SCOBES평가제도 계획서의 특징
1. SCOBES는 새 평가 개발 자체에 참여하는 것이 아니라, 기존의 평가제도 및 기관을 리뷰, 평가, 선택, 인가하는 과정에 참여할 것이다. 이에 참여할 평가 및 평가 기관들은 공통교과목표와 일관성이 있어야 한다.

2. 여기에는 교과과정, 교과서, 학습자료, 형성평가, 종합평가, 교사 지침서, 채점제도를 포함하고 있어야 한다.
3. 고등학생들의 진로는 크게 5분야(2년제 대학/ 2년제 혹은 기술학교, 자격증 프로그램/ 일부 4년제 대학/ STEM에 관계된 4년제 대학/ 취업 및 자격증 취득, 기술직)로 나뉜다. <ul style="list-style-type: none"> • 이 컨소시엄은 고등학교 하위, 상위 분야를 구분하여, 모든 학생들이 기본적으로 하위 분야에 해당되는 기본 지식 및 기술 습득을 요구한다. • 하위 분야에 있어 평가제도는 영어, 수학, 과학, 역사, 예술(10학년이후) 과목에서 패스할 것을 의무화한다. • 그 패스 기준은 2년제와 4년제 대학 기초과목에 필요한 읽기/쓰기 능력을 기본으로 한다.
4. 하위 분야 완수 후, 학생들은 위에서 언급한 5분야의 진로에 따라 그에 해당하는 상위 분야 과정을 이수한다.
5. 하위, 상위 분야 진로에 따라 평가제도가 마련될 것이며, 주로 작문이나 서술형 문항, 프로젝트 완수로 구성될 것이다. <ul style="list-style-type: none"> • 채점은 학교 외부 단체에 의해 행해지되, 컴퓨터나 훈련을 받은 교사에 의해 행해질 것이다. • 특히, 학생들의 성장 정도를 장기적으로 측정하는 방법들이 연구 개발될 것이다.

□ 2010 교육과정·교육평가 국제동향 연구 사업 -핀란드- (한국교육과정평가원, 2010)

※ 본 연구는 2010년에 행해졌으며, 핀란드에서 발표한 자료 중 2009년 자료들을 통해 핀란드의 평가에 대한 방향을 살펴보고자 한다.

● 핀란드, 국제 교육평가 도구 개발에 참여할 예정(09.6.7)

- ✓ 핀란드, 오스트레일리아, 영국, 포르투갈, 싱가포르와 함께 Melbourne대학, 컴퓨터 공학 분야의 회사들인 Cisco, Intel, Microsoft가 참여하는 Assessing and Teaching of 21st Century Skills 연구개발 프로젝트에 참여함(*자세한 사항은 atc21s.org 사이트를 참조*)
- ✓ 이 프로젝트의 목표는 교육평가 부문에서 국내와 국제적인 차원 모두에서 이용할 수 있는 새로운 과학 기술을 창조하는 것으로, 평가 기술의 도움으로 여러 다른 학교 형태도 앞으로 광범위하게 비교 연구할 수 있을 것이다(ex. PISA 연구)
- ✓ 이 프로젝트의 또 다른 중심목표는 창의성, 혁신 능력, 문제 해결 능력, 상호 영향, 협동 능력과 과학 기술 사용 능력을 평가하는 것인데, 이들은 미래의 직장 생활에서 필요하고 교육의 도움으로 개발하고자 하는 기술, 능력들이다.
- ✓ 이 연구결과들은 핀란드에서 수업시간 분배 개선과 교과과정 기본 원칙들을 개정하는데 활용할 수 있을 것이다.

● 초·중등교육 발전을 위한 교육의 질에 관한 기준 완성(09.6.11)

- ✓ 초·중등교육의 질에 대한 평가기준들이 처음으로 완성됨
- ✓ 평가기준들은 10개의 평가 영역에 초점을 두며, 각각의 영역은 교육의 질에 대한 자체의 평가기준들을 가지고 있음
- ✓ 핵심 평가기준들은 다음과 같이 *구조적인 질(4 criteria)*과 학생들과 관련한 교육의 질(6 criteria)이 포함된다.
 - 1) 감독운영(Governance)
 - 2) 인사(Personnel)
 - 3) 자원(Economic resources)
 - 4) 평가(Evaluation)
 - 5) 교과과정 실행(Implementation of the curriculum)

- 6) 수업·교수 계획(Instruction and teaching arrangements)
- 7) 배움, 성장, 복지 지원(Support to learning, Growth and Well-being)
- 8) 학생 참여와 영향력(Inclusion and influence)
- 9) 학교와 가정 간의 협력(School-home cooperation)
- 10) 학습 환경의 안전성(Safety of the learning environment)

Ⅲ. 교육과정과 교육평가의 연계방안(곽병선, 1998)

※ 교육평가는 교육과정의 실현을 돕는 하나의 과정이라고 설명되고 있지만, 실제로 교육평가는 교육적 성취 기준으로 작용하여 학습자의 학습에 결정적 영향을 미침으로써 교육과정을 규율하고 있다. 특히 한국의 학교 교육에서 대학입학시험제도는 학교 교육과정과 수업 전반을 규율하는 엄연한 현실이었음에도 불구하고, 이러한 선발시험이 교육과정에 미치는 영향을 심층적으로 분석하지 않고 교육평가를 마치 교육과정의 하위 변수인 것처럼 단순히 취급하여 논의하는 경우를 많이 본다.

※ 대학입학이나 고시처럼 선발기능이 강한 고부담 시험¹⁾이 무엇을 성취기준으로 삼고, 어떤 양태로 실시되는가는 교육과정에 결정적 변수로 작용하고 있다.

1. 외적 고부담 시험제도의 교육 파괴적 영향

한국의 대표적인 고부담 시험
<ul style="list-style-type: none"> • 대학수학능력시험, 사법·행정고시, 기업체(방송, 언론사, 은행 포함) 입사 시험 등 • 이들의 시험 결과는 수험생의 장래에 갈림길과 같은 영향을 미친다.

- ① 단편적 지식·정보 중심의 정답 찾기 평가
 - ✓ 교과서 내용에 대한 반복적 암기학습을 요구하기 때문에 창의적 발상이나 출제자의 의도와는 다른 새로운 구상을 펼치는 능력이 제한을 받게 된다.
- ② 지필 시험 위주의 평가방법
 - ✓ 수험자의 체득된 능력을 총체적으로 평가하지 못한다. 예를 들면 수험자의 수영 능력, 과학 실험 수행능력, 피아노 연주능력, 토론능력, 타인을 배려할 줄 아는 능력, 기본 생활습관과 태도 등은 지필시험으로 측정할 수 없는 것이다.
- ③ 국가 주도의 외적평가
 - ✓ 외적 시험은 학교 수업을 그 시험의 기준에 맞추도록 작용하여 학교 자체의 내생적 교육 계획 추진을 어렵게 한다. 즉, 학교가 독자적으로 교육을 추진하기 어렵다.
- ④ 고부담 시험제도의 학교교육에 대한 영향
 - ✓ 공식적인 교육과정이 설정되어 있을지라도, 수험자 개개인에게 막대한 영향력을 미치는 고부담 시험이 있다면 학교 교육을 좌우하는 것은 교육과정이 아니라 바로 고부담 시험이다.
 - ✓ 누구에게나 동등한 기회가 주어진 점에서 평등성이 보장된 제도이고, 국가가 직접 관장

1) 여기서 고부담 시험이란 필자가 사용한 용어로서, 시험의 결과가 수험자의 학교 성적이나 순위 결정, 선발 과정에서 합·불합격 등 수험자의 이익과 불이익에 작용하는 평가로서, 수험자에게 큰 부담을 주는 평가라는 뜻에서 고부담이라는 말을 붙였다. 영어의 high stakes tests에 해당하는 말이다. 마다우스(Madmen, 1988)는 high stakes tests를 시험의 결과가 학생, 교사, 학부모, 행정가 등이 옳던 또는 옳지 않던 중요한 결정에 사용되는 시험이라고 정의하였다. 그렇지 않은 시험은 저부담 시험(low stakes tests)이라고 이름 붙였다.

하여 공정성과 객관성에 있어서는 사회 전반의 신뢰를 받고 있다고 볼 수 있다. 즉 선발 기능이라고 하는 시험의 목적에만 비추어 보면 한국의 고부담 시험은 성공적인 제도인 것이다.

- ✓ 하지만 우리의 고부담 시험제도에는 깊은 자폐적 문화 결핍증이 있다. 복합적으로 작용하는 다면적 현상을 한두 가지의 관점이나 방법론으로 해결하려는 데에서 생기는 결핍증이다. 거기에는 종합적으로 고려하고 예외가 있을 수 있다는 것을 인정하지 않는 사고방식이 가져오는 폐쇄성도 작용한다. 선다형 시험 성적 하나만으로 전국의 학생들을 등급화하고, 대입 당락을 결정했던 과거 수십 년간의 대학입학제도가 바로 대표적인 예이다. 여기에는 크게 두 가지가 관련된다. 하나는 교육이라는 것을 기존 신념체계나 가치체계를 수용하는 것으로 생각하는 교육의 이해에 대한 결핍증이다. 즉 기존의 전통, 사상, 기술, 경험을 받아들이는 것이 교육이지, 새로운 자기의 것을 생성하는 능력을 기르는 것을 강조하지 못하는 교육인 것이다. 교과서대로 따르도록 하는 것이 잘 가르치고 잘 배우는 것이라고 생각하여 모든 지필검사가 학생들에게 요구하는 것은 기존 질서에 대한 수용을 요구하고 있는 것이다. 산 경험의 학습이 없는 서적(교과서) 중심의 죽은 지식 암기교육이 바로 그것이다. 다른 하나는 고부담 시험에 대한 제도 설정에서 평가를 너무 단순하게 생각해 왔다는 것이다.
- ✓ 지금까지의 고부담 시험제도는 “선발”이라는 데에만 초점을 두었지 “교육을 위한” 평가여야 한다는 데에는 관심을 두지 않은 제도이다. 교육을 도외시한 평가는 엄정한 의미에서 교육평가라 볼 수 없으며, 학교 교육 역시 제대로 설 수 없고, 교육과정도 제자리를 찾을 수 없다.

2. 새로운 교육평가의 필요

● 최근 교육과정 개발자들은 지필시험에 변화가 생기지 않는 한 교육과정에 변화가 일어나지 않는다고 결론지었다(Gipps and Murphy, 1994). 우리는 이미 과거의 대학입학학력고사나 현재의 대학수학능력시험제도를 통해 이를 현실에서 확인하고 있다. 이 시험의 성적이 학생들의 진로 결정을 좌우하고 있는 한 이 시험에 철저히 대비하는 것보다 중요한 교육과정은 없는 것이다.

● 현재와 같은 국가 주도 지필시험방식의 고부담 시험제도를 운영해서는 안 되는 이유

- ① 외적 고부담 시험 위주의 사회가 교육을 왜곡시키고 장기적으로 사회의 진보를 가로막고 있다는 현실에 대한 깊은 통찰
 - ✓ 교육을 철저히 수단화하고, 피상적 학습을 조장한다.
 - ✓ 총체적으로 교육이 학습자의 인간됨이나 실제적인 역량을 길러내는데 실패하여 학습과 인격의 소외 현상이 발생하고 있는 것이다.
- ② 지금의 시험제도를 가지고서는 다가오는 지식기반사회에 살아남을 수 있는 능력을 가진 사람을 길러내기 어려울 것이라는 전망
 - ✓ 지식기반사회는 지식이 삶의 중심에 서는 사회로, 각성된 의식, 풍부한 지식과 정보가 새롭고 가치 있는 삶을 만들어가는데 중점적 역할을 하는 사회인 것이다. 이 새로운 문명에 참여하려면 우리의 교육은 지식수용교육에서 지식 생성교육으로 그 근본을 바꾸어야 한다. 교사마다 전문성을 자율적으로 발휘할 수 있는 조건을 만들어주어야 하는 것이다.

3. 교육평가의 개선방향

① 학교가 시험 준비 위주의 교육에서 벗어나 그 본연의 인간교육에 충실하도록 하려면 사회 각 분야에서 실시되고 있는 외적 고부담 시험의 비중을 줄여야 한다. 우선 대학수학능력 시험 점수에 대입당락을 결정하는 제도가 수정되어야 한다. 그러기 위해서는 학교가 누구에게나 교육적으로 설득력을 가질 수 있도록 학생생활기록부를 심층기록자료가 되도록 개선해 나가는 것이 필요하다.

② 학교 학생 성취평가를 측정평가 중심 사고에서 교육의 질 중심 사고로 전환하여야 한다. 측정평가 중심 사고란 심리검사의 기법을 교육평가에 적용할 수 있다고 믿는 사고방식이다. 우선 평가는 교육적이어서 학습자에게 교육적 성취를 가져오게 해야 한다. 이를 위해 교사와 학습자가 같이 학습의 목표가 무엇이며, 그것을 어떻게 도달할 것인지에 대해 상호 문제 의식을 공유하고 있어야 한다. 교수-학습과정에서 도움을 얻기 위해 그 진도에 관한 정보를 생성하는 과정이 평가인 것이며, 학습과정에서 학습자가 보인 진도에 관한 정보가 평가정보가 되는 것이다. 학생 각자의 삶과 선택에 의미를 가질 수 있도록 학습내용을 학습자의 인격에 연결시키는 심층학습이 필요하며, 그러한 심층학습을 촉진하기 위해서는 논술시험, 프로젝트 시험, 수행평가, 소집단 과제법 등을 부과하여 이 결과를 학생 개개인에 대한 심층 기록으로 정리하여야 한다. 이러한 심층학습을 자극하기 위해 교육평가는 신뢰도나 변별력보다 타당도를 중요하게 고려하여야 하며, 외적 평가는 자제하고, 자체평가, 교사 상호평가 등이 장려되어야 한다. 외적 평가의 남용은 고부담 시험제도와 유사한 부정적 영향을 남긴다.

③ 평가업무에 있어 교사의 역할을 증대하여야 하며, 학교생활기록부가 공정하고 신뢰받는 자료가 되도록 계속 발전시켜야 한다. 이를 위해 수업의 일과로 바쁜 교사들을 대신하여 교육과정평가원이 교과별 평가기준안을 개발하여 보급하여야 하며, 학생 개인별 심층기록인 학교생활기록부가 학교별, 지역별 자료와의 비교성을 갖출 수 있도록 이에 대한 후속 보완 연구가 있어야 한다.

④ 학교의 학생생활기록을 공정하고 객관성 있게 조정하고 중재할 전국-지역 단위의 교과 중심 전문조직을 형성하여야 한다. 시·도 지역 수준에서 학교의 학생생활기록이 상호 비교성을 갖출 수 있도록 조정해주는 조정관(moderators)제도를 적어도 교과별로 설치, 운영하여야 한다. 성적 인플레이션 현상 같은 폐단을 막고 항상성 있는 평가가 이루어지도록 하기 위해 교육부와 시·도 교육청에 교과별 조정관을 설치하고 이 조정관들의 협의체를 구성하여 해당 교과의 학생 성적 등급 판정의 기준을 마련해야 한다. 또한 이 협의체를 통해 해당 교과의 평가와 관련하여 지역차원 또는 학교별로 발생하는 현실적 문제를 취합하고 이들에 대해 어떻게 대응하도록 할 것인가를 심의 결정하여 지침으로 제시하여야 한다. 시·도 교육청의 교과별 조정관은 해당 관할 지역 내의 교과 담당교사들이 상호 유기적으로 정보를 교환하며, 동일한 원칙을 가지고 학생들을 평가하고 그 기록을 공정하게 관리할 수 있도록 관찰하고 조정하는 역할을 해야 한다. 이러한 제도는 독일의 아비투어(abitur)제도나 영국의 중등교육 자격시험(GCSE)에서 참고할 수 있다.

4. 교육과정을 위한 교육평가의 자리

● 국가교육과정을 운영하고 있는 우리의 교육과정 정책 하에서는 대학신입생선발을 위하여 현재의 대학수학능력시험과 같은 국가 관리의 외적 고부담 평가는 필요하지 않다고 생각한다. 국가 교육과정이 있다는 말은 학교가 지향해야 할 교육목표와 성취기준이 국가 차원에

서 마련되어있다는 것을 의미하며 학교의 역할은 바로 이 교육과정에 제시된 정신과 교육목표, 성취기준이 실현되도록 충실한 교육과정을 펼치는 것이다. 학생이 최대한 발달을 이루어나가도록 돕고, 학생이 어떤 발달 정도와 특징을 보이는지를 학생생활기록부에 심층 기록하여 학교교사의 전문적인 지도를 받게 하고, 학생이 진로 결정을 하는데 돕는 이러한 활동을 충실히 하는 것 이외에 더 중요한 일은 학교에 없다고 생각한다.

- 따라서 학교의 교육과정에 직접적 영향을 미치는 외적 고부담 평가를 설정하는 것은 국가 교육과정 시책에 역행하는 것이다. 어떠한 형태로든 선발기능을 가질 수 있는 고부담 시험이 필요하다면 이는 외적 평가의 비중을 줄이고 내적 평가의 비중을 늘려나가는 방향으로 가야할 것이다. 외적 평가는 교사와 학생에게 수업과 학습에서 내재적 동기보다 외적 동기를 자극하고 학교교육을 시험에 종속시키는 악영향을 가져오기 때문이다.

- 교육평가의 기본원리는 학습을 돕는 것이다. 우리 한국의 상황에서는 창의력을 길러낼 수 있도록 교육평가가 제 역할을 하여야 한다. 한국 교육은 수용교육체질을 벗어 던지고 신지식 생성력을 기르는 교육으로 탈바꿈하여야 하며, 여기에 교육과정의 과제가 있고 교육평가의 역할이 있다. 지금과 같은 지필시험 위주로 고급인력을 선발하는 외적 고부담 시험제도는 없어져야 하며, 교사들의 학생생활기록을 사회가 믿어주어야 하며 교사들은 평가 전문성을 높여야 한다.

IV. 미국의 고부담 시험의 가능성과 한계(박세훈, 2010)

※ 지식기반 사회의 도래와 더불어 세계 각국은 교육경쟁력을 높이기 위한 교육개혁에 더욱 박차를 가하고 있고, 이는 교육경쟁력이 국가 경쟁력을 결정하는 중요한 요인이라는 공통된 인식에 기반을 두고 있다.

1. 고부담 시험의 의의와 논쟁

- 레이건 행정부 당시 수월성 교육 전국위원회에서 '위기에 처한 국가'(A Nation at Risk)라는 보고서를 발간한 이후부터 교육개혁의 차원에서 학교가 높은 수준의 성취기준을 충족하고 책무성을 지녀야 한다는 요구가 많이 제기되었으며, 그 이후부터 고부담 시험이 중요한 교육문제로 부각되고 있다. 이 보고서 이후에 많은 주들은 졸업시험을 최소한의 능력시험으로 대체하게 되었다(Heubert, 2000)

- NCLB법의 시행으로 미국 공립학교에서 시험과 책무성 운동은 새로운 전기를 맞이하였다(Ravitch, 2010). 주 정부가 주요 교과에 대한 학력 성취기준을 제정하고 표준화 시험을 통해 확인할 필요가 생겼으며, 특히 유색인종이나 영어학습자(ELL) 및 장애학생들이 미국의 학교에서 교육 수준이 낮은 집단으로 분류되어 그들도 높은 수준의 교수학습 표준 운동의 혜택을 받아야 할 필요성이 제기되었다(Ravitch, 2010). 이런 점에서 표준화 운동과 고부담 시험은 특히 이들에게 큰 영향을 미치고 있다.

2. 고부담 시험의 기원과 역사

- 미국 국민들이 기존에 학교에 대해 가졌던 긍정적인 관점이 바뀌기 시작한 것은 1970년

대 초반으로 이는 SAT 시험 성적의 저하된 것과 맞물려 있다. 성적저하가 소수인종이나 빈곤자녀의 대학진학으로 SAT 시험 지원자가 늘었기 때문이라는 분석결과도 있었지만 1970년대 이후에 미국에서 시행된 책무성 노력은 미국의 공교육 질이 저하되고 있다는 인식에서 비롯되었다고 할 수 있다.

- 1970년대에 시행되었던 최소능력시험(minimum competency test)운동은 시험이 교육개혁의 효과를 측정하는 수단뿐만 아니라 교육개혁의 주요 추진 동인으로 활용되는 최초의 교육개혁 일환이라고 할 수 있다. 의회 의원 등이 주축이 되어 최소능력시험을 고등학교 졸업장과 연계시키기 시작했으며, 1980년에 37개 주에서 최소능력 성취기준을 의무적으로 규정하여 이를 통과하지 않으면 학년 진급이나 고등학교 졸업을 할 수 없도록 하는 조치를 취했다.

- 하지만 1983년에 발표된 '위기에 처한 국가'보고서에서 최소능력시험은 근본 해결책이 될 수 없다는 주장이 제기되었다. 학생들에게 요구하는 최소의 성취기준이 학생들이 도달해야 하는 최고 수준으로 잘못 인식되어 결과적으로 모든 학생을 위한 성취기준을 저하시키는 결과를 초래하기 때문이라는 것이다(National Commission on Excellence in Education, 1983).

- 이런 점에서 미국의 공립학교에서의 고부담 시험 기원은 레이건 행정부 당시의 '위기에 처한 국가'라는 보고서를 발간한 시점으로 거슬러 올라간다고 보는 관점이 일반적이다. 이 보고서는 미국의 공립학교들에게 엄격한 성취기준이 있어야 하며, 현재의 공교육은 실패하고 있다고 주장하였다. 또한 기업의 대표들이 모여 소위 기초교육으로의 교육과정 회귀를 주창하고, 학교가 높은 수준의 표준을 충족하고 책무성을 지녀야 한다고 요구하기 시작하였다. 이러한 교육개혁의 움직임은 당시의 경제사정을 잘 이해하고 있는 경영 전문가 집단에 의해 주도되었다(Johnson, Johnson, Farenga, & Ness, 2008).

- 루이지애나 주 의회가 학교책무성자문위원회를 1998년 위촉한 것이 미국에서 가혹한 결과를 초래할 고부담 시험제도를 처음으로 도입한 것으로 알려져 있다. 4학년과 8학년이 시험의 대상이 되었으며, 미리 설정된 성취기준에 도달하지 못한 학생들은 유급을 당하게 되었고, 성적이 저조한 학교들에게도 점점 심한 제재조치가 가해졌다. 2008년까지 24개 주들이 고등학교를 졸업하기 위하여 주정부 수준의 시험에 합격해야 할 것을 요구하고 있다(Education Week, 2006). 또한 16개 주에서 시험성적이 우수한 학교에 대해 금전적인 보상을 해주고 있으며, 14개 주는 실패한 학교의 운영권을 기업경영자에게 위임했으며, 28개 주는 성적이 부진한 학교에 대해 다양한 벌을 부과하였으며, 10개 주는 학교의 폐쇄를 허가하였다(Johnson & Johnson, 2006).

- 2002년 공표된 NCLB법은 표준화된 시험 성적을 근거로 학생의 하위집단별 적정수준향상도(AYP)에 도달하지 못한 실패한 학교를 제재하도록 하였다. 극단적인 경우에는 학교의 전 직원을 교체하는 경우도 있다. 오바마 정부는 NCLB법을 존속시키되, 그간 지적되어왔던 문제를 수정·보완하여 지속적으로 추진할 계획임을 밝혔다. 수정안의 골자는 자원 분배 방식의 변화, 학업성적 도달 기한의 폐지, 교사의 자격과 평가기준의 재정의로 요약된다(염철현, 2010). 특히 오바마 정부는 주마다 다르게 설정된 학업 표준을 전국적으로 통일하여 전국

공통의 전국학력평가기준(common academic standards)을 만들어 모든 학생들이 고등학교를 졸업하면 대학진학준비 혹은 취업준비가 될 수 있는 연계체제를 구축할 계획이다. 이는 앞으로 미국의 공립학교에서 실시하는 평가가 고부담 시험의 성격을 더욱 강화하는 방향으로 전개될 것임을 시사한다.

3. 고부담 시험의 가능성과 한계

※ 고부담 시험의 효과에 대해서는 아직도 논쟁이 많다.

□ 고부담 시험의 가능성

● 고부담 시험 지지자들은 막대한 인센티브와 위협적인 처벌이 성적에 따라 주어지면, 교육자와 학생들이 열심히 노력하여 결과적으로 학생의 성적이 높아질 것이라는 생각에 근거하여 학생의 표준화 시험 성적을 근거로 보상과 제재를 가하는 체제를 도입해야 미국 공교육 질이 높아질 수 있다고 믿는다(Raymond & Hanushek, 2003).

● Bishop 등(2000)은 학생의 학교교육, 학습, 소득에 미친 영향을 분석한 결과, 학생들의 기초 기술을 측정하는 졸업시험을 포함한 학교개혁 관련 정책이 많은 학생들에게 대학진학을 제고나 소득의 증가를 가져왔다고 보고하였다. Heubert와 그의 동료들은(2009) 표준 기반 교육개혁과 시험이 시험 성적이 저조한 이른바 미국에서 교육적으로 불이익을 받고 있는 집단의 학생들에게 도움을 주고 있다고 주장하였다.

● 최소능력시험에 탈락하는 학생이 계속해서 감소하고 있다는 통계도 고부담 시험의 효과를 지지하는 증거로 활용되고 있다.

● 최근의 연구는 고부담 시험이 학습내용에 미치는 영향이 여전히 상당함을 보여주고 있으며, 평가의 내용이 개선되는 정도로 수업과 교육과정의 개선이 나타나고 있음을 확인시켜 주고 있다. 그 긍정적인 예로 글쓰기 수업이 있다.

□ 고부담 시험의 한계

● 고부담 시험의 부정적 요과를 종합한 연구 보고서(Shepard, 2010)

- ✓ 학생의 학습능력이 제고되지 않았음에도 불구하고 시험 점수는 상승하는 이른바 시험 점수 부풀리기를 가져왔다.
- ✓ 시험과목이 아닌 교과목은 교육과정에서 빼버리고 대신 필요한 과목에 많은 시간을 투자하고 기계적인 암기 위주의 교육이 이루어져 교육과정의 왜곡을 불러오기도 하였다.
- ✓ 시험성적에 근거하여 인센티브나 처벌을 하는 것은 학습동기의 관점에서 심각한 문제가 있다는 것이 학계의 상식이다. 시험성적을 올리기 위해 공부하는 학생은 학습흥미를 유발하는 내적인 동기유발이 덜 일어나며, 학습도 적게 하며, 어려운 과제 해결을 위해 지속적으로 노력하지도 않는다는 것이다.

● 고부담 시험과 책무성이 오히려 학생들이 도달해야 하는 수행 표준을 낮추는 결과를 초래하였다는 평가를 하고 있다(Hursh, 2009). 예를 들어 New York주의 표준화 시험문제는 그 문제의 질이 떨어지고, 학생들을 호도하는 문제들도 있으며, 때로는 문제가 잘못 출제되는 경우도 있다.

● 학업성취율과 졸업률을 높이기 위해 주(州)들이 보다 쉽게 졸업시험을 출제하며, 낮은 성취도를 보이는 학생들을 시험대상 집단에서 제외하는 현상도 일어나고 있음을 지적하고 있다(Huebert, 2009). 또한 Huebert는 유색인종이나 영어학습자 및 장애 학생들은 더욱 교육

에서 멀어져 졸업시험을 보기 전에 학교를 그만 두거나 학년을 유급당하는 고통을 받게 되는 희생자가 되고 있음을 지적하였다. 즉 미국 사회에서 교육적 배려 대상자에게 균등한 교육기회를 제공하기 위해 시도된 교육개혁 조치가 오히려 교육 불평등을 초래하는 역효과를 보이고 있으며, 실제로 그들의 중도 탈락률도 대폭 증가하고 있다는 것이다.

4. 논의 및 한국의 교육을 위한 시사점

□ 고부담 시험이 교육과정과 수업에 어떠한 영향을 미치는가?(Madaus, 1988)

- ① 학생, 교사, 행정가가 어떤 시험에 대해 그 결과가 중요하다고 믿고 있다면 그 사실여부와 관계없이 그렇게 믿는 만큼 영향을 미친다.
- ② 어떤 양적 사회적 지표를 사회적 결정에 사용하면 할수록 그것을 관리하는 사회적 과정은 왜곡되거나 부패할 가능성이 높다. 시험 결과 역시 중요한 사회적 결정에 이용할수록 시험의 긍정적 효과는 감소하고 결국 교육의 본질을 훼손하고 부패하기 쉽게 만든다.
- ③ 중요한 결정에 시험결과를 사용할수록 교사는 시험만을 가르치게 된다.
- ④ 고부담 시험이 제도로 확립되어 있고 이 시험의 통과가 하나의 중요한 전통으로 서있는 경우, 이 시험은 실질적으로 교육과정을 규정한다. 여기서 크게 주목할 점은 교사가 학생들에게 이 시험의 통과에만 열중하도록 함으로써 교원의 전문성을 강조하는 온갖 논리들을 무의미하게 만든다는 것이다.
- ⑤ 교사들은 고부담 시험의 출제 방식에 관심을 기울이고 그 방식에 맞는 교수법을 사용하게 된다. 즉 시험 방식에 부적합한 내용이나 교수방법은 도외시되는 부정적 결과가 초래되는 것이다.
- ⑥ 시험결과가 미래의 교육이나 진로 선택에 사용되는 경우 고부담 시험 준비가 학교 교육의 중요한 목표가 되어야 한다고 사회에서는 여기게 되어 고부담 시험의 성취가 때로는 교육에 해가 될 수도 있다는 사실에 대해 별로 관심을 기울이지 않게 된다.
- ⑦ 고부담 시험은 교육과정을 관할하는 기능을 소관 부처나 학교가 아닌 시험관리 기관으로 전이시키는 결과를 가져온다. 이는 우리나라에서도 매년 대학수학능력시험이 치러지고 나면 사설 입시기관들의 시험문제 분석과 입시 예측을 언론기관들이 앞 다투어 보도하고 있는 상태에서 잘 드러난다.

□ 고부담 시험이 교육과정 운영과 교수학습에 미치는 영향을 바람직한 방향으로 유도하기 위한 미국의 다각적 노력

- 미국 교육학회(AERA), 심리학회(APA), 전미교육측정위원회(NCME), 국가연구위원회(NRC) 등이 연합하여 고부담 시험의 사용을 위한 전문적 지침을 개발함
- 미국교육학회는 이러한 지침 중 고부담 시험 프로그램의 시행을 위해 충족되어야 하는 12개의 조건을 추출하여 발표함(AERA, 2000)
- ✓ 이러한 전문적인 표준지침은 고부담 시험의 한계를 극복하고 이점을 살리는데 유용한 틀을 제공해 준다고 볼 수 있다.

12개의 전문적 표준지침	
1	매우 중요한 결정을 위한 유일한 준거로 사용될 시험으로부터 학생을 보호해야 한다.
2	충분한 자원과 학습기회가 제공되지 않았다면 학생과 교사들이 새로운 성취기준을 충족하지 못하였다 해도 제재를 가해서는 안된다.
3	시험의 타당도를 시험의 목적에 맞게 조정하여 활용하여야 한다.

4	시험 프로그램은 시험의 부작용 효과를 충분히 밝혀야 한다.
5	시험은 교육과정과 연계되어야 하며 교육과정 중에서 가장 쉬운 부분으로 한정해서는 안 된다.
6	시험 그 자체의 타당도뿐만 아니라 합격점수와 성취수준의 타당도도 함께 분석되어야 한다.
7	고부담 시험에 실패한 학생들에게는 의미 있는 교정기회가 제공되어야 하며, 그것이 모의시험을 반복하여 연습하는 이상이 되어야 한다.
8	영어가 모국어가 아닌 학생을 평가할 때는 시험에 사용되는 언어가 내용 지식의 평가에 지장을 주지 않도록 특별히 편의가 제공되어야 한다.
9	장애를 가진 학생에게도 시험의 형식에 의해 방해받지 않고 시험의 대상이 되는 내용의 숙달도를 보여줄 수 있도록 조치가 취해져야 한다.
10	영어학습자나 장애아를 배제하기 위한 분명한 규칙을 만들어 학교나 교육구 및 주정부가 점수를 올리기 위해 몇 학생을 제외시키지 않도록 해야 한다.
11	시험의 결과가 의도한 목적을 위해 사용되기 위해서는 충분히 신뢰할 수 있어야 한다.
12	연구 프로그램을 만들어 고부담 시험 프로그램의 의도한 결과나 의도하지 않은 역효과를 평가해야 한다.

□ 한국교육을 위한 시사점

● 지금까지 한국의 고부담 시험은 주로 수험자인 학생에게 더욱 중대한 영향을 미쳤다고 볼 수 있다. 하지만 최근 국가수준의 학업성취도 평가가 실시되고 수능시험의 성적이 학교나 지역별로 공개되면서부터 시험의 결과에 대해 학교나 교육청에서 민감하게 받아들이고 있는 상황이다. 아직 시험 성적의 결과에 따른 보상이나 처벌이 미국처럼 주어지지 않고 있지만, 미국에서 수십 년간에 걸쳐 나타난 고부담 시험의 폐해가 한국에서는 바로 나타나고 있음을 알 수 있다. 성적 조작 사건이 그 대표적인 사례이다.

● 미국의 고부담 시험 정책에서 나타난 시사점을 통해 한국의 고부담 시험이나 평가의 정책이 나아가야 할 바람직한 방향 검토

- ① 국가와 교육청은 학생들이 성취해야 할 최소 기준의 목표를 성취기준의 형식으로 분명하게 제시해 줄 필요가 있다. 미국은 철저하게 분권화된 교육과정과 내용체계를 가지고 있는 나라라서 전국 공통의 학력기준으로 바꾸기 위해 많은 노력을 기울이고 있다. 그에 비해 한국은 국가 수준의 획일화된 교육과정을 도입하고 있기 때문에 상대적으로 최소 기준의 학습 표준을 설정하는 것이 어렵지 않을 것이다. 이와 같은 표준이 마련되어 있고, 이를 중심으로 학업성취도 평가가 실시되는 것이 바람직하다고 생각된다.
- ② 시험의 내용과 방법이 변화되지 않으면 안 될 것이다. 선발기능이 강한 시험이니 만큼, 시험의 객관성과 공정성을 우선으로 고려하지 않으면 안 되는 안타까운 현실을 어느 정도 인정하지만, 시험의 내용과 방법을 미래지향적으로 변화시키지 않으면 미래의 국가 교육경쟁력을 높이는 어려울 것이다. 미국에서 글쓰기 과목이 교육과정에 추가되고 평가되고 있는 것처럼 평가의 내용이 개선되는 정도로 수업과 교육과정의 개선이 나타나고 있음을 확인할 수 있다. 이는 아직도 구태의연한 평가 모습을 벗어나지 못하고 있는 우리에게도 시사하는 바가 크다.
- ③ 평가의 주체에 있어 지나친 국가 주도의 외적 시험에 의존하는 경향을 탈피하지 않으면 안 될 것이다. 국가 관리 하에 시행되는 시험은 효율성과 객관성 및 공정성만을 중요한 요소로 생각하고, 교육의 본질을 외면하고 있다. 지역 교육청이나 단위 학교의 실정을 고

려한 종합적인 목표 접근이 필요하며, 미국처럼 우리도 고부담 시험에 대한 전문적인 표준지침을 만들어서 시험이 교육적 효과를 제고할 수 있도록 노력하여야 한다.

- ④ 지식기반 사회에 걸맞은 인재를 양성하기 위하여 새로운 평가체제를 고안해야 한다. 이것의 출발은 고부담 시험의 비중을 줄이는 것이며, 교육적 평가, 심층적 학습평가, 참평가 등의 대안적 평가방법에서 그 해법을 찾을 수 있을 것이다.

V. 고부담 시험 시대의 사회과교육: 학교 책무성의 신화, 학력, 그리고 교사 전문성의 관계(조대훈, 2011)

□ 미국과 한국의 고부담 시험의 제도 간의 유사점과 차이점

유 사 점	<ul style="list-style-type: none"> • 학생 개개인 및 개별 단위 학교의 학업성취 수준을 국가가 구체적으로 파악할 의무가 있다. • 고부담 시험은 불리한 교육환경에 처한 학교와 학생들의 학력을 끌어올리려는 목적을 가진다. • 교육정책 결정자들은 고부담 시험의 결과를 경쟁의 원리에 입각한 학교개혁의 효과적인 도구로 인식하는 경향이 있다.
차 이 점	<ul style="list-style-type: none"> • 각 나라마다 특징적인 학교 구조와 문화의 차이로 인해 고부담 시험에서 사용되는 핵심개념들은 서로 다른 의미를 내포할 수 있다. • 미국의 고부담 시험은 유급제도와 졸업 인증 제도와 불가분의 관계를 지니며, 이는 학생들에게 고부담을 전가하는 요인으로 작용하지만, 진급이나 졸업에 영향을 미치지 않는 한국의 국가 수준 학업성취도 평가는 실질적으로 학생들에게 고부담으로 인식되지 않는다. • 고부담 시험 결과의 활용 정도: 미국의 고부담 시험 결과는 학교 및 교사에 대한 성과급은 물론이고, 교원의 해임이나 학교 폐쇄라는 강력한 제재 조치를 수반하지만, 한국의 고부담 시험은 이제 막 교원 성과급에 부분적으로 반영되는 수준에 머물고 있다.

1. 고부담 시험의 주요 주장에 대한 비판적 검토

- 미국의 NCLB Act에서 발견할 수 있는 고부담 시험 타당성의 유일한 근거는 '책무성(accountability)' 개념이다. 전국의 모든 학생들의 현재 학력 수준 및 과거 대비 학력 향상 정도를 교육 당국이 매년 파악하는 것은 교육적으로 옳은 일이며, 이를 가능하게 해주는 유일한 도구가 바로 고부담 시험이라고 말한다. 하지만 이 같은 합리화는 고부담 표준화 시험이 가지는 '정당성'에 대한 근거일 뿐, *평가 도구로서의 '타당성'을 입증하지는 못한다.*

● 고부담 시험의 타당성 검토(사회과교육의 목표와 관련하여)

- ① 미국과 한국의 고부담 시험은 특정 교과 영역을 전제로 선택형이나 단답형 중심의 문항 구조를 보이는데, 1회의 표준화 시험만으론 비판적/창의적 사고 기능, 탐구의 태도, 다문화적 감수성 등의 성취 정도를 평가할 수가 없다. 또한 한국과 달리 미국은 주별로 학업성취도 평가 도구의 내용과 방식 자체가 다르기 때문에 난이도 차이나 기술적인 문제들로 인해 시험점수가 허위적인 점수로 변질되는 상황도 발생하여 평가 도구로서의 고부담 시험이 지니는 타당성을 약화시킨다.
- ② 시험은 긍정적·부정적 모습을 모두 지니고 있어 어떠한 시험도 그 자체로 교육적으로 선한 행위로 정당화될 수는 없다.
- ③ 학업성취도 평가 점수에 영향을 미치는 다양한 요인이 있다는 것을 알면서도 그 점수의 원인을 상당부분 교사라는 단일 요인의 영향으로 환원하여 설명하는 방식 역시 고부담

시험을 단위 학교 및 교사의 상벌에 활용하는 정책의 논리적 타당성을 근본적으로 흔들
어 놓는다.

- 과거 축적된 다양한 연구 문헌의 결과들을 종합적으로 분석해보면 고부담 시험은 사회과
교실 안의 학생들 학업성취도 향상에 그다지 기여하지 못하였고, 교육적으로 불리한 위치에
있는 학생들의 전반적인 학력에 특히 부정적인 영향을 미쳤으며, 이전보다 교육적 불평등을
더욱 심화시키는 결과를 가져왔다. 즉, 고부담 시험은 학생들의 중도탈락률을 높이는 요인
중 하나이며 성적향상을 근거로 학교 책무성의 성공 사례라고 치켜세운 몇몇 모범 사례들은
일종의 조작된 '환상(illusion)'에 가깝다는 것이다.

- 관련 연구들은 고부담 시험이 궁극적으로 교사 전문성의 핵심인 '주체적이고 반성적인 실
천가로서의 교수 능력'을 현저히 약화시킬 수 있다는 점을 공통적으로 지적한다(Au, 2007;
Crocco & Costigan, 2007; Daniels, 2002; Gerwin & Visone, 2006; Grant, 2001, 2006;
McNeil, 2000; Sirotnik, 2002). 또한 교사 전문성을 약화시키는 고부담 시험 시대의 수업 환
경은 교사들의 사기저하와 교직이탈 현상을 가중시키는 결과를 낳는다. 그중에서 가장 큰
피해를 입은 교사집단이 가장 능력 있고 열성적인 교사들이다(Crocco & Costigan, 2007;
Yeager & Davis, 2005; McNeil, 2002). 이처럼 교수의 자율성에 침해가 일어날 때 나타나게
되는 치명적 문제점은 교사의 '탈숙련화(deskilling)'현상이다. 그리고 교사의 탈숙련화와 전
문성 약화는 학교 외부로부터 주어진 준거와 표준들에 의해 학교교육이 관리·평가되는 것을
교육적으로 바람직하다고 여기게 만든다. 하지만 반대로 고부담 시험의 압력 속에서도 지혜
롭게 적응하여 창의성과 고급 사고력을 지향하는 야심찬 수업을 하는 교사들도 존재하는데
(Bolgatz, 2006; Gradwell, 2006; Yeager & Davis, 2005) 이는 변화의 동인으로서 교사가 가
지는 힘을 잘 보여주는 것이고, 고부담 정책의 효과를 최종적으로 결정하는 것은 현장 교사
들이라는 점을 말해주는 것이다.

2. 지난 20여 년 동안 축적된 미국의 고부담 시험 관련 경험연구 문헌에서 얻은 결론

- 한국과 미국의 고부담 시험 체제는 상당히 다른 맥락에서 전개되고 있으며, 핵심 개념들
(예: 표준화와 통제)은 두 나라 사이에서 서로 다른 의미로 해석되는 경향이 있었다.
- 현행 고부담 시험은 미국 교육부가 제시한 '수월성' 또는 '수준 높은 성취표준'을 측정하
는 평가도구로서의 기술적·논리적·직업윤리적 타당성을 모두 결여하고 있었다.
- 미국 교육 당국이 주장한 고부담 시험 성적과 책무성의 상관관계는 허위적임이 드러났다.
- 고부담 시험은 전문성을 갖춘 훌륭한 교사들을 지원하고 고무시키는 기능을 하는 것이
아니라, 오히려 교사 전문성을 약화시키고 탈숙련화를 가속화시키는 평가체제였다. 하지만
고부담 시험이 전적으로 교사 전문성을 결정하는 것은 아니었으며, 일부 현장 교사들은 영
향력 있는 교육적 변화의 동인으로서 새로운 교육 환경에 적응·타협하면서 고부담 시험의
의미를 능동적으로 해석해 나가기도 하였다.

3. 한국의 국가수준 학업성취도평가에 시사하는 점

□ 점수로 환원되는 '좋은 학교'의 의미

- 학업성취도 평가의 방식인 표준화 시험은 한 나라의 교육의 질을 관리하는데 필수적인
정보를 제공한다. 고부담 시험의 장점은 충분히 인정하지만 우리가 문제 삼는 것은 표준화

시험으로서의 고부담 시험 자체가 아니라, 학교 교육의 질을 높인다는 목적으로 그 진단 결과를 활용하는 방식인 것이다.

- 앞에서 미국 사회에서 시험 평균 점수의 적정연간향상도(AYP)로 대변되는 학업성취도의 의미가 대중을 기만하고 바람직한 교육의 의미를 쉽게 왜곡할 수 있다는 점을 살펴보았다. 한국에서도 이미 이와 같은 현상이 벌어지고 있다. 학생들의 학업성취도에 대한 국민의 알 권리도 중요하지만, 그보다 더 중요한 교육 당국의 과제는 '공개된 수치가 어떠한 교육적 의미를 담고 있는지'를 시민들이 제대로 이해할 수 있도록 도와주는 일이다. 교육정책결정자들은 관련 통계 수치가 표방하는 의미에 대해 의도적으로 침묵하고 있는 듯하며, 이러한 방관은 시민들이 고부담 시험의 점수를 좋은 학교와 나쁜 학교를 구분하는 판단 준거로 활용하도록 유도하고 있다.

□ 학력 논쟁 및 교육의 위기 담론의 신화적 허구성

- 미국의 경우 1950년대 구소련의 스푸트니크 인공위성 발사를 배경으로 하는 학문중심의 교육과정의 등장과, 1980년대 대일 무역적자 심화와 국가 간 학력진단평가에서의 저조한 성적을 근거로 당시 상황을 교육의 위기 상황으로 규정한 '위기의 국가'보고서의 발간 등은 21세기 학교 책무성과 고부담 시험의 담론과 여러 가지 면에서 유사한 접근을 보인다. 이 모든 개혁정책들은 공통적으로 학력의 위기를 강조함과 동시에, 그 학력이 한 국가의 경쟁력을 잘 대변해 준다는 믿음을 반영하고 있다.

- 하지만 이러한 믿음은 타당한 근거를 결여한 교육신화에 가깝다. 학력고사의 평균 점수와 국가 경쟁력 사이의 관계를 인과관계로 설정하는 것은 논리적 비약의 위험을 안고 있다. 그 관계를 검증할 현실적 방법이 존재하지 않기 때문이다.

- 교육사회학의 연구 성과가 이미 말해주고 있듯이, 단위 학교의 학업성취 수준에는 교사 요인, 학교장의 리더십 요인, 교육과정 요인 이외에도 지역사회의 교육 수준과 경제 수준, 그리고 소수민족의 문화적 요인 등 많은 부분이 복합적으로 영향을 미칠 수 있다. 따라서 한국과 미국의 교육 당국이 낮은 학업성취도 점수에 대한 원인을 학교장과 교사가 엄격한 관리 감독을 하지 않은 탓으로 돌리는 것은 마찬가지로의 논리적 비약을 범하는 것이다.

□ 효율성의 이데올로기와 교사 전문성

- 표준화와 고부담 시험을 양 축으로 하는 학교 책무성 운동은 미국 역사상 그 어느 때보다 강력한 중앙집권적 교육 정책을 상징하기도 하지만, 다른 한편으로 20세기 전반 '과학'의 이름으로 학교 교육의 모든 것을 풀어나가려고 했던 효율성(efficiency) 이데올로기의 화려한 부활로도 해석할 수 있다(Kliebard, 1995).

- 앞서 언급했듯이, 효율성의 이데올로기에 의해 지배받는 학교 책무성 개념은 교사의 탈숙련화(deskilling)를 조장하고, 이는 다시 교사 전문성의 핵심인 교실 수업의 능력을 감퇴시키는 결과를 가져온다. 수치화되고 측정할 수 있는 것만이 교육적인 증거로 간주되는 환경에서 가르친다는 것에 자부심을 느끼는 교사는 아마 존재하지 않을 것이다.

VI. 서울지역 학교교육 실태조사 분석 결과(전교조 서울지부, 2010.06)

※ 2010년 7월 교육학 심포지엄으로 열린 '평가를 평가한다'에 실린 자료로 학교현장을 중심으로 조사·분석이 이루어졌다.

실태 조사 방법

- 조사 대상 : 서울지역 초, 중, 고등학교(급별, 설립자별로 전체학교의 약 20% 수준의 학교 수 조사)
- 조사 기간 : 2010. 5월(약 한달 간 진행)
- 조사 방법 : 전교조 서울지부 주관으로 온라인/오프라인 설문 조사 병행

참여현황

- 총 256개교 : 초등학교 106개교, 중학교 76개교, 고등학교 74개교(공립교 38개교, 사립교 36개교)

1. 초등학교 실태조사 결과 분석

- 일제고사 실시 이후 성적 향상에 절대다수 학교에서 신경을 쓰고 있었으며(약 95%), 일제고사 이전 이후에 별다른 차이를 못 느끼는 경우는 약 3%에 불과한 것으로 조사됨
- ✓ 중학교와 비교해 볼 때, 상대적으로 성적 압박이 더 심한 것으로 조사되어 그나마 자유로웠던 초등교육이 일제고사 이후 성적 위주, 입시 경쟁의 영향을 더 받고 있음을 알 수 있음

4. 일제고사 실시 이후 학교 현장은 어떤 변화가 있다고 생각하십니까?

- ① 성적 향상을 위한 압박이 점점 커지고 있다. (76개교 / 71.7%)
- ② 성적 향상에 신경을 쓰는 편이지만 압박이 심한 편은 아니다. (25개교 / 23.58%)
- ③ 일제고사 실시 이전이나 이후나 별 차이가 없다. (3개교 / 2.83%)
- ④ 기타 (1개교 / 0.94%)



- 일제고사 이후 부진아 지도 강화(약 76%)와 성적 위주 교육활동 강제(약 27%), 시험대비 문제풀이식 수업 강요(약 28%) 등이 나타나고 있고, 이로 인해 정상적인 교육활동에 부정적 영향을 미치는 것으로 조사됨(약 72% 응답)

5. 일제고사 이후 학교 현장에서 벌어지는 파행 사례가 있다면 어떤 것들입니까? (해당하는 것을 모두 선택)

- ① 방학 중 보충 수업 실시(11개교 / 10.38%)
- ② 시험 대비 문제풀이식 수업 강제(30개교 / 28.3%)
- ③ 교과 보충수업 전면화(12개교 / 11.32%)
- ④ 부진아 지도 강화(81개교 / 76.42%)
- ⑤ 성적 위주 교육활동 강제(29개교 / 27.36%)
- ⑥ 아침이나 야간 자율학습 강제(4개교 / 3.77%)
- ⑦ 기타 (4개교 / 3.77%)



2. 중학교 실태조사 결과 분석

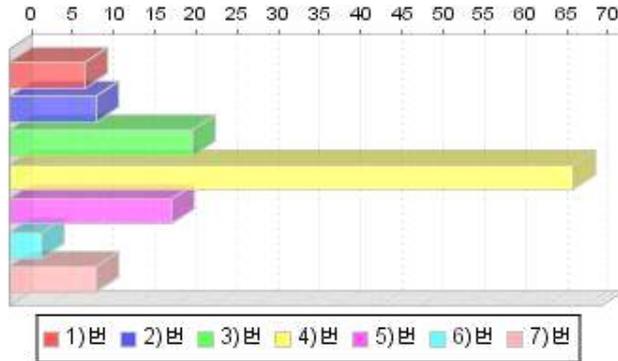
● 중학교 역시 절대다수 학교에서 성적 향상에 신경을 쓰고 있었으며(약 87%), 일제고사 이전 이후에 별다른 차이를 못 느끼는 경우는 약 11%에 불과한 것으로 조사되어 일제고사로 인한 학교교육 왜곡 정도를 가늠할 수 있음



● 중학교 역시 일제고사 이후 부진아지도 강화(약 68%)와 교과보충수업 전면화(약 22%), 성적위주 교육활동 강제(약 20%) 등으로 인해 정상적인 교육활동에 부정적인 영향을 미치는 것으로 조사됨(약 64% 응답)

5. 일제고사 이후 학교 현장에서 벌어지는 파행 사례가 있다면 어떤 것들입니까? (해당하는 것을 모두 선택)

- ① 방학 중 보충 수업 실시(7개교 / 9.21%)
- ② 시험 대비 문제풀이식 수업 강제(8개교 / 10.53%)
- ③ 교과 보충수업 전면화(17개교 / 22.37%)
- ④ 부진아 지도 강화(52개교 / 68.42%)
- ⑤ 성적 위주 교육활동 강제(15개교 / 19.74%)
- ⑥ 아침이나 야간 자율학습 강제(3개교 / 3.95%)
- ⑦ 기타 (8개교 / 10.53%)



7. 학교에서 실시하는 부진아 지도나 교과 보충수업으로 인해 교사들의 정상적인 수업이나 교육활동에 어느 정도 영향을 미친다고 생각하십니까?

- ① 매우 부정적인 영향을 주고 있다.(17개교 / 22.37%)
- ② 부정적인 영향을 주는 편이다.(32개교 / 42.11%)
- ③ 그저 그렇다. (12개교 / 15.79%)
- ④ 별로 부정적인 영향은 없는 편이다. (9개교 / 11.84%)
- ⑤ 부정적인 영향은 전혀 없다. (2개교 / 2.63%)

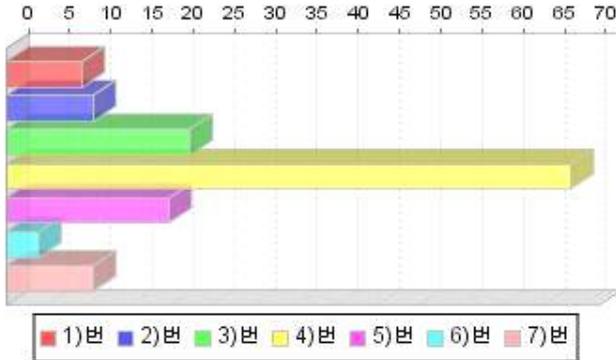


● 특히 방과후 활동이 교과보충수업의 형태로 바뀌고 있고, 1교시 수업을 오전 8시 30분 전후에 시작하는 학교(D여중, D중)는 물론 야간자율학습(H중, Y중, D중 등)을 진행하는 학교 까지 등장하는 등 학교교육의 비정상적인 운영이 확산되고 있는 모습을 보이고 있음. 특히 야간자율학습은 대부분 사립중학교에서 시행되고 있어 사립학교의 학사파행이 더욱 심각한 모습으로 나타나고 있음.

2. 다음 중 귀 학교에서 실시하고 있는 것들을 모두 골라주십시오.

- ① 아침 자율학습을 영어듣기 등 30분 이상 진행 (8개교 / 10.53%)
- ② 1교시 정규수업을 8시 30분 전후로 실시 (2개교 / 2.63%)
- ③ 방과후 활동을 교과 보충수업으로 실시 (41개교 / 53.95%)

- ④ 교과 보충수업을 2~3개 과목을 묶어서 진행 (20개교 / 26.32%)
- ⑤ 야간자율학습 실시 (9개교 / 11.84%)
- ⑥ 주말 또는 놀토에도 교과 보충수업을 실시 (6개교 / 7.89%)
- ⑥ 아침이나 야간 자율학습 강제(3개교 / 3.95%)
- ⑦ 기타 (8개교 / 10.53%)



VII. 일제고사 파행 전국 사례 요약(전교조, 2011.07)

지역	파행사례
경기	<ul style="list-style-type: none"> • “2011 수원 학력향상 지원 계획”을 자체 제작하여 지침으로 관내 모든 초·중학교에 시행 • 지침의 주요내용: 2008, 2009, 2010년도 기출문제 분석 및 지도, 학생보상제도 실시, 2010년 시험방법과 동일한 모의시험 연습, 지침 이행여부 점검하여 성적향상 우수학교 인센티브 제공” • 일제고사 대비비용 지침으로 관내 대부분의 초등학교에서 0교시, 문제풀이식 수업 등의 파행적인 교육과정 운영 • 학업성취도 평가 대비 문제집을 학교에서 제작하여 정규수업 시간에 교과서 진도를 중단하고 문제풀이만 하고 있음
경남	<ul style="list-style-type: none"> • 0교시: 도내 88개 초등학교 중 25개교, 55개 중학교 중 15개교 실시 • 3학년 국·영·수 위주 8교시 방과후수업. 10교시까지 시험 준비 학습 • 3학년 수업시간에 문제풀이, 일제고사 대비 3학년 모의고사 실시
경북	<ul style="list-style-type: none"> • 학생들이 일정 점수에 이르면 무조건 2~3만원에 상당하는 상품권을 시상하도록 하거나, 학교 혹은 학년 평균이 몇 점 이상 넘으면 모든 학생에게 3~5만원에 상당하는 상품권 시상 • 일부 학교는 6학년 학생뿐만이 아니라 전교생이 100% 야간 수업에 참여하는 학교도 있음 • ‘학력중점학교’(2010년 일제고사 기준 기초미달률 5%이상인 초등학교)의 대부분은 10교시(18:00분)까지 수업을 하는 학교가 많았으며, 밤 9시 30분까지 야간 수업을 하는 학교도 다수 • ‘학력 향상을 위한 컨설팅단’이라는 이름으로 교육장과 장학사들이 각급 학교를 직접 순시 • 일요일과 토요일에 등교하는 학교도 있음 • 학력중점학교의 경우 학생 인원수와 상관없이 수천만 원의 지원금이 배정 • 일부 기초자치단체의 경우 ‘국·영·수 중심의 학력향상을 위한 보충수업’을 진행할 것을 조건으로 수천만 원의 예산 지원 • 도교육청의 방과후 학교 지원비, 튼튼교실 지원비, 튼튼캠프 지원비, 기초부진학생 지도비(부진아 1인당 20만원) 등 각종 명목의 ‘학력 향상 지원비’를 무차별적으로 배정 • 돈을 쓰기 위해 밤 9시까지, 토요일, 일요일까지 수업을 해야 하는 비극이 벌어짐

	<ul style="list-style-type: none"> • '2011년 학력 마일리지제 운영 계획'을 발표 • 일제고사 준비에 실적이 있는 교사를 대상으로 10만원 상당의 상품권과 해외 연수 기회를 우선 제공
대구	<ul style="list-style-type: none"> • 2, 3, 4학년 7교시 주 5회 60분씩 / 5, 6학년 7교시 주 5회 80분씩 공부 • 6학년은 아침시간에 운동장에 노는 것 금지, 계발활동도 하지 않음. • 학력 관련 평가뿐만 아니라 각종 대회 입상자에게 현금으로 상금 지급 • A4 용지 4,000쪽 분량의 문제집 제작 활용(아침자습, 점심시간, 방과후, 과제로도 시간이 모자라 교과 수업시간까지 침해), 예산 확보의 문제도 따름 • 동부 교육지원청에서 제공한 문제지를 관내 학교에서 0교시와 방과후 및 수업시간을 활용해 문제 풀이 진행
대전	<ul style="list-style-type: none"> • 관내 고등학교에 대전학력신장 TF팀이 개발한 2011년 고2 학업성취도 국어, 영어, 수학 형성평가 문제지 공문을 내려 보내고 그 활용 결과를 보고 하라고 지시(7월1일까지 총8차례의 형성평가 문제지를 제공 계획) • 중학교의 경우 전체 87개교 중 무려 47곳에서(54%) 정규수업 전 0교시를 운영하여 전국 최고 • 80% 이상 학생을 강제로 참여시키고 있는 학교 비율이 동부는 무려 90.5%, 서부도 65.4%에 달했다. 대전 전체로 보면 10곳 중 7곳 이상이(76.6%) 사실상 0교시에 강제로 참여시킴 • 학업성취도평가에서 좋은 성과를 내기 위해 '학력 우수반'을 선정해 상품권을 시상하거나, 초등 6학년이나 중학교 3학년의 경우 성적에 따른 우열반을 편성하는 등 비교육적 파행 사례 속출 • 70억여 원이라는 막대한 돈이 사실상 0교시, 강제 보충수업, 문제풀이용 책자 구입 등 '교육과정 파행 후원금'으로 이용 • '학업성취도 향상 프로그램 운영 지원'이라는 명목으로 학교 현장에 교부하고 있는 돈이 32억 2천 7백여만 원 • 학력평가 관련 예산 26억 원 • '교육복지우선지원사업학교' 지원 예산 43억여 원 중 시청보조금 6억 원은 '학력 신장'에만 쓰이고 있고, 나머지 37억여 원 중 상당 부분이 '교육 복지'와는 별 상관없는 '성적 올리기'에 사용
인천	<ul style="list-style-type: none"> • 초등학교 6학년 교과부진학생 수업시간은 33시간이 아닌 48시간!! 강제불법 전체 7교시와 맞물리면 50시간이 넘어 학생의 학습권, 건강권을 침해하고 있음. • 모 초교는 주당 2~3회씩 문제풀이를 했지만 평가가 임박한 1달 전 부터는 매일 진행하고 있음. • 7교시(또는 6교시) 시행을 내부결재를 얻어 시행한 학교는 한곳도 없었음. • 전체학생을 대상으로 한 7교시(또는 6교시) 수업을 마치 교과부진이나 학습부진아 지도한 것으로 하여 불법,편법적으로 '강사비'를 지급 • 국어, 수학, 사회, 과학 관련 부서가 개설되면서 방과후학교의 학원화 현상이 일어나고 있음. • 수학 1교과 수강료가 5만원으로 일반 학원 전과목대비 12~ 15만원보다 훨씬 비싼 것으로 조사됨. • 부교재 선정 시 교재선정위원회, 학교운영위원회 심의조차 받지 않고 부교재를 구입한 경우가 있었음. • 특활, 재량, 예체능 교과 시간에 시험지 풀이하는 학교가 많음. • 실과, 도덕 등 주요교과가 아닌 시간에 6월 3주차부터 1일 1시간 문제풀이

	<ul style="list-style-type: none"> 수학 특별보충 수업 편성예산으로 6학년 전아동 부교재비용(1인 1만원, 총120만원), 6학년 담임 보충수업비(15시간 30만원) 지급
제주	<ul style="list-style-type: none"> 도 일제고사 대비 문제풀이 수업 진행 휴무토요일에 등교시켜 일제고사 대비 문제풀이 0교시 및 방과후 오후 5시까지 문제풀이 진행
충남	<ul style="list-style-type: none"> 2011년: 4.13(90일전) 2010년 성취도문제 시험 → 5.24(50일전) 2009년 성취도문제 시험 → 6.13(30일전) 토, 일요일 특별수업 → 7.11(1일전) 시험출정식(다집대회) 성취도평가시상 ✓ 개인: 1등 20,000, 2등 15,000, 3등 10,000원, 성적향상자 5,000원 문화상품권 시상 ✓ 학급: 최우수 30만원, 우수 20만원, 장려 10만원 ✓ 교사: 우수반 교사 50,000 / 기초미달 ZERO화 교과목 교사에게 인센티브(상품권 및 동계 스키 무료 참여) 열공증공 주말 캠프: 토요일요일 9:00~13:00 6학년 전체 예상문제 집중지도 1일 5시간 (15:20~20:00) 학업성취도평가 마무리 캠프 : 6학년 전원 6.27~7.8(10일간), 17:00~20:10(4시간) 예상문제 및 정리문제 학습 모의고사 실시: 5월부터 7회 이상 실시 토요일무일 전일제: 짜파게티반(상위권) 20명, 스파게티반(하위반) 20명 드림캠프 (16:50~20:55) 200명: 수준별 소수 정예 기출예상 문제집 '꿈이 영어 가는 우리' 발간 - 3학년 전체 학력중점학교 교감 회의: 충남교육청 담당자가 “7월 전까지 예체능 과목을 교체해서라도 학력을 향상시켜라.” “학급 담임에게 학습부진아 Zero화 서명을 시켜라.”라는 불법과 탈법을 조장하는 말을 서슴없이 요구
충북	<ul style="list-style-type: none"> 방과 후 학교에 참여하지 않는 학생들을 별도의 교실에 모아 두어 강제적으로 자율학습을 실시한 사례 강제적 특별 보충과정 운영(일과 전 1시간, 일과 후 2시간: 살인적 학습 부담 가중) 학생 및 학부모의 동의 과정도 없이 6학년 전체를 대상으로 강제 보충 실시 방과후학교 운영을 위해 점심시간 40분으로 단축 학교평가 및 학교 간 차등성과급 지급 지표상에 '방과 후 학교 운영 및 참여실적'을 명문화하여 교육과정 파행 조장

Ⅷ. 초등학교 중간·기말고사 폐지 찬반논란 및 그 이후(관련자료, 2011)

□ 초등학교 시험 폐지 바람직한가?(연합뉴스 2011.01.12.)

※ 2010년 말 광노현 서울시 교육감이 초등학교의 중간 및 기말 고사를 폐지하겠다고 밝힘
 - 평가권을 일선교사에게 주고 과제 수시평가를 통해 학생성적을 내도록 하라는 취지로, 교사의 자율권 확대를 위해 학급별 평가로의 전환을 추진하며, 이는 사교육을 없애는 가장 확실한 방법이라고 밝혔다.

● 폐지찬성론 “사교육을 줄일 수 있는 가장 확실한 방안이다”

- ✓ 교사마다 서로 다른 시험도구로 평가하기 때문에 사교육이 끼어들 여지가 줄어든다.
- ✓ 시험과 경쟁에 짓눌린 아이들에게 단순한 시험대비 공부가 아닌 전인교육을 시킬 수 있는 여지가 그만큼 많아져서 교육 본래의 목적을 달성하기 위한 출발점이 될 수 있다.

- ✓ 오효숙(서울시 교육청 초등교육정책과장): 학생들을 성적으로 줄 세우는 평가방식에서 벗어나 학생의 부족한 부분을 파악해 보완해주는 식의 대안을 마련하고 있으며, 교사에게 평가를 맡겼을 때 업무가 가중되지 않도록 하는 방안도 함께 고민 중이다.

● **폐지반대론 “학생의 성취도를 확인해볼 수 있는 객관적 준거가 사라진다”**

- ✓ 한국교원단체총연합회가 '11년 8~10일 초등학교원 445명을 대상으로 벌인 설문조사

Q1. 초등학교 중간기말고사는 폐지되어야 한다.

반대 (총 62.24 %)		찬성 (총 24.05 %)	
반대한다.	42.02%	찬성한다.	14.61%
적극 반대한다.	20.22%	적극 찬성한다.	9.44%

Q2. 시험 없이 수행평가만으로 학생의 실력을 평가하는 것이 가능하다.

불가능 (총 67.64 %)	가능 (총 21.35 %)
-----------------	----------------

Q3. 중간기말고사 폐지가 학력 저하 등 부작용을 초래할 것이다.

그렇다 (총 74.15 %)

- ✓ 권오남(서울대 수학교육학과 교수)
- 학년별 정기고사는 학습동기 유발의 효과가 있으며, 이를 통해 학생 자신의 성취도가 어느 정도인지 확인해볼 수 있는 기회를 제공한다. 또한 학부모에게는 자녀의 학업 성취도, 자녀의 학습 이해 수준을 알려줄 수 있는 객관적 준거가 될 수 있고, 교사에게는 학급을 운영하며 적용한 교수·학습 방법의 객관적 타당성을 검증하는 중요한 도구이기도 하다. 더욱이 수능이 중시되는 현실에서 초등학교 정기고사의 폐지는 큰 의미가 없다.
- 고등학교에서 국가 수준의 학업성취도 평가만을 볼 수 있도록 제한한 이후 학원 등에서 사설 모의고사를 보는 학생 수가 증가한 것처럼 초등학교에서의 중간기말고사 폐지는 학부모로 하여금 자녀의 객관적 학업능력 평가를 받게 하기 위해 오히려 더 사교육에 의존하게 할 가능성도 있다.

□ **사실상 월말고사 부활- 서울 초교들 부글부글(교육희망 2011.08.26)**

- 서울시교육청이 서울지역 591개 초등학교 6학년을 대상으로 '수시평가 실시 현황'을 조사한 결과 중간기말고사를 폐지하고 수시평가를 한 첫 학기인 2011년 1학기에 서울지역 초등학교의 29%가 20회 이상의 평가를 치른 것으로 나타났다.

- ✓ 평균 수시평가 횟수: 15.5회

• 20회~25회: 94개교 (16.0%)	• 25회 이상: 75개교 (12.7%)	• 40회~60회: 일부 존재
-------------------------	------------------------	------------------

- 서울H초 교사: “시교육청이 수시평가 예시문항에서 문제풀이 형태를 제시함에 따라 대부분의 초등학교가 월말고사 식으로 국어, 수학, 사회, 과학에 대해 반별, 학년별 고사를 실시했다”

- 서울지역 초등학교 교사들을 대상으로 조사한 서울시교육청 자료

- ✓ 개별 출제로 인한 업무과중 (49.4%)
- ✓ 잦은 평가 실시에 따른 수업 결손 (29.7%)
- ✓ 현행과 같은 수시평가 실시 찬성의견 (10.2%)

- 서울시교육청 교육과정과 담당자의 해명: “지난해보다 시험횟수가 늘어난 것은 맞지만 과목별 시험을 1회로 계산하다보니 더 많이 늘어난 것으로 착각할 수 있으며, 시교육청도 학

교의 어려움이 크다는 것을 잘 알고 있어 학교에 자율성을 주고 교원의 평가부담을 줄이는 방향으로 개선안을 마련 중이다.”

□ 서울 초등학교 수시평가 부담 줄인다.(연합뉴스, 2011.09.06.)

- '11년 1학기에서 드러난 문제점에 대한 개선으로 서울시교육청은 '초등 교과학습 평가시행 계획'을 발표
- ✓ 각 초등학교는 지필·수행평가를 실시할 때 평가 영역, 방법, 기준, 성적처리 방법, 결과 활용 등 구체적인 방법을 학교구성원의 심의를 거쳐 자율적으로 결정할 수 있게 됨
- ✓ 교사들은 수시평가 문제를 학교장에게 결재 받지 않아도 되고, 쪽지시험 등 간편한 형태로 대체할 수 있음
- 이 개선안이 시행된 '11년 2학기에 관련된 조사는 현재 나오지 않은 상태임

□ 참고자료1. Wales의 사례(2개의 Articles)

1. What happens when high stakes testing stops? The impact on teachers of compulsory national testing in science of 11 year olds in England and its abolition in Wales.(Collinsa, S., Reiss M. & Stobart G., 2010)

- 고부담 시험(11세, 14세 대상 과학시험 폐지)을 없애면 어떻게 될까?
 - ✓ 웨일즈는 11세와 14세를 대상으로 한 수학, 과학, 영어 국가시험을 폐지했으며, 2004년 웨일즈 국회는 1991년부터 도입된 국가시험제도를 폐지하고 이러한 결정을 사회적 실험 (social experiment)으로 간주하였다.
 - ✓ 영국에서는 11세의 경우 국가시험을 지속시켰으나 14세는 2008년부터 시험을 폐지하였다. 교사종합평가(summative teacher assessment)에 찬성하며, 다른 대안적 시험들과 함께 학생들을 평가하도록 하였다.
 - ✓ 국가시험을 지속하는 영국의 사례와 국가시험을 폐지한 웨일즈의 국가적 사례를 통해, 11학년 과학 교사들의 경험과 의견을 양적/질적 분석하였다.
- ※ 잉글랜드와 웨일즈에서 행한 600건의 전화조사와 74명의 관계자 인터뷰(8개 집단으로 구분)를 통해 양적, 질적 데이터를 모으고 분석해서 다면적 접근을 하였다.

England (참여자수)		Wales (참여자수)	
Y6 교사와 과학 코디네이터 (2집단)	25	Y6 교사와 과학 코디네이터 (2집단)	18
학교장	8	학교장	11
중등학교 과학교사	8	중등학교 과학교사	4

(양측 t검정 사용, 유의수준 5%)

※ 분석결과 요약(5가지)

- ① **과학을 가르치는 교사들이 직면하는 가장 큰 어려움:** 두 집단의 교사 모두 과학시수 부족과 과학수업 자료부족을 꼽았다. 영국의 교사들은 시험에 나오지 않는 내용들을 전략적으로 수업에서 제외시키는 자신의 모습을 발견하고, 그러한 수업에 대한 회의적 느낌을 언급하였다. 웨일즈 교사들은 같은 과학내용을 다루더라도 교사종합평가와 대안적 평가에 대한 압박이 있지만, 국가시험 압박에서 벗어나 더 많은 체험적 수업을 할 수 있음을 강조하였다.
- ② **수업에서 사용되는 과학 평가 전략들:** 국가시험이 없어졌을 뿐 웨일즈 교사들은 종합평가와 다른 대안평가를 활용하고자 지속적인 노력을 하고 있었다. 영국과 웨일즈 모두 학생들의 성적향상에 대한 정보를 제공하는데 있어 학생들의 assessed work를 검토하였고,

영국은 웨일즈와 달리 국가시험 기출문제나 과학실전문제집 등을 평가에 이용하였다.

- ③ **시험 준비와 다른 대안적 시험자료들:** 웨일즈 교사들은 교사에 의한 종합평가, 대안평가 등의 신뢰도와 정확성에 대해 일부 회의적인 반응을 보였고, 정확성을 갖춘 평가 마련에 대한 요구가 높았다.
- ④ **국가과학시험 폐지가 실제로 미치는 영향:** 교사의 직접평가 방법을 고안하고 향상시키는 것 외에도 대다수의 웨일즈 교사들은 과학수업전략의 변화에 대한 필요성을 주장하였는데, 개별학생의 지식수준을 판단하기 위해 학생에 대한 focused questioning을 바탕으로 한 일관된 관찰, 토론, 소집단 활동 등이 필요하였음을 이야기하였다. 대집단, 교사주도 과학수업이 더 이상 실제적이지 않음을 제안하며, 교사는 개념 맵핑 등을 통해 필요한 지식과 개념을 유도하고 학생들 스스로 디자인한 표 등을 통해 과학적 탐구의 가설, 결과, 예측 등을 기록할 수 있도록 유도하였다.
- ⑤ **과학에서 평가와 표준화와의 관계:** 두 집단 모두 국가과학시험이 학생들의 성취도에 대한 정확한 정보를 제공하지 못함을 강하게 동의하였다.

● 양국의 과학수업에 대한 정리

① 영국

- ✓ 영국의 교사들은 학생들이 더 높은 점수를 얻을 수 있도록 하기 위한 전략을 개발하고, 시험 준비를 위한 수업에 많은 시간을 쏟았으며, 그 과정에서 교사들은 교육과정을 축소시켜 과학에서 다룰 수 있는 여러 내용 중 지필고사를 통해 답을 할 수 있는 부분만 한정해서 가르치게 되었다.
- ✓ 영국이 웨일즈보다 과학시수가 더 많았지만 고부담 시험 준비에 수업시간을 활용함으로써 실제 과학수업시수가 많은 것은 아니었다.
- ✓ 과학교육과정을 실제생활 중심으로 바꾸더라도 고부담 시험의 영향은 수업을 더 지시적이고 대집단 수업으로 만드는 경향이 있었다.

② 웨일즈

- ✓ 웨일즈의 교사들은 다양한 시험자료의 사용 등 효과적인 교사평가 전략개발에 힘을 쓰고 있었으며, 학생들에게 종합판단(summative judgment)에 따른 평가를 제공하려고 하였다. 탐구 등을 비롯한 실제 과학 활동들이 과학수업의 주요 특징이 되었고, 국가시험준비에 대한 압박에서 서서히 벗어나 학생들의 과학 학습에 초점이 맞추어지게 되는 현상이 나타났다.
- ✓ 웨일즈 교사들은 부모들에게 자녀의 성취도를 알려줄 국가시험점수가 없을 때, 효과적인 의사소통 제도와 전략을 개발할 필요가 있음을 알게 되었다.

● 논의 및 제언

- ✓ 지역의 평가전문가와 교사들이 협력하여 교사종합평가가 학생들의 성취도를 정확하게 반영할 수 있도록 명료한 준거를 만드는데 노력을 해야 한다.
- ✓ 교사가 직접 쓴 개별성적표를 통해 학생들의 향상 Tracking, 핵심교과의 성취도 수준을 알 수 있게 하고, 향후 발달을 위한 목표를 제시하여야 한다.
- ✓ 결국, 이와 같은 성적보고체제와 의사소통체제가 개발되어야 한다.

2. Wales Eliminates National Exams for Many Students: Move breaks tradition with

England, which inspired U.S. policies.(Jeff Archer, 2006)

- 미국에서는 최근 몇 년간, 시험기반의 책무성제도가 교육과정 축소화, 교사판단의 평가절하, 학생들이 학습으로부터 소외되는 현상 등 심각한 문제를 야기 시킨다는 비판을 받아왔다.
- 하지만 웨일즈는 이와 달리 영국으로부터 평가제도에 있어 상당한 자율성을 얻어 표준화 평가를 덜 강조하기 시작하였다. 과거의 웨일즈는 고부담 시험의 압력이 상당한 국가 중 하나였지만, 이제는 14세의 학생들이 응용학습능력(applied learning)을 보이는 것을 강조하며, 유아교육은 다시 놀이를 강조하기 시작하였다.
- 보스턴 대학 교육학 교수인 Andy Hargreaves는 웨일즈가 제공하는 교육평가에 대한 대안적 미래를 주의 깊게 살펴보아야 한다고 하였다.
- **비판**
 - ✓ 일부 전문가들: 학생의 성적 향상에 대한 표준화된 측정 방법을 개발하지 않고는 학생과 부모, 교사 및 대중들에게 필요한 정보를 제공하기 힘들며, 교사들의 평가에 대한 부담도 커질 것이다.
 - ✓ David Hopkins(A former official at the Department of Education and Skills of the United Kingdom): 초등학교에서 중등학교 사이의 전환기는 아주 중요한 시기이므로 학생 성적에 대한 객관적 자료가 없다는 것은 아주 위험한 일이다.
- **웨일즈의 대처방안**
 - ✓ 표준화 시험을 없애는 대신, 국가수준의 평가지도를 바탕으로 교사들에게 평가를 개발하여 사용하도록 하고, 평가 결과를 매년 부모, 교육청 및 정부에 보고하도록 하고 있다.
 - ✓ 학생들은 자신이 해야 할 일에 대해 조직하고 분석하는 능력을 갖게 되었고 더 독립적인 존재가 되었으며 학습을 하는 방식도 아주 달라졌다.
 - ✓ 다양한 검사제도가 있을 뿐만 아니라 꾸준히 개발 중에 있으며, 영국과 마찬가지로 웨일즈도 정기적으로 학교탐방을 통해 장학지도를 하고 있다. 또한 외부 검사(outside check) 중 하나로 PISA평가에도 참여하고 있다.
 - ✓ 국가시험제도가 폐지된 상황에서 초등, 중등교사는 교육과정 및 교수방법의 연계 등을 고려하여 더 긴밀히 협조할 필요가 있다.

□ 참고자료2. 「OECD Education Working Papers No.58」(2011.04)

- Integrating Formative and Summative Assessment

● 평가의 개념

- ✓ **총괄평가(Summative assessment):** 교육 프로그램이 끝난 다음 학생의 수행능력을 당초 목표와 비교하여 종합적으로 평가하는 것으로 시험, 연말시험 등을 포함한다. 이는 상급학년으로의 진급, 증명(검정), 입학에 하는데 사용되며, 학습평가라고도 한다.
- ✓ **형성평가(Formative assessment):** 교수·학습이 진행되는 과정에서 학생의 부족한 부분을 점검하고 수업 방법을 개선시키기 위해 실시하는 평가로 학습을 위한 평가라고도 한다.
- *많은 교육자들과 교육평가 전문가들은 평가로부터 얻은 데이터를 활용하여 교실에서의 교수·학습 방법을 구축할 수 있도록 형성평가와 총괄평가를 통합하기를 원함*
- ✓ 국가수준의 표준화평가(총괄평가)는 교육체계, 개별학교와 학생집단들의 전반적인 성취를 감독하는데 유용한 데이터를 제공한다. 또 그 데이터는 교육정책 방향과 필요한 교과과

정, 재정의 분배, 적합한 학습전략 등의 결정을 하는데 도움을 준다.

- ✓ 또한 이 평가는 학업성취도 관리, 국가나 지역의 교육표준 강화에 도움이 되기도 한다.
- ✓ 하지만 이 큰 규모의 표준화된 평가에서 얻어진 데이터로 특정 학생이나 개별교실이 필요한 것을 반영하는 데에는 현실적 한계가 있다. 표준화된 평가로는 문제해결능력이나 협력학습능력 같은 복합적 능력을 판단하는데 어려움이 있으며, 그 결과가 학교로 전달 되는데도 몇 주가 걸리기 때문에 피드백의 적시성도 떨어진다. 이 외에도 교육과정의 협소화, 시험대상인 학생의 편파적 분류 등의 문제가 존재한다.
- ✓ 이러한 문제점 때문에 형성평가가 필요한데 이 둘의 결합에는 비용문제, 결합된 평가로 얻는 점수에 대한 신뢰성과 일반성 등의 문제가 여전히 남아있다.

● 하지만 교실 중심의 평가는 학교정책 결정자들에게도 유용한 정보를 제공해 줄 수 있기 때문에 현재 두 평가를 통합하는데 있어 중대한 기술적 장벽이 있음에도 불구하고 이 분야에 있어 많은 발전이 있었음

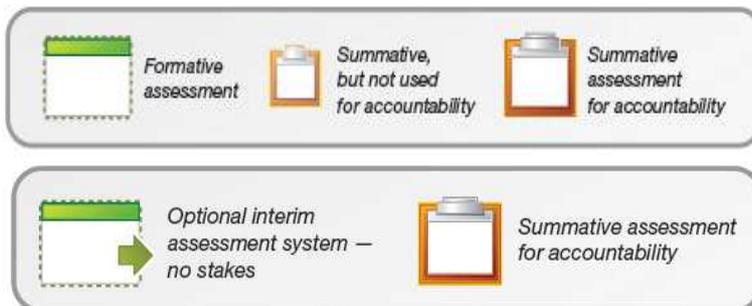
● 현재 추진 중인 연구개발 목적은 평가기술의 발전뿐만 아니라 교실 중심의 형성평가 강화에도 중점을 두고 있음. 보다 향상된 양상으로 형성평가와 총괄평가의 통합을 하려면 새로운 측정 기술, 교사 재교육 및 전문성 함양, 보다 나은 연구개발 등을 위한 투자가 요구됨

● 대규모의 표준화 평가를 통해 학생의 학습적 요구와 성장을 돕기 위한 노력

- ✓ 교사의 평가 역할 강화
- ✓ 교원평가의 강화
- ✓ 형성평가와 총괄평가의 질을 높이기 위해 인지과학의 발전 필요
- ✓ 커리큘럼 임베디드 혹은 수요자 중심의 평가 개발
- ✓ 전문적 기술 없이도 학생들이 어려워하는 특정학습을 식별할 수 있는 보완적 진단평가의 개발
- ✓ 전수보다 표집을 통해 주로 감독용으로 개발된 대규모 평가의 관리를 고려하는 것
- ✓ 테크놀로지 기반의 평가를 개발하여 그 장점을 취하는 것

□ 참고자료3. 「PARCC / SBAC」에 대한 브리프 (www.12center.org, 2011.07 update)

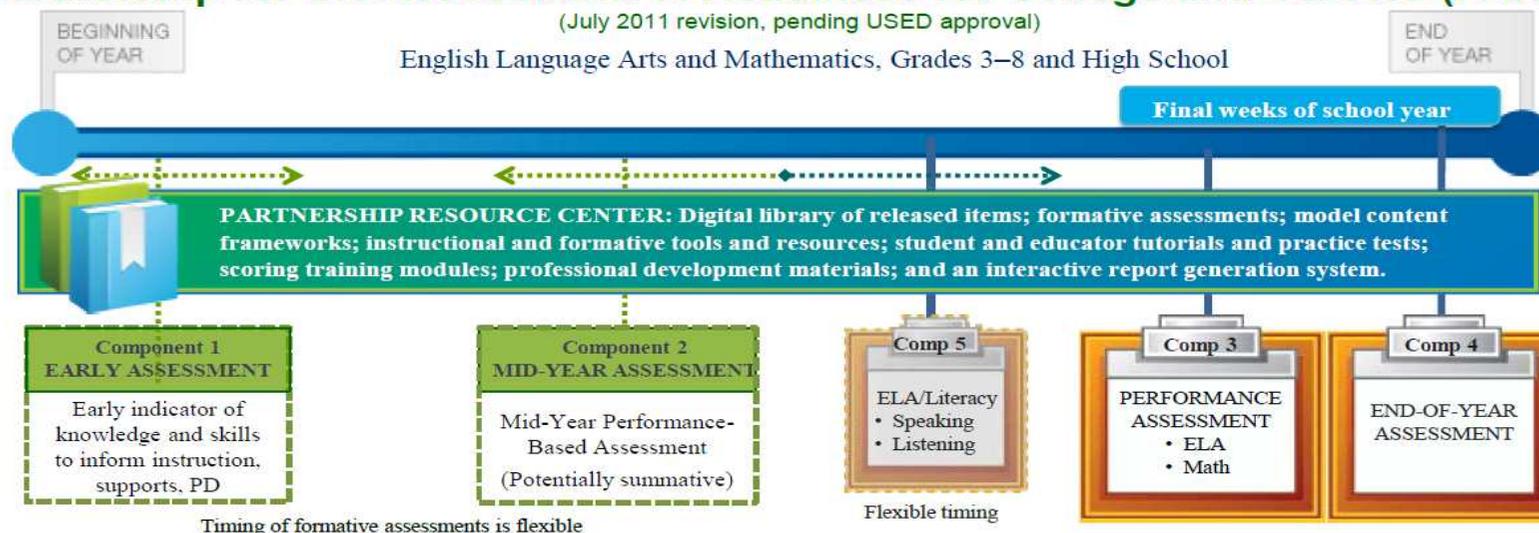
- Partnership for the Assessment of Readiness for College and Careers(PARCC)
 - ▶ 사이트: <http://PARCCOnline.org>
- The SMATER Balanced Assessment Consortium(SBAC) Design
 - ▶ 사이트: www.k12.wa.us/SMARTER



Partnership for the Assessment of Readiness for College and Careers (PARCC)

(July 2011 revision, pending USED approval)

English Language Arts and Mathematics, Grades 3–8 and High School



Description of Major PARCC Features

Optional Assessment Components:

Component 1: Formative early assessment designed to be an indicator of student knowledge and skills so that instruction, supports, and professional development can be tailored to student needs. (Use is optional.)

Component 2: Formative mid-year performance tasks that provide teachers and students with instructionally useful feedback and prepare them for the summative performance tasks to be given in Component 3. This component will focus on hard-to-measure standards. Teachers will be given an online score training tool to score these tasks and improve their understanding of the CCSS expectations. (Use is optional.)

Summative Assessment Components:

Component 3: Composed primarily of performance tasks and taken over several sessions/class periods, the ELA assessments will focus on writing effectively when analyzing text and the mathematics assessments will focus on math practices and solving multi-step problems. Scores returned within two weeks.

Component 4: These End-Of-Year comprehensive assessments will be computer-based and consist of innovative, machine-scorable item types.

High School: In mathematics, both traditional and integrated math sequences will be supported; in ELA, literacy skills in ELA, science, and social studies will be assessed, as defined in the CCSS.

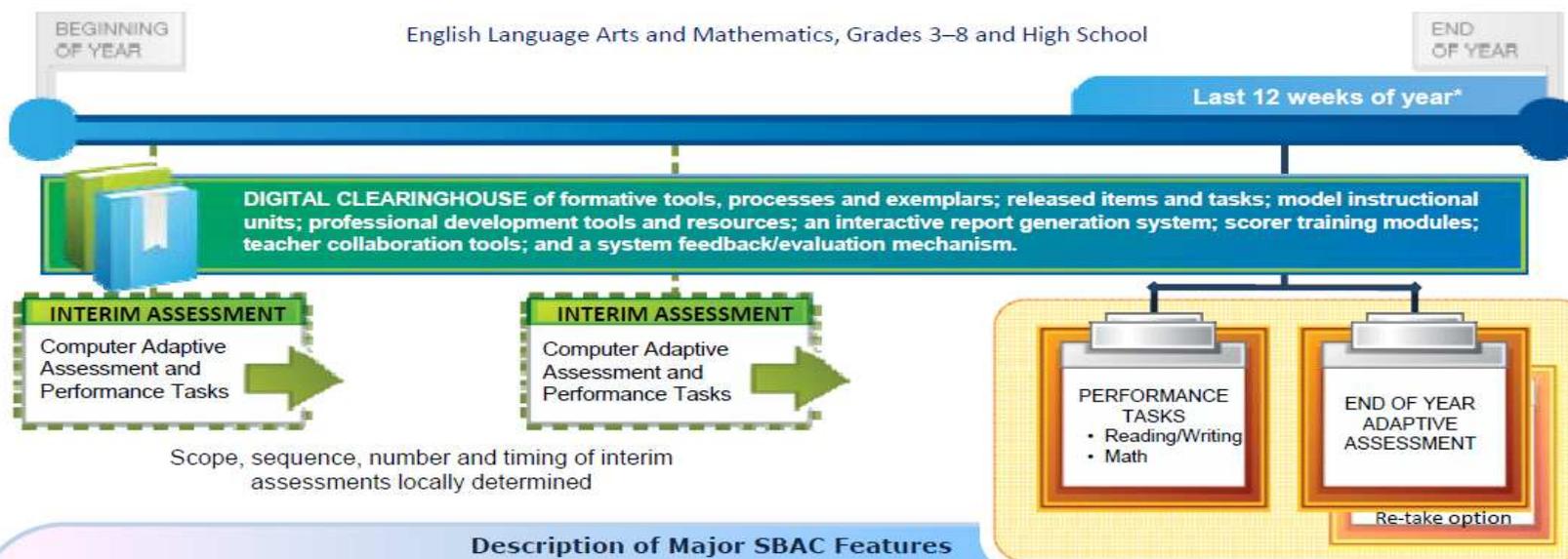
Component 5: (Required, not used for accountability) This component is being redefined to allow it to be given any time between mid-year and end-of-year.

Scoring: Scores for Components 3 and 4 will be combined for the student's annual accountability score.

Partnership Resource Center: A digital library of optional resources to support teachers and students, grades K–12, including formative assessments for students in grades K–2 and a Text Complexity Diagnostic Tool.

Implementation: Field testing in 2012–2014; operational by fall 2014.

The SMARTER Balanced Assessment Consortium (SBAC) Design



Description of Major SBAC Features

Interim assessments: These optional computer adaptive assessments will provide results on the same scale as the computer adaptive component of the summative assessment. This flexible system can assess either the full set of grade level standards or a smaller set of standards, based on local curricula, at a deeper level to identify individual student needs. Teachers will have access to items and student responses. Reports will link teachers to appropriate formative strategies and professional development resources. Available in 2012-13.

Summative assessments:

- **Performance tasks will be completed annually** (2 tasks in ELA and 2 in mathematics) during Consortium-defined testing windows.* Tasks will generally require two one-hour periods and involve student-initiated planning, management of information and ideas, interaction with other materials and/or people, and production of an extended response such as an oral presentation, exhibit, product development or extended written piece.
- **Computer adaptive assessments** will include approximately 40 to 65 questions per content area and be taken during the last 12 weeks* of the school year. They will include selected response, constructed response and technology-enhanced constructed response items. Additional administration opportunities will be available, as locally determined.

Scoring: A combination of machine and teacher scoring will be used. Scores for each summative component will be returned within two weeks and will be combined for the student's annual accountability score.

Digital Clearinghouse: A digital library of optional resources to support teachers and students, grades K–12.

Implementation: Field testing in 2013; Operational by January 2015.

* Time windows may be adjusted based on results from the research agenda and final implementation decisions.