

학업성장 정보 산출을 위한 수직척도 활용 방안 고찰¹⁾

송 미 영(한국교육과정평가원 연구위원)*

이 현 숙(건국대학교 부교수)**

《 요 약 》

2008년부터 국가수준 학업성취도 평가가 전수평가로 전환되면서 동일 학생들에 대한 초6·중3·고2, 세 시점의 학업성취도 자료를 종단적으로 수집할 수 있는 틀이 마련되었다. 이와 같이 수집된 학생수준의 종단자료를 이용하면 적어도 두 시점 이상에서 학생들의 학업성취도 변화를 종합적으로 파악하여 교육의 질적 향상과 교육의 책무성 제고를 위한 교육적 지원 및 정책 결정의 근거 자료를 산출할 수 있다. 학생수준의 종단자료를 활용하여 학업성장 정보를 다양한 접근 방법으로 유연하고 정교하게 분석하기 위해서는 평가시점별 학업성취도 점수가 수직적으로 연계된 점수체계 기반이 요구된다. 본 연구에서는 향후 국가수준 학업성취도 평가의 초-중-고 간 수직척도가 개발되어 본격적으로 적용되면 수직척도화된 종단자료를 어떻게 분석하고 해석하여 보고할 것인지와 각 모형을 적용할 경우 고려해야 할 사항들을 이론적으로 고찰하고 학생 및 학교수준에서 효과적인 학업성장 정보를 산출할 수 있는 방안을 제안하였다. 또한 국가단위의 대규모 평가에 대한 수직척도 개발 연구의 초기 단계인 현 시점에서 향후 수직척도의 효과적인 활용을 위해 고려해야 할 사항들에 관해 논의하였다.

주제어 : 수직척도, 학업성장, 국가수준 학업성취도 평가

1) 이 논문은 한국교육과정평가원과 한국교육평가학회가 공동주최한 '개별 학생의 학업성장과 학교교육 성과 분석 방안'(2011. 12.) 세미나(RRE 2011-6-2)에서 발표된 원고를 수정·보완한 것임

* 제1저자

** 교신저자, hyunsi@konkuk.ac.kr

I. 서론

학교교육의 질을 관리하고 학생들의 학업에 관한 유용한 정보를 제공하기 위하여 우리나라에서는 국가수준에서 대규모 학업성취도 평가를 주기적으로 실시하고 있다. 매년 초등학교 6학년과 중학교 3학년, 고등학교 2학년 학생을 대상으로 시행하고 있는 국가수준 학업성취도 평가(이하 학업성취도 평가)는 학교교육의 결과를 체계적으로 점검하고 관리하는 평가 체제이다. 2008년부터 학업성취도 평가가 전수평가로 전환되면서 평가의 성격과 목적에 있어서 변화가 있었다. 기존의 표집평가에서는 학업성취도의 연도별 변화 추이를 파악함으로써 국가수준에서 교육과정의 실행을 점검하고, 이를 바탕으로 학교교육의 질을 개선하고자 하는 데 주요한 관심을 두었던 데에 반하여, 전수평가 시행 후에는 개별 학생의 성취수준을 진단, 보정하고, 개별 학교의 교육 책무성을 점검, 지원하는 목적이 추가되었다(정은영 외, 2009; 김성숙 외, 2010). 이와 같이 평가의 성격과 목적이 변화함에 따라, 개별 학생의 학업 능력에 대한 진단을 통하여 학력향상을 위한 정보를 제공하거나 이러한 정보를 바탕으로 교육정책의 효과에 대한 직접적인 점검을 하는 것이 가능하게 되었다(김양분, 임현정, 이광현, 신혜숙, 2009).

매년 초등학교 6학년과 중학교 3학년, 고등학교 2학년의 전체 학생을 대상으로 학업성취도 평가가 실시되면서, 학업성취도 평가의 종단자료는 학교수준 및 학생수준에서 수집되고 있다. 2011년부터는 2009년에 중3이었던 학생들이 고2가 되면서 중학교와 고등학교 간에 개별 학생 수준에서의 데이터가 종단적으로 연계되었고, 2009년에 초6이었던 학생들이 고2가 되는 시점인 2014년부터는 초-중-고 3개 학교급을 잇는 학생수준의 종단자료가 축적된다. 이와 같은 학생수준 종단자료의 축적으로, 기존에 학교 차원에서 성취수준별 비율에 대한 연도별 변화 추이를 코호트 간에 비교하던 것을 넘어서, 개별 학생이 초-중-고 교육과정을 거치면서 이룬 학업성취의 변화 정도를 파악하는 것이 가능하게 되었다(김경희 외, 2011). 또한 학교별 향상도를 산출하는 데 있어서도 여러 시점에 걸친 개별 학생들의 성취도 변화를 학교수준에서 통합하여 파악하는 것이 가능하게 됨으로써, 코호트 간 변화에 기초한 향상도에 비하여 더 정교한 학업성장 정보를 제공할 수 있게 되었다.

학생수준의 종단자료가 축적되더라도 현재 학업성취도 평가의 점수 척도가 학교급별로 각각 개발되어 서로 독립적이기 때문에, 절대적인 의미에서 학생 개인별 성장을 추적하거나 이를 기반으로 학교별 향상도를 분석하는 데에 제한점이 발생할 수 있다. 현재의 학업성취도 평가 점수를 이용하여 학교별 향상도를 산출하고 학교교육의 성과를 분석할 수는 있다. 그러나 학년간 점수 척도를 연계하는 수직척도화가 이루어지지 않은 경우에는 평가시점별 점수간 단순 차이로서 개별 학생의 학업성장을 측정할 수 없기 때문에 다소 복잡한 통계적 과정을 통해 학업성장 정보를 산출해야 하며, 학업성장의 크기에 대한 질적 의미를 해석하는 것 역시 간단한 문제가 아니

다. 학생수준의 종단자료를 활용하여 다양한 접근 방법에 의해 유연하고 정교하게 학업성장 정보를 분석하기 위해서는 평가시점별 학업성취도 점수가 수직적으로 연계된 점수체계, 즉 수직척도(vertical scale) 기반이 요구된다. 학업성장을 측정하는 데 적용할 수 있는 모형에 따라 수직척도의 필요 여부는 다르지만, 수직척도가 존재한다면 개별 학생의 학업성장 및 학교의 평균적인 향상도를 산출하는 데 있어서 통계적 분석 모형 활용의 유연성이 높아진다는 장점이 있다. 이에 여러 학년에 걸쳐 학생들의 성취도를 상호 비교할 수 있도록 학년 간 학업성취도 평가 점수를 측정학적으로 연계하는 수직척도화(vertical scaling) 연구가 수행되고 있다(예: 이상하, 김수진, 신선희, 김완수, 2010; 송미영, 2011; 송미영, 남민우, 김창일, 강태훈, 2011).

여러 학년에 걸친 수직척도를 구축하기 위한 가장 기본적인 전제 조건은 각 학년에서 측정하는 구인이 동일한 선상에서 비교 가능한 내용상의 연속성을 갖추어야 한다는 점이며, 이와 관련해서 전문가의 식견에 의한 분석을 기초로 학년별 평가내용의 수직연계성(vertical alignment; Wise & Alt, 2006) 확보를 바탕으로 초6-중3간 수직척도 개발 연구가 수행된 바 있다(송미영 외 2011). 본 연구에서는 평가 내용상의 연속성이라는 전제 조건이 확보되었다는 가정 하에, 향후 학업성취도 평가의 점수체계가 초-중-고 간 수직척도를 갖추고 본격적으로 적용될 경우 수직척도화된 종단자료를 어떻게 분석하고 해석하여 보고할 것인지에 대해 이론적으로 고찰함으로써, 학업성장 정보를 산출하기에 적절한 방안을 모색하고, 향후 경험 자료를 실증적으로 분석하기 위한 기초적인 틀을 제공하고자 한다. 이를 위해, 학생수준의 종단자료를 바탕으로 개별 학생의 학업성취도 변화 및 학교교육 성과를 분석하는 데 적용할 수 있는 다양한 성장모형들을 탐색하고, 학업성취도 평가에 각 모형을 적용할 경우 고려해야 할 사항들을 검토함으로써 학업성장 정보를 산출할 수 있는 효과적인 방안을 제안하고자 한다.

Ⅱ. 학업성장 정보 산출 모형

먼저 여러 시점의 평가 결과로 축적된 학생수준 종단자료로부터 개별 학생의 학업성장에 대한 정보를 산출하는 데에 적용 가능한 모형의 특징과 적용 사례를 고찰하였다. 이 장에서 다루고 있는 모형 중 단순 성장모형과 단순 성장점수를 종속변수로 한 성장모형, 다층 성장모형, 중층적 무선효과 모형 등은 수직척도화된 종단자료에만 활용 가능한 모형이다. 반면 맥락 부가가치모형과 학생성장 백분위모형은 수직척도화 여부와 관계없이 모두 적용 가능하지만, 수직척도화된 종단자료에 적용할 경우 성장의 크기에 대한 절대적인 해석이 가능하다는 장점이 있다.

1. 단순 성장모형

일반적으로 학생 개인이나 학교의 학업성장을 논의할 때 주요 관심은 '동일한 학생 또는 동일한 학교의 학업성취도가 두 시점 간에 어느 정도 변화하였는가'와 같이 동일 대상에 대한 학업성취 변화의 절대적인 '크기'를 알아보는 데 있다(Yen, 2010). 학업성취도를 평가한 점수체계가 수직척도화가 되어 있다면 특정 기간 동안 동일 학생이 보여준 학업성취 변화 혹은 차이를 두 시점의 점수 차이로 나타내는 것이 가능하기 때문에 성장의 크기를 가장 간편하게 산출할 수 있다는 장점을 가진다(Blank, 2010; Goldschmidt et al., 2005; Patz, 2007). 예를 들어, 어떤 학생이 초등학교 6학년 때 수직척도 상에서 1,070점(중학교 입학 당시 성적으로 간주함)을 받고 3년 후에 1,190점을 받았다면 이 학생이 3년 동안 이룬 성장의 크기는 120점이 된다. 이와 같이 학생 개인별로 산출된 단순 성장점수를 평균하면, 그 학교 내 학생들의 평균적인 성장에 관한 정보를 산출할 수 있다. 예를 들어, Texas 주에서 사용하고 있는 수직척도성장(Vertical Scale Growth: VSG) 지표는 개별 학생들의 단순 성장을 기초로 한 VSG 값을 총 학생 수로 나누어 산출한 것이다(Texas Education Agency, 2010).

단순 성장점수로 두 시점 간에 학생 및 학교가 어느 정도 성장했는지를 기술할 수는 있지만 두 가지 이유에서 이를 학교의 학업성장을 나타내는 지표로 사용하는 것은 적합하지 않다. 첫째, 단순 성장점수에 내포된 성장의 원인을 구분하여 해석하는 것이 불가능하다. 어떤 중학교에 재학 중인 3학년 학생들이 초등학교 이후 3년간 평균 200점의 성장점수를 나타내었다고 할 때, 200점의 성장 중에서 학생의 자연스러운 발달에 의한 전형적인 성장과 학교와 교사의 노력으로 인한 성장을 구분하는 것은 불가능하다. 둘째, 단순 성장점수는 척도점수의 위치와 관계없이 성장의 크기만 산출하는 것이기 때문에 학교 간 수행을 비교하는 데에는 주의가 필요하다. 저성취 학생들과 고성취 학생들에 있어서 기대되는 성장의 폭이 다를 수 있으므로 점수의 위치에 따라 성장의 의미가 다를 가능성이 있기 때문이다. 이러한 점을 고려하여, 학교의 학업성장 정보를 파악하는 데에는 부가가치의 개념에 기초한 모형을 활용하는 것이 더 바람직할 것으로 보인다.

2. 부가가치의 개념을 기반으로 한 성장모형

부가가치모형(value-added model)은 학생들의 학업성취를 예측하는 데 있어서 개별 학생의 사전 성취도 점수와 배경변인 등을 통제함으로써 학생들 간에 이미 존재하는 학업성취도 차이를 조정한 상태에서 학력향상을 위해 교사나 학교가 고유하게 기여한 정도를 산출하는 통계적 모형을 의미한다(National Research Council and National Academy of Education, 2010). 부가가치모형에서는 개별 학생의 학업성취도가 종단적으로 추적된 자료를 활용하여 학

생의 학업성장 정보, 교사나 학교의 교육적 노력 정도를 추정한다.

이 때 수집된 종단자료가 수직적도화된 경우에는 3시점 이상의 반복측정자료를 이용한 성장 모형(Raudenbush & Bryk, 2002), 2시점의 자료와 측정의 표준오차를 이용한 성장모형(Bryk, Thum, Easton, & Luppescu, 1998), Tennessee 주의 교사 책무성 평가에 활용되고 있는 중층적 무선효과 모형(Sanders & Horn, 1994), 두 시점의 차이점수를 종속변수로 하여 각 학교의 부가적인 성장을 산출하는 모형(임현정, 김양분, 신혜숙, 이광현, 2010) 등 다양한 모형을 사용하는 것이 가능하다. 수직적도화가 이루어지지 않은 종단자료에는 현재 학업성취도 평가의 학교 향상도 산출에 사용하고 있는 맥락부가가치모형(Ray, Evans, & McCormack, 2009; 김경희 외, 2011)을 적용하는 것이 가능하다.

가. 다층 성장모형

다층 성장모형은 반복 측정된 종단자료를 바탕으로 일정 기간 동안 학생 개인의 학업성장을 예측 설명할 수 있는 모형으로, 이를 활용하면 개인의 성장 곡선은 어떠한지, 성장 곡선에 있어 개인차가 유의한지, 개인차에 영향을 미치는 요인은 무엇인지 등에 대한 탐색이 가능하다(김양분 외, 2009).

일반적인 다층성장모형은 모형 내에서 측정오차를 통제하기 때문에 적어도 3시점 이상의 반복측정 자료가 있어야 한다. 즉, 개인별로 서로 다른 초기수준과 성장률을 무선효과로 보고 시점별 학업성취도를 개인의 고유한 시간의 함수로 모형화(Raudenbush & Bryk, 2002)하여 분석하는 것이다. 이 때, 시간에 따른 성장의 양상이 선형적인 패턴을 보인다면 개별 학생에 대한 초기수준과 성장률을 추정함으로써 학생들 간의 성장 패턴의 차이를 파악할 수 있다. 4시점 이상의 자료가 종단적으로 축적되었다면, 성장의 패턴에 대하여 비선형적인 함수 또는 비연속적 시간설계(piecewise modeling)를 적용할 수도 있다. 그러나 수직적도 상에서 2시점에 측정된 학생들의 성취도만 활용가능한 상황이라면, 측정오차의 통제를 위하여 표준오차 정보를 모형에 포함시킴으로써 2시점의 잠재변인 성장모형(Bryk, Thum, Easton, & Luppescu, 1998)을 적용하여 성장 정보를 산출하는 것이 가능하다(김경희 외, 2011에서 재인용). 이 모형은 다음과 같은 식으로 나타낼 수 있다.

$$1\text{수준 모형: } y_{tij} = \pi_{0ij} + \pi_{1ij} \times t_{ij} + r_{tij}, \quad r_{tij} \sim N(0, SE(y_{tij})^2)$$

$$2\text{수준 모형: } \begin{aligned} \pi_{0ij} &= \beta_{00j} + r_{0ij}, \quad r_{0ij} \sim N(0, \tau_{\pi 00}) \\ \pi_{1ij} &= \beta_{10j} + b_j(\pi_{0ij} - \beta_{00j}) + r_{1ij}, \quad r_{1ij} \sim N(0, \tau_{\pi 10}) \end{aligned}$$

$$3\text{수준 모형: } \begin{aligned} \beta_{00j} &= \gamma_{000} + u_{00j}, \quad u_{00j} \sim N(0, \tau_{\beta 00}) \\ \beta_{10j} &= \gamma_{100} + b_2(\beta_{00j} - \gamma_{000}) + u_{10j}, \quad u_{10j} \sim N(0, \tau_{\beta 10}) \\ b_{1j} &= \eta_0 + \eta_1(\beta_{00j} - \gamma_{000}) + u_{bj}, \quad u_{bj} \sim N(0, \tau_b) \end{aligned}$$

1수준 모형에서 y_{tij} 는 시점 t 에 측정된 학교 j 의 학생 i 에 대한 성취도 점수, π_{0ij} 와 π_{1ij} 는 각각 학생 i 의 초기수준과 성장률을 의미한다. t_{ij} 는 측정 시점을 나타내는 변수로, 각 시점 간의 상대적인 차이를 시간 점수로 나타내어 분석에 사용한다. $SE(y_{tij})^2$ 은 학생들의 성취도 점수에 대해 추정된 오차 분산을 의미하므로, 이 식에서 추정해야 할 모수는 π_{0ij} 와 π_{1ij} 만 남게 되어 두 시점의 성취도 정보만으로도 모수 추정이 가능하게 된다. 2수준 모형에서는 각 학교의 초기수준(β_{00j})과 학교 j 에 속한 개별 학생의 초기수준을 통제한 학교 j 의 평균 성장률(β_{10j}), 그리고 학생의 초기수준과 학교의 초기수준의 차이가 학생 i 의 성장률에 미치는 효과(b_j)가 추정된다. 이 때, 학생수준의 실제 성장률과 기대 성장률의 차이를 나타내는 잔차(r_{1ij})를 학생의 학업성장 지표로 사용할 수 있다. 3수준 모형에서는 각 학교의 초기수준(β_{00j})이 학교 j 의 평균 성장률(β_{10j})의 예측변수로 투입되었으므로 γ_{100} 은 전체 학교의 평균 성장률, $\gamma_{100} + b_2(\beta_{00j} - \gamma_{000})$ 는 각 학교의 초기수준을 고려할 경우 각 학교에 대해 기대되는 성장률로 정의된다. 잔차(u_{10j})는 학교의 실제 성장률과 기대 성장률의 차이를 나타낸 것이며, 이를 학교수준 학업성장의 지표로 사용할 수 있다(김경희 외, 2011).

나. 중층적 무선효과 모형

중층적 무선효과 모형(Layered Mixed Effect Model)은 Sanders & Horn(1994)이 학생의 학업성장에 대한 누적적 정보를 바탕으로 학교효과를 추정하고자 제안한 모형으로, 학생의 성취도를 설명하는 데 있어서 시점에 따라 학교효과가 누적되는 것으로 가정함으로써 한 학생에 대한 두 시점 간의 성취도 점수 차이를 전체 평균 차이점수와 학교수준의 성장분, 그리고 학생의 고유한 차이점수로 분리하여 설명하는 모형이다. 이 모형을 사용하기 위해서는 두 시점 이상의 측정 자료가 필요하며, 각 시점별로 측정된 성취도 점수가 수직척도화 되어 있어야 한다. 중층적 무선효과 모형에서 각 시점별 성취도 점수는 다음과 같이 학교효과가 누적된 형태로 표현될 수 있다(김준엽, 2009).

$$Y_{ij}^1 = \beta_0^1 + u_j^1 + e_{ij}^1, \quad Y_{ij}^2 = \beta_0^2 + u_j^1 + u_j^2 + e_{ij}^2, \quad Y_{ij}^3 = \beta_0^3 + u_j^1 + u_j^2 + u_j^3 + e_{ij}^3$$

여기서 Y_{ij}^t 은 학교 j 에 재학 중인 학생 i 에 대하여 시점 t 에 측정된 학업성취도를, β_0^t 는 시점 t 에서 전체 학교 평균을 의미하며, u_j^t 는 시점 t 에서 학교 j 의 효과, e_{ij}^t 는 시점 t 에서 학교 j 에 속한 학생 i 의 잔차를 의미한다. 성취도 점수가 수직척도화 되어 있기 때문에 학생수준의 학업성장 정보는 다음과 같이 두 시점 간 점수 차이를 바탕으로 산출할 수 있다.

$$Y_{ij}^2 - Y_{ij}^1 = (\beta_0^2 - \beta_0^1) + u_j^2 + (e_{ij}^2 - e_{ij}^1), \quad Y_{ij}^3 - Y_{ij}^1 = (\beta_0^3 - \beta_0^1) + u_j^2 + u_j^3 + (e_{ij}^3 - e_{ij}^1)$$

이 때 $Y_{ij}^2 - Y_{ij}^1$ 는 시점 1과 시점 2 간에 학생이 이룬 성장을 의미하며, $Y_{ij}^3 - Y_{ij}^1$ 은 시점 1과 시점 3 간의 성장을 의미한다. 두 시점에서 학생 i 의 성취도 차이는 전체 학교의 평균 차이 ($\beta_0^2 - \beta_0^1$)와 두 시점 간 잔차의 차이 ($e_{ij}^{t2} - e_{ij}^{t1}$), 그리고 학교효과의 누적적인 증가분 $\sum_{k=2}^K u_j^{t_k}$ 으로 구성된다. 여기서 $\sum_{k=2}^K u_j^{t_k}$ 가 시점 t_k 까지 누적된 그 학교의 부가가치 지수이며(차성현 외, 2010), 이를 학교수준의 성장 정보로 활용할 수 있다.

중층적 무선효과 모형의 경우 이전 시점에 측정한 점수를 현재 학교의 교육효과가 반영되지 않은 사전 점수로 사용함으로써 두 시점 간 개별 학생의 성취도 증가분에 있어서 해당 학교가 고유하게 기여한 성장분을 산출할 수 있고, 학생 개인의 성적에 있어 결측치 유무에 관계없이 모든 학생들의 자료를 사용할 수 있기 때문에 학생 선택의 편향에 대해 크게 염려하지 않아도 된다는 점(Sanders, 2006), 학생 개인의 성취도가 현재 학교뿐 아니라 이전 학교에도 연계되어 있기 때문에 학교수준의 교육효과와 관련 없는 요인의 영향을 자연적으로 통제할 수 있다는 점, 그리고 학생의 학교 간 이동을 모형에서 유연하게 반영할 수 있다는 장점을 가진다(김준엽, 2009; Ballou, Sanders, & Wright, 2004). 그러나 모든 학생들에 대하여 학교급 간에 교차 분류를 실시해야 하고, 여기에 학생의 학교 간 이동을 반영하여 모형을 설정하는 작업이 더 해지면 모형의 통계적 복잡성이 매우 커진다는 점에서 실제 적용을 위해서는 현실적으로 충분한 경험적 검증이 필요하다.

다. 단순 성장점수를 종속변수로 한 부가가치 모형

두 시점 간 성취도 점수가 수직척도화되어 있으면 차이점수, 즉 단순 성장점수를 종속변수로 하여 학생 개인의 성장 정도를 이용한 학교의 성장을 추정할 수 있다. 어떤 학생이 두 시점에서 얻은 수직척도 점수의 차이를 d_{ij} 는 다음과 같이 학생수준과 학교수준에서 모형화할 수 있다(임현정 외, 2010)

$$\begin{aligned} \text{1수준 모형: } d_{ij} &= \beta_{0j} + \beta_{1j}(X_1)_{ij} + \dots + \beta_{pj}(X_p)_{ij} + r_{ij} \\ \text{2수준 모형: } \beta_{0j} &= \gamma_{00} + \gamma_{01}(Z_1)_j + \dots + \gamma_{0q}(Z_q)_j + u_{0j}, \quad u_{0j} \sim N(0, \tau_{00}) \\ \beta_{pj} &= r_{p0} \end{aligned}$$

여기서 $X_1 \sim X_p$ 는 학생수준의 독립변수, $Z_1 \sim Z_q$ 는 학교수준의 독립변수를 의미한다. β_{0j} 는 학생수준 독립변수를 통제된 상태에서 산출된 학생별 단순 성장점수의 학교 평균, r_{ij} 은 학생수준의 잔차이다. 이 모형은 수직척도를 전제로 하기 때문에 차이점수 자체가 단순 성장의 개념에 기초한 학생의 학업성장 정보로 활용될 수 있으나, 소속 학교 학생들에게 평균적으로 기대되는 성장에 비하여 학생이 부가적으로 성취한 학업성장을 산출하는 데 관심이 있다면 학생수준의 잔차를 활용하는 것이 가능하다. 또한 2수준 모형에서는 학교의 평균 차이점수 β_{0j} 가 학교수준의

배경변수를 통제한 상태에서 산출된 전체 학교의 평균 차이점수와 잔차(u_{0j})로 개념화되며, 이때 잔차는 학생 개인의 변화와 성장 정도를 고려하여 산출한 학교수준의 성장이라고 볼 수 있다.

라. 맥락 부가가치모형

맥락 부가가치모형(Contextualized Value-added Model: CVA)은 학생의 사전 성취도 및 맥락변인을 독립변수로 통제하여 얻은 사후 성취도에 대한 기대 점수와 실제 점수의 차이에 근거하여 학생 및 학교의 성장정보를 산출하는 모형으로(Ray, Evans, & McCormack, 2009), 다음과 같은 다층모형으로 정의된다.

$$y_{ij(t_2)} = \beta_0 + \beta_1 y_{ij(t_1)} + \beta_2 x_{1i} + \dots + \beta_{p+1} x_{pi} + u_j + e_{ij}$$

여기서 $y_{ij(t_2)}$ 는 학교 j에 재학 중인 학생 i에 대하여 시점 2에서 측정한 성취도를, $y_{ij(t_1)}$ 은 그 학생에 대해 시점 1에서 측정한 사전 성취도를 의미한다. $x_{1i} \sim x_{pi}$ 는 모형에서 통제하고자 하는 학생수준의 맥락변인, u_j 는 학교수준의 잔차, e_{ij} 는 학생수준의 잔차를 각각 나타내며, 잔차는 0을 중심으로 정규분포를 따르는 것으로 가정된다. 이 모형에서는 사전 성취도와 맥락변인을 투입하여 모형에서 기대되는 시점 2의 성취도와 실제 그 학생의 시점 2의 성취도 간 차이인 학생 수준의 잔차가 부가가치 개념에 기초한 학생의 성장정보로 정의되며, 학교수준의 잔차가 해당 학교의 성장정보로 정의된다.

영국의 학교별 향상도 산출에 사용하기 시작한 맥락 부가가치모형은 학생수준에서 2시점의 종단자료가 수집되면 수직척도화 여부에 관계없이 사용 가능하기 때문에, 우리나라에서도 2011년부터 학교 향상도 산출을 위해 적용하고 있다(김경희 외, 2011). 이 모형이 수직척도화된 점수에 적용될 경우, 부가가치 지수로서 산출된 잔차는 '다'항에서 소개한 단순 성장점수를 종속변수로 한 모형의 잔차와 수리적으로 동일하게 된다.

3. 학생성장 백분위 모형

학생성장 백분위(Student Growth Percentile: SGP) 모형은 학생들의 성장을 상대적인 관점에서 설명하는 모형으로, 사전 성취도가 동일한 학생들의 사후 성취도 분포 내에서 한 학생의 성취도가 어느 위치에 해당하는지를 백분위로 나타냄으로써 능력이 유사한 집단 내에서 해당 학생의 성장을 상대적으로 비교하는 모형이다. 개별 학생의 학생성장 백분위를 산출하려면, 먼저 두 시점의 성취도 점수가 이루는 이변량 결합 분포를 구한 후, 시점 1에서 동일한 수행을 보인 학생들이 시점 2에서 얻은 점수에 대한 조건 분포에서 해당 학생의 점수가 위치한 백분위를

구하면 된다. 학생성장 백분위 모형은 이처럼 상대적인 위치를 바탕으로 성장의 의미를 해석하므로 수직척도화가 요구되지는 않는다. 그러나 수직척도화된 종단자료에도 적용하는 것이 가능하며, 이러한 경우 두 시점 간의 성장 점수 차이가 가지는 크기가 절대적인 성장의 크기라는 점에서 성장에 대한 해석이 보다 용이해진다는 장점을 가진다.

학생성장 백분위(SGP)는 숙달수준(proficient level) 도달 비율과 함께 학교교육의 성과를 평가하는 근거로 활용되고 있다(Carey & Manwaring, 2011). Colorado 주에서는 학교의 책무성을 평가할 때, 해당 학교에 속한 개별 학생들의 SGP에 대한 중앙값과 그 학교의 학생들 중 당해 연도에 숙달수준에 도달한 학생들의 비율을 동시에 고려한다. 즉, SGP의 중앙값과 숙달수준 도달 비율이 둘 다 높은 학교를 효과적인 학교로 보고, 반대로 SGP의 중앙값과 숙달수준 도달 비율이 둘 다 낮은 학교를 적극적인 개입이 필요한 학교로 구분하는 것이다.

Ⅲ. 학업성장 정보의 해석 및 보고 방안

1. 학업성장에 대한 해석 기준

수직척도화된 점수체계가 마련되면 학년이 올라감에 따라 학생이 어느 정도의 성장을 이루고 있는지에 관한 정보를 수직척도점수 상의 점수 변화로 정의하고 성장의 절대적인 크기를 측정하는 것이 가능하다. 그러나 두 시점간 성장의 크기를 해석하는 데에는 해석의 기준이 필요하다. 학생 및 학교가 일정 기간 동안 달성해야 할 ‘적정(adequate)’ 수준의 성장을 보인 것인지, 다른 학생이나 학교에 비하여 어느 정도의 성장을 한 것인지와 같은 질문에 대해 의미 있는 답변을 제공하기 위해서는 준거참조적 해석이나 규준참조적 해석과 같이 성장의 크기에 대한 해석의 기준이 필요하다.

가. 학업성장의 준거참조적 해석

수직척도 상에서 각 점수대별로 학생들에게 기대할 수 있는 학력 수준을 단계적으로 기술한 정보가 있으면, 이를 통해 학생의 성장 혹은 발달에 관한 ‘절대적인’ 의미의 피드백을 제공할 수 있다. 예를 들어, Connecticut 주의 CMT(Connecticut Mastery Test) 읽기 시험의 경우 <표 1>과 같이 200점에서 700점 사이의 수직척도를 읽기 교과에서 요구되는 이해의 수준에 따라 5단계로 구분하여 단계별 특성을 정의한 수직척도 기술표(vertical scale descriptor)를 사용하고 있다. 여기서 각 단계는 특정 학년에 국한된 것이 아니라 일반적으로 해당 점수대의 학생

들이 읽기 교과에서 전형적으로 할 수 있는 것과 알고 있는 것에 대해 기술한 것이다(Connecticut Department of Education, 2010). 수직척도 기술표는 학년별 성취수준의 특성을 기술한 성취수준 기술표(performance level descriptor)와는 다른 것으로, Connecticut 주에서는 학년 간에 공통적으로 정의된 구인 상에서 발달 수준을 판정·해석하는 데에는 수직척도 기술표를 사용하고, 동학년 내에서 학생들의 성취수준을 판정·해석하는 데에는 성취수준 기술표를 사용하고 있다.

〈표 1〉 Connecticut 주의 CMT 읽기 시험의 수직척도 기술표

이해 수준	단계별 특징	수직척도 점수
5단계	이 단계의 학생들은 도움 없이도 문학 및 비문학 텍스트를 읽고 반응하는 데 뛰어난 능력을 보인다. 또한 텍스트를 이해, 해석, 평가하기 위한 복잡한 전략들을 텍스트를 읽기 전과 읽는 도중, 읽은 후에 효과적으로 사용할 수 있다.	529 이상
4단계	이 단계의 학생들은 최소한의 도움으로 문학 및 비문학 텍스트를 읽고 반응하는 데 안정된 능력을 보인다. 또한 텍스트를 이해, 해석, 평가하기 위한 전략들을 텍스트를 읽기 전과 읽는 도중, 읽은 후에 효과적으로 사용할 수 있다.	495~528
3단계	이 단계의 학생들은 어느 정도의 도움으로 문학 및 비문학 텍스트를 읽고 반응하는 데 적합한 능력을 보인다. 또한 텍스트를 이해, 해석, 평가하기 위한 전략들 중 일부를 텍스트를 읽기 전과 읽는 도중, 읽은 후에 효과적으로 사용할 수 있다.	460~494
2단계	이 단계의 학생들은 문학 및 비문학 텍스트를 읽고 반응하는 데 제한된 능력을 보이므로 읽기 과제를 수행하는 데 도움을 필요로 한다. 또한 텍스트를 이해, 해석, 평가하기 위한 일부 전략들을 텍스트를 읽기 전과 읽는 도중, 읽은 후에 제한적으로 사용할 수 있다.	425~459
1단계	이 단계의 학생들은 문학 및 비문학 텍스트를 읽고 반응하는 데 매우 제한된 능력을 보이므로 읽기 과제를 해결하기 위해서는 상당한 도움을 필요로 한다. 또한 텍스트를 이해, 해석, 평가하기 위한 전략들을 제대로 사용하지 못한다.	425 이하

학생수준의 성장뿐만 아니라 학교수준의 성장에 대한 절대적인 의미를 부여하는 데에도 이와 같은 수직척도 기술표를 이용할 수 있다. 수직척도 기술표는 각 발달 단계에 있는 학생들이 전형적으로 할 수 있는 것에 대해 기술한 것이므로, 어떤 학교에 속한 학생들의 평균적인 성장 점수의 수직척도상에서의 변화를 통해 해당 학교의 학생들이 발달 단계 상에서 ‘적정’한 성장을 보였는지 파악할 수 있다. 이와 더불어, 정부가 정의한 성장의 목표 수준이 수직척도 상에서 정해져 있는 경우에는, 이를 각 학교가 이루어야 할 ‘적정’한 성장으로 파악하여 절대적인 의미에서 성장을 해석할 수 있다.

향후 학업성취도 평가의 수직척도가 개발·적용되면 각 교과에서 측정하고자 하는 구인 정의에 대해 학생들이 어느 정도의 능력 수준 또는 이해 수준을 나타내고 있는지를 점수대별로 기술

하는 표를 준거참조적인 정보로서 제공하는 방안을 고려해볼 수 있다. 이를 위해서는 학년에 관계없이 각 교과별로 정의된 구인이 시간에 따라 어떻게 발달하는 지에 대한 단계를 수직척도 상에서 연결함으로써 수직척도 상에서 각 단계별 특성을 기술하는 것이 필요하다. 이러한 기술표가 제공되면 절대적인 관점에서 각 학년별 ‘적정’ 성장을 정의하는 것이 가능하며, 해당 학생 또는 학교가 각 교과에서 요구하는 능력 중 어느 단계에서 시작하여 어느 단계로 변화하였는지에 대한 정보를 통해 준거참조적인 관점에서 보다 풍부한 발달적 정보를 얻을 수 있을 것이다.

나. 학업성장의 표준참조적 해석

학업성장 정보를 표준참조적으로 해석하는 것은 한 학생 또는 학교의 성장점수를 해석하는데 있어서 ‘타인 준거’에 기초하여 해석하는 것을 의미한다. 학생이나 학교의 학업성장 정보를 산출하는 데 있어서 단순 성장점수가 국가 표준집단이나 여러 하위 표준집단 내에서 어느 정도에 위치해 있는지를 상대적으로 파악하거나, 부가가치의 개념에 근거하여 학생 또는 학교에 기대되는 성장에 비해 부가적으로 성장한 정도를 파악하는 것이 학업성장에 대한 표준참조적 해석의 대표적인 예라고 할 수 있다.

단순 성장점수를 표준참조적으로 해석하려면 학생 또는 학교가 소속된 집단의 규준에 대한 정보가 필요하다. 즉, 소속 집단의 단순 성장점수 평균 및 표준편차에 대한 정보를 바탕으로 해당 학생이나 학교의 성장점수에 대한 표준화 점수를 계산하거나, 규준 집단 내에서 해당 학생이나 학교의 성장점수보다 낮은 점수를 받은 학생의 비율을 바탕으로 백분위 점수를 산출하면 그 학생이나 학교의 성장점수가 가지는 의미를 집단 내에서의 상대적인 위치를 바탕으로 산출할 수 있다. 이 때 규준이 되는 집단은 국가, 지역, 학교, 성별 등 제공하고자 하는 정보에 따라 다양하게 정의될 수 있다. 또한 같은 크기의 성장이라도 1,100점에서 600점의 성장과 1,400점에서 600점의 성장은 그 의미가 다를 수 있으므로 출발점이 유사한 수준에 있는 학생이나 학교들끼리 비교하는 것 역시 단순 성장점수에 대한 표준참조적 해석의 다른 예라고 볼 수 있는데, 이전 시점에서 동일한 성장점수를 가진 학생 또는 학교들을 규준집단으로 하여 조건부 백분위에 해당하는 성장 백분위(growth percentile)를 산출하면 출발점의 차이를 고려한 규준적 정보를 얻을 수 있다.

부가가치의 개념을 기반으로 한 학생 및 학교의 성장점수는 모형에 따라 산출하는 방식이 다르지만 대체로 기대되는 성장에 비한 상대적인 성장의 개념으로 해석된다는 점에서는 유사하다. 예를 들어, 다층 성장모형을 적용하는 경우 학생에게 기대되는 성장률과 그 학생이 실제로 나타낸 성장률 간의 차이를 학생수준의 학업성장 지표로 사용할 수 있으며, 3수준 모형에서 개별 학교와 전체 학교의 초기수준 차이를 통제한 상태에서 전체 학교의 평균 성장률을 바탕으로 기대되는 성장률과 실제 그 학교의 성장률 간 잔차를 학교수준의 학업성장 지표로 사용할 수 있다.

단순 성장점수를 종속변수로 한 성장모형과 맥락 부가가치 모형의 경우도, 모형에 의해 예측되는 점수와 실제 점수 간의 차이를 바탕으로 산출된 학생수준 잔차와 학교수준 잔차를 각각 학생 및 학교의 성장정보로 사용할 수 있다는 점에서 역시 규준참조적이라고 할 수 있다. 중층적 무선폭도 모형은 학생의 학업성장 정보를 산출하는 데에는 수직척도 상에서의 두 시점 간 단순 차이점수를 이용한다는 점에서 규준참조적이라고 볼 수 없다. 그러나 학생수준의 차이점수를 전체 평균 차이점수와 학교수준의 성장분, 그리고 학생의 고유한 차이점수로 분리하여 모형화함으로써 학교수준의 성장분을 해당학교의 성장에 대한 지표로 사용할 수 있으며, 이 때 학교수준의 성장분이 전체 평균에서 그 학교가 부가적으로 성취한 점수로 정의되기 때문에 규준참조적 기준에 의한 성장 정보라고 할 수 있다.

이와 같이 규준참조적 해석 기준에 의해 학생 및 학교의 성장 정보를 해석하면 학생 및 학교가 소속된 집단 내에서의 상대적 위치에 대한 정보를 바탕으로 성장에 대한 부가적인 해석을 제공할 수 있다는 점에서, 상대적 '비교'에 의한 교육의 수월성 추구라는 관점에서는 매력적인 정보로 기능할 수 있다. 그러나 평균적으로 기대되는 성장에 비해 더 많은 혹은 적은 성장을 이루었다는 것 자체로 그 학생 또는 학교가 적정한 성장을 이루었는지에 대한 근거가 될 수는 없으며, 전체 학생 또는 학교의 절반은 항상 기대되는 성장보다 부족한 성장을 이룬 것으로 평가되어 불필요한 무한경쟁을 초래할 수 있다는 점에서 규준참조적 해석 기준에 의한 성장정보만으로 학생 및 학교의 학업성장을 파악하는 것은 제한적이라고 할 수 있다. 따라서 이와 같은 규준참조적 해석과 더불어 해당 학생 또는 학교가 구인발달 단계상에서 '적정'한 성장을 이루었는지에 대한 준거참조적 해석이 뒷받침되어야 수직척도상의 점수를 의미 있게 해석할 수 있을 것이다.

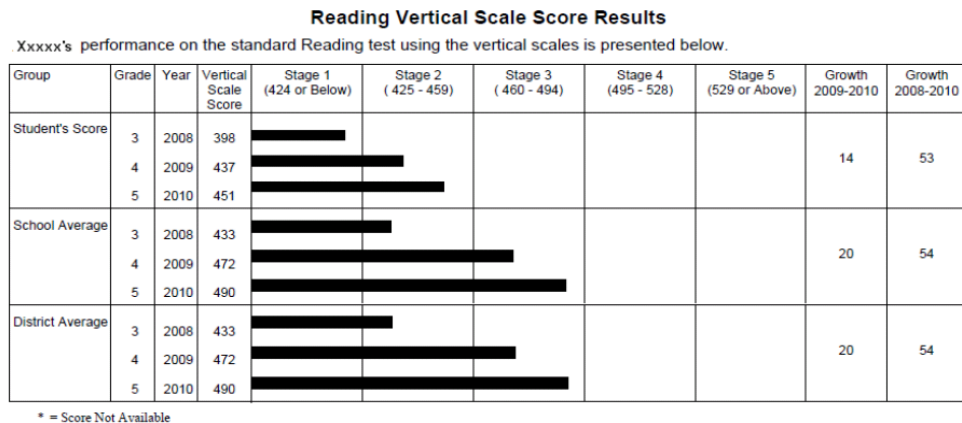
2. 학업성장 정보 보고 방안

가. 학생수준의 학업성장 정보 보고

수직척도화된 종단자료를 바탕으로 개별 학생의 학업성장을 기술하고자 할 때 사용할 수 있는 가장 명료하고 직접적인 정보는 단순 성장을 바탕으로 한 정보이다. 학생수준에서 보고되는 학업성장 정보는 기본적으로 단순 성장점수에 대한 준거참조적 정보 및 규준참조적 정보를 기초로 하는 것을 고려할 수 있다.

[그림 1]은 Connecticut 주에서 실시하는 CMT 읽기 시험의 수직척도 상에서 보고된 학생 성적표의 예시이다(Connecticut Department of Education, 2010). 학생 성적표에는 해당 학생의 각 학년별 수직척도 점수가 5단계의 발달 단계 상에서 상대적인 위치로 표현되어 있으며, 비교 연도 간에 학생이 달성한 성장점수가 수직척도 상에서의 차이점수로 산출되어 학교 및 지역 평균과 함께 제시되어 있다. 각 학생의 현재 수직척도 점수 및 두 시점 간 성장점수는 학

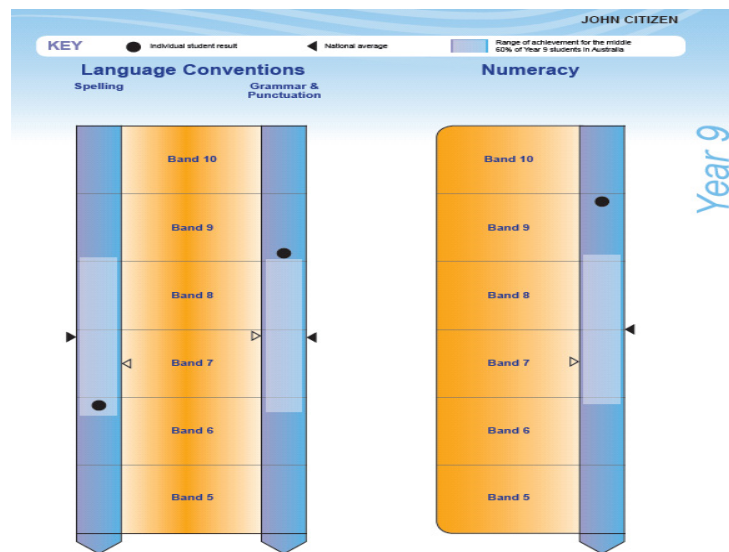
교평균 및 지역평균과 비교하여 규준참조적으로 해석된다. 1단계(stage 1)에서 5단계(stage 5)는 각각 해당 과목에 대한 학생들의 이해 수준을 의미하는 것으로, 수직척도의 단계별로 학생들에게 기대되는 성취를 기술한 수직척도 기술표(〈표 1〉 참조)에 근거하여 준거참조적으로 해석할 수 있다.



〔그림 1〕 Connecticut 주 CMT 읽기 시험의 수직척도점수에 대한 학생 성적표 일부

호주의 국가수준 학업성취도 평가(National Assessment Program - Literacy and Numeracy: NAPLAN)에서도 유사한 형태의 규준참조적 정보를 제공하고 있다. [그림 2]는 NAPLAN의 학생 성적표 예시로, 시험 영역별 척도점수, 동일 학년 학생들의 성취도와 비교한 학생의 성취도, 기초학력 수준 도달 여부 등이 시각적으로 제시되어 있다(Australian Curriculum, Assessment and Reporting Authority, 2010). 그림에서 검은색 원으로 표시된 부분은 학생이 받은 점수의 위치이며, 검은색 삼각형으로 표시된 부분이 국가평균, 흰색 삼각형으로 표시된 부분이 학교평균이다. 음영으로 된 구간은 9학년 국가평균을 중심으로 전체 학생 중 60%가 속해 있는 구간을 의미한다. 이 학생은 'Grammar & Punctuation'과 'Numeracy'에 있어서 국가 및 학교평균에 비해 매우 높은 성취를 보이고 있음을 알 수 있으며, 9학년 기초학력 수준으로 설정된 Band 6에 비교할 때, 'Spelling'을 제외한 나머지 영역에서 기초학력에 도달했음을 알 수 있다.

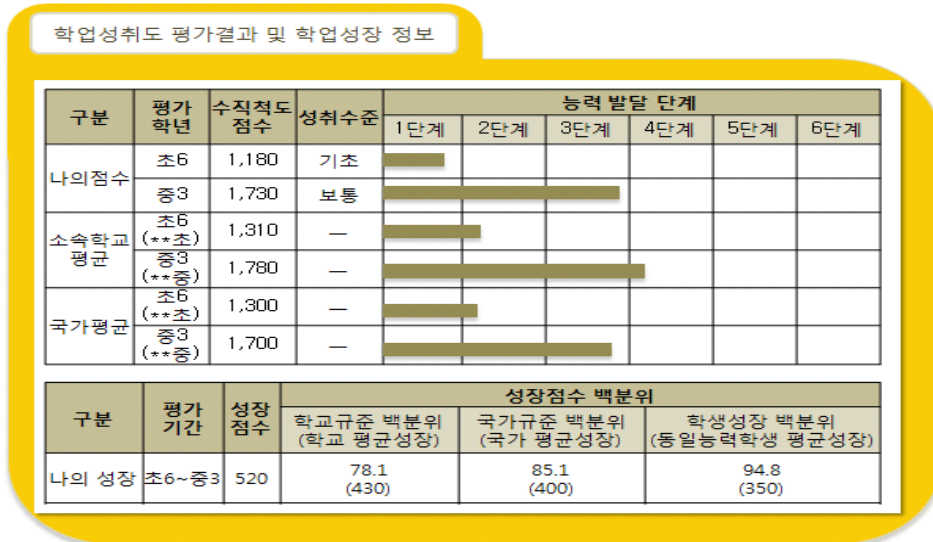
학업성취도 평가 결과에 대해 학생수준에서 보고되는 학업성장 정보도 이와 같이 현시점에서의 학업 성취수준과 과거로부터의 성장에 대한 진단적 정보를 상세히 기술하는 것에 초점을 두어 제시하는 것이 바람직할 것으로 보인다. 초등학생은 성장에 대한 정보가 없기 때문에 기존에 제공되던 학업성취도 평가 결과 제시 방식을 그대로 사용하고, 중학생과 고등학생은 기존에 제공되던 정보 외에 학업성장에 대한 정보를 추가로 제시하는 방식이 적합할 것으로 판단된다.



[그림 2] 호주 NAPLAN 시험의 학생 성적표 일부

[그림 3]은 본 연구에서 제안한 중학생용 학업성장 보고 방안을 예시적으로 나타낸 것이다. 학생에게 제공되는 성적표는 학생의 현 시점 및 과거 시점에서의 성취도 정보를 누가적으로 제공하는 부분과 성장정보를 제시하는 부분으로 구성하였다. 학업성취도를 누가 기록하는 부분에는 과거와 현 시점에서 얻은 학업성취도의 수직척도점수와 성취수준이 학교 및 국가 평균과 함께 기록된다. 각 시점별 수직척도점수는 ‘능력발달 단계’상에서 시각적으로 제시된다. 이 때 ‘능력발달 단계’는 각 단계에 속한 학생들이 어떠한 능력을 갖추고 있는지를 해당 구인에 대한 발달의 측면에서 정의하여 제공한 것으로 <표 1>의 수직척도 기술표 상에 정의된 발달 단계와 유사한 것이다. 구인 정의에 기초한 학업 능력 발달 단계는 추후 각 교과별 심층 연구에 기초하여 결정되어야 하므로, 본 연구에서는 예시적인 목적으로 편의상 초6에서 고2까지 6개 학년에 걸쳐 전형적으로 기대되는 능력 수준에 기초하여 6단계가 정의된 것으로 가정하였다. 각 단계는 각 평가 시점에서 학생들에게 전형적으로 기대되는 ‘적정’ 발달을 정의한 것이므로 과거와 현재에 해당 학생이 ‘적정’하게 성장하고 있는지에 대한 절대적인 해석의 준거를 제공한다.

[그림 3]의 하단에 제시된 학업성장 정보는 학교 및 국가규준 내에서 개별 학생의 성장에 대한 위치를 백분위로 나타내었으며, 동일한 크기의 성장이라 하더라도 출발점 정보에 따라 성장의 의미가 다를 수 있음을 고려하여 학생성장 백분위(SGP) 정보를 통해 과거에 동일한 성취수준을 가진 학생들 중 자신의 성장이 어느 정도의 위치에 있는지를 상대적으로 파악할 수 있는 정보도 함께 제시하였다.



[그림 3] 국가수준 학업성취도 평가에 대한 학생수준 학업성장 보고 방안 예시(중학생용)


나. 학급 또는 학교수준의 학업성장 정보 보고

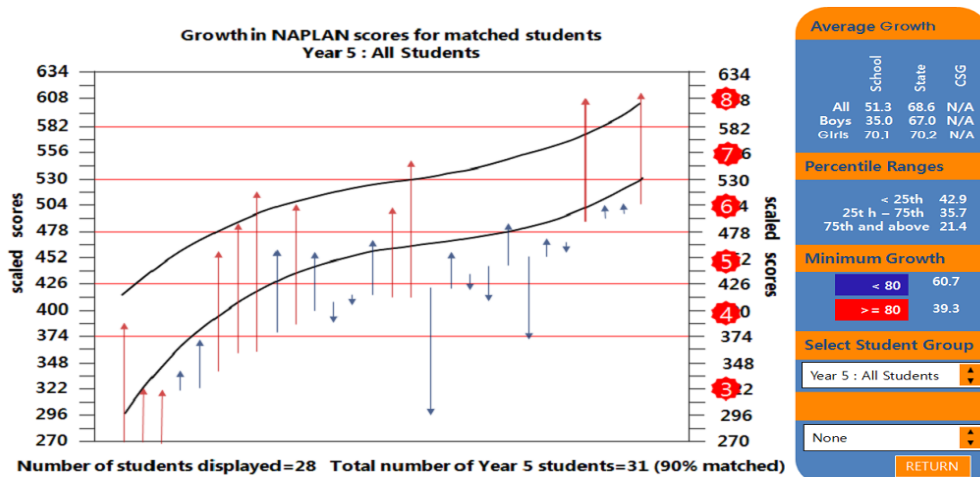
학급 또는 학교수준에서 보고되는 학업성장 정보는 학생수준의 종단자료를 바탕으로 생성할 수 있으며, 학생수준의 학업성장 정보와 유사하게 준거참조적 기준과 규준참조적 기준에 근거하여 해당 학교내 학생들의 현시점의 성취수준에 기초한 수행 및 과거로부터의 성장에 기초한 수행이라는 두 가지 측면에서 제공하는 것을 고려할 수 있다.

학생수준의 종단자료에 기초하여 학생들의 학업성장을 개별 학교에 보고하는 국외 사례를 보면, 대체로 단순 성장에 기초한 정보와 부가가치 개념에 근거한 정보를 둘 다 포함하고 있다. 예를 들어 호주 NAPLAN 시험의 교사용 보고서 중 일부를 발췌한 [그림 4](Australian Curriculum, Assessment, and Reporting Authority, 2010)의 상단에 제시된 바와 같이 교사들은 전체 학생들의 연도별 성적을 수직적도에서 나타낸 것과 두 시점 간의 차이점수를 학생별로 볼 수 있도록 되어 있다. 또한 그림의 하단에 제시된 바와 같이, 시점 1에서의 성취도 점수를 기준으로 정렬된 학생들에 대한 두 시점 간 성장의 크기와 방향을 화살표로 나타내어 학생별로 이룬 성장을 시각화한 정보도 제공된다. 오른쪽에 제시된 학교 평균과 주 평균, 백분위 정보 등은 개별 학생들의 성장 및 학급의 성장점수에 대한 규준참조적인 해석 기준으로써 제시되고 있다.

미국 Pennsylvania 주에서는 [그림 5]의 예시와 같이 연도별로 해당 학교에 대한 수직적도 점수의 평균과 백분위, 학교효과 지수(부가가치 지수)와 표준오차, 그리고 동일한 시험을 치른

집단의 평균과 비교할 때 해당 학교의 부가가치 점수가 차지하는 위치를 3등급으로 구분한 값 등을 보고한다(McCaffrey & Hamilton, 2007).

 Local School - Year 5				
2008				
Table 6LI: Overall Literacy Progress of Year 5 Students				No. of Students: 27
Students in Order of Year 3 Scores				
OVERALL LITERACY				
Student Name		2006	2008	Difference
Hayley	Sanders	207	312	105
Seth	Whiteley	212	380	168



[그림 4] 호주 NAPLAN 시험의 교사용 보고서 일부

Test	Grade	Year	N	Mean Student Score	Mean Score %tile	Mean Pred Score	Pred. Score %tile	School Effect	Effect Std Err	School vs Testing Pool Avg	Testing Pool
Math	5	2001	20	1513.0	87	1433.8	73	71.9	25.55	Above	TerraNova
		2002	21	1540.1	86	1474.4	76	54.6	24.06	Above	TerraNova

RAND TR506-1.2

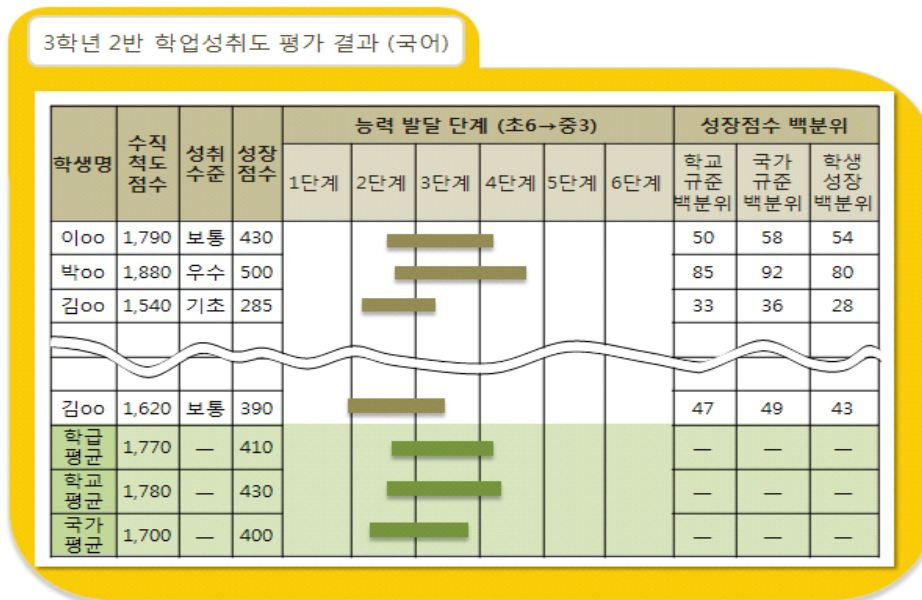
[그림 5] 미국 Pennsylvania 주 학교 부가가치 평가 보고서 일부

[그림 6]과 [그림 7]은 학업성취도 평가 결과에 수직척도가 도입될 경우, 본 연구에서 제안하는 학급 또는 학교 성장 정보의 보고 방안을 각각 예시적으로 나타낸 것이다. [그림 6]은 담

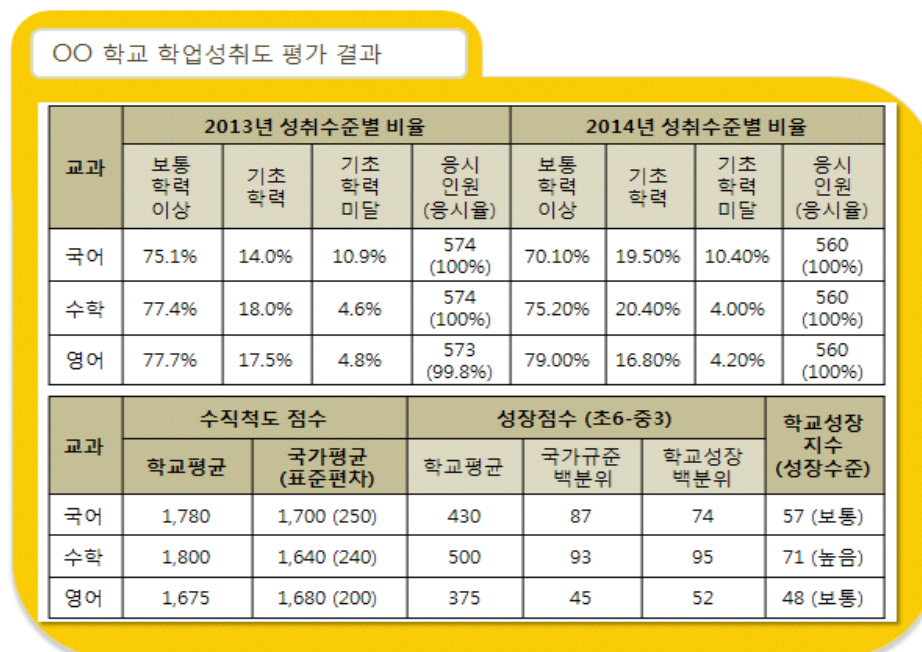
임교사에게 제공될 학급별 보고서로 개별 학생들의 수직척도 점수 및 성장정보, 그리고 이를 학급 전체에 대해 요약한 부분으로 구성되어 있다. 그림의 상단에 제시된 정보는 단순 성장점수에 기초한 정보로서 먼저 학생 개인별 성취도 및 성장점수가 수직척도 상에 제시되어 있고, 성장의 절대적인 의미를 파악할 수 있도록 평가 시점 (초6~중3) 간의 변화가 해당 교과와 구인에 대한 6단계의 능력 발달단계 상에 시각적으로 표시되어 있다. 그림의 좌측 하단에는 해당 학급 및 학교, 국가의 수직척도 평균과 성장점수 평균이 제시되어 있어 학급의 평균적인 성장에 대한 기준 참조적 정보로 활용할 수 있도록 하였으며, 성장평균을 6단계의 능력 발달단계 상에 시각적으로 표시함으로써 준거참조적 관점에서 적정 성장 여부를 해석할 수 있도록 하였다. 그림의 우측에는 개별 학생의 성장점수에 대한 학교 및 국가 기준 내의 백분위와 초등학교 6학년때 동일한 수직척도 점수를 받았던 학생들을 대상으로 산출한 학생성장 백분위 정보를 제공함으로써 기준집단 내에서 해당 학급 학생들의 상대적인 위치를 파악할 수 있도록 하였다.

개별 학교에 제공될 학업성장 보고서를 예시한 [그림 7]의 상단에 제시된 표는 현재 학업성취도 평가 결과에 대해 공시되고 있는 성취수준별 비율을 나타내며, 그림의 하단에 제시된 표는 해당 학교의 수직척도 점수 및 성장점수에 대한 교과별 평균 및 기준집단의 평균과 표준편차, 그리고 성장점수에 대한 기준참조적 해석을 할 수 있도록 국가기준 백분위와 학교성장 백분위 정보 및 부가가치에 기초한 학교성장 지수 등을 포함하고 있다. 학교성장 백분위는 이전 시점에서 측정한 점수가 비슷한 학교들 중에서 해당 학교의 위치를 파악하는 것이므로 전년도 수직척도 점수가 동일한 학교들을 기준집단으로 하여 산출한 백분위를 의미한다. 하단에 제시된 표의 우측에 제시된 학교성장 지수 및 성장 수준은 앞서 제시한 부가가치 모형 중 2시점 잠재변인 모형이나 단순 성장점수의 차이점수를 종속변수로 한 부가가치 모형 중 하나를 적용하여 산출한 정보라고 가정하였다. 이와 같은 모형에서는 학교의 학업성장 정보가 0을 중심으로 양수 또는 음수의 값을 가지므로, 평균 50, 표준편차 10으로 선형 변환하였다고 가정하고, 평균에서 2표준편차를 기준으로 학교성장 지수를 성장이 높은 학교(높음), 평균적인 성장을 보인 학교(보통), 더 많은 향상이 필요한 학교(향상 요함) 등 3단계로 구분하여 예시하였다.

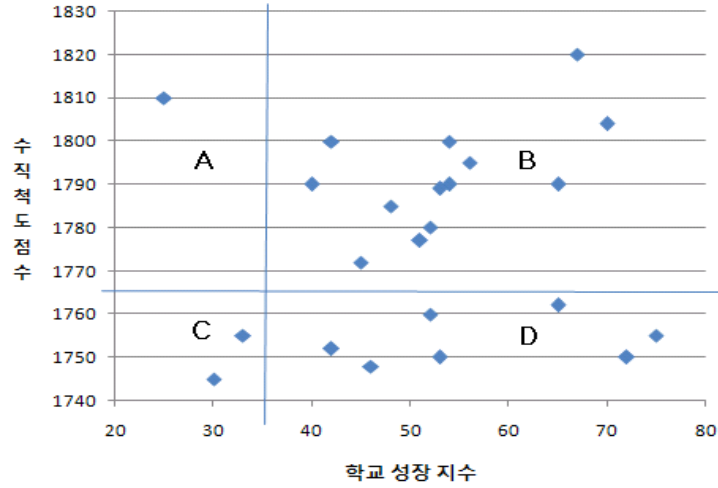
이밖에, '학력향상형 창의경영학교'와 같이 학력향상에 보다 집중해야 할 학교를 선정하는 데에 학교별 학업성장 지수와 학교의 전반적인 학업성취 수준을 동시에 고려하는 방안을 제안하고자 한다. [그림 8]은 개별 학교의 학업성장 지수와 수직척도 상의 성취도 점수 평균 간의 관계를 나타낸 2원 도표의 예시이다. 그림에서 A, B구역과 C, D 구역을 구획하는 수평선은 수직척도 상에서 해당 학년의 적정 발달단계를 나타내는 구분선이며, A, C 구역과 B, D 구역을 구획하는 수직선은 학교의 학업성장 지수에 기초한 수준 중 '향상 요함'과 '보통'을 구분하는 구분선이다. 따라서 C구역에 속한 학교들과 같이, 현재 성취수준이 해당 학년의 적정 발달단계에 미치지 못하면서 학교가 이론 부가적인 성장이 상대적으로 낮은 학교를 교육 개선을 위한 적극적인 개입이 필요한 학교로 지정하는 방식을 고려할 수 있을 것이다.



(그림 6) 국가수준 학업성취도 평가에 대한 개별 학급의 학업성장 보고 방안 예시(학급 담임교사용)



(그림 7) 국가수준 학업성취도 평가에 대한 개별 학교의 학업성장 보고 방안 예시(학교용)



(그림 8) 학교의 성장 지수와 수직척도 점수 간의 2원 도표

IV. 논의 및 제언

본 연구에서는 국가수준 학업성취도 평가의 초-중-고 간 수직척도 개발 연구가 진행되고 있는 현 시점에서 학생수준 종단자료로부터 학생 및 학교의 학업성장 정보를 산출할 수 있는 다양한 모형을 탐색하고, 이를 향후 수직척도화된 학업성취도 평가에 적용할 경우 고려해야 할 사항들에 관해 고찰하였다. 또한 이러한 작업을 토대로 학업성장 정보의 의미 및 해석기준을 제시하였으며, 학생 및 학교수준에서 효과적으로 학업성장 정보를 보고할 수 있는 방안을 제안하였다. 현재는 국가수준 학업성취도 평가의 수직척도의 개발이 완료되지 않았기 때문에 경험 자료를 이용하여 본 연구에서 제시한 학업성장 정보 산출 방안을 실증적으로 확인하는 데에는 한계가 있었으며, 이러한 점에서 본 연구는 향후 경험 자료를 실증적으로 분석하기 위한 기초 자료로서 유용성을 가진다고 본다. 또한 수직척도가 구축되면 학년 간 성장을 계량화하여 보고할 수 있다는 점에서 교육적 활용도가 높지만, 해당 교과에서 측정하는 구인이 학년간에 연속성을 가져야 한다는 기본 가정이 만족되지 않으면 수직척도화를 통해 수리적으로 산출된 값을 타당하게 해석하기는 쉽지 않다. 따라서 각 교과별로 수직척도 적용의 타당성에 대한 검토가 면밀히 이루어져야 하며, 향후 경험 자료의 분석을 통해 수직척도의 타당성을 검증할 수 있는 다양한 연구가 뒷받침되어야 한다. 본 연구는 수직척도가 타당하게 개발되었다는 전제 하에, 수직척도를 바탕으로 학생 및 학교의 성장을 산출하는 방안에 대해 검토하여 학업성취도 평가에 적용할 수 있

는 가능성을 모색하였으며, 향후 수직척도의 효과적인 활용을 위하여 다음과 같은 점을 고려할 필요가 있음을 제안하고자 한다.

첫째, 학업성취도 평가의 교육적 활용도를 높이기 위해서는 평가 결과에 대한 정보를 더 구체적으로 제공할 필요가 있다. 우선 학업성장에 대해 준거참조적 해석이 가능한 정보를 제시하기 위해서 수직척도 상에서 각 점수대별 구인 정의의 특성을 기술한 수직척도 기술표의 개발이 이루어질 필요가 있다. 수직척도에서 산출된 점수에 대해 의미 있는 해석을 하기 위해서는 수직척도 상의 각 점수가 의미하는 바가 무엇인지에 관해 정의되어 있어야 하며, 수직척도 기술표가 개발되면 수직척도 상의 발달단계에 따라 학생들에게 적절한 교육적 피드백을 제공할 수 있는 해석의 틀로 기능할 수 있을 것이다. 이와 더불어 학업성취도 평가의 각 문항별로 요구되는 인지적 요소들을 구체적으로 정의한 상태에서 학생이 각 인지요소를 숙달했는지에 대한 정보를 함께 제공하면 해당 학생이 무엇을 알고 무엇을 모르는지에 대한 구체적인 진단적 정보로서 유용하게 활용될 수 있을 것이며, 이는 개별 학생의 성취수준을 진단, 보정하고자 하는 학업성취도 평가의 취지에 부합함으로써 평가 결과의 활용도를 높이는 데 중요한 역할을 할 수 있을 것으로 본다.

둘째, 수직척도 상의 점수 변화가 절대적인 성장의 의미를 가지기 위해서는 각 점수대별로 변화의 정도가 동일한 의미, 즉 엄밀한 의미의 등간성을 가져야 한다. 그러나 척도의 등간성은 동일 학년 내에서 구축된 일반적인 척도 상에서도 확보되기 어려운 가정이며, 수직척도를 구축할 경우 학년 내 뿐 아니라 학년 간에 점수 변화의 의미가 동일하다는 가정을 만족하기 어렵다는 점에서 척도 제작 및 해석시 주의가 필요하다. 수직척도의 실제적 등간성이 확보되지 않은 상태에서 학생 및 학교의 성장정보를 산출하고 이를 학생이나 학교의 수행에 대한 중요한 의사결정 자료로 사용한다면 학생 또는 학교가 위치해 있는 점수대에 따라 유·불리가 발생할 수 있다. 이를 방지하기 위하여 수직척도 제작시 원점수와 척도점수 간에 비선형적 변환을 가함으로써 원점수 상의 1점 상승에 따른 척도점수 상의 변화가 점수대별로 다르게 척도화하는 것을 고려해 볼 수 있으나, 점수대별로 성장의 크기를 차별적으로 반영할 수 있는 경험적 근거를 확보하는 작업 역시 쉽지 않다. 따라서 수직척도 상의 점수를 바탕으로 학교의 수행에 대한 효과를 파악하는데 사용할 경우 이와 같은 점을 고려한 해석상의 주의가 요구된다.

셋째, 본 연구에서는 학교수준에서 학업성장 정보를 산출하는 데 있어서는 단순 성장모형보다 부가가치의 개념에 기초한 모형을 사용하는 것이 더 적합함을 시사하였다. 그러나 부가가치 모형을 바탕으로 학교의 학업성장 정보를 산출하는 데에도 역시 주의가 요구된다. 먼저 부가가치 모형에서 산출된 학교의 학업성장 정보를 학교의 노력에 의한 효과로 가정하기 위해서는 학교의 여건을 설명하는 맥락변인과 학교의 수행변인이 독립이라는 강한 가정이 필요하다(김준엽, 2009). 그러나 실제로 학교의 여건에 따라 학교의 학생 구성이나 학부모 특성이 다를 수 있고 학교가 투입하는 교육적 노력의 수준에서도 차이가 날 수 있기 때문에 이러한 가정이 만족되는

데에는 현실적으로 제약이 있을 수밖에 없다(Raudenbush, 2004). 학교에 재학중인 학생수에 따라서도 특정 학생의 성취가 그 학교의 효과에 미치는 영향력이 다를 수 있으므로, 각 학교별 모수 추정의 정확성 차이도 학교 간 비교에 있어 한계 요소로 작용한다. Braun(2005)은 이러한 점들을 고려할 때, 부가가치모형을 인과적으로 해석하여 교사 또는 학교에 상벌을 제공하거나 고부담적인 조치를 취하는 것은 매우 위험하며, 따라서 성장의 의미를 인과적인 의미보다 기술적인(descriptive) 의미에서 해석하는 것이 더 바람직함을 강조하였다. 이러한 관점에서, 학교의 성장정보를 행·재정적 지원 및 제재나 과도한 경쟁을 유발할 수 있는 정보 공개의 목적보다는 학교의 수행에 대해 학교 자체적으로 점검하는 목적으로 사용하도록 하기 위하여 개별 학교에만 보고하는 것이 바람직할 것으로 보이며, 학교의 성장정보에 대한 공시 범위에 대해서는 추가적인 논의가 필요하다고 본다.

넷째, 학업성취도 평가를 바탕으로 중학생 및 고등학생의 성장정보를 산출하는 데 있어서 단순성장점수를 종속변수로 한 부가가치모형이나 2시점 종단자료를 이용한 다층성장모형을 적용하는 것이 적합할 것으로 보인다. 학업성취도 평가는 초6-중3-고2의 3시점에서 수집되므로 향후 수직적도 개발이 완료되면 중학교의 성장정보를 산출하는 데에는 2시점, 고등학교의 성장정보를 산출하는 데에는 2시점 또는 3시점의 종단 자료를 이용할 수 있다. 초6-중3과 중3-고2의 2시점 자료를 사용하는 경우, 단순성장점수를 종속변수로 한 부가가치모형이나 2시점 종단자료를 이용한 다층성장모형을 이용하는 것은 적용 및 해석의 용이성 등의 측면에서 학업성취도 평가에 적용하는 데 별다른 무리가 없어 보인다. 고등학교의 경우 3시점의 정보를 이용하는 것이 가능하기는 하지만, 학업성취도 평가가 각 학교급 별로 한 시점의 정보만 측정하는 구조로 되어 있으므로, 다층 성장모형이나 중층적 무선효과 모형을 적용할 경우 초-중-고를 연계한 교차 분류방식의 학교효과 모형을 구축해야 하는 통계적 복잡성으로 인하여 현실적으로 모형을 구현하는 데 어려움이 있기 때문에(김준엽, 2009), 모형 적용의 용이성 및 해석의 명료성의 측면에서 고등학교의 경우도 직전 시점의 자료에만 근거한 2시점 모형을 적용하는 것이 현실적인 방안으로 보인다. 부가가치 개념에 기반한 학교 성장정보를 산출하기 위한 최적의 모형을 선택하는 데에는 이밖에도 평가 데이터 구조와의 적합성 여부, 평가 데이터 저장 관리 체계가 특정 모형의 요구 사항을 충족하는지의 여부, 모형을 적용한 결과가 정부 정책과 부합하는지의 여부, 교육 관련 이해 당사자들 간 의사소통의 용이성, 그리고 모형을 통해 산출되는 정보가 학교 간 비교의 ‘공정성’을 얼마나 잘 확보하고 있는지 등에 대한 경험적인 평가 등을 포괄적으로 고려해야 하며, 향후 학업성취도 평가에 가장 적합한 모형을 선택하기 위해 실제 데이터를 활용한 다양한 경험 연구가 필요하다.

마지막으로, 학업성취도 평가 실시와 함께 수집되는 맥락변인에 대한 재검토가 필요하다. 개별 학교의 교육적 노력을 공정하게 평가하기 위해서는 학교 간의 비교가 학생의 배경변인이나 사전 성취도, 학교의 여건 등에 영향 받지 않아야 하며, 이를 위해 학업성취도를 가장 정확하게

예측하는 학생수준의 배경변인들을 학교효과 산출 모형에서 통제할 필요가 있다. 본 연구에서 소개한 2시점 잠재변인 성장모형이나 중층적 무선효과 모형 등에서는 시간에 따른 학생수준의 성취도 정보만 사용하고 있지만, 이밖에 학생 및 학교수준에서 수집한 맥락변인을 모형에 투입하는 것도 가능하다. 현재 학업성취도 평가에서는 학생수준의 맥락변인이 매우 제한적으로 수집되고 있다. 학교의 여건과 관련된 학교수준 변인 중에서는 기초생활수급자비율과 다문화가정비율이 있으나 이러한 정보가 학생수준에서는 조사되지 않고 있어, 학교 향상도 지표 산출시에도 맥락변인을 투입하지 않고 사전 성취도만 통제할 상태에서 각 학교별 부가가치 지수를 산출하고 있다(김경희 외, 2011). 학생수준에서 사전 성취도를 통제할 경우 학업 성취도에 영향을 미치는 개인의 특성을 상당 부분 통제하는 것이 가능하다는 연구 결과(예: Sanders & Horn, 1994)들이 있기는 하나, 우리나라의 상황에서 이러한 가정에 대한 경험적 확인이 필요하며, 사전 성취도 외의 맥락변인을 통제하는 것이 보다 정교한 학교효과 정보를 산출하는 것으로 나타난다면 우리나라 상황에 적합한 맥락변인에 대한 연구를 통해 학생수준에서 맥락변인을 수집할 필요가 있다.

참 고 문 헌

- 김경희, 김수진, 김희경, 송미영, 시기자, 신진아, 김준엽, 신인수(2011). **국의 학업성취도 평가 비교와 교육정보 공시를 위한 학교 향상도 지표 개발**. 한국교육과정평가원 연구보고 CRE 2011-11.
- 김경희, 김완수, 최인봉, 상경아, 김희경, 신진아, 김성훈(2011). **국가수준 학업성취도 평가 결과 향상도 지표 개발**. 한국교육과정평가원 연구자료 ORM 2011-53-24.
- 김성숙, 송미영, 최인봉, 김희경, 김준엽, 이현숙(2010). **우리나라 초·중·고 학생의 학업성취 특성 분석 - 2009년 국가수준 학업성취도 평가 전수 결과 중심-**. 한국교육과정평가원 연구보고 RRE 2010-7-1.
- 김양분, 임현정, 이광현, 신혜숙(2009). **국가수준 학업성취도 평가 활용 체제 구축 연구**. 한국교육개발원 연구보고 RR 2009-28.
- 김준엽(2009). 대규모 학업성취도 평가 자료를 활용한 학교 부가가치 지수 설정 방법. **대규모 학업성취도 평가 자료의 활용 방안 탐색** 세미나 자료집. 한국교육과정평가원 연구자료 ORM 2009-31.
- 김준엽(2010). 대규모 학업성취도 평가 자료를 활용한 학교 부가가치 추정모형의 비교. **교육평가연구**, 23(1), 101-124.
- 송미영(2011). 학업성장 분석을 위한 국가수준 학업성취도 평가의 종단척도 구축. **개별 학생의 학업성장과 학교교육 성과 분석 방안** 세미나 자료집. 한국교육과정평가원 연구보고 RRE 2011-6-2.
- 송미영, 남민우, 김창일, 강태훈(2011). **학교교육 성과 평가를 위한 학업성취 지표와 평가모형 개발 - 국가수준 학업성취도 평가의 종단척도 구축-**. 한국교육과정평가원 연구보고 RRE 2011-6-1.
- 이상하, 김수진, 신선희, 김완수(2010). **학교교육 성과 평가를 위한 학업성취지표와 평가모형 개발: 학교효과 추정 모형과 평가 체제를 중심으로**. 한국교육과정평가원 연구보고 RRE 2010-10.
- 이현숙(2011). 종단척도를 활용한 학업성장 정보 산출 방안. **개별 학생의 학업성장과 학교교육 성과 분석 방안** 세미나 자료집. 한국교육과정평가원 연구보고 RRE 2011-6-2.
- 임현정, 김양분, 신혜숙, 이광현(2010). **학교교육 실태 및 수준 분석(III): 중학교 연구**. 한국교육개발원 연구보고 RR 2010-23.
- 정은영, 최인봉, 김희경, 김소영, 유진은(2009). **국가수준 학업성취도 평가 결과 추이(2003년~2008년)-고등학교 1학년-**. 한국교육과정평가원 연구보고 RRE 2009-8-3.

차성현, 민병철, 김성식, 김준엽, 박소영, 변중석(2010). **고교 유형에 따른 학교효과 분석을 위한 패널조사 설계: 자율형 고교를 중심으로**. 한국교육개발원 연구보고 RR 2010-01.

Australian Curriculum, Assessment, and Reporting Authority (2010). *2010 NAPLAN Principals Handbook*. Sydney, New South Wales.

Ballou, D., Sanders, W., & Wight, P. (2004). Controlling for student background in value-added assessment of teachers. *Journal of Educational and Behavioral Statistics*, 1(2), 127-155.

Blank, R. K. (2010). *State Growth Models for School Accountability: Progress on Developing and Reporting Measures of Student Growth*. Paper presented at the Council of Chief State School Officers, Washington, D.C.

Braun, H. I. (2005). *Using Student Progress To Evaluate Teachers: A Primer on Value-Added Models*. Educational Testing Service.

Bryk, A. S., Thum, Y. M., Easton, J., & Luppescu, S. (1998). Assessing school academic productivity: The case of Chicago school reform. *Social Psychology of Education*, 2, 103-142.

Carey, K., & Manwaring, R. (2011). *Growth Models and Accountability: A Recipe for Remaking ESEA*. Education Sector.

Connecticut Department of Education (2010). *CMT Interpretive Guide*. Available at www.sde.ct.gov.

Goldschmidt, P., Roschewski, P., Choi, K., Auty, W., Hebbler, S., Blank, R., & Williams, A. (2005). *Policymakers' Guide to Growth Models for School Accountability: How do Accountability Models Differ?* The Council of Chief State School Officers, Washington, DC.

Lissitz, R. W., & Huynh, H. (2003). Vertical equating for state assessments: issues and solutions in determination of adequate yearly progress and school accountability. *Practical Assessment, Research & Evaluation*, 8(10). Retrieved from <http://PAREonlin.net/getvn.asp?v=8&n=10>.

McCaffrey, D. F., & Hamilton, L. S. (2007). *Value-Added Assessment in Practice: Lessons from the Pennsylvania Value-Added Assessment System Pilot Project*. RAND Corporation.

National Research Council and National Academy of Education (2010). *Getting Value Out of Value-Added: Report of a Workshop*. Committee on Value-Added Methodology for

- Instructional Improvement, Program Evaluation, and Accountability, Henry Braun, Naomi Chudowsky, and Judith Koenig, Editors, Center of Education, Division of Behavioral and Social Sciences and Education. Washington, DC: The National Academies Press.
- Patz, R. J. (2007). *Vertical Scaling in Standards-Based Educational Assessment and Accountability Systems*. Washington, D.C.: the Council of Chief State School Officers. Stanford Learning First. San Antonio, TX: Pearson Education.
- Raudenbush, S. W. (2004). What are value-added models estimating and what does this imply for statistical practice? *Journal of Educational and Behavioral Statistics*, 29(1), 121-129.
- Ray, A., Evans, H., & McCormack, T. (2009). The use of national value-added models for school improvement in English schools. *Revista de Educación*, 348, 47-66.
- Sanders, W. L. (2006). Comparison among various educational assessment value-added models. Paper presented at the power of two national value-added conference. Columbus, Ohio.
- Sanders, W. L., & Horn, S. P. (1994). The Tennessee value-added assessment system (TVAAS): Mixed model methodology in educational assessment. *Journal of Personnel Evaluation in Education*, 8, 299-311.
- Texas Education Agency (2010). 2010 Accountability Manual: The 2010 Accountability System for Texas Public Schools and School Districts.
- Yen, W. M. (2010). *Measuring Student Growth With Large-Scale Assessments in an Education Accountability System. Paper presented at the Exploratory Seminar: Measurement Challenges Within the Race to the Top Agenda*. Center for K-12 Assessment & Performance Management.
- Wise, L., & Alt, M. (2006). *Assessing vertical alignment*. Washington, DC: Council of Chief State School Officers.

· 논문접수 : 2012-09-01/ 수정본접수 : 2012-10-11/ 게재승인 : 2012-10-17

ABSTRACT

Theoretical Reviews on Interpreting and Reporting Students' Academic Growth Based On Vertically Scaled Scores

Mi-Young Song

(Research Fellow, Korea Institute for Curriculum and Evaluation)

Hyun Sook Yi

(Associate Professor, Konkuk University)

Transition in administration of the National Assessment of Educational Achievement (NAEA) from a sample-based to the entire population-based assessment for target grades since 2008 has provided opportunities to collect longitudinal achievement data at an individual student level. In order to maximize the utility of the panel data, research has been conducted on feasibility and validity of constructing vertical scales for the NAEA Korean and Math subjects so that student achievement could be compared across a single continuum of the construct being measured. At the outset of constructing the vertical scales linking scores on the 6th-9th-11th grades, the present study theoretically investigated how student-level longitudinal data can be analyzed and interpreted before empirical studies can be carried out. For this purpose, the present study explored several growth models applicable to student-level longitudinal data, such as models based on simple gain scores, multilevel growth models, layered mixed effect model, and discussed issues that should be considered in application of each model to the NAEA vertical scale scores. In addition, exemplary score reports for students and schools were presented in order to propose appropriate information to be included and recorded. Finally, suggestions that should be considered were made in developing valid vertical scales and making effective uses of the vertically scaled scores.

Key Words : Vertical scaling, Academic growth, National Assessment of Educational Achievement