

# 검사 조건에 따른 저성취 · 학습장애 학생의 점수 변화: 검사 조정의 필요성과 타당성을 중심으로

최종근(崔種根)\* · 김동일(金東一)\*\*

## 논문 요약

본 연구에서는 저성취(LA)나 학습장애(LD) 학생을 위한 검사 조정의 필요성과 타당성을 조사하였다. 저성취학생, 학습장애, 일반학생을 대상으로 각각 네 가지 검사 조건(표준, 시간연장, 대독(녹음테이프), 시간연장+대독(녹음테이프))에서 2차례 검사를 실시하여 검사 조건의 변화 즉, 검사 조정에 따른 점수 변화를 분석하였고, 이를 통해 각각의 검사 조정의 효과를 추정해내었다. 특히, 학습장애 학생의 피검 행동에 비추어 볼 때 학습장애 학생에게 검사 조정이 필요하다고 할 수 있는지를 논의하였다.

■ 주요어 : 검사 조정, 변별적 점수 향상, 저성취, 학습장애

## I. 연구의 필요성 및 목적

검사 조정(test accommodations)이란 '장애를 가진 학생이 평가에 참여할 수 있도록 하기 위해 검사(measures)를 적응하거나 수정(modification)하는 것(Salvia & Ysseldyke, 1995)'을 말한다. Friend와 Bursuck(1996; Burns, 1998에서 재인용)은 IDEA(미국장애인교육법) 조항을

\* 서울대학교 시간강사

\*\* 서울대학교 교육학과 교수

인용하면서 '장애를 가진 학생을 평가할 때 가장 중요한 것은 그 검사 결과가 그들의 장애가 아니라, 그들의 지식과 기술을 제대로 반영하도록 하는 데 있다'고 하였다. 즉, 검사 조정의 목표는 어떤 학생이 장애를 가졌다는 이유로 검사점수를 인위적으로 부풀려 주거나 검사 상황을 유리하게 해주려는 데 있는 것이 아니라, '검사 결과가 불공평하게 왜곡되는 것을 막고 그 검사가 당초에 측정하려고 했던 것을 정확하게 측정'하려는 데 있다. 요컨대, 검사 조정이란 검사의 본래 목적을 해치지 않는 범위 내에서 최대한 장애로 인한 왜곡 부분을 제거해내려는 일체의 노력을 말한다고 할 수 있다(김동일·최종근, 2003에서 재인용).

그런데, 1992년에서 1997년 사이에 SAT I에서 검사 조정을 요구하는 장애학생의 수는 2배로 증가하였고, 매년 약 14%씩 증가하고 있다(Camara, Copeland, and Rothschild, 1998; Camara & Schneider, 2000에서 재인용). 그리고 이들 장애학생들 중에서도 학습장애 학생이 거의 90%를 차지하고 있으며, 시간 연장 조건하에서 치러진 ACT 시험의 숫자도 1989-90학년도에 8,519명이던 것이 1995-96학년도에는 23,463명으로 대폭 증가하였다(Ziomek & Andrews, 1998; Camara & Schneider, 2000에서 재인용).

그리고 검사 조정의 형태에도 제시형태, 반응형태, 검사 시간, 검사 환경 등 다양한데, SAT I에서 검사 조정 요구의 약 90%를 차지하는 학습장애 학생에 대한 검사 조정의 2/3 이상이 단순한 시간 연장 조정이고, 따라서 최근의 학습장애 학생의 증가 및 이들의 검사 조정 요구의 증가와 더불어 이들이 검사 조정 조건에서 얻은 점수가 표준 조건에서 얻어진 점수와 동등한 가치를 갖는다(equivalent)고 할 수 있는가에 관한 조사연구의 필요성이 심각한 연구과제로 대두되고 있다(Camara & Schneider, 2000).

특히, 학습장애 학생에 대한 검사 조정에 대하여는 외국에서도 찬반 입장이 팽팽히 맞서고 있다. 예컨대, Ofiesh et al.(2004)은 다음과 같은 이유로 학습장애 학생에 대한 시간 연장 등의 검사 조정이 필요하다고 주장하였다. 즉, 학습장애는 그 정의에 대한 논쟁과 더불어 다양한 조작적 정의 방법들이 제시되고 있어 명확하게 합의된 부분이 많지 않지만, 한 가지 일치되는 견해는 학습장애가 학습 또는 학업활동(work)에 상당한 영향을 미치는 인지적 처리과정 또는 신경심리학적 결손(deficits)을 지칭한다는 점이다. 그리고 이러한 합의점과 관련하여, 학습장애를 가진 학생들은 빠른 속도로 검사를 수행하지 못한다는 공통적인 문제를 안고 있으며, 학습장애 학생들은 학습장애 문제를 가지지 않은 학생들이 시간제한이 있는 여러 가지 검사과제들(예, 독해, 수학 계산)을 완수하는 데 걸리는 시간보다 많은 시간을 요구한다는 특징이 있다는 사실이 여러 연구들(Bell & Perfetti, 1994; Deuel, 1992, 1995; Geary & Brown, 1990; Hayes, Hynd, & Wisenbaker, 1986; Rubin & Johnson, 2002 등)을 통해 밝혀졌다(Ofiesh et al., 2004에서 재인용). 뿐만 아니라, 심지어 학습장애 학생들의 학업성취 점수가 평균 수준인 경우에도 그와 같은 (인지적/신경심리학적) 결손으로 인하여 자동적이고 유창한 읽기와 학업관련 지식

의 유창한 적용이 지속적으로 방해될 수 있다(Gregg, Jordan et al., 2003; Lerner, 2003)는 사실이 여러 연구들을 통하여 밝혀졌다(Ofiesh et al., 2004에서 재인용). 바로 이러한 사실과 연구 결과들이 학습장애 학생에 대한 시간 연장 검사 조정의 근거가 된다는 것이다.

반면, Willingham et al.(1988)은 당시까지의 검사 조정 관련 연구들의 검사 조정에 따른 비교가능성(점수/과제)을 8가지 기준을 중심으로 검토한 결과, 일반학생에 대한 표준조건에서의 검사와 장애학생에 대한 검사조정 조건에서의 검사는 ‘대체적으로’ 비교가능성이 있다고 결론내릴 수 있지만, 시간 연장은 검사 점수의 비교가능성을 해치는 1차적 원인이 될 수 있다고 하였다. 특히, 학습장애 학생에 대한 시간 연장은 그 자체 문제뿐만 아니라 학습장애 학생의 이질성 및 진단의 불명확성과 같은 다른 방해 요인과의 얽혀 있어 더욱 복잡하다(김동일·최종근, 2003에서 재인용)고 하면서 학습장애 학생에 대한 시간 연장 검사 조정의 적용에 대하여 경계와 우려를 표하였다.

이렇게 볼 때, 근본적으로 이들 학습장애 학생들에게 검사 시간 연장 등의 검사 조정을 제공할 필요가 있는지 그 적격성(eligibility) 판단 문제와 구체적인 의사결정 방법 등에 대하여는 쉽게 결론내리기 어려운 문제이고, 그런 점에서 이에 관한 연구가 시급히 요청된다고 할 수 있다. 그래서 본 연구에서는 향후 우리 나라에서 학생들에 대한 검사 상황의 조정을 고려함에 있어서 미국에서 학습장애 학생들에게 가장 일반적으로 제공되는 시간 연장 형태의 검사 조정을 포함하여 저성취 학생(LA: Low Achievement)이나 학습장애(LD: Learning Disability) 학생에게(도) 검사 조정이 필요한지, 이들에 대한 검사 조정이 타당한가에 관한 질문에 답을 해보려고 한다.

한편, 장애학생에 대한 검사 조정은 장애 당사자가 지닌 장애로 인한 영향을 최대한 제거하여 공평하게 검사에 임할 수 있도록 하는 인권적/공평성 차원과 그러한 검사 조정이 지나치거나 부족하지 않아 표준조건에서의 검사 점수와 비교가능한 점수를 산출해주는 타당성/비교가능성을 모두 갖춰 그야말로 ‘적절하다’는 평가를 받을 수 있어야 한다. 그래서 Phillips(1994)는 검사 조정과 관련된 핵심적인 쟁점은 장애학생의 인권 보호와 검사도구의 타당성 확보라는 두 가지 요구간의 균형을 얼마나 잘 맞추는가에 있으며, 결국 공평성(fairness)과 타당도(validity)가 검사 조정의 적절성(appropriateness)을 결정하는 핵심 요소라고 하였다. 그리고 어떠한 검사 조정이 ‘적절하다’고 할 수 있는지는 다음과 같은 5가지 질문을 제기하고 이에 응답해봄으로써 검토할 수 있는데, 이들 질문에 대한 답변이 하나라도 “예”인 경우 그 검사 조정은 적절치 않다고 하였다(김동일·최종근, 2003에서 재인용).

- ① 검사 형태(format)(제시 형태 혹은 반응 형태)의 변화나 검사 조건(환경)의 변경으로 인해 측정되는 기술(skill)이 변하는가?

- ② 표준화 검사 조건에서 측정된 피험자의 검사 점수가 검사 조정이 제공된 상태에서 측정된 피험자의 검사 점수와 다른 의미를 가지는가?
- ③ 비장애 피험자에게 동일한 검사 조정을 제공하면 이익(benefit)이 되는가?
- ④ 장애 피험자가 표준화검사에 적응(adapt)할 능력을 조금이라도 가지고 있는가?
- ⑤ 장애 증명 또는 검사 조정 관련 정책이 타당도나 신뢰도가 불확실한 절차에 기초하여 이루어지는가?

본 연구에서는 구체적으로 Phillips(1994)의 위와 같은 5가지 질문 중에서 주로 세 번째 질문과 관련하여 저성취 및 학습장애 학생에 대한 검사 조정의 필요성을 검토하고자 하였다. Phillips(1994)의 세 번째 질문은 달리 표현하자면, 어떠한 검사 조정이 장애학생과 비장애학생에게 동일하게 주어졌을 때, 장애학생에게서는 그 효과가 드러나지만 비장애학생에게는 그 효과가 미미하거나 없어야만 그 검사조정이 비장애학생에게 주어지는 것이 타당하고 허용될 수 있다는 것이다. 달리 말하면, 위에 제시한 Phillips(1994)의 5가지 질문 중 세 번째를 제외한 나머지 질문들에 대한 응답이 모두 '아니오'라면, 비장애학생에 비해 장애학생에게서만 '변별적인 점수 향상'(differential boost)이 나타나는 검사 조정 방법이 '적절한' 것이라고 할 수 있다는 것이다. 그래서 본 연구에서는 저성취 및 학습장애 학생이 표준조건에서와 달리 일반학생에 비해 일정한 검사 조정 조건(시간연장, 대독(녹음테이프), 시간연장+대독)에서 특별히 검사 점수가 향상되어 그와 같은 검사 조건이 이들에게 보다 적절한 검사 환경임을 암시해주는지 여부를 확인해보고자 한다.

## II. 연구 문제

위와 같은 필요성을 바탕으로 본 연구에서는 저성취 및 학습장애 학생에게(도) 시간 연장 등의 검사 조정이 필요한지 여부와 근거를 일반학생들과는 달리 이들 학생들에게서만 검사 점수에 변별적인 향상이 나타나는 검사 조정 조건이 존재하는지를 분석하는 방법으로 확인하고자 한다.

본 연구에서 살펴 볼 연구문제는 저성취 및 학습장애 학생을 위한 검사조정의 필요성과 타당성이며 이를 구체적으로 정리하면 다음과 같다.

<문제 1> 저성취 및 학습장애 학생은 검사 조정 조건에서 일반학생에 비해 더 많은 점수

향상을 보이는가?

<문제 2> 저성취 및 학습장애 학생집단에서 반복/연습 효과를 제거한 순수한 검사 조정의 효과는 어느 정도인가?

<문제 3> 저성취 및 학습장애 집단과 일반학생 집단에서 각각 점수 향상이 가장 많은 검사 조정 조건은 무엇이고 집단간에 차이가 있는가?

### III. 연구 방법

#### 1. 연구 대상 및 자료수집 방법

본 연구를 위하여 서울 시내 2개 중학교의 3학년 일반학급 각 8개 반(총 16개 반) 527명과 서울 시내 10개 중학교의 특수학급 3학년 학생 23명을 합하여 총 550명의 학생들이 2차례 검사에 참여하였다. 구체적으로 본 연구에 참여한 학생 현황은 <표 1>과 같았다. 저성취 여부는 1차 검사 점수를 기준으로 판정하였고, 저성취 학생(LA)은 일반학급에 배치되어 있지만 본 연구에 사용된 기본읽기 소검사에서 부진 기준에 해당하는 '기본학습부진' 학생이다. 학습장애(LD) 학생은 특수교육준거와 부진준거에 모두 해당되는 학생을 말한다. 즉, (1) 현재 학업 문제로 인하여 학교별 진단 과정을 통하여 특수교육대상자로 선정되어 특수학급에 입학되어 있고 (2) 본 연구에 사용된 기본읽기 소검사에서 부진 기준에 해당하여 '기초학습부진'으로 판단되는 학생을 지칭한다.

전체 참여 학생 550명중 420명은 1차 검사에서 학습부진 기준 이상의 점수를 얻었고, 130명은 학습부진 판정 기준 미만의 점수를 얻었다. 그리고, 2차 검사에서 3가지 중 한 가지 검사 조정 조건에서 검사를 실시한 실험집단에는 404명이, 2차 검사에서 문형만 달리하여 1차 검사에서와 같이 표준조건에서 검사를 실시한 비교집단에는 146명이 배치되었다. 한편, 본 연구에 참여한 학생의 성별 구성은 <표 2>와 같이, 남학생이 296명(53.8%)으로 여학생(254명,

<표 1> 연구 대상 학생 현황

구분	일반학급		특수학급		계
	보통	저성취(LA)	보통	학습장애(LD)	
실험집단	313	79	-	12	404
비교집단	105	30	2	9	146
소계	418 (79.3%)	109 (20.7%)	2 (8.7%)	21 (91.3%)	550
	527		23		

&lt;표 2&gt; 연구 대상 학생의 성별 현황

구분	일반학급		특수학급		계
	보통	저성취(LA)	보통	학습장애(LD)	
남자	210	73	1	12	296 (53.8%)
여자	208	36	1	9	254 (46.2%)

46.2%)보다 약간 많았고, 남학생(85명, 전체의 15.5%, 남학생의 28.7%)이 여학생(45명, 전체의 8.2%, 여학생의 17.7%)보다 약 2배 가까이 부진학생이 많았다.

자료 수집은 학교 교육과정 일정 및 여건을 고려하여 1차 검사(표준조건)의 경우 2004년 7월 2일(금)에서 7월 7일(수) 사이에, 2차 검사(조정 조건)의 경우 2004년 7월 15일(목)에서 7월 21일(수) 사이에 실시하였고 1-2차 검사 간격은 약 2주가 되도록 하였다. 일반학급 학생을 대상으로 한 검사는 해당 학교 국어과 교사가 수업시간에 실시하였고, 특수학급 학생을 대상으로 한 검사는 연구자와 특수교육전공 대학원 석사과정 학생 2명이 직접 실시하였으며, 참여 학생들의 성실한 응답을 유도하기 위하여 응답 학생에게 2색 볼펜을 제공하였다.

표준조건(비교집단)에서는 문형만 달리하여 30분간 지필검사 형태로 검사를 실시하였고, 시간 연장 조건의 경우 문형만 달리한 검사지를 45분 동안 풀도록 하였으며, 대독(녹음테이프) 조건의 경우 연구자가 사전에 검사문항을 녹음한 테이프(30분 분량)를 녹음기를 이용하여 틀어놓고 검사를 실시하였으며, 시간연장+대독(녹음테이프) 조건의 경우 제시문, 각 문항 및 답지를 읽는 간격을 더 늘여 45분 분량으로 천천히 녹음한 테이프를 녹음기를 이용하여 재생시킨 상태에서 검사를 실시하였다. 구체적인 실험 배치 및 검사 조정 방법은 <표 3>과 같았다.

&lt;표 3&gt; 실험 배치 및 검사 조정 방법

구분	검사조건	학급구분	학급수	인원수	소계	총계
실험집단	조건1	일반학급	4	122	126	404
		특수학급	2	4		
	조건2	일반학급	4	132	136	
		특수학급	2	4		
	조건3	일반학급	4	138	142	
		특수학급	2	4		
비교집단	조건4-1	일반학급	2	67	74	146
		특수학급	2	7		
	조건4-2	일반학급	2	68	72	
		특수학급	2	4		

\* 조건1 : 표준조건(30분, 지필) → 시간연장(45분, 지필)

\* 조건2 : 표준조건(30분, 지필) → 대독(카세트)(30분, 카세트)

\* 조건3 : 표준조건(30분, 지필) → 연장+대독(45분, 카세트)

\* 조건4-1 : 표준조건(30분, 지필)(A형) → 표준조건(30분, 지필)(B형)

\* 조건4-2 : 표준조건(30분, 지필)(B형) → 표준조건(30분, 지필)(A형)

## 2. 검사 도구

본 연구에서는 학생들의 기본적인 학습능력을 측정하기 위해 한국교육개발원에서 개발한 「기본학습 부진학생 판별도구」의 중학교 3학년용 국어 읽기 소검사를 사용하였다(<표 4> 참조).

<표 4> 기본학습 부진학생 판별도구(중3용) 국어 읽기검사의 구성

하위 영역	문항번호	문항수	학습부진 판정기준	비고(*)
읽기의 본질(읽기 방법)	1, 2, 3	20	13문항 이하 (70점 미만)	13.6%
단어 이해의 기본원리	7, 15			
내용 이해의 기본원리	4, 8, 11, 12, 13, 17, 18, 19, 20			
평가 및 감상의 기본원리	5, 6, 9, 10, 14, 16			
기초적인 읽기 태도/습관	-			

(\*) 검사 개발·표준화 당시 일반학생 중 학습부진 판정기준 해당 비율

이 검사는 중학교 3학년 학생을 대상으로 직전 학년도인 중학교 2학년 교육과정의 성취도를 측정하는 일종의 집단 지필형 학업성취도 검사로서 4지 선택형 총 20문항으로 구성되어 있고, 원칙적으로 3학년 1학기 초에 실시하도록 되어 있으며, 검사 시간은 약 30분이 소요되고, 검사 점수가 총점의 70% 미만(13문항 이하)인 경우를 보충학습 대상자로 분류할 수 있고, 표준화 당시 피험자의 13.6%가 이에 해당한 것으로 보고되어 있다(박효정 외, 2001). 그리고, 이 검사의 개발자들(박효정 외, 2001)에 따르면, 재검사 신뢰도(검사-재검사 상관)가 .78로 비교적 양호한 것으로 나타났고, 검사의 문항내적 합치도(Cronbach's  $\alpha$ ) 계수도 .83으로 신뢰로운 것으로 나타났다고 보고하였다(최종근, 2005).

한편, 본 연구에서는 동일한 검사를 2회에 거쳐 반복 실시하게 됨에 따라 답지의 순서만 달리한 2가지 검사지(A형, B형)를 제작하여 활용하였다. 실험집단의 경우 1차 검사(표준조건)는 A형 검사지를, 2차 검사(검사 조정 조건)에서는 B형 검사지를 사용하였다. 그래서, A형과 B형 검사가 동형이라고 할 수 있는지를 점검하였다. 분석 결과, <표 5>와 같이 일반학급과

<표 5> 1차 검사 A형, B형 검사간 점수 차이 검증

구분	문형	사례수	평균	표준편차	t	df	유의도
일반학급	A형	459	15.81	3.71	-.577	525	.564
	B형	68	16.09	3.41			
특수학급	A형	19	7.84	3.22	-.607	21	.550
	B형	4	9.00	4.69			

특수학급 모두 1차 검사에서 A형 검사지 사용자와 B형 검사지 사용자간의 점수 차이는 통계적으로 유의하지 않은 것으로 나타났다.

그리고 검사지 유형만 달리하여 표준조건에서 반복하여 검사를 실시한 비교집단(표준→표준)의 1-2차 검사간 상관은 <표 6>과 같이 일반학급의 경우 .715, 특수학급의 경우 .773으로 비교적 높은 편으로 나타나 본 연구에서 검사지 유형에 따른 영향은 거의 없다고 해석할 수 있는 것으로 나타났다. 즉, 본 연구에서 사용한 두 가지 검사지 유형(A형과 B형)간의 동형성은 검사-재검사 신뢰도(상관) 계수에 비추어 볼 때 비교적 높다고 할 수 있다.

<표 6> 비교집단(표준→표준)의 1-2차 검사간 상관

구분	검사 차수	사례수	평균	표준편차	상관
일반학급	1차	135	15.33	4.06	.715**
	2차	135	15.30	4.71	
특수학급	1차	11	9.36	3.72	.773**
	2차	11	8.36	3.01	

\*\*  $p < .01$

### 3. 분석 방법

2회에 걸쳐 실시된 검사 중 표준조건에서 시행된 1차 검사에서 본 연구에서 사용된 검사 도구의 개발자들이 학습부진 판정 기준으로 제시한 검사 총점의 70%(14문항)을 기준으로 크게 3가지 집단으로 나누어 분석하였다. 즉, 일반학급 학생중 1차 검사(표준조건)에서 학습부진 판정 기준 이상의 점수를 획득한 학생을 일반학생, 그 미만의 점수를 획득한 학생을 저성취 학생(LA), 그리고 특수학급 학생중 학습부진 판정기준 미만의 점수를 획득한 학생을 학습장애(LD) 학생으로 조작적으로 정의하였다. 특히, 본 연구의 목적이 일반학생에 비하여 저성취 및 학습장애 학생들에게 검사 조정이 필요한지 여부를 분석하는 데 있었고, 특수학급 학생중 1차 검사에서 학습부진 기준 이상의 점수를 획득한 학생은 2명에 불과하여 이들은 집단간 비교 분석에서 제외하였다.

본 연구의 목적을 달성하기 위하여 우선 각 집단별 검사조정 여부 및 세부 검사 조건에 따른 1-2차 검사 점수의 변화를 면밀히 분석하고, 각 집단에서 가장 점수 변화가 큰 검사 조정 조건을 확인해내는 한편, Camara와 Schneider(2000)가 SAT I 반복 응시 자료에서 검사 시간 연장의 효과를 추출해내기 위하여 적용하였던 방법 즉, 표준→연장 조건에서의 점수 증가량에서 표준→표준 조건에서의 점수 증가량을 빼주는 방법을 준용하여 반복/연습 효과를 제거한 순수한 검사 조정의 효과를 각각 계산해 보고자 하였다.



## IV. 분석 결과 및 해석

### 1. 검사 조정 여부에 따른 점수 차이

검사 조정 여부에 따른 점수 증가 여부와 그 정도를 분석하기 위하여 각 집단별로 실험집단과 비교집단의 1-2차 검사점수의 평균을 비교한 결과 <표 7>과 같았다.

<표 7> 실험 배치에 따른 집단별 1-2차 검사 평균과 점수 증가

구분	실험 배치	사례수 (n)	1차		2차		점수증가 (gain)	조정효과 (*)
			평균	표준편차	평균	표준편차		
일반 학생	실험집단 (표준→조정)	313	17.48	1.74	17.33	2.58	-0.15	-0.01
	비교집단 (표준→표준)	105	17.11	1.80	16.97	2.84	-0.14	-
저성취 (LA)	실험집단 (표준→조정)	79	10.27	2.80	11.44	4.64	1.17	0.81
	비교집단 (표준→표준)	30	9.07	3.54	9.43	5.25	0.36	-
학습 장애 (LD)	실험집단 (표준→조정)	12	6.83	2.73	7.92	3.78	1.09	1.53
	비교집단 (표준→표준)	9	8.11	2.71	7.67	2.87	-0.44	-

(\*)조정효과= 조정조건에서의 점수증가(조건1,2,3)-표준조건에서의 점수증가(조건4)  
= 실험집단의 점수증가(gain) - 비교집단의 점수증가(gain)

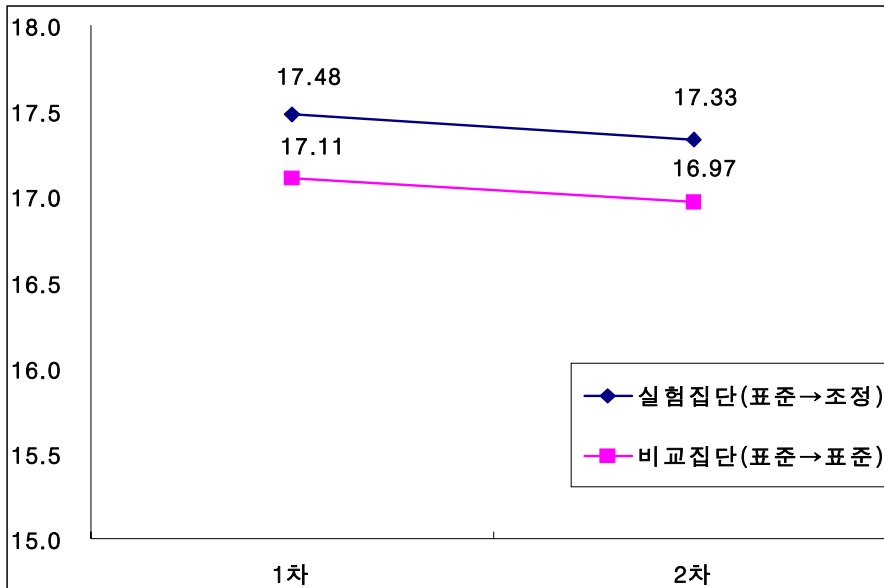
일반학생 집단의 경우 실험집단이나 비교집단 모두 1-2차 검사 점수에 거의 변화가 없거나 약간 낮아진 것으로 나타났다. 반면에 저성취 학생집단의 경우에는 일정한 검사 조정이 이루어지는 경우 1차 검사에 비해 1.17점(20점 만점) 정도의 점수 향상이 있는 것으로 나타났다. 그리고 학습장애 학생집단의 경우에는 실험집단의 경우 1차 검사에 비해 1.09점(20점 만점)의 점수 향상을 보여 2차 검사에서 오히려 점수가 하락한 비교집단과 1.53점의 차이가 발생하는 것으로 나타났다. 전체 20문항중 저성취 학생집단의 1.17문항과 학습장애 학생집단의 1.53문항의 차이는 100점 만점으로 환산하면 각각 6점, 7.5점 정도에 불과하여 그 크기가 큰 편이라고 말하기는 어렵다.

이러한 원점수 차이를 비교집단(표준→표준)의 표준편차로 나누어 각각을 효과 크기(Es: Effect Size) 개념으로 바꾸어 계산해 보면, 일반학생 집단은 .0056, 저성취(LA) 집단은 .2288, 그리고 학습장애(LD) 집단은 .5646이었다. 김계현(2000)에 따르면 효과 크기(ES)가 .2 정도이면 작은 크기의 효과, .5 정도이면 중간 크기의 효과, .8 정도이면 큰 크기의 효과라고 본다.

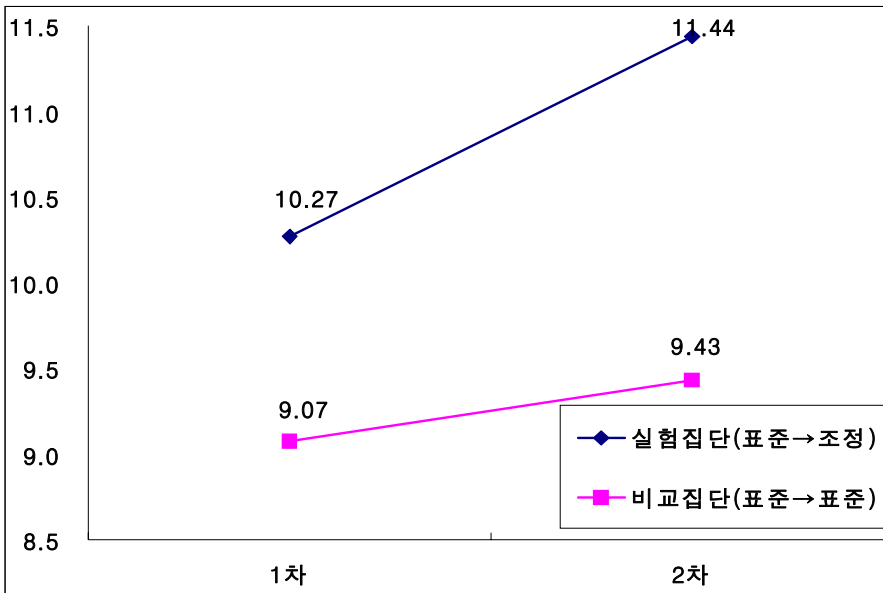
이러한 기준에 비추어 해석한다면 일반학생과 저성취 집단에서는 검사 조정의 효과가 작았고, 학습장애 집단에서는 검사 조정의 효과가 중간 정도였다고 할 수 있다. 바꾸어 말하면 학습장애 집단에서 나타난 효과 크기 .56 은 표준조건에서 평균 이하의 점수를 얻었던 피험자 중에서 21% 정도가 검사 조정으로 인해 평균 이상의 점수를 얻게 되는 정도의 효과(변화)를 의미한다고 해석할 수 있다.

[그림 1]~[그림 3]은 세 집단의 실험집단과 비교집단의 1-2차 검사 점수의 변화를 그래프로 표현한 것으로서, 이들 세 집단 내에서 각각 표준조건→표준조건에서 검사를 치른 비교집단에 비해 표준조건→조정조건에서 검사를 치른 실험집단이 어느 정도의 점수 변화를 나타내는지를 시각적으로 확인할 수 있고 또 이를 통해 검사 조정의 효과를 가늠할 수 있다.

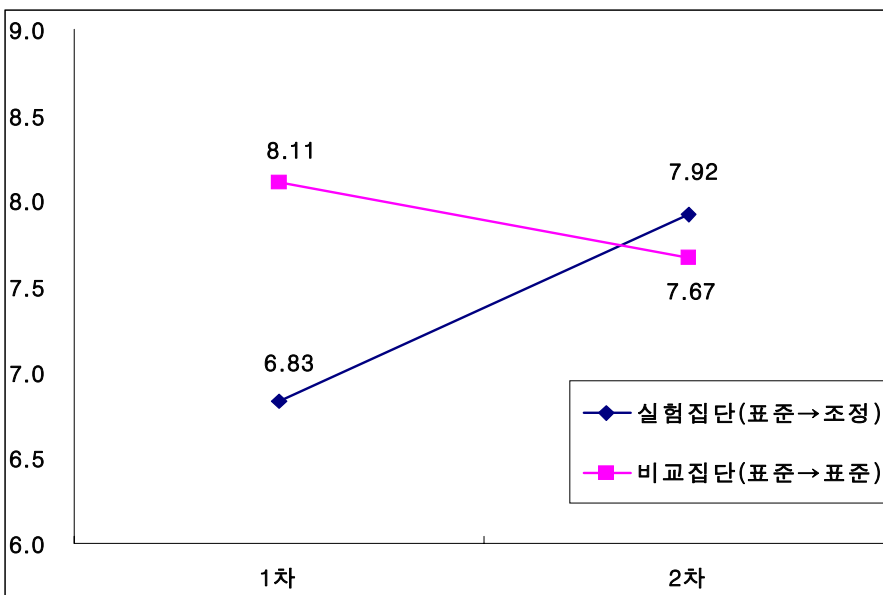
한편, 세 집단(일반학생, 저성취, 학습장애) 모두 1차 및 2차 검사에서 실험집단과 비교집단간의 평균 차이가 통계적으로 유의( $p < .05$ )하지 않은 것으로 나타났다. 다만, 저성취 집단에서 2차 검사의 경우 유의수준 .1 에서 실험집단과 비교집단간에 통계적으로 유의한 차이 ( $t=-1.946$ ,  $df=107$ , 유의도 .054)가 있는 것으로 나타났다.



[그림 1] 일반학생의 1-2차 검사 점수 변화



[그림 2] 저성취 학생의 1-2차 검사 점수 변화



[그림 3] 학습장애 학생의 1-2차 검사 점수 변화

## 2. 검사 조건별 점수 변화와 검사 조정의 효과

<표 8>은 세부 검사 조건에 따른 집단별 1-2차 검사점수 평균과 점수 증가 상황을 정리한 것이다. 대체적으로 보아 저성취 학생집단의 경우 세 가지 검사 조정 조건에서 고루 점수 향상을 보인 반면, 일반학생 집단은 검사 조정에도 불구하고 점수 변화가 거의 없었고, 학습장애 학생집단은 조건3(표준→연장+녹음)에서 주목할 만한 점수 향상을 나타내었다. 그리고 세 집단 모두 검사 조건3(연장+녹음)에서 표준조건에 비해 점수 증가가 큰 편이라는 공통점이 발견되었다. 다만, 저성취 학생 집단에서는 검사 조건3(연장+녹음)의 효과가 검사 조건1(표준→연장)의 효과에 다소 못미치는 예상 밖의 결과가 나타났다.

구체적으로, 일반학생 집단에서는 조건3(표준→연장+녹음)이 조정 효과가 가장 컸고(0.50점), 저성취 집단에서는 조건1(표준→연장)이 조정 효과가 가장 컸고(0.98점), 학습장애 집단에서는 조건3(표준→연장+녹음)이 조정 효과가 가장 큰 것(3.94점)으로 나타났다. 이러한 차

<표 8> 세부 검사조건에 따른 집단별 1-2차 검사 평균과 점수 증가

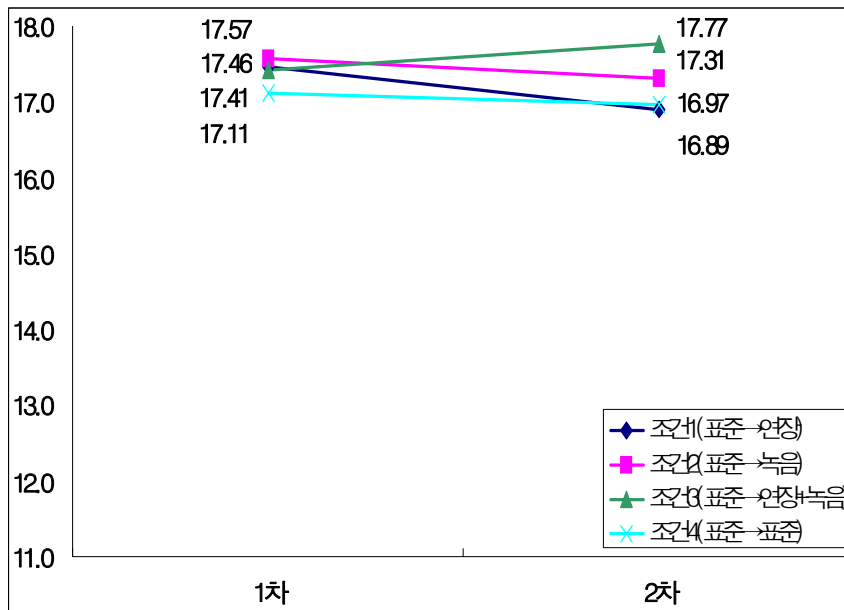
구분	실험 배치	사례수 (n)	1차		2차		점수 증가 (gain)	조정 효과 (*)
			평균	표준편차	평균	표준편차		
일반 학생	조건1 (표준→연장)	101	17.46	1.82	16.89	3.00	-0.57	-0.43
	조건2 (표준→녹음)	108	17.57	1.78	17.31	2.60	-0.26	-0.12
	조건3 (표준→연장+녹음)	104	17.41	1.62	17.77	2.01	0.36	0.50
	조건4 (표준→표준)	105	17.11	1.80	16.97	2.84	-0.14	-
저성취 (LA)	조건1 (표준→연장)	21	10.33	2.42	11.67	4.39	1.34	0.98
	조건2 (표준→녹음)	24	10.33	3.09	11.33	5.33	1.00	0.64
	조건3 (표준→연장+녹음)	34	10.18	2.89	11.38	4.41	1.20	0.84
	조건4 (표준→표준)	30	9.07	3.54	9.43	5.25	0.36	-
학습 장애 (LD)	조건1 (표준→연장)	4	8.00	0.82	7.50	2.65	-0.50	-0.06
	조건2 (표준→녹음)	4	5.25	3.30	5.50	3.87	0.25	0.69
	조건3 (표준→연장+녹음)	4	7.25	3.20	10.75	3.40	3.50	3.94
	조건4 (표준→표준)	9	8.11	2.71	7.67	2.87	-0.44	-

(\*)조정효과= 조정조건에서의 점수증가(조건1,2,3)-표준조건에서의 점수증가(조건4)  
= 실험집단의 점수증가(gain) - 비교집단의 점수증가(gain)

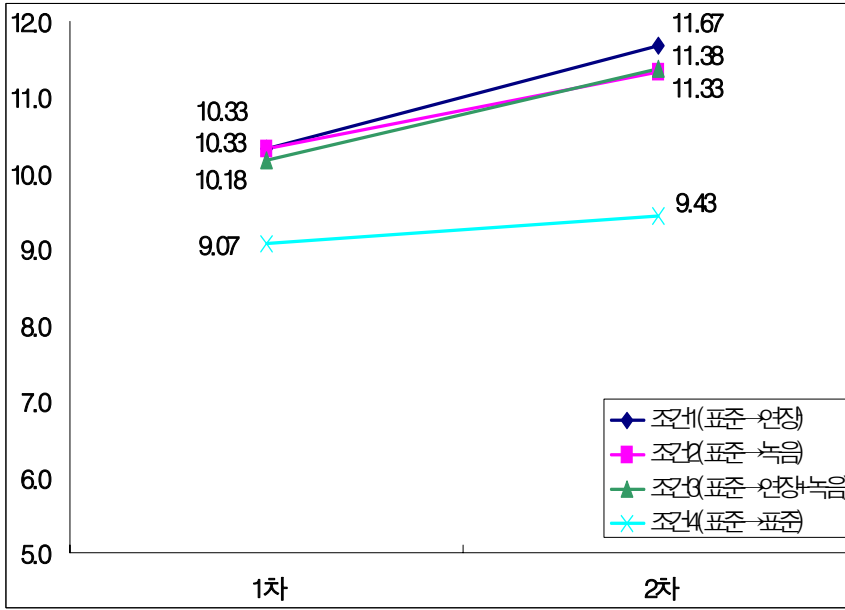
이를 100점 만점으로 환산하여 보면, 일반학생 집단의 2.5점과 저성취 집단의 약 5점은 그 크기가 큰 편인지를 가늠하기 어렵지만, 학습장애 집단의 약 20점은 상대적으로 큰 편이었음을 알 수 있다.

이러한 원점수 차이를 비교집단(표준→표준)의 표준편차로 나누어 각각을 효과 크기(ES) 개념으로 바꾸어 계산해 보면, 일반학생 집단은 .2778, 저성취 집단은 .3391, 그리고 학습장애 집단은 1.4539이었다. 효과 크기(ES)에 대한 일반적인 해석 기준에 비추어 볼 때, 일반학생과 저성취 집단에서는 검사 조정의 효과가 작았고, 학습장애 집단에서는 검사 조정의 효과가 컸다는 것을 확인할 수 있다. 특히, 학습장애 집단에서 나타난 효과 크기 1.45는 표준조건에서 평균 이하의 점수를 얻었던 피험자 중에서 약 43%가 검사 조정으로 인해 평균 이상의 점수를 얻게 되는 정도의 효과 크기에 해당한다. 즉, 학습장애 학생에게 연장+녹음 검사 조정을 제공할 경우 전체 학습장애 학생의 93%가 당초의 표준조건에서의 평균 점수 이상을 받게 된다는 것을 의미한다.

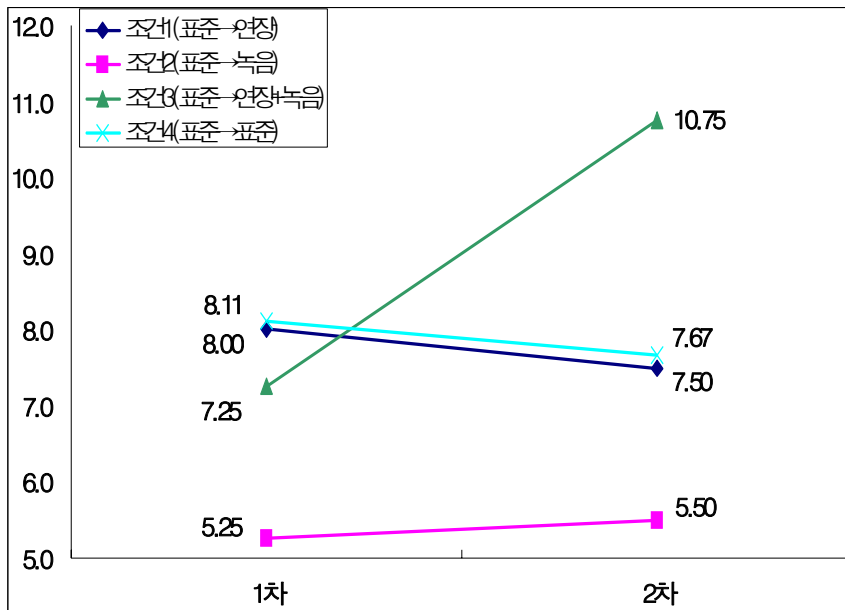
한편, [그림 4]~[그림 6]은 세 집단의 각 검사 조건에서의 1-2차 검사 점수의 변화를 그래프로 표현한 것으로서, 이들 세 집단 내에서 각각 표준조건→표준조건에서 검사를 치른 비교집단에 비해 표준조건→조정조건에서 검사를 치른 실험집단들이 어느 정도의 점수 변화를 나타내는지를 시각적으로 확인할 수 있고 또 이를 통해 각각의 검사 조정의 효과를 가늠할 수 있다.



[그림 4] 일반학생의 검사조건별 1-2차 검사 점수 변화



[그림 5] 저성취 학생의 검사조건별 1-2차 검사 점수 변화



[그림 6] 학습장애 학생의 검사조건별 1-2차 검사 점수 변화

## V. 논의 및 결론

### 1. 검사 조정 여부에 따른 점수 차이와 검사 조정의 필요성

분석 결과 일반학생의 경우 표준조건에 비해 시간연장, 대독(녹음테이프), 시간연장+대독(녹음테이프)과 같은 검사 조정이 제공되어도 점수 증가가 거의 없거나 오히려 떨어졌다. 이는 일반학생에 대해서는 본 연구에서와 같은 검사 조정이 점수 증가에 아무런 영향을 미치지 않는다는 것을 의미하고 따라서 이들 일반학생들에게는 그와 같은 검사 조정이 필요치 않을 수 있음을 시사한다.

반면에 저성취 학생과 학습장애 학생의 경우, 개개 검사 조건에 따라 다르지만, 전체적으로 표준조건에 비해 시간연장, 대독(녹음테이프), 시간연장+대독(녹음테이프)과 같은 검사 조정이 제공될 경우 대체적으로 점수가 증가하는 것으로 나타났는데, 이는 이들 저성취 및 학습장애 학생들의 경우 검사 조정이 점수 증가에 일정한 영향을 미친다는 것을 의미하고, 따라서 이들 집단에게 있어서 그와 같은 검사 조정 환경이 표준조건에서는 드러나지 않았던 '진짜 지식 또는 기술'을 보다 잘 드러내주는 검사 환경이 될 수 있음을 시사한다.

이러한 결과는 일반학생과 학습장애 학생의 SATI 재응시에 따른 점수 변화를 분석한 Camara와 Schneider(2000)의 연구 결과와도 일치하는 것이다. 즉, 표준 조건에서 검사를 반복 응시한 일반학생과 검사 조건(시간 연장)을 달리하여 반복 응시한 학습장애 학생들의 점수 변화를 비교 분석한 이들의 연구에서 일반학생(표준-표준)이나 연장-연장 조건의 학습장애 학생의 검사 점수는 13점 내외가 향상된 데 그친 반면, 표준-연장 조건으로 검사 조정(시간 연장)을 한 경우에는 그와 같은 '반복/연습 효과'의 3배 가까운 점수 향상(수학(math) 38.1점, 언어(verbal) 44.6점)을 보였으며, 연장-표준 조건에서 시험을 치른 경우에는 오히려 약 7-8점 정도 점수가 하락되는 것으로 나타났다(김동일·최종근, 2003, <표 7> 참조).

### 2. 검사 조정의 효과와 시사점

검사 조정 여부에 따른 점수 차이의 분석 결과를 종합적으로 보면, 검사 조정의 효과는 대체적으로 일반학생에게서는 거의 나타나지 않은 반면, 저성취 학생과 학습장애 학생 집단에서만 존재하는 것으로 나타났다. 그리고 저성취 집단에 비해 학습장애 학생 집단에서의 전체 검사 조정의 효과가 더 큰 것으로 나타났다.

이러한 결과를 세부 검사 조건에 따라 구체적으로 살펴본 결과, 세 집단 모두 검사 조건3

(연장+녹음)에서 표준조건에 비해 점수 증가가 큰 편이라는 공통점이 발견되었다. 이러한 결과는 시간 연장의 효과와 대독(녹음테이프)의 효과가 합해져 그 효과가 가장 클 것이라는 당초의 예상과 대체적으로 일치하는 것이었다. 다만, 저성취 학생 집단에서는 검사 조건3(연장+녹음)의 효과가 검사 조건1의 효과에 다소 못 미치는 예상과 다른 결과가 나타났다. 검사 조건1에 배치된 저성취 학생집단(평균 10.33)에 비해 검사 조건3에 배치된 저성취 학생집단(평균 10.18)이 1차 검사 점수가 약간 낮다는 점을 감안하더라도, 검사 조건3이 검사 조건1과 검사 조건2의 검사조정의 효과를 합한 효과( $0.98+0.64=1.62$ )에 근접하기는커녕, 검사 조건1의 검사 조정 효과(0.98)보다도 그 효과가 작은 것(0.84)은 실험상의 오차 이외에 대독(녹음테이프) 검사 조정이 저성취 학생들의 독립적인 검사 수행을 간섭하는 요인으로 작용했기 때문일 수도 있어 추가적인 연구와 분석이 요청된다.

한편, 본 연구에서 일반학생에게는 천장 효과(ceiling effect)가, 저성취 집단과 학습장애 집단에는 평균으로의 회귀(regression to mean) 효과가 작용했을 가능성을 완전히 배제할 수는 없다. 표준조건에서 치러진 1차 검사(20점 만점)에서 일반학생은 평균 17.11점, 저성취 학생은 평균 9.07점, 그리고 학습장애 학생은 8.11점을 나타내어 일반학생의 85% 내외의 정답률과 저성취 및 학습장애 집단의 40-45%에 불과한 정답률은 매우 대조적일 뿐만 아니라 그 차이도 매우 컸다는 점에서 더욱 그렇다.

그리고 이러한 천장 효과 및 회귀 효과와 관련하여 검사 조정 방법과 검사 자체의 특성간에 상호작용이 존재할 가능성도 동시에 고려할 필요가 있다. 본 연구에서 일반학생이 평균 17.11점(약 85% 정답률)을 얻은 데서 알 수 있듯이 본 연구에서 사용된 것과 같은 비교적 검사 곤란도가 낮은 읽기 학업성취도 검사(독해력)와 달리 검사 곤란도가 높은 검사 도구를 사용할 경우 연구 결과가 본 연구 결과와 다를 수도 있다. 요컨대, 개개 검사 조정 방법들의 보다 정확한 효과를 알아내기 위해서는 다수의 피험자를 대상으로 다양한 검사 조정 방법들과 다양한 검사 난이도 수준, 나아가서는 다양한 교과영역 및 검사 시간, 검사 도구의 길이 등을 종합적으로 고려하여 연구할 필요가 있다.

다만, 적정수의 사례 표집의 어려움으로 인해 결합표집 등의 방법으로 이러한 천장 효과와 회귀 효과를 통제하지 못한 점을 고려하여 지나친 일반화와 확대 해석을 하지 않을 것을 전제로 본 연구의 결과를 해석하면, 전체적으로 일반학생에게는 검사 조정의 효과가 거의 없는 것으로 나타난 반면, 저성취/학습장애 학생집단의 경우 검사 조정의 효과가 있는 것(특히, 저성취 집단의 경우 3가지 검사 조건에서 고루 검사 조정의 효과가 발견됨)으로 나타나, 검사 환경의 불충분이 이들의 학업성취도 점수에 개입되어 이들의 실제 지식/기술 수준을 왜곡하고 있을 가능성이 확인되었다고 말할 수 있다.

물론 이와 같은 결과가 모든 저성취 학생에게, 개개 검사들의 목적이나 의도와 관계없이



모든 검사 상황에게까지 일반화될 수 있는 것은 아닐 것이다. 저성취 학생으로 동일하게 범주화될 수는 있어도 일부 학생들은 검사 조정에도 불구하고 전혀 변화가 없는 경우도 있을 수 있고, 무엇보다도 제한된 시간 안에 일정한 과제를 완수해낼 수 있는지 여부가 검사의 중요한 목적인 경우, 허용하는 검사 시간 또는 시간 연장의 정도는 매우 신중하게 결정되어야 할 것이다. 또한, 효과 크기(ES)에 비추어 볼 때, 학습장애 집단의 연장+녹음 조건을 제외하고는 그 효과 크기가 모두 중간 이하였다는 점도 고려해야 할 것이다.

특히, 저성취 및 학습장애 학생들에 대한 시간 연장 문제와 관련해서는 앞서 예시한 Camara와 Schneider(2000)의 SAT I 반복 응시에 관한 연구에서 연구자들이 '시간 연장(혹은 추가 시간)이 학습장애 학생들에게 보상(compensating) 요인으로 기능함에는 분명하지만, 1.5배 또는 2배의 시간은 일부 학생들로 하여금 훨씬 여유롭게 검사에 임할 수 있게 하고 결과적으로 대학에서의 학업수행(수준)을 과대추정하는 오류를 범할 수 있어 지나친 보상(overcompensation)이 될 수 있다고 결론내린 바 있다는 점도 간과해서는 안될 것이다.

### 3. 학습장애 학생의 피검행동과 검사 조정의 필요성

한편, 학습장애 집단에서는, 사례수가 적다는 제한점을 고려하지 않는다면, 검사 조건2(녹음)와 검사 조건3(연장+녹음)에서는 검사 조정의 효과가 있는 반면, 검사 조건1(시간 연장)에서는 검사 점수가 오히려 낮아져 검사 조정의 효과가 없거나 오히려 역효과가 있는 것으로 나타났고, 반대로 검사 조건3(연장+녹음)에서는 그 효과가 주목할 만한 수준인 것으로 나타났다. 이러한 결과는 예컨대, 이들은 읽기능력(해독과 독해)이 부족하여 시간만 연장해주는 것으로는 점수가 향상될 수 없고, 녹음테이프 등을 이용하여 누군가 대신 읽어 주어야만 점수 향상이 나타날 수 있고, 이러한 대독 서비스와 함께 검사 시간이 연장될 경우 그 향상 폭이 매우 커질 수 있다는 가능성을 보여준다고 해석할 수 있다.

다시 말해, 학습장애 학생의 경우 읽기능력(특히, 배경지식 등을 포함한 독해 능력)과 같은 기초학습능력의 부족으로 인해 대독 서비스가 병행되지 않는 시간 연장은 검사 점수에 아무런 영향을 미치지 못할 가능성도 있다. 실제로 이와 같은 가능성은 이미 선행연구에서 어느 정도 확인된 바 있는 다음과 같은 학습장애 학생의 몇 가지 특징적인 피검행동들(최종근, 2005에서 발췌·재인용)과 관련해서도 짐작이 가능하다.

첫째, 본 연구에 사용된 읽기 검사는 속성상 제시된 지문을 읽고 전체 주제를 파악하거나 제시문의 일부에 관한 이해도를 측정하는 문항들이 대부분인데, 학습장애 학생들은 제시된 지문을 읽지 않고 곧바로 문제를 풀이하는 경우가 많았다. 그렇다고 이 학생들이 문제를 먼

저 읽고 문항의 요구사항을 미리 파악한 뒤 제시문을 나중에 읽고 풀이하는 '문제의 요구를 먼저 파악하는 전략'을 구사한다고 보기도 어려웠다. 왜냐하면 문항의 요구상 전체 지문을 찬찬히 읽고 주제를 파악해야 하는 문항임에도 불구하고 제시문을 아예 읽지 않고 문제를 풀이하기도 하기 때문이다.

둘째, 학습장애 학생들이 대체적으로 검사 기술(test taking skill)이 부족하고 오히려 부적절하다는 사실을 쉽게 발견할 수 있었다. 한 학교에서는 대상 학생이 한 명뿐이어서 문제를 풀면서 말을 하도록 하였는데(일종의 사고발성법(think-aloud)이라고 할 수 있을 것이다), 어려운 문항을 접하면 학생들은 일반적으로 오답지를 하나씩 제거해 가다가 남은 선택지 중에 정답을 고르는 '지우개 전략'을 사용하는데 비해, 이 학생은 오히려 선택지를 무작위로 하나씩 골라잡고 '이게 정답인가? 아닌가?'를 계속 반복하는 무작위선택이나 찍기에 가까운 검사 행동을 보였다. 이러한 피검 행동은 겉으로는 지우개 전략을 구사하는 것처럼 보이지만 실상은 '가짜 지우개전략'인 셈이다.

셋째, 녹음테이프를 제공한 상태에서 검사에 임한 학생 중 한 명은 녹음테이프에서 다음 문항이 들려오는 데도 이전 문항에 대한 답변을 포기하고 현재 문항에 집중하지 못하고 계속하여 다음 문항을 놓쳐 결과적으로 녹음테이프와 거의 무관하게 검사를 수행하기도 하였다. 이는 녹음테이프의 진행 속도에 비해 이 학생의 인지적 정보처리 속도가 늦기 때문이라고도 해석할 수 있고, 놓친 문항에 대한 포기과 새로운 문항에 대한 주의집중 전략을 구사하지 못하기 때문으로 해석할 수도 있을 것이다.

넷째, 제시문에 나온 단어를 포함하고 있거나 유사한 선택지가 정답이라고 응답하는 학생이 있었고 심지어는 일치하는 음절을 포함하고 있는 선택지를 정답이라고 반응하는 경우도 있었다. 예컨대, 4번 문항의 경우 선택지에 제시된 단어들 모두가 본문에 포함되어 있지만, 본문에서 가장 먼저 나와 가장 눈에 잘 띄는 '구 조선총독부 칙탁 철거'라는 선택지를 선택하여 우연히 정답반응을 하는 경우가 있었다. 또 다른 예로 8번 문항과 같이 제시문 안에서 밑줄 쳐진 문장('또 한겹 더 겹쳐')과 관련 깊은 속담이 무엇인지를 묻는 문항에서 정답 선택지('엷친 데 덮친 격이다')가 우연히도 제시된 문장과 매우 유사하여 우연히 정답하게 되는 경우도 있었다.

다섯째, 학습장애 학생의 배경지식이나 어휘력이 부족하다는 점을 쉽게 알 수 있었다. 검사가 끝나고 질문이 있으면 하라고 했을 때, '서사', '논증', '묘사'와 같은 단어의 의미를 묻는 경우도 있었지만, 중학교 3학년 정도면 이미 알고 있을 만한 '6하 원칙'의 의미를 묻거나 쉬운 속담의 의미를 질문하는 경우가 있었다. 예컨대, 8번 문항은 결합표집된 일반학생 전체의 정답률이 87.8%로 전체 검사 문항중 정답률이 가장 높은 문항이었고 학습장애 학생집단도 이 문항의 정답률이 가장 높았지만, 실제로는 선택지 중 '빈 수레가 요란하다', '벼는 익

을수록 고개를 숙인다'와 같은 속담의 비유적 의미가 무엇인지를 몰라 질문하는 학습장애 학생이 있었다. 이는 학습장애 학생의 낮은 검사점수가 누적적인 학습 결손에서 비롯되었음을 시사한다고 할 수 있을 것이다.

여섯째, 검사가 종료된 후 일부 학습장애 학생에게 문제지의 제시문을 소리내어 읽어 보도록 한 결과, '창공'→'창고', '왜놈들이'→'웬놈들이'와 같은 대치 오류를 자주 보였다. 또한, 방금 읽은 본문 내용에 관한 질문에 대해서도 정확한 대답을 하지 못하였다. 이는 해당 학생의 읽기 수준이 단순히 글자 읽기 즉, 해독(또는 해부호화) 수준에 머물러 있어 단어나 문장의 의미나 전체적인 주제 및 구조 파악이 가능한 의미론적 읽기 단계 즉, 독해 수준으로 발전하지 못했음을 의미한다고 해석할 수 있다. 다시 말해 학습장애 학생의 기초학습능력 즉, 읽기 능력이 학년 수준에 비해 매우 부족한 상태임을 시사한다.

#### 4. 연구의 제한점 및 제언

이상과 같은 본 연구의 시사점과 의의에도 불구하고, 본 연구에서 학습장애(LD) 학생을 조작적으로 정의한 방식을 고려하여 결과 해석에 유의해야 할 것이다. 즉, 일반적으로 학습장애가 읽기 영역뿐만 아니라, 쓰기, 셈하기 등 다양한 영역에서 발생할 수 있다는 점에서 본 연구에서는 학습장애의 전체(영역)를 포괄하지 못하였다. 또한, 일반적인 학습장애 정의에 포함되는 '기본적인 심리과정상의 장애'라는 학습장애 발생 원인을 밝혀 그에 부합하는 학습장애 학생을 연구 대상으로 삼은 것도 아니다. 덧붙여서 IQ와 같은 능력 수준에 비추어 (읽기)성취도가 교육적으로 심각한 불일치를 보이는가를 판별 기준으로 한 불일치 정의나 일정 기간의 효과성이 입증된 중재에도 불구하고 진전을 보이지 않았는지를 판별 기준으로 한 중재-반응 모델에 입각한 정의와 같은 연구 목적을 위한 조작적 정의에 기초한 '연구 목적으로 판별된 학습장애(Research-identified LD) 학생'을 연구 대상으로 했다기보다는 일선 학교에서 이미 학업 문제로 인하여 특수교육대상자로 선정되어 특수학급에 배치되어 있는 학생 중에서 부진 기준에 해당하는 학생을 연구 대상으로 삼은 일종의 '학교에서 판별한 학습장애(School-identified LD) 학생'에 가깝다고 할 수 있다. 그런 점에서, 본 연구에서의 학습장애는 엄밀하게 말하면 '특수학급 저성취 학생'을 지칭한다고 할 수 있다.

또한, 최근 우리나라 중·고등학교 특수학급의 중증·중복화 경향을 고려할 때 본 연구의 '학습장애' 학생에 학습장애 정의의 배제기준에 해당하는 정신지체 학생이 포함되었을 가능성도 높고, 누적적인 학업결손으로 인해 배경지식을 요하는 본 연구에서 사용된 검사와 같은 독해력 검사에서 별다른 점수 변화를 보이기 어려운 학생들이 포함되었을 가능성도 높다.

그런 점에서 본 연구의 결과는 현재 학업상의 어려움으로 인해 특수학급에 배치되어 있는 학생들이 검사 조건의 변화에 따라 어떠한 변화상을 보이는지에 관한 시사적인 정보로서 의의가 있다. 다만, 향후 국가 차원에서 엄밀한 기준에 의하여 '학습장애' 학생에 대하여 일정한 검사 조정을 제공하고자 할 경우 그 대상 범위를 한정하는 문제를 논의할 때 본 연구의 결과를 결정적인 근거로 단순히 활용하는 것은 과잉 일반화의 오류를 범하는 셈이 될 수 있으므로 유의해야 할 것이다.

그리고 특정한 검사 조정 조건에서 비장애학생에 비해 장애학생에게서 보다 많은 점수 향상 또는 '변별적인 점수 향상'이 나타났다는 이유만으로 해당 검사 조정이 '적절'하다거나 정당화되는 것은 아니라는 점도 간과해서는 안될 것이다. 서두에서 밝힌 바와 같이 본 연구는 Phillips(1994)의 5가지 질문 중 세 번째 질문에만 초점을 맞추어 수행되었다. 다시 말해서 검사 조정의 필요성과 타당성과 관련하여 Phillips(1994)가 제시한 5가지 질문 중 세 번째 질문을 제외한 나머지 질문들에 대하여 모두 '아니오'라고 대답할 수 있어야 비장애학생에 비해 장애학생에게서만 나타나는 '변별적인 점수 향상'을 이유로 그 검사 조정 방법이 '적절'한 것이라고 말할 수 있을 것이기 때문이다. 하물며 당사자에게 미치는 영향이 크고 중대한 의사결정과 관련되는 고부담 시험(high-stakes test)에서는 검사 조정의 타당성 못지않게 해당 검사 조정 방법의 하위 집단간 공평성에 대한 요구가 커진다고 볼 때, 비장애학생이나 그 학부모가 납득할 수 있도록 장애학생에게 허용되는 검사 조정의 구체적인 방법(예, 검사 시간의 연장, 대필, 대독 등)이 적절하게 이루어져야 한다는 것은 굳이 부연할 필요조차 없다고 할 것이다.

또한, 본 연구의 결과를 학습장애 학생들에게는 무조건 대독이나 대독+시간연장과 같은 검사 조정이 필요하다는 식으로 지나치게 확대해석하는 것도 바람직하지 않다. 본 연구에서와 같이 특정 검사의 목적이 피험자의 읽기 능력 특히 독해 능력을 측정하는 데 그 목적이 있는 경우, 대독(녹음테이프)과 같은 검사 조정은 검사의 본래 목적을 해치고 검사 결과의 타당성을 저해하는 요소로 작용할 수도 있음을 간과해서는 안된다. 왜냐하면 당초에 검사의 목적이 독해력의 측정에 있는 검사에서 대독 서비스를 제공해주는 경우, 그 검사의 성격이 독해력 검사가 아니라 청해 능력 검사로 변질될 가능성이 있어 결과적으로 당초에 측정하려고 했던 것과는 다른 능력을 재는 셈이 될 수 있기 때문이다. 그래서 최종근(2005)은 학습장애 학생에 대한 읽기 검사에서 대독의 남용이 가져올 수 있는 문제점에 대하여 이미 다음과 같이 경계한 바 있다.

학습장애 학생들은 대체적으로 주의집중력이 떨어지고, 지속시간도 짧은 편이었다. 선생님이나 동료의 사소한 움직임에도 언어적 반응을 하거나 주위가 산만해지는 경향이 있었다. 심지어는 녹음

테이프를 통해 제시문이 들려오는 동안 현재의 제시문과 무관한 다음 문제를 풀이하는 학생도 있고, 녹음테이프에서 제시문을 읽는 동안 시험지를 보지 않고 듣기만 하는 학생도 있었다. 그런 점에서 녹음테이프로 문제를 읽어주는 검사 조정 방법은 주의집중력이 떨어지는 학습장애 학생이 검사에 집중하게 하는 장점이 있다. 그러나 본문을 읽는 동안 자연스럽게 띄어 읽기를 해주는 셈이 되어 문장의 의미를 파악하기 쉽게 해주어 검사의 본래 의도를 훼손할 가능성이 크고, 심하게는 독해능력을 측정하는 검사를 청해능력 검사로 변질시킬 우려도 있음을 간과해서는 안 될 것이다(최종근, 2005, p. 89).

현재 우리나라에서는 학습장애(LD)의 개념적, 조작적 정의가 분명하게 합의되어 있다고 볼 수도 없다. 학습장애가 워낙 이질적이고 다양한 학업상의 문제를 포괄하고 있기 때문에 그와 같은 정의의 합의가 어려울 수밖에 없다고 말할 수도 있을 것이다. 문제는 그 이유야 어떻든 그러한 합의가 없기 때문에 ‘학습장애’ 학생에 대한 적절한 검사 조정 방안을 마련하기가 어렵고, 그로 인해 대학수학능력시험과 같은 대규모 고부담시험에서조차 ‘학습장애’ 학생에 대한 검사 조정이 공식적으로 이루어지지 못하고 있다는 것이다. 하지만, 적어도 학교 현장에서 저성취(특히, ‘중추신경계통의 장애’ 또는 ‘심리과정상의 장애’ 등 그 원인이 분명치 않아 ‘학습장애’ 판정이 어려운 경우) 및 학습장애 학생들이 국가와 학교가 정한 교육과정 목표를 어느 정도 달성하고 있는지 그 지식이나 기술(skill) 수준을 있는 그대로 파악하고자 한다면, 각 교과나 검사의 목적을 해치지 않는 범위 내에서 해당 학생이 지닌 ‘장애’의 영향은 받지 않은 진정한 의미의 ‘학업성취도(검사 점수 또는 수행 능력)’를 파악할 수 있는 환경으로서 적절한 검사 조정이 신중하게 검토될 필요가 있을 것이다. 이러한 적극적인 평가와 점검 활동을 통하여 이들에게 이루어진 일정한 교육 활동의 성과를 가늠하고 개선할 점을 발견하여 해당 학생에 대한 추후 교육 활동의 개선을 위한 자료로 활용할 수 있을 것이기 때문이다.

본 연구는 바로 이러한 필요성을 실제 데이터를 통해 확인하고자 시도했다는 데 그 의의가 있다. 그러나, 본 연구의 이와 같은 의의에도 불구하고, 앞서 언급한 바와 같이, 본 연구를 기초로 읽기뿐만 아니라, 쓰기, 수학, 사회 및 과학교과 등 다양한 교과와 다양한 학년, 다양한 문항 유형들에 대한 이와 같은 연구가 추후에도 많이 이루어질 필요가 있다는 것은 아무리 강조해도 여전히 부족함이 많을 것이다.

## 참고문헌

- 김계현 (2000). 상담심리학 연구 : 주제론과 방법론. 서울: 학지사.
- 김동일, 최종근 (2003). 시간 연장을 중심으로 살펴 본 대입시험 검사 조정의 현황과 연구 전망. *교육학연구*, 41(4), 107-130.
- 박효정, 이재분, 현 주 (2001). 중학교 기초·기본 학습 부진학생 판별도구 개발 연구. 한국교육개발원 수탁연구 CR 2001-13.
- 최종근 (2005). 장애학생을 위한 학업성취도 검사 조정의 타당성 연구: 시각장애, 청각장애, 지체장애, 학습장애 학생과 일반 학생의 검사 비교가능성을 중심으로. 박사학위 논문. 서울대학교.
- Burns, Edward (1998). *Test accommodations for students with disabilities*. Illinois: Charles C Thomas Publisher LTD.
- Camara, W. J., & Schneider, D. (2000). *Testing with extended time on the SAT I : effects for students with learning disabilities*. The college board research notes RN-08. New York: College Entrance Examination Board-Office of Research and Development.
- Ofiesh, N. S., Hughes, C., & Scott, S. S. (2004). Extended test time and postsecondary students with learning disabilities: a model for decision making. *Learning Disabilities Research & Practice*, 19(1), 57-70.
- Phillips, S. E. (1994). High-stakes testing accommodations: Validity versus disabled rights, *Applied Measurement in Education*, 7(2), 93-120.
- Salvia, J., & Ysseldyke, J. E. (1995). *Assessment*(6th Ed.). Boston: Houghton Mifflin.
- Willingham, W. W., Ragosta, M., Bennett, R. E., Braun, H., Rock, D. A., & Powers, D. E. (1988). *Testing handicapped people*. Boston, MA: Allyn and Bacon, Inc.
- Ziomek, R. L., & Andrews, K. M. (1998). *ACT assessment score gains of special-tested students who tested at least twice*. ACT research report series 98-8. Iowa City, IA: ACT.

- \* 논문접수 2005년 8월 5일 / 1차 심사 2005년 8월 23일 / 2차 심사 2005년 9월 2일
- \* 최종근: 서울대학교 교육학과를 졸업하고, 동대학원에서 석사학위를 취득하였으며, 동대학원에서 특수교육전공으로 박사학위를 취득하였다. 현재 서울대학교 입학관리본부 전문위원으로 재직중이고, 연구 관심 분야는 학습장애, 국기수준의 학업성취도 평가, 장애학생의 평가 참여 및 검사 조정, 통합교육 등이며, 주요논문으로 '학습장애 조기선별을 위한 읽기 전 능력검사의 현황과 개발방향' 등이 있다.
- \* e-mail: jkchoi@snu.ac.kr
- \* 김동일: 서울대학교 교육학과 및 동대학원을 졸업하고 미국 미네소타대학교 교육심리학과에서 학습장애전공으로 석사, 박사학위를 취득하였다. 현재 서울대학교 교육학과 교수로 재직중이며, 주요 저서로는 『학습장애아동의 이해와 교육』, 『특수아동상담』 등이 있다.
- \* e-mail: dikimedu@snu.ac.kr

Abstract

## Impacts of Test Accommodations for the Students with Learning Disabilities and Low Achievers

Choi, Jong-Keun\* · Kim, Dong-Il\*\*

The purpose of this study was to investigate the necessity of test accommodations for the students with learning disabilities(LD) or low achievers(LA). In pursuit of this purpose, 2nd grade middle school 527 regular class students and 23 special class students took reading comprehension test twice. At first in condition of standard paper-pencil test, at second in standard or one of three accommodated test condition, which involved time extension, reading(tape-recorded test), and time extension + reading(tape-recorded test).

Comparing the gain scores of general students, low achievers, and students with LD, possibility of 'differential boost' was detected in some accommodated test condition. Although the effect size was no so big, partially the bigger gain score for the LA or students with LD than that of general students was interpreted that test accommodation like time extension or reading could be positively considered for the LA or students with LD. Finally, several implications of several test-taking behavior of students with LD, several limitations and matters to be attended to was discussed.

Key words: test modification or accommodation(TA), differential boost, low achievers (LA), learning disabilities(LD)

---

\* Lecturer, Seoul National University

\*\* Professor, Dept. of Education, Seoul National University