

## 교실평가에 대한 고등학생의 인식 유형 분석<sup>1)</sup>

조 영 민(경북대학교 대학원생)\*

손 원 숙(경북대학교 교수)\*\*

### <요 약>

본 연구의 목적은 고등학생의 교실평가인식 유형을 분석하고, 학생의 배경 특성 및 성취목표지향성이 평가인식 유형을 예측하는 정도를 파악하고, 이와 더불어 각 유형에 따른 시험불안과 학업참여의 차이를 알아보는 것이다. 이를 위해 D시와 G지역의 고등학교 1, 2학년 학생 662명의 자료에 대하여 2단계 군집분석을 실시한 결과, 평가가 학습향상에 도움이 된다고 생각하면서도 평가 결과에 대해 강한 책임을 느끼는 학습향상-책무성 유형(39.7%), 평가를 부적절하게 여기면서도 강한 책무성을 느끼는 부적절성-책무성 유형(39.3%), 평가의 유용성을 전혀 느끼지 못하는 부적절성 유형(21.0%)으로 분류되었다. 다항로지스틱 회귀분석을 통한 교실평가인식 유형에 대한 예측요인의 효과를 살펴본 결과, 학교 유형, 사전성취도, 및 접근지향적 성취목표지향성이 통계적으로 유의하게 나타났다. 한편 시험불안은 부적절성 유형에서 가장 낮게 나타났고, 학업참여는 학습향상-책무성 유형에서 가장 높고, 부적절성 유형에서 가장 낮게 나타났다. 본 연구 결과를 통해 교실평가의 목적에 대한 학습자의 긍정적인 인식의 필요성과 더불어 접근지향적인 성취동기 향상을 위한 교수학습 및 평가 전략 마련을 위한 시사점을 논의하였다.

주제어 : 교실평가, 학생평가인식, 학습향상을 위한 평가(AFL), 학습결과에 대한 평가(AOL), 책무성

## I. 서론

교실평가는 학습의 질을 결정하는 주요한 학습 환경 요인이며 따라서 교실평가에 대한 학생들의 인식은 학습과정의 기제를 이해하는데 중요한 역할을 수행할 것으로 가정된다. 인식

1) 이 논문은 제1저자의 2017년 경북대학교 석사학위 논문을 수정·보완한 것이며, 2017년 대한민국 교육부와 한국연구재단의 지원을 받아 수행된 연구임 (NRF-2017S1A5A2A01023770)

\* 제1저자, fgjym@naver.com

\*\* 교신저자, wsohn@knu.ac.kr

(conceptions)이란 실체에 대한 정신적인 표상(Brown & Hirschfeld, 2007)을 의미하며, 개인의 의견, 태도 또는 신념 등을 포괄하며 복잡하고 다양한 개인의 경험을 설명할 수 있는 기능을 갖고 있다. 본 연구에서는 교수학습 과정에서 평가의 역할 및 학생들의 평가에 대한 인식의 중요성에 주목하여 학생들의 평가인식 유형을 분석하고 이러한 인식유형의 예측 및 결과요인과의 관련성을 탐색하고자 한다.

교실상황에서 평가는 학교교육의 책무성 점검과 교수개선 및 학습향상이라는 목적을 위하여 시행되어 왔고, 기능면에서는 총합(summative) 또는 형성(formative) 평가의 형태로 구분되기도 한다. 최근까지도 우리나라를 포함해 전 세계적으로 교실평가는 학습지향적인 형성평가 보다는 결과지향적인 총합평가에 보다 많은 비중을 두어왔다(Black & Wiliam, 1998; Black, Wilson, & Yao, 2011). 평가의 목적을 학생의 학습 향상 및 교사의 교수 개선에 두는 ‘학습향상을 위한 평가(Assessment For Learning: AFL)’(김명숙, 2008; 지은림, 2010; Black & Wiliam, 1998)는 학습자의 현재 상태나 도달해야 하는 목표를 점검하고 이러한 목표 도달방법을 결정할 수 있도록 증거를 수집하고 해석하는 과정으로 정의된다. 이와 같은 AFL은 평가의 목적을 성취수준 확인 및 책무성 점검 등에 두는 ‘학습결과에 대한 평가(Assessment Of Learning: AOL)’와는 구분된다. 하지만 이처럼 교실평가의 목적을 ‘학습 vs 책무성’ 평가 혹은 ‘과정중심 vs 결과중심’ 평가 등으로 분류하여 이들을 양립할 수 없는 양극단의 목표로 간주하기 보다는 이들은 학교교육을 위하여 각기 고유한 역할을 수행하고 있음을 인지할 필요가 있다. 교실평가는 교실 밖에서 시행되는 다양한 책무성 평가와 연계성을 가져야 하며 최근 강조되고 있는 21세기 핵심역량, 성취기준기반 교육 및 학습동기 이론 등을 반영하여 형성적인 역할을 수행할 수 있어야 한다(McMillan, 2014). 따라서 교실평가는 어느 하나의 목적에만 초점을 두는 평가이기 보다는 교실상황에서 요구되는 다양한 평가목적들을 조화롭게 연계시킬 필요가 있을 것이다. 이는 궁극적으로 교수개선 및 학습향상을 위한 교실평가의 기능을 극대화할 수 있는 방안이며 학생들의 성취도 및 학습동기 향상에 기여할 것으로 가정된다.

일반적으로 교실 안에서 평가실체는 이를 직접 수행하는 교사의 평가인식과 밀접한 관련을 가진다. 교사의 평가에 대한 인식은 평가가 앞으로 어떤 방향으로 이루어져야 하는지 생각해볼 수 있도록 하여 평가 방식을 수정하고, 교수·학습 방법을 발전시키는 기회를 제공한다. 그러나 교사들이 평가를 계획하고 시행하는 주체로서 학생들보다 학습목표를 더 명확하게 인식하고, 학습문제를 더욱 깊이 이해하기 때문에 교사는 자신의 평가 실제보다는 의도된 평가 목적에 맞추어 인식할 가능성이 높다(박민애, 손원숙, 2017). 그러나 학생의 경우에는 교사가 어떠한 목적과 의도를 가지고 평가를 진행하는지 파악하기 어렵고 동일한 평가를 수행한다고 하더라도 학생 개인의 주관적 해석에 따라서 교사의 평가 실제에 대한 효과는 차별적으로 나타날 수 있다. 최근 들어 국내에서도 수업과 평가를 연계하여 평가의 형성적 기능을 강화할 수 있는 교육정책들이 발표되고 있다. 대표적으로 ‘과정중심평가’라는 용어를 도입하여 평가결과

산출에 집중하기 보다는 평가를 통하여 학생의 강점과 약점을 파악하고 이를 수업 및 학습 개선을 위한 유용한 정보로 활용하는데 관심을 두고 있다(교육부, 2016). 교실평가에 대한 학습자의 인식을 알아보는 것은 이와 같이 의도된 교실평가의 실재를 확인할 수 있는 하나의 방안이 될 수 있을 뿐 아니라 실제 의도하지는 않았으나 평가가 학생들에게 미치는 잠재적인 효과(McMillan, 2014) 역시 파악할 수 있을 것이다. 예를 들어서 교실평가의 방향이 AFL의 관점에 기반 하여 이루어지고 있는 경우, 학생들의 평가인식을 탐색해 봄으로써 이러한 평가 정책이 실제 의도된 방향으로 시행되고 있는지를 경험적으로 파악해 볼 수 있을 것이다. 만약 실제 학생들은 의도된 평가의 목적과는 상이한 인식을 나타낸다면 교사의 평가 활동 및 수업 전반을 검토하는 등 그 원인을 탐색해 볼 필요가 있을 것이다. 교육 수요자인 학생들을 대상으로 한 이와 같은 평가인식에 대한 경험적 결과는 교실평가 정책의 개선에 있어서 주요한 기초자료로 활용될 수 있을 것이다.

기존의 학생의 평가인식에 대한 국내 연구는 실태조사가 대부분이었으나, 본 연구에서는 평가인식에 대한 이론적 틀을 기반으로 한 타당화된 척도를 활용하여 학생의 평가인식 유형을 탐색하고자 한다. 교실평가 정책의 차별성 및 학업부담의 차이 등으로 인하여 초, 중, 고 학교 급별 학생의 평가인식에는 차이가 있을 것으로 가정되나, 본 연구에서는 특히 고등학생의 평가인식에 초점을 두고자 한다. 다른 학교급 학생들에 비하여 고등학생들은 다양한 교실평가 경험이 축적되어 있을 뿐 아니라 교실평가의 목적이나 특성에 대한 자신의 인식을 비교적 객관적으로 보고할 수 있다는 면을 고려하였다. 한편 뉴질랜드의 Brown과 Hirschfeld(2007)가 뉴질랜드 중등학생을 대상으로 개발하고 타당화한 학생의 평가인식 척도(Students' Conceptions of Assessment: SCoA)를 활용하여 학생의 평가인식을 측정하고 유형화하고자 한다. 이러한 학생의 평가인식의 유형화는 평가에 대한 다양한 인식들을 개별적이고 독립된 개념으로 간주하며, 학생이 평가의 목적을 다차원적으로 인식하는 것이 가능하다는 점을 가정한 것이다. 따라서 다양한 평가목적들이 실제 교실장면에서 학습자들에 의하여 어떻게 인식되고 있으며 이러한 인식유형별 특성을 탐색하고자 한다.

한편 학생의 평가인식은 실제 학생들에게 제공되는 교실평가실제뿐 아니라 이에 대한 학생의 주관적 해석이 관여되기 때문에 평가인식 유형을 발달시키는데 있어서는 학습자의 개인적 특성도 직접적인 관련성이 있을 것으로 예상된다. 본 연구에서는 대표적인 동기이론인 기대-가치이론(Eccles & Wigfield, 2002)과 Brookhart(1997)의 교실평가 모형에 기반 하여서 사전 성취도 및 성취목표지향성이 학생의 평가인식 유형에 미치는 영향력을 탐색하고 또한 평가인식 유형이 후속적인 학습과정에 해당하는 학업참여와 시험불안에 미치는 효과를 살펴보고자 한다. 이들 변인에 대한 구체적인 설명은 2장 이론적 배경에 추가하였다.

마지막으로 본 연구를 통해 학생들의 학업 증진을 위해서 교육현장에서 학교와 교사가 지향해야 할 평가 유형을 파악하고 학생들의 학습에 부정적인 영향을 미치는 평가인식을 개선하

는 방안을 탐색해보고자 한다. 요약하면 본 연구에서는 교실평가에 대한 학생 인식 유형을 분석하고, 관련 변인과의 관계를 분석하고자 하며 구체적인 연구문제는 다음과 같다.

1. 교실평가인식의 군집 유형은 어떻게 구성되는가?
2. 교실평가인식의 군집 유형에 영향을 주는 요인은 무엇인가?
3. 교실평가인식의 군집 유형에 따른 시험불안 및 학업참여의 차이는 어떠한가?

## II. 이론적 배경

### 1. 교실평가에 대한 인식 연구

기존의 대다수 평가인식에 관한 연구는 주로 교사 대상으로 이루어져왔고, 이와 관련된 국내 연구(강대중, 염시창, 2014; 김신영, 2014; 박정, 이진주, 2015)에 의하면 대부분의 교사들은 학생평가에 대한 개념을 잘 숙지하고 있고 그 필요성을 잘 인식하고 있지만, 인식하고 있는 만큼 실제로 옮기지는 못하고 있는 것으로 나타났다. 이러한 국내 연구들은 대체로 실태조사의 형태였지만, 타당화된 척도를 통하여 교사의 평가인식을 측정하는 일부 국외 연구(Brown et al, 2011)도 있었다. 한편 학생의 평가인식은 학생의 학습에 대한 주관적인 해석으로 볼 수 있으므로, 학생의 학습에 대한 기대와 가치에 영향을 줄 수 있고(Eccles & Wigfield, 2002), 학생의 학습 태도에도 영향을 미칠 수 있다(Brookhart, 1997)는 점에서 교사의 평가인식과 구분하여 연구할 가치가 있다고 볼 수 있다. 국내에서 수행된 학생의 평가 인식 연구는 주로 수행평가에 관한 인식 조사(김진수, 김석우, 2008; 홍정립, 최은경, 2002)에 집중된 경향이 있었다. 또한 시험 상황에서 학생들이 느끼는 불안 수준이나 정서를 측정하는 척도(도승이 외, 2011; 박병기, 임신일, 김어진, 2008)는 일부 개발되어 있으나, 평가의 목적이나 특성에 대한 인식 수준을 측정하고, 더불어 타당화된 척도의 사용은 거의 전무하다고 볼 수 있었다.

교실평가는 교사의 의사결정을 돕기 위해 필요한 정보를 수집하고, 해석하며, 활용하는 것과 관련된 모든 활동을 의미한다(McMillan, 2014). 교실평가의 목적은 학생의 강, 약점 진단, 성적 부여, 수업의 효율성 결정, 학생들에게 피드백 제공 등으로 요약될 수 있다. 이러한 평가의 목적을 제대로 인식하는 것은 평가의 성격, 시행 및 결과 활용을 위한 초석이 되는 활동이라고 할 수 있다. 이와 같은 교실평가의 목적과 기능은 다양하게 정의될 수 있으나 Brown과 Hirschfeld(2007)는 책무성, 학습향상, 부적절성, 유용성 등 4개 요인으로 나누어 척도를 개발하였다. 이러한 척도는 평가의 목적 및 특성에 대한 학생의 인식을 측정하기 위한 것으로 뉴질

랜드 고등학생을 대상으로 타당화되었다는 면에서 고등학생을 대상으로 한 본 연구에도 적절한 척도이다. 먼저 책무성(accountability)은 평가의 목적을 학생 및 학교로 하여금 학업에 대한 책임감을 갖게 만드는 것이라고 인식하는 것이다. 즉, 평가의 목적을 학생 개인의 성취수준을 확인하거나 학교의 수준이나 가치를 점검하기 위한 것으로 인식하며, 이는 전통적인 평가관인 AOL 관점과 일관적이라고 할 수 있다. 평가는 학습목표 도달확인 또는 서열화를 위하여 사용되며, 수업이 종료된 이후 시행되는 총괄평가의 기능을 수행한다. 이러한 평가는 학교나 학생의 책무성을 확인하려는 목적에 잘 부합되지만 학생들의 학습에 대한 환류효과가 상대적으로 제한적이라는 단점이 있다(Stiggins, 2005). 둘째, 학습향상(learning improvement)은 평가가 학생의 학습에 도움이 되며 학습개선에 유용한 정보를 제공하기 위한 것이라는 인식이다. 이는 교수·학습을 연계하는 평가의 순기능을 활용하여 학생의 학습 능력을 향상시키고 긍정적인 학습 지향적인 성향을 발달시키는데 목적을 두는 AFL 관점(James & Pedder, 2006)과 일치한다. 셋째, 부적절성(irrelevance)은 평가는 공정하지 못하며 오히려 학습을 방해하기 때문에 평가를 부정적으로 인식하는 것이다. 따라서 평가결과를 무시하게 되는데, 대체로 평가에 대한 부정적인 경험 예컨대 낮은 성취나 불공정한 결과를 받았다고 인식하는 경우 부적절성 인식이 발달될 가능성이 높으며, 이러한 인식은 다시 평가에 대한 부정적인 정서 및 낮은 성취감으로 이어질 가능성이 파악되고 있다(Brown et al., 2011). 마지막으로 유용성(usefulness)은 평가가 타당하고 신뢰로우며 또한 자신의 학습과정이나 결과를 제대로 반영한다고 생각하는 인식이다. 최근 들어 교수개선 및 학습향상을 위해 구안된 평가실체는 학생들의 인지 및 정서적인 측면 모두에서 긍정적인 효과를 지닌다는 다수의 경험적 연구결과(박민애, 손원숙, 2017; William & Thompson, 2007)가 보고되고 있다. 이는 학습향상이나 유용성 등과 같은 평가인식과 연계되는 결과이긴 하지만 교실평가는 국가나 시도교육청에서 교육시스템의 질을 점검하려는 목표로 수행되는 책무성 평가와는 무관할 수 는 없을 것이다(McMillan, 2014). 즉, 교실평가는 학생들의 학습향상이나 자율적인 학습자 양성뿐 아니라 책무성 이행 등 다양한 목적(multiple purpose)으로 수행되게 된다. 이에 실제 교실장면에서 학생들은 교실평가의 목적을 어떻게 인식하고 있는지 알아볼 필요가 있으며, 이는 현재 우리나라 교실평가의 실체에 대한 점검뿐 아니라 개선방향에 대한 아이디어를 제공할 것으로 기대한다.

## 2. 교실평가인식과 관련 요인들

본 연구에서는 Eccles와 Wigfield(2002)의 기대-가치이론 모형과 Brookhart(1997)의 교실평가 모형을 기반으로 하여 학생의 평가인식과 관련요인과의 관계를 탐색하고자 한다. 먼저 학생들의 사전 성취경험에 해당하는 학업성취도 및 공부를 하는 이유에 해당하는 성취목표지향성에 따라서 학생들의 평가인식은 어떠한 차이를 보일 것인가를 알아보하고자 한다. 사전성취도

와 평가인식 발달과의 관련성은 이전 성취관련 경험이 경험에 대한 학생의 해석과 정서적 기억을 통하여 후속되는 학습상황에서 발생하는 주관적 과제 가치 및 성공에 대한 기대에 영향을 미친다는 기대-가치이론(Eccles & Wigfield, 2002)을 토대로 지지된다. 또한 일부 평가인식 관련 선행연구(Brown & Hirshfield, 2007; Brown, et al., 2011)에 따르면, 성취도가 높은 학생들은 그렇지 않은 학생들에 비하여 평가에 대한 이해도나 유용성에 대한 인식이 높은 것으로 나타났으며, 대체로 평가에 대한 부정적 인식이 높은 학생들의 성취도는 낮은 경향이 나타났다. 따라서 학생들이 경험한 사전 성취도 수준이 평가에 대해 가지는 주관적 인식에 미치는 영향을 파악함으로써 성취수준에 따른 평가방법의 적정성을 검토하는 등 교실평가의 개선에 시사점을 제안할 수 있을 것이다. 한편 성취목표지향성(Elliot & McGregor, 2001)은 개인이 성취를 추구하고, 경험하고, 해석하는 틀을 만들도록 하는 성취동기로써 이는 학습자의 자기조절, 학습전략, 내적동기 등과 같은 심리적 특성에 영향력이 있는 요인이다. 성취목표지향성은 유능감을 정의하는 방식(competence-defined framework)에 따라 초기에는 유능감을 개발하거나 과제를 숙달하는 것에 중점을 두는 ‘숙달목표’와 타인과 비교하여 자신의 유능감을 나타내는 것에 중점을 두는 ‘수행목표’ 등 2요인으로 측정하였지만, 추후 이들 목표를 접근 및 회피 측면으로 추가적으로 분화시켜 4요인 모형으로 정의되고 있다(하위요인 설명은 3장 연구방법 참고). 따라서 학습자들은 자신이 지닌 성취목표지향성에 따라서 학업상황에서 서로 다른 학습전략(이주화, 김아영, 2005)을 보이며 또한 평가에 대한 주관적 해석에 해당하는 평가인식 역시 차이를 보일 것으로 가정된다.

한편 학생의 평가인식은 평가 장면에서 영향을 미치는 것뿐만 아니라 학생의 학업 전반에 영향을 미칠 것이라고 가정할 수 있다. Brookhart(1997)는 AFL 관점에서 학생들의 학업성취를 설명하고 있는데, 교사의 평가 활동이 학생들이 지각하는 과제 가치와 자기 효능감 등에 영향을 미치며 이는 학업정서나 학업참여 및 노력 등을 통해 학업성취에 영향을 준다고 가정하였다. 먼저 학업참여(추헌택, 손원숙, 2011)는 학생들의 학업성취와 수업 및 학습의 질을 예측하는 중요한 변인이다. 따라서 평가에 대해 가지는 학습자의 인식이 학업참여에 미치는 효과를 살펴봄으로써 학업참여 및 학업성취를 향상시킬 수 있는 평가 방안 마련에 기여할 수 있을 것이다. 마지막으로 시험불안(Owens et al, 2014)은 자신의 능력을 평가 받는 상황에서 나타나는 시험에 실패할 가능성에 대한 걱정, 타인과의 비교나 자신의 능력에 대한 낮은 자신감으로 인한 두려움 등 인지적, 정서적 측면이나 가슴이 두근거리는 등의 신체적 반응을 포함한다. 학생들이 평가에 대해 가지는 기대와 가치는 학업상황에서 나타나는 다양한 학업정서와 관련을 가질 것이며 특히 시험불안은 모든 학교급 학생들이 겪게 되는 주된 정서로 간주된다. 우리나라에서는 평가라는 제도가 타인들과의 비교, 서열화의 목적으로 인식되는 경향이 높으며, 타인과의 경쟁에서 인정을 받지 못하거나 시험에서 실패한 경험 등으로 인해 시험불안은 높아질 수 있을 것이다. 더욱이 학년이 올라갈수록 평가에 대한 불안을 경험할 기회가 더 많아지기 때문

에 고학년일수록 시험불안의 수준은 높은 것으로 나타나고 있다(김빛나, 2014). 따라서 평가에 대한 인식 유형에 따라서 시험불안의 차이를 알아보는 것은 향후 교실상황에서 학생들에게 평가의 목적과 특성에 대해서 어떻게 인식시키고 시행해 나가야 할 것인가에 대한 정보를 제공할 수 있을 것으로 기대한다.

### III. 연구방법

#### 1. 연구대상

본 연구에서는 D 광역시 및 K 지역에 소재한 6개 고등학교 1, 2학년 학생 726명을 연구 대상으로 선정하였다. 이 중 불성실하게 응답하여 연구에 적합하지 않는 64명을 제외한 결과, 662명을 대상으로 최종 분석을 실시하였다(<표 III-1> 참고).

<표 III-1> 학교 유형, 학년, 성별에 따른 연구대상 분포(%)

학교 유형		일반고		특목고		전체		전체
학년		1학년	2학년	1학년	2학년	1학년	2학년	
성별	남	203 (35.2)	108 (18.7)	38 (44.6)	10 (11.8)	241 (36.4)	118 (17.8)	359 (54.2)
	여	139 (24.1)	127 (22.0)	6 (7.1)	31 (36.5)	145 (21.9)	158 (23.9)	303 (45.8)
	전체	342 (59.3)	235 (40.7)	44 (51.8)	37 (43.5)	386 (58.3)	276 (41.7)	662 (100)
전체		577 (87.2)		85 (12.8)		662 (100)		662 (100)

#### 2. 측정도구

##### 가. 학생의 교실평가 인식

본 연구에서는 학생의 평가인식 측정을 위하여 Brown과 Hirschfeld(2007)가 뉴질랜드 중등 학생을 대상으로 개발한 Students' Conceptions of Assessment(SCoA) 척도의 일부를 타당화 하여 사용하였다. SCoA는 총 35개의 문항, 4개의 하위요인(책무성, 학습향상, 부적절성, 유용성)으로 개발되었으나 후속적인 타당화 연구(Hirschfeld & Brown, 2009)를 통해 문항 및 하위요인이 축소되어 사용되고 있다. 한편 본 연구에서는 전문가의 내용 검토 및 탐색적 요인분석

을 통해 중복성이 관찰되는 학습향상과 유용성 하위요인을 통합하여 총 3개의 하위요인, 14개의 문항을 사용하였다. 총 14문항을 활용하여 3요인 확인적 요인분석모형을 분석한 결과,  $\chi^2 = 312.828$  ( $p < .001$ ),  $df = 71$ ,  $TLI = .930$ ,  $CFI = .945$ ,  $RMSEA = .072$  등으로 수용 가능한 수준의 적합도로 판단되었다. 3개의 하위요인은 평가가 학습향상에 도움을 준다는 인식을 묻는 ‘학습향상’(5문항), 평가가 학생 혹은 학교로 하여금 학업에 대한 책임감을 갖게 만드는 것이라는 인식인 ‘책무성’(6문항), 그리고 평가는 유용성이 없다고 생각하는 ‘부적절성’(3문항) 요인이다. 응답척도는 6점 Likert 척도를 활용하였으며 문항 구성과 하위척도별 신뢰도는 <표 III-2>와 같다.

<표 III-2> 교실평가인식 척도의 하위요인별 문항구성과 신뢰도

하위요인	문항예시	문항수	신뢰도
책무성	·평가란 정해진 기준에 내 학업 수행을 비교하는 것이다.	6	.783
	·평가는 나의 학업에 성적(점수, 등급, 백분위, 성취수준 등)을 부여하는 것이다.		
	·평가는 성취목표와 비교하여 나의 향상 정도를 확인하는 것이다.		
	·평가는 학교를 투명하고 만족할만한 수준으로 이끈다.		
	·평가는 학교가 얼마나 잘하고 있는지에 대한 정보를 제공한다.		
학습향상	·평가는 학교의 수준과 가치를 측정한다.	5	.882
	·평가는 나의 학습 향상에 도움을 준다.		
	·평가는 내가 최선을 다하게 만든다.		
	·평가는 나에게 적절하고 이롭다.		
	·평가는 나의 수행에 관한 피드백을 제공한다.		
부적절성	·평가는 학습의 일부분이다.	3	.884
	·나는 평가에 대한 정보를 무시한다.		
	·나는 평가 결과를 무시한다.		
전체	·나는 평가를 받기는 하지만 그 결과를 거의 활용하지 않는다.	14	

## 나. 사전성취도

학생의 사전성취도는 평가인식을 측정하기 전 1년 간 학교에서 시행된 모든 평가에 대하여 자신이 생각하는 성취수준에 가장 가까운 번호에 체크하도록 하였다. 9점 척도를 활용하였으며 자신이 1등급 이내라고 생각할 경우 ①, 1~2등급 정도라고 생각하면 ②를 선택하는 방식으로 9등급까지 표기하도록 하였다. 평가의 범위를 학교 교육과정 내의 모든 평가로 설정하였기 때문에 학생이 자신의 등급을 명확하게 대답하기 곤란할 경우를 감안하여 급간을 겹치게 선택할 수 있도록 고안하였다. 설문답안을 역채점하여 수치가 높을수록 사전 성취수준이 높음을 의미한다.

## 다. 성취목표지향성

이주화, 김아영(2005)가 2×2성취목표지향성 이론에 기초하여 한국 중·고등학생을 대상으로 타당화한 성취목표지향성 척도를 사용하였다. 4개의 하위요인 별로 5문항씩 총 20문항으로 이루어져 있으며, 응답 양식은 5점 Likert 척도를 활용하였다. 성취목표지향성 척도의 하위요인별 신뢰도는 수행접근(타인과의 상대적인 유능성을 획득하려는 경향, 예를 들어, 나의 목표는 다른 학생들보다 좋은 성적을 받는 것이다)이 .850, 수행회피(상대적인 무능력을 회피하려는 경향, 예를 들어, 나의 목표는 다른 학생들과 비교하여 나쁜 성적을 받지 않는 것이다)가 .850, 숙달접근(과제의 숙달이나 능력의 발전을 달성하려는 경향, 예를 들어, 나는 새로운 것을 배울 수 있는 도전적인 숙제를 좋아한다)이 .796, 숙달회피(자신의 능력감소나 숙달의 실패가능성을 회피하려는 경향, 예를 들어 나는 수업 내용을 완벽하게 이해하지 못할까봐 두렵다)가 .825로 나타났다.

## 라. 시험불안

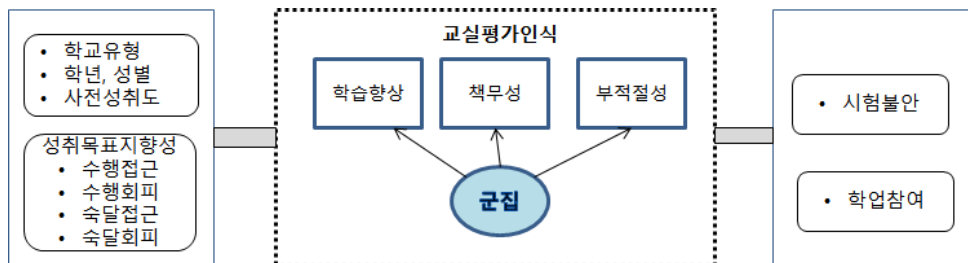
본 연구에서 활용한 시험불안 척도는 Spielberger(1980)의 Test Anxiety Inventory(TAI)를 고혜영(1992)이 수정 보완해서 타당화한 척도이다. 하위 요인은 일반불안, 정서 및 걱정 3가지로 구성되어 있다. 응답 양식은 5점 Likert 척도를 활용하였으며, 각 문항의 점수를 합산했을 때 점수가 높을수록 시험불안이 높은 것으로 설명된다. 하위요인별 문항 수는 일반불안이 4문항이며 정서와 걱정이 각각 8문항으로 총 20문항으로 이루어져있다. 하위요인별 신뢰도는 일반불안(시험점수 생각에 시험공부가 잘 안될 때가 있다)이 .634, 정서(시험은 나에게 좌절감과 패배감을 갖게 한다)가 .911, 걱정(이제 시험이나 성적 걱정에서 벗어났으면 좋겠다)이 .894로, 전체 신뢰도는 .948로 나타났다.

## 마. 학업참여

학생들의 학업에 대한 인지 및 정서적 참여를 의미하는 학업참여는 UWES-S(The Utrecht Work Engagement Scale) 척도를 고등학생을 대상으로 국내 타당화한 척도(추헌택, 손원숙, 2011)를 사용하여 측정하였다. 응답 양식은 5점 Likert 척도를 활용하였으며, 총 15문항이 3개 하위요인으로 구분된다. 과업에 대한 높은 수준의 에너지인 ‘열정(6문항)(나는 학생으로서 공부할 때, 에너지가 넘친다)’, 자신의 과업이 가지는 의미를 알고 학업에 대한 높은 열의 및 충성심을 뜻하는 ‘헌신(4문항)(나에게 공부는 힘들지만, 도전할 만한 것이다)’, 자신이 하는 일에 완전히 집중하는 ‘몰두(5문항)(나는 공부할 때, 공부에 완전히 빠져든다)’ 등이다. 하위요인별 신뢰도는 열정이 .890, 헌신이 .870, 몰두가 .872로, 전체 신뢰도는 .942로 나타났다.

### 3. 분석방법

본 연구에서 수집된 자료는 PASW Statistics 및 AMOS 18을 이용하여 다음과 같은 방법으로 분석되었다. 첫째, 자료의 내적 합치도와 타당도 분석을 위해 신뢰도 분석과 요인분석을 실시하였다. 둘째, 학생의 교실평가인식의 유형 분석을 위하여 학습향상, 책무성과 부적절성을 각각 군집변인으로 설정하여 이단계 군집분석(a two-step cluster analysis)을 실시하였다. 군집변인은 모두  $z$  점수로 변환하여 분석하였고, 군집 수 결정을 위하여 상대적 적합도 지수인 Bayes Information Criterion(BIC)를 활용한 ‘자동결정’ 및 이론적 가정과 군집의 해석 가능성을 고려한 ‘사전 결정’ 방법 등을 상호보완적으로 활용하였다. 또한 도출된 군집의 질을 판단하기 위하여 결합 및 분리의 실루엣 측도(silhouette measure of cohesion and separation)(.5이상은 양호, .2~.5는 보통, .2이하는 나쁨 등)를 검토하였다. 셋째, 학생의 교실평가인식을 예측하는 요인으로 학교 유형, 학년, 성별, 사전성취도, 성취목표지향성을 가정하고 다항 로지스틱 회귀분석(multinomial logistic regression analysis)을 실시하였다. 마지막으로 도출된 학생의 교실평가인식 유형별 시험불안과 학업참여의 차이를 살펴보기 위해 일원배치 분산분석 및 Scheffé 사후분석을 실시하였다. 본 연구의 분석모형을 도식화한 것은 [그림 III-1]과 같다.



[그림 III-1] 분석 모형

## IV. 연구결과

### 1. 주요 변인 간 상관 및 기술통계

고등학생의 평가인식, 사전성취도, 성취목표지향성, 시험불안, 학업참여 변인의 일반적인 경향을 알아보기 위해 평균, 표준편차를 포함한 기술통계치와 상관관계를 알아보았다(<IV-1>참고). 학생의 평가인식 하위요인의 평균은 학습향상은 3.81, 책무성은 4.03, 부적절성은 2.88로 나타나 책무성 인식이 가장 높았다. 학생의 평가인식 하위요인들 간의 상관관계를 살펴보면

학습향상과 책무성은  $r=.55(p < .01)$ 의 정적 상관을 보였고 학습향상과 부적절성은  $r=-.41(p < .01)$ , 책무성과 부적절성은  $r=-.16(p < .01)$ 의 부적상관을 보였다. 한편 학생의 평가인식과 성취목표지향성의 하위요인들과의 상관을 살펴보면, 학습향상과 책무성 요인 모두 성취목표지향성의 모든 하위요인들과 정적 상관을 가지면서도 접근목표 요인들과 더 강한 상관이 있었다. 반면 부적절성 요인은 모두 부적상관을 갖고 있었으며 수행접근, 숙달접근 요인과 비교적 더 강한 부적상관을 보였다. 마지막으로 학생의 교실평가인식과 시험불안은 책무성 인식( $r=.11, p < .01$ )을 제외하고는 통계적으로 유의한 관련성을 보이지 않았고, 학업참여와는 학습향상과 책무성은 정적( $r=.46, .22, p < .01$ )인, 부적절성과는 부정적( $r=-.31, p < .01$ )인 상관을 나타내었다.

&lt;표 IV-1&gt; 주요 변인 간 상관 및 기술통계치(N=662)

변인명	1	2	3	4	5	6	7	8	9	10
1 학습향상	1									
2 책무성	.55**	1								
3 부적절성	-.41**	-.16**	1							
4 사전성취도	.21**	.14**	-.24**	1						
5 수행접근	.39**	.32**	-.28**	.25**	1					
6 숙달회피	.19**	.24**	-.08*	.05	.66**	1				
7 숙달접근	.45**	.31**	-.31**	.22**	.48**	.18**	1			
8 숙달회피	.21**	.20**	-.09*	.06	.48**	.57**	.43**	1		
9 시험불안	.03	.11**	-.04	.06	.31**	.46**	.15**	.52**	1	
10 학업참여	.46**	.22**	-.31**	.30**	.27**	.01	.57**	.20**	.00	1
평균	3.81	4.03	2.88	5.37	3.51	3.25	3.43	3.00	2.95	2.87
표준 편차	1.00	.80	1.17	1.59	.82	.88	.71	.81	.83	.75

\* $p < .05$ , \*\* $p < .01$ 

## 2. 학생의 교실평가인식 유형 분석

학생의 교실평가인식의 유형 분석을 위하여 학습향상, 책무성과 부적절성을 각각 군집변인으로 설정하고 이단계 군집분석을 실시하여 최적의 군집수를 도출하였다. 연구자의 사전 가정에 의거하여 2~5개의 군집 모형이 비교되었고, 이와 더불어 BIC 기준에 의거하여 자동적으로 군집 수를 결정한 결과, 총 3개의 군집이 적절한 것으로 파악되었다. 3개의 군집의 경우 결합 및 분리의 실루엣 측도 역시 .5이상으로 양호한 것으로 판단되었고, 각 군집의 차별적 특성이 가장 두드러져 해석 가능성 역시 높은 것으로 나타났다. 따라서 본 연구에서는 교실평가인식 유형을 3개로 결정하였고, <표 IV-2>에 군집유형에 따른 군집변인의 양상을 제시하였다. 도출된 군집들의 양상을 해석하기 위하여 신이나와 손원숙(2012)의 기준을 적용하였는데 군집 변인의 점수(z점수)가 .75이상은 매우 높음(VH), .25~.75는 높음(H), -.25~.25는 평균(A),

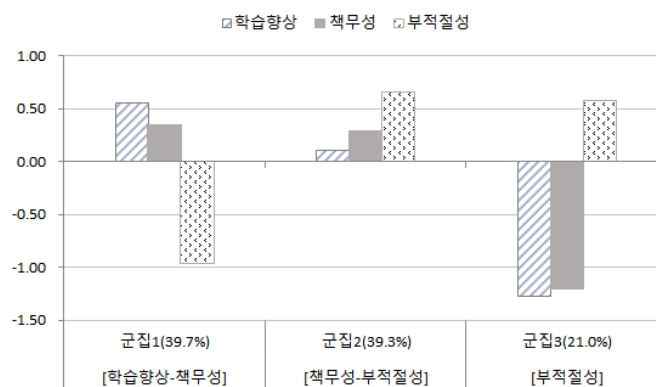
-.75~-.25는 낮음(L), -.75이하는 매우 낮음(VL)으로 해석하였다.

<표 IV-2> 학생의 교실평가인식 군집분석 결과

군집 유형	빈도 (%)	군집변인		
		학습향상	책무성	부적절성
1 학습향상-책무성	263 (39.7%)	H 0.56	H 0.35	VL -0.96
2 부적절성-책무성	260 (39.3%)	A 0.11	H 0.29	H 0.66
3 부적절성	139 (21.0%)	VL -1.27	VL -1.20	H 0.58

VL=매우낮음, L=낮음, A=평균, H=높음

도출된 군집변인의 양상에 기반 하여 군집별 명칭을 부여하였는데 첫 번째 군집은 부적절성 인식은 ‘매우 낮음’이며, 학습향상과 책무성 인식이 각각 ‘높음’으로 나타났고, 전체 학생의 39.7%가 이 군집에 소속되었다. 세 개의 평가인식 중 학습향상과 책무성 인식이 높은 이 군집을 학습향상-책무성(Learning Improvement-Accountability) 유형으로 명명하였다. 한편 학습향상 인식은 평균적인 수준을 보였으나 상대적으로 평가에 대한 책무성과 부적절성 인식이 높은 두 번째 군집이 도출되었다. 이 유형에는 전체 학생의 39.3%가 소속되었고, 이를 부적절성-책무성(Irrelevance-Accountability) 유형으로 명명하였다. 마지막으로 세 번째 군집에 속한 학생들은 교실평가의 유용성을 전혀 인식하지 못하고 평가에 대한 부정적인 인식이 높은 경우로써 이들은 학습향상 및 책무성 인식은 매우 낮고, 부적절성 인식만 높은 것으로 나타났다. 따라서 이 군집은 부적절성(Irrelevance) 유형으로 명명하였고, 여기에는 전체 학생의 21.0%를 차지하였다. 참고로 [그림 IV-1]에 각 군집별 학생의 평가인식 유형 양상이 제시되어 있다.



[그림 IV-1] 교실평가인식 유형 양상

### 3. 학생의 평가인식 유형 예측 요인 분석

다음으로 학생의 교실평가인식을 예측하는 요인들을 탐색하기 위하여 다항 로지스틱 회귀 분석을 실시하였다(<표 IV-3> 참고). 먼저 학습향상-책무성 유형과 부적절성 유형을 통계적으로 유의하게 분류하는 변인은 성취목표지향성 중 수행접근과 숙달접근이었고, 학교유형이나 학생 특성, 사전 성취도 등의 배경 변인들은 이 두 집단을 유의하게 구분하지 않았다. 구체적으로 부적절성 유형과 비교하여 학습향상-책무성 유형에 속할 승산비는 수행접근에서는  $\exp(b)=2.37$ , 숙달접근에서는  $\exp(b)=3.78$ 로써 각각 137%, 278% 정도 높았다. 즉, 접근지향적 성취목표를 지향할수록 부적절성 유형보다 학습향상-책무성 유형에 속할 확률이 높았다.

둘째, 학습향상-책무성 유형과 부적절성-책무성 유형을 유의하게 분류하는 변인은 학교 유형, 사전성취도, 수행접근, 숙달접근으로 나타났다. 부적절성-책무성 유형과 비교하여 학습향상-책무성에 속할 특목고의 승산비는  $\exp(b)=2.48$ 로, 일반고와 비교하여 148% 높았다. 즉, 특목고 학생이 일반고 학생보다 부적절성-책무성 유형에 비해 학습향상-책무성 유형에 속할 확률이 높았다. 또한 사전성취도( $\exp(b)=1.19$ )가 높을수록, 수행( $\exp(b)=2.42$ ) 또는 숙달( $\exp(b)=1.66$ ) 접근지향적 성취목표를 지향할수록 부적절성-책무성 유형에 비해 학습향상-책무성 유형에 속할 확률이 높았다.

<표 IV-3> 다항 로지스틱 회귀분석 결과

참조집단	부적절성 유형			부적절성-책무성 유형			부적절성 유형		
비교집단	학습향상-책무성 유형						부적절성-책무성 유형		
독립변인	B	SE	exp(B)	B	SE	exp(B)	B	SE	exp(B)
학교 유형 (1=특목고)	-.03	.34	.97	.91**	.31	2.48**	-.94*	.36	.40*
학년 (1=2학년)	-.06	.25	.94	-.06	.20	.94	.00	.23	1.00
성별 (1=여자)	.39	.25	1.48	.02	.19	1.02	.37	.23	1.45
사전성취도	.15	.08	1.16	.17**	.06	1.19**	-.03	.07	.97
수행접근	.86***	.22	2.37***	.88***	.19	2.42***	-.02	.21	.98
수행회피	.20	.21	1.22	-.37	.17	.69	.57**	.20	1.77**
숙달접근	1.33***	.22	3.78***	.51**	.18	1.66**	.82***	.21	2.28***
숙달회피	-.31	.21	.73	-.25	.16	.78	-.06	.20	.95

\*\* $p < .01$ , \*\*\* $p < .001$

마지막으로 부적절성 유형과 부적절성-책무성 유형을 통계적으로 유의하게 분류하는 변인으로는 학교 유형, 수행회피, 숙달접근으로 나타났다. 부적절성 유형과 비교하여 부적절성-책

무성에 속할 학교 유형(1=특목고)의 승산비는  $\exp(b)=.40$ 으로, 일반고에 비교하여 특목고의 경우 60% 낮았다. 즉, 일반고 학생이 특목고 학생보다 부적절성 유형에 비해 부적절성-책무성에 속할 확률이 높은 것으로 나타났다. 또한 수행회피목표( $\exp(b)=1.77$ )를 지향할수록, 숙달접근 목표( $\exp(b)=2.28$ )를 지향할수록 부적절성 유형보다는 부적절성-책무성 유형에 속할 확률이 높은 것으로 나타났다.

#### 4. 학생의 평가인식 유형에 따른 시험불안 및 학업참여의 차이 분석

학생의 평가인식 유형에 따른 시험불안 및 학업참여의 차이를 알아보기 위해 일원배치 분산 분석 및 Scheffé 사후분석을 실시하였으며, 그 결과는 <표 IV-4>에 제시되어 있다. 일원배치 분석 결과, 시험불안 및 학업참여는 교실평가인식 유형에 따라 통계적으로 유의한 차이를 보였다( $p < .05$ ,  $p < .001$ ). 사후검증 결과, 학습향상-책무성 유형과 부적절성-책무성 유형의 시험불안은 각각 .06과 .05로 유의한 차이를 보이지 않았고, 이들은 부적절성 유형(-.21) 보다는 통계적으로 유의하게 높은 수준의 시험불안을 나타내었다. 한편 학습향상-책무성 유형(.36)에 속한 학생들이 통계적으로 유의하게 가장 높은 수준의 학업참여를 보였고, 부적절성 유형(-.50)에 속한 학생들이 가장 낮은 학업참여 수준을 보였다.

<표 IV-4> 교실평가인식 유형에 따른 시험불안 및 학업참여의 차이

	교실평가인식 유형			<i>F</i> ( <i>df</i> =2)	사후검증
	1. 학습향상-책무성 ( <i>SD</i> )	2. 부적절성-책무성 ( <i>SD</i> )	3. 부적절성 ( <i>SD</i> )		
시험불안	.06 (.99)	.05 (.91)	-.21 (1.13)	3.976*	1,2>3
학업참여	.36 (.89)	-.10 (.96)	-.50 (1.02)	39.827**	1>2>3

\* $p < .05$ , \*\* $p < .01$

## V. 논의

본 연구에서는 우리나라 고등학생들의 평가인식 유형을 분석하고, 이를 예측하는 변인을 살펴보고 각 유형에 따른 학업참여와 시험불안의 차이를 탐색하였다. 이에 따른 연구 결과를 요약하고, 논의하면 다음과 같다.

첫째, 학습향상, 책무성 및 부적절성이라는 평가인식 하위요인에 기반 하여 교실평가에 대한 학생인식 유형을 분류해 본 결과, 총 세 개의 유형이 도출되었다. 가장 많은 비중을 차지하는 유형으로는 학습향상-책무성(39.7%) 및 부적절성-책무성(39.3%) 유형이었고, 그 다음으로는 부적절성 유형(21.0%)이 탐색되었다. 이는 현재 우리나라 고등학생들에게 다양한 평가인식 유형이 경험적으로 존재할 수 있음이 확인된 결과로써 ‘학습향상과 책무성’ 또는 ‘부적절성과 책무성’ 인식 등이 서로 공존할 수 있다는 점이 탐색되었다. 특히 본 연구를 토대로 우리나라 고등학생들은 평가의 목적을 학습향상과 책무성 모두로 인식하는 경향이 높다는 점이 파악되었다. 교실평가의 주된 목적에 해당하는 책무성 점검 및 학습향상의 목적이 학생들에 의하여 모두 인식되고 있다는 점은 현 교실평가 실제에 대한 긍정적인 결과라고 볼 수 있다. 그러나 이와 더불어서 평가가 부적절하다고 인식하면서도 평가의 목적이 학생 및 학교의 책무성을 확인하는 것이라고 인식하는 유형(부적절성-책무성) 역시 높은 비중을 차지하였다. 이는 교실평가가 학습개선을 위한 목적 보다는 학습 결과를 확인하려는 목적으로 주로 사용된다는 인식이 강하다는 점을 의미한다. 또한 이러한 책무성 인식이 평가에 대한 부정적인 인식과 공존할 수 있다는 면이 특징적이다. 이는 평가가 학습결과를 확인하는 것과 같이 책무성 점검을 목적으로 수행될 경우, 학생들은 자신의 학습 개선을 위한 유용한 정보를 평가로부터 받고 있다고 느끼기 보다는 자신의 성취수준이나 서열 등을 확인하는 용도로 인식되며 또한 이러한 과정에서 평가에 대해 부정적인 인식이 축적되었을 가능성을 예측해 볼 수 있다. 또한 고등학교의 평가는 대학 입시의 틀에 묶여 있으므로 학생의 강점과 약점, 교육적 요구를 진단할 수 있는 학습향상을 위한 평가를 구현해 내기 힘들다는 점이 반영된 결과로도 해석된다. 또한 평가가 학습에 아무런 도움이 되지 않는다고 생각하는 부적절성 유형의 학생들 역시 상대적으로 낮은 비중이기는 하지만 전체 21% 가량 존재한다는 점은 주목할 만하다. 이들은 현재 경험하고 있는 교실평가에 대하여 부적절하며 평가결과를 무시하는 경향을 보이고 있으며 책무성 점검이나 학습향상과 같은 평가의 순기능에 대하여 부정적인 인식을 보이고 있는 것으로 나타났다. 이처럼 평가에 대한 부정적 인식은 학습 태도와 학업성취에 부적인 효과를 보인다는 선행연구(Brown & Hirschfeld, 2007) 결과를 고려해볼 때, 부정적 인식을 야기 시키는 원인을 보다 종합적으로 탐색해 보아야 할 것이다.

둘째, 본 연구에서 도출된 학생의 교실평가 인식의 세 가지 유형은 개인의 성취목표지향성에 따라 모두 구분되었고, 학교유형 및 사전 성취도 역시 일부 유형을 예측하는 요인으로 나타났다. 먼저 학교 유형의 경우, 특목고 학생이 일반고 학생에 비하여 부적절성-책무성 유형보다는 학습향상-책무성 유형, 부적절성-책무성 유형보다는 부적절성 유형에 속할 확률이 높은 것으로 나타났다. 이는 일반적으로 수업참여나 성취도 측면에서 우수한 것으로 알려진 특목고 학생들이 학습향상과 책무성과 같은 긍정적인 인식도 높았지만, 더불어서 평가에 대한 부정적인 인식 역시 동시에 높음을 알 수 있다. 물론 본 결과의 일반화를 위해선 모집단을 잘 대표하

는 표본자료가 필요하다는 점에서는 보수적인 해석이 필요하지만 학교 유형에 따른 평가인식의 차이 및 원인에 대한 추후 연구는 제안된다. 다음으로 사전성취도 역시 평가인식 유형을 유의하게 예측하였는데, 사전성취도가 높을 경우 다른 유형보다는 학습향상-책무성 유형에 속할 확률이 높았다. 이는 학생의 사전 성취경험이 긍정적인일 경우 평가에 대하여 긍정적 인식을 가질 가능성이 높고, 반대로 사전 성취도가 낮을 경우 평가를 부적절하다고 인식하는 경향이 높음을 의미한다. 따라서 교사들은 성취수준이 낮은 학생들의 평가인식을 개선하기 위하여 수업 중 형성평가를 지속적으로 실시하여 긍정적인 평가 결과의 경험을 제공할 필요가 있을 것이다. 또한 학생 개인차를 고려한 구체적이고 피드백을 제시해주고 적절한 수업 조정의 노력을 통해 평가의 유용성을 지각하게 할 수 있을 것이다.

한편 성취목표지향성과 평가인식과의 관련성을 살펴보면, 학습향상-책무성 유형은 부적절성-책무성 또는 부적절성 유형과는 접근지향적 성취목표지향성 측면에서 차이를 보였다. 즉 수행 또는 숙달접근적 성취동기가 높은 학생일수록 학습향상-책무성 유형에 포함될 확률이 높았다. 접근지향적인 성취목표지향성을 지닌 학생들은 다양한 학습전략을 사용하는 경향이 있고, 또한 학습의 인지 및 정서적 몰입이 강하다는 선행연구(Diseth, 2011)와 일관적으로 이러한 학습자들은 학습향상과 책무성이라는 평가의 목적을 비교적 균형있게 인식하고 있다고 볼 수 있다. 한편 수행회피목표 지향성이 높을수록 부적절성 유형보다는 부적절성-책무성 유형에 속할 확률이 높았다. 이는 평가를 부적절하게 여기면서도 평가 결과에 대한 책임감은 강하게 인식하는 학생들이 타인의 시선을 의식하고 자신의 부족함을 드러내는 것을 회피하려는 수행회피경향이 강하다고 해석할 수 있다. 따라서 학습 결과에 대한 평가의 지나친 강조를 통해 학생들이 평가에 대한 부정적인 인식이 축적되게 하기 보다는 평가가 실제 학생들의 학습향상에 유용하게 활용될 수 있도록 하는 것이 필요하다. 예를 들어, 학생들에 대한 지속적인 형성평가를 토대로 학생의 강점과 약점을 파악하여 학생의 학습향상을 위한 구체적인 방법에 대한 조언을 제공하거나, 자신의 평가 결과에 대한 자기평가를 시행하는 것이 한 방법이 될 수 있을 것이다.

셋째, 학생의 평가인식 유형이 학생의 평가 수행과 학습 태도에 어떠한 영향을 미치는가를 알아보기 위해 학생의 평가인식 유형별 시험불안과 학업참여의 차이를 살펴보았다. 학습향상-책무성 유형과 부적절성-책무성 유형에 속한 학생들의 시험불안은 서로 유의한 차이가 없었지만, 부적절성 유형보다는 통계적으로 유의하게 높았다. 이는 평가결과에 대한 책임을 인식하는 학생들이 그렇지 않은 학생들 보다 시험불안이 높은 것으로 해석된다. 시험불안과 학업성취간의 연구(Owens, et al., 2014)를 참고해 볼 때, 시험불안이 지나치게 높거나 낮은 경우 보다는 적절한 수준의 시험불안이 높은 학업성취와 관련성을 갖는다고 하였다. 따라서 평가를 부적절하게 인식하는 학생들에게는 평가 결과에 대한 책임을 인식시켜 줌으로써 시험불안을 적절한 수준으로 조절하여 학업을 촉진시키는 방안을 제안할 수 있을 것이다. 한편 학업참여는 학습향상-책무성 유형에서 가장 높은 수준을 보였고, 부적절성-책무성 유형, 부적절성 유형

순으로 나타났다. 이러한 결과는 학생들이 평가를 부적절하다고 인식하기 보다는 학습향상과 책무성과 같은 긍정적 인식을 가질 경우 학업에 보다 적극적으로 참여함을 시사한다. 따라서 학생들이 적극적으로 학업에 참여하도록 독려하기 위해서는 평가에 대한 긍정적인 인식을 도모할 수 있는 방안이 요구된다. 예를 들어 학생들의 성취수준에 맞는 평가를 제공하고, 평가결과에 대한 구체적인 피드백을 제공하며, 이에 맞는 수업을 구안하는 등의 노력은 평가에 대한 긍정적인 인식을 함양시키는 데 도움이 될 것으로 생각된다.

요약하면 본 연구는 교실평가의 다양한 목적 및 성격에 대한 학생들의 인식을 유형화하여 각 유형별 특성을 분석함으로써 현재 우리나라 교실평가의 역할 및 기능에 대한 시사점을 도출하고자 하였다. 구체적으로 본 연구 결과를 토대로 긍정적인 교실평가 인식의 중요성이 확인되었다. 이러한 긍정적인 인식 함양은 실제 교실평가 방향을 개선함으로써 가능할 것이며, 이를 위해서 교사들을 대상으로 한 평가 교육을 활성화하고 교사들이 참고할 수 있는 다양한 평가 자료 등을 제공해 줄 필요가 있을 것이다. 특히 교수개선 및 학습향상에 도움이 될 수 있는 구체적인 평가방안에 대한 기초연구를 수행하여, 학교급 및 교과별 특수성을 반영한 교실평가 모형을 구안할 필요성이 제안된다. 평가는 수업의 일부로써 평가방안의 적합성은 학교급이나 교과에 따라서 차별성을 보일 것이므로 이에 대한 면밀한 검토를 토대로 학교 현장에서 유용하게 참고할 수 있는 평가모형 마련을 제안한다.

본 연구의 제한점을 토대로 추후 연구에 대한 제언을 추가하면 다음과 같다. 첫째 본 연구에서는 국외에서 개발된 SCoA척도를 활용하여 학생들의 평가인식을 측정하였다. 이는 이론적 틀에 기반 하여 타당화된 척도라는 면에서는 유용성이 있으나 국내 표본을 대상으로 한 평가 인식 척도의 개발 및 적용을 위한 후속연구가 제안된다. 둘째, 본 연구의 대상은 고등학생으로 제한되었으나 학교급별 교실평가 정책 및 실제의 차이로 인한 학생들의 인식 역시 차별성을 보일 것으로 예상되므로 연구대상의 확대 및 학교급별 비교를 위한 연구가 요구된다. 셋째, 학생인식 유형 도출을 위하여 본 연구에서 적용한 군집분석 이외에 추후연구에서는 측정오차를 통제하며 영향요인과의 관련성을 보다 종합적으로 탐색할 수 있는 잠재변인 접근법에 기반 한 혼합모형의 사용이 제안된다(Asparouhov, & Muthén, 2014). 넷째, 본 연구에서는 학생들의 평가에 대한 인식의 예측요인으로 사전성취도 및 성취동기 변인을 주요하게 탐색하였으나 추후 연구에서는 교수학습 및 학습동기 이론 등을 반영하여 보다 다양한 예측요인을 포함할 필요가 있을 것이다. 한편 자기보고식으로 측정된 사전성취도의 경우 신뢰도 측면에서 한계가 있으므로 추후 연구에서는 객관적인 성취정보의 활용을 제안한다. 마지막으로 학생과 교사의 평가인식의 차이 및 원인을 파악할 수 있는 연구도 추후 제안되며, 이는 교수학습과정에서 평가 실제의 개선 및 평가에 대한 교사 및 학생의 긍정적인 인식 발달을 위한 주요한 개입점이 될 수 있을 것이다.

## 참 고 문 헌

- 강대중, 염시창(2014). 초등학교 교사들의 형성평가에 대한 인식 및 실태 분석. **학습자중심교과교육연구**, 14(2), 27-43.
- 고혜영(1992). 교사기대지각 · 학문자아개념 · 시험불안과 학업성취간의 분석연구. 숙명여자대학교 대학원 석사학위 논문.
- 교육부(2016. 1. 25). 2016년 교육부 업무계획. 교육부 보도자료.
- 김명숙(2008). **교육과정, 교수-학습 및 평가의 연계: 학습을 위한 평가. 교수-학습 혁신 세미나: 교수-학습과 평가의 연계**. 한국교육과정평가원 연구자료. ORM 2008-18, 29-48.
- 김빛나(2014). 긍정적 정신 건강, 시험 관련 정서 평가 및 학업 소진 간 관계. 고려대학교 대학원 석사학위 논문.
- 김신영(2014). 학생평가에 대한 교사의 인식과 실천. **교육평가연구**, 27(1), 141-161.
- 김진수, 김석우(2008). 수학과 수행평가에 관한 고등학생의 인식연구. **교육연구**, 18(1), 1-21.
- 도승이, 손수경, 변준희, 임지윤(2011). 한국어판 성취정서 질문지 (K-AEQ) 개발 및 타당화. **교육심리연구**, 25(4), 945-970.
- 박민애, 손원숙(2017). 초등 교실수업에서의 학습을 위한 평가와 정의 및 인지적 성취와의 관계. **초등교육연구**, 30(3), 21-42.
- 박병기, 임신일, 김어진(2008). 시험불안의 재개념화에 의한 척도개발 및 요인구조 탐색. **교육심리연구**, 22(1), 87-109.
- 박정, 이진주(2015). 학생 자기평가에 대한 초등교사의 인식과 실제. **학습자중심교과교육연구**, 15(11), 601-620.
- 신이나, 손원숙(2012). 영재와 일반 중학생의 성취목표지향성 프로파일 분석. **한국심리학회: 학교**, 9(1), 65-83.
- 이주화, 김아영(2005). 학업적 성취목표지향성 척도개발. **교육심리연구**, 19(1), 311-325.
- 지은림(2010). 교사의 형성평가 피드백 수행 척도 개발 및 타당화: Rasch 모형의 적용. **교육평가연구**, 23(1), 79-100.
- 추헌택, 손원숙(2011). 한국판 학업참여 척도(UWES-S)의 타당화. **교육평가연구**, 24(4), 897-920.
- 홍정림, 최은경(2002). 과학과 수행평가에 관한 중등학생의 인식 및 자아효능감 조사. **한국과학기술정보연구원**, 22(2), 230-239.
- Asparouhov, T. & Muthén, B.(2014). Auxiliary Variables in Mixture Modeling:

- Three-Step Approaches Using Mplus. *Structural Equation Modeling*, 21, 329 - 341.
- Black, P. & Wiliam, D. (1998). Assessment and classroom learning. *Assessment in Education*, 5(1), 7-74.
- Black, P., Wilson, M., & Yao, S. Y. (2011). Road maps for learning: A guide to the navigation of learning progressions. *Measurement: Interdisciplinary Research & Perspective*, 9(2-3), 71-123.
- Brookhart, S. M. (1997). A theoretical framework for the role of classroom assessment in motivating student effort and achievement. *Applied Measurement in Education*, 10, 161-180.
- Brown, G. T. L., & Hirschfeld, G. H. F. (2007). Students' conceptions of assessment and mathematics: Self-regulation raises achievement. *Australian Journal of Educational and Developmental Psychology*, 7, 63 - 74.
- Brown, G. T. L., Hui, S. K., Flora, W. M., & Kennedy, K. J. (2011). Teachers' conceptions of assessment in Chinese contexts: A tripartite model of accountability, improvement, and irrelevance. *International Journal of Educational Research*, 50(5), 307-320.
- Diseth, Å. (2011). Self-efficacy, goal orientations and learning strategies as mediators between preceding and subsequent academic achievement. *Learning and Individual Differences*, 21(2), 191-195.
- Eccles, J. S. & Wigfield, A. (2002). Motivational Beliefs, Value, And Goals. *Annual Review of Psychology*, 53, 109-132.
- Elliot, A. J., & McGregor, H. A. (2001). A 2× 2 achievement goal framework. *Journal of Personality and Social Psychology*, 80(3), 501.
- Hirschfeld, G. H., & Brown, G. T. (2009). Students' conceptions of assessment: Factorial and structural invariance of the SCoA across sex, age, and ethnicity. *European Journal of Psychological Assessment*, 25(1), 30-38.
- James, M., & Pedder, D. (2006). Beyond method: Assessment and learning practices and values. *The Curriculum Journal*, 17(2), 109-138.
- McMillan, J. H. (2014). *Classroom assessment: Principles and practice for effective standards-based instruction. 6<sup>th</sup> edition*, New York: Pearson Higher Ed. [손원숙, 박정, 강성우, 박찬호, 김경희 역 (2015). **교실평가의 원리와 실제: 기준참조수업과의 연계**. 서울: 교육과학사].
- Owens, M., Stevenson, J., Hadwin, J. A., & Norgate, R. (2014). When does anxiety help

- or hinder cognitive test performance? The role of working memory capacity. *British Journal of Psychology*, 105(1), 92-101.
- Spielberger, C. D. (1980). *Test Anxiety Inventory*, Preliminary Professional Manual for the Test Anxiety Inventory(「Test Attitude Inventory」) TAI. Palo Alto: Consulting Psychologists Press.
- Stiggins, R. J. (2005). From Formative Assessment to Assessment FOR Learning: A Path to Success in Standards-Based Schools. *Phi Delta Kappan*, 87(4), 324-328.
- Wiliam, D, & Thompson, M. (2007). *Integrating assessment with instruction: What will it take to make it work?* In C.A. Dwyer(ed) *The Future of Assessment: Shaping Teaching and Learning*, Mahwah, N.J.: Lawrence Erlbaum Associates.

· 논문접수 : 2017.12.31. / 수정본접수 : 2018.02.05. / 게재승인 : 2018.02.19.

## ABSTRACT

### Identifying Types of High School Students' Conceptions of Assessment

**Young-Min Jo**

Graduate Student, Kyungpook National University

**Won-Sook Sohn**

Professor, Kyungpook National University

The main purpose of this study was to examine the types of high school students' conceptions of assessment. Also the effects of previous academic achievement and achievement goal orientations on them and students' test anxiety and academic engagement by their types were explored. The subjects were 662 Korean high school students (10 to 11th graders) sampled from six high-schools located in D-city and K-region. The results showed that three groups of students with unique conceptions of assessment were identified by the two-step cluster analyses such as Learning-Accountability (39.7%), Irrelevance-Accountability (39.3%), and Irrelevance (21.0%) groups. Secondly, the variables predicting these types have been found to be school types, previous academic achievement, and performance or mastery approach goal orientations. Additionally, test anxiety of the learning-accountability group was as almost high as that of the irrelevance-accountability group, which was higher than that of the irrelevance group. Academic engagement was the highest in the learning-accountability group, and the irrelevance-accountability group showed higher academic engagement than the irrelevance group, which was ranked the lowest. Finally students' conceptions of assessment was articulated and implications for teachers' conceptions of assessment as well as further studies were discussed.

*Key Words : Classroom assessment, Students' Conceptions of Assessment, Assessment for Learning(AFL), Assessment of Learning(AOL), Accountability*