

교사의 학생평가 역량 연구

임 은 영(한국교육과정평가원 부연구위원)*

<요 약>

최근의 학생평가 동향인 과정 중심 평가, 고교학점제, 교사별 학생평가, 대입전형에서의 수시 비율의 확대 등에 의해 교실 단위의 학생평가에 대한 중요성이 높아지고 있고, 이와 함께 교사의 학생평가 역량 중요성이 더욱 부각되고 있다. 이에 본 연구에서는 교사의 학생평가 역량을 측정할 수 있는 18개 문항으로 구성된 자기보고식 척도를 개발하였다. 척도에는 최신 학생평가 동향인 과정 중심 평가와 성취평가제의 특성을 측정할 수 있는 문항들을 포함하여, 최신 학생평가와 관련된 역량을 실태를 파악할 수 있도록 개발하였다. 학생평가 역량 조사는 온라인 설문으로 진행되었고, 총 1,307명의 교사가 응답에 참여하였다. 이 중 300명의 교사의 응답을 척도 타당화를 위해 사용하였고, 나머지 1,007명의 교사 응답을 바탕으로 실태조사를 하였다. 척도 타당화를 위해 탐색적 및 확인적 요인 분석을 실시하였고, 척도는 평가시행($\alpha=0.919$), 평가계획 및 도구개발($\alpha=0.913$), 윤리성 및 법적절차 준수($\alpha=0.938$), 결과환류($\alpha=0.876$), 4개의 요인으로 구성되어 있었다. 개발된 척도를 활용하여 학교 급에 따라, 교사의 교육 경력에 따라, 성별에 따라 교사의 학생평가 역량 실태조사를 하였다. 초등 교사가 모든 요인에서 다른 학교 급 교사에 비해 학생평가 역량을 높이 인식하고 있었고, 중학교, 고등학교 순으로 역량을 인식하고 있었다. 남녀 교사의 학생평가 역량은 모든 요인에서 초·중·고 모두 차이가 크지 않았으며, 교육 경력별로는 경력이 오래될수록 자신의 역량이 높다고 인식하고 있었다. 또한 최근 학생평가 정책에 대한 교사의 학생평가 역량을 살펴보면 과정 중심 평가와 관련된 항목들이 성취평가와 관련된 항목보다 높은 평균을 보였다.

주제어 : 교사의 학생평가 역량, 척도 개발 및 타당화, 교사의 학생평가 역량 실태, 과정 중심 평가, 성취평가제

* 제1저자, 교신저자 elim@kice.re.kr

I. 서 론

학교 교육 내실화를 위한 교사의 평가 역량 또는 전문성에 대한 논의는 2000년대 초반부터 지속적으로 이루어졌다(예, 김신영, 2002; 이인제 외, 2004; 김재춘 외 2005; 김경희 외, 2006 등). 최근 교실 단위의 학생 평가의 중요성이 부각되면서 교사의 학생평가 역량 또는 전문성에 대한 관심과 중요성이 커지고 있다. 교실 단위의 학생평가 중요성이 부각된 이유는 3가지로 정리될 수 있다. 첫째, 2015 개정 교육과정에서 과정을 중시하는 평가를 강조하고 있으며, 둘째, 성취평가제와 고교학점제 같은 학교 및 학급 단위의 학생평가 내실화 필요성 증가하고, 셋째, 대학입학전형의 수시 비율 증가로 인한 교사가 작성하는 학생부에 대한 중요성이 높아지고 있기 때문이다.

2015 개정 교육과정 총론(교육부, 2015)의 ‘교육과정 구성의 중점’을 살펴보면 학생평가에 대해, ‘학습의 과정을 중시하는 평가를 강화하여 학생이 자신의 학습을 성찰하도록 하고, 평가 결과를 활용하여 교수·학습의 질을 개선한다.’를 교육과정의 중점으로 기술하고 있다. 학생이 자신의 학습을 성찰하도록 하고, 평가 결과를 활용하여 교수 학습의 질을 개선한다는 것을 통해 교실에서 이루어지는 평가가 강조되고 있음을 알 수 있다. 둘째, 성취평가제 확산이나 고교학점제 도입 같은 정책은 학교 또는 학급 단위 학생평가의 내실화 필요성을 강조하고 있다. 성취평가제는 2012년 중학교 1학년을 시작으로 2016년에 고등학교 3학년까지 시행되어, 현재 중학교와 고등학교 모든 학교급에서 성취평가제가 시행 중이다. 고교학점제는 교육부가 2025년에 학점제를 본격적으로 시행하겠다는 학점제 3단계 도입 계획을 발표하였다(교육부, 2018). 고교학점제는 ‘진로에 따라 다양한 과목을 선택·이수하고, 누적 학점이 기준에 도달할 경우 졸업을 인정받는 교육과정 이수·운영(교육부, 2017)’제도로, 학점제를 통하여 다양한 과목 개설이 이루어질 것이다. 이와 같은 상황에서는 교과별 평가보다 교사별 평가가 확대될 것이다. 또한 학점이 기준에 도달해야만 졸업이 인정되므로 교사가 시행하는 평가가 더욱 중요해질 것이다. 이에 교육부(2017)가 제시한 학점제형 학사제도 운영 체계의 최종 완성 모형은 평가를 ‘수업과 연계한 과정 중심 교사별 평가’, ‘성취평가제 적용’으로 제안하고 있다(한혜정 외, 2018, p.10, p.133).

마지막으로 대입 전형에서 수시 입학 비율의 증가로 인해, 교사의 학생부 기록이 중요해지고 있다. 수시가 시작된 2000년에 수시 비율이 3.4%이었지만, 2018학년도에는 73.7%로 확대되었고, 이 중, 학생부 전형(교과 종합)의 비율이 63.6%를 차지하고 있다(박형수, 중앙일보,

2018.5.10.). 사회적으로 민감한 대학입학에서 학생부 전형의 비율이 커지면서 학생부를 작성하는 교사의 학생평가에 대한 권한이 커짐에 따라, 교사의 평가 전문성에 대한 요구가 높아지고 있다.

이처럼 학생평가의 민감성과 중요성이 증대되고 있는 상황에서 교사의 학생평가 역량 강화에 대한 중요성도 같이 커지고 있다. 이에 본 연구에서는 교사의 학생평가 역량과 전문성에 대한 기존 이론을 바탕으로 최근 한국의 학생평가 정책을 반영한 척도를 개발하고, 요인분석을 통해 교사의 학생평가 역량의 구인을 탐색하고 확인하였다. 나아가 타당화 된 척도를 사용하여 교사들의 학생평가 역량에 대한 기초조사를 하고 최근 학생평가의 정책(성취평가제 및 과정 중심 평가)과 관련된 평가 역량에 대한 기초분석을 통해 학생평가 정책과 관련된 교사의 학생평가 역량에 대한 실태조사를 하였다.

II.이론적 배경

1. 교사의 학생평가 역량

교사 역량과 전문성에 대한 대부분의 연구들은 교수학습, 교과전문성, 대인관계, 학생 이해, 의사소통, 상담, 행정/경영 등 교사의 직무와 관련된 영역에 대한 연구들이 주를 이루고 있다(조대연, 2009; 박용호 외, 2012; 김희규, 주영호, 엄준용, 2016). 교사의 학생평가 역량 및 전문성은 교수학습의 중요한 부분임에도 불구하고 많은 연구가 이루어지지 않고 있다.

교사의 학생평가 역량에 본격적으로 관심을 두기 시작한 것은 1990년 AFT(American Federation of Teachers), NCME(National Council on Educational Association), NEA(National Education Association)에서 교사의 학생평가 역량에 대한 7가지 기준을 제시한 것부터 볼 수 있다. 7가지 기준은 “①교사는 교수적 결정을 위해 적절한 평가 방법 선택에 능숙해야 한다. ②교사는 교수적 결정을 위해 적합한 평가 방법 개발에 능숙해야 한다. ③교사는 외부에서 제작한 평가도구와 교사가 제작한 평가도구 모두 시행, 채점, 해석에 능숙해야 한다. ④개별 학생, 수업 계획, 교육과정 개발, 학교 향상을 위한 결정을 할 때, 평가 결과를 사용하는 데 능숙해야 한다. ⑤타당한 채점 절차를 개발하는 데 능숙해야 한다. ⑥학생, 학부모, 평가 결과에 대한 특별한 지식이 없는 일반인, 다른 교육자 또는 교사들과의 의사소통에 능숙해야 한다. ⑦교사는 비윤리적이고, 불법적이고, 적절하지 않은 평가 도구와 평가 정보 사용을 인지하는 데 능숙해야 한다(AFT, NCME & NEA, 1990)”이다. 이 기준 중 ①, ②, ⑤는 주로 평가계획에서 이루어지는 역량이다. 또한 기준 ③은 외부 평가에 대한 시행 채점 해석, ④, ⑥

은 평가결과 활용, ⑦은 평가에 대한 윤리성과 법적 절차 관련으로 기준들을 범주화할 수 있다.

Brookhart(2011)은 1990년 7가지 기준이 형성평가(formative assessment)와 성취기준 기반 평가(standard-based assessment)를 포함하고 있지 않다고 지적하였고, 형성평가와 성취기준에 대한 내용을 포함하면서 1990년 기준보다 더욱 구체화된 교사에게 필요한 교육평가 지식과 기술을 11가지 기준으로 제시하였다.

11가지 기준은 “①교사는 가르치는 내용 영역에 대한 학습을 이해해야 한다. ②교사는 성취기준과 교육과정 목표에 근거한 내용과 생각의 깊이 모두 부합하는 학습 의도를 달성 가능하고 평가 가능한 방법으로 명확하게 말할 수 있어야 한다. ③교사는 학습 의도에 따른 성취가 무엇인지 학생들과 의사소통할 수 있는 전략 레퍼토리가 있어야 한다. ④교사는 사용 가능한 평가 방법 범위의 목적과 내용을 이해해야 하고 그것들을 활용할 때, 능숙해야 한다. ⑤교사들은 학생들이 학급 단위 시험 문항, 수행평가 과제들을 수행하는 데 필요한 구체적인 지식 사고력을 알아보기 위해 그것들을 분석하는 기술을 가지고 있어야 한다. ⑥교사는 학생의 과제에 대해 효과적이고 유용한 피드백을 제공하는 기술을 가져야 한다. ⑦교사는 교실 평가에서의 학생 수행을 학생, 교실, 학교, 지역에 관한 결정을 하기 위한 유용한 정보로 활용하기 위한 수량화 하는 채점 계획을 구성할 수 있어야 한다. 이와 같은 결정은 학생의 학습, 성장, 발달 향상을 이끌어야 한다. ⑧교사는 학생, 학급, 학교, 지역과 관한 결정을 위해 외부 평가를 시행하고 점수를 해석할 수 있어야 한다. ⑨교사는 평가 결과에 대한 해석과 평가결과에 기반을 둔 교육적인 결정에 관한 추론을 교육관계자에게 명확하게 말할 수 있어야 한다. ⑩교사는 학생들이 평가결과를 활용하여 건전한 교육적 결정을 만들 수 있도록 도울 수 있도록 해야 한다. ⑪교사들은 자신이 시행하는 평가에 대한 법적 윤리적 책임을 이해하고 실행해야 한다.”이다.

Brookhart(2011)의 기준은 평가계획, 시행, 채점, 해석, 활용에 앞서, 가르치는 내용 영역, 학생의 성장, 성취기준과 교육과정에 따른 학습 의도 등을 알아야 한다는 것을 강조하였다. ①, ②, ③은 평가가 이루어지기 전에 교사가 알아야 하는 지식과 기술이면서 동시에 평가 계획을 수립하기 위해 필요한 것들이다. ④의 경우 평가 계획 단계에서 평가 방법 선정 및 평가 도구 개발을 위해 필요한 지식과 기술이다. ⑤, ⑦은 채점 및 해석과 관련된 기준이고, ⑥은 결과에 대한 의사소통, ⑧은 외부 평가에 대한 시행, 해석, 채점, ⑨와 ⑩은 결과 활용 ⑪은 평가 결과 활용에 대한 법과 윤리적 책임과 관련된 영역으로 범주화할 수 있다.

한국에서 교사의 학생평가 역량 및 전문성 연구를 살펴보면, 2000년대 초부터 이루어졌다. 김신영(2002)은 현장 교사의 평가 전문성 연구를 위해 AFT, NCME, NEA에서 개발한 35개 문항을 번안하여 교사들의 학생평가 전문성 실태조사를 하였다. 이 연구에서 교사들이 가장 높은 평균을 보인 것은 윤리적 문제에 대한 의식이었고, 가장 낮은 것은 평가 결과에 대한 의

사소통이었다. 김경희 외(2006)는 교사의 학생평가 전문성 기준을 개발하기 위해, 개념을 분석하고 능력요소를 추출하였다. 이 연구에서는 한국 교사들의 평가 관련 직무 영역과 미국의 학생평가 전문성 기준 영역 등을 고려하여 5가지 능력으로 범주화하였다. 5개 영역은 평가방법의 선정 능력, 평가도구 개발 능력, 평가의 실시, 채점, 성적 부여 능력, 평가 결과의 분석, 해석 활용, 의사소통 능력, 평가의 윤리성 인식 능력이다. 이를 바탕으로 송미영, 김경희(2007)는 5개의 영역을 5점 척도를 사용하는 60개의 문항으로 구성된 척도를 개발하였다. 이 연구에서는 구인을 확인하기 위해 요인분석을 실시하였고, 구인은 단일요인으로 나타났다.

교사의 학생평가 역량 이론에 나타난 역량 영역은 <표 1>처럼 정리될 수 있다. 각 기준들은 평가 시행을 중심으로 시행 전 단계, 시행, 시행 후 단계로 구분될 수 있다. 시행 전 단계의 경우는 AFT, NCME, NEA 기준(1990)과 김경희 외(2006)에서는 평가 방법 선정과 평가도구 개발이, Brookhart(2011)의 기준에서는 평가 계획, 방법 선정, 도구 개발, 채점에 필요한 지식과 기술에 대한 것이 제시되고 있다. 시행 이후에는 모든 연구에서 채점, 결과 분석, 결과 환류, 의사소통이 제시되고 있다. 그리고 평가 전반적인 절차와 결과에 대한 법적 절차 및 윤리성에 대해 제시하고 있다. 본 연구에서는 문헌 조사를 바탕으로 교사의 학생평가 역량을 5~6개의 영역으로 정리하여 교사의 학생평가 역량 구인 탐색에 활용하고자 한다. 다음으로 탐색된 구인에 대해 정의하고 이에 적합한 교사의 학생평가 역량 진술문을 개발하고, 역량 진술문에 대한 교사 본인의 역량 정도를 묻는 척도를 개발하고자 한다.

〈표1〉 교사의 학생평가 역량 영역

구분	영역
AFT, NCME, NEA 기준 (1990)	평가 방법 선정, 평가도구 개발, 외부 평가 시행, 채점 결과 해석, 평가 결과 활용, 결과에 대한 의사소통, 윤리적 법적 문제 해결
Brookhart(2011)	평가계획, 채점 및 해석, 평가결과 활용, 평가 결과에 대한 의사소통, 평가 결과에 대한 법과 윤리적 책임
김경희 외(2006) 및 송미영, 김경희(2007)	평가 방법 선정, 평가도구 개발, 평가의 실시, 채점, 성적 부여, 평가의 결과 분석 및 해석 활용, 의사소통 능력, 평가의 윤리성 인식

2. 한국의 최근 학생평가 방향

교사의 학생평가 역량에 대한 문헌 조사는 교사의 학생평가 역량 구인에 대한 근거를 마련하기 위한 것이었다면, 최근 학생평가 방향에 대한 문헌 조사는 역량 진술문에 최근 학생평가 방향이 반영될 수 있도록 개발하기 위해서이다. 이는 Brookhart(2011)의 기준에 성취 기준 평가와 형성평가가 독립된 기준으로 제시된 것이 아닌, 각 기준안에 반영되어 제시된 것처럼 역

량 진술문에 반영되도록 개발하기 위해서이다.

한국의 교실 또는 학교 단위의 학생평가에 영향을 주고 있는 정책 또는 방향은 성취평가제와 과정 중심 평가라고 할 수 있다. 성취평가제는 교육부의 ‘중등학교 학사관리 선진화 방안(교육부, 2011)에 맞춰 도입된 학생평가 제도로 2012학년도에 중학교 1학년에 도입되어, 2014학년도에 중학교 전 학년에 적용되었다. 고등학교 보통교과의 경우 2012~2013년도에 시범운동을 거쳐, 2014학년도에 1학년부터 도입되었으며 2016학년도에 고등학교 전 학년에 도입되었다.

성취평가제는 학생 점수를 평가할 때, ‘학생이 무엇을 어느 정도 성취하였는가’를 성취기준과 성취수준에 비추어 학생의 성취 정도를 판단하는 준거참조평가이다. 성취평가제의 성취기준(achievement standard)은 각 교과목에서 학생들이 학습을 통해 성취해야 하는 지식, 기능, 태도에 대한 진술로 교수·학습 및 평가의 실질적 근거가 된다. 성취평가제에서는 성취기준에 부합한 평가계획을 수립하고 성취기준에 맞게 학생의 성취수준을 판별할 수 있도록 채점 기준을 만들어야 하며, 성취수준 및 평가기준에 따라 학생의 성취정도를 평가할 수 있어야 한다(김지영, 안유민, 홍수진, 2014, p. 8; 서민희 외, 2015, p. 3). 이와 같은 성취평가제의 특징은 Brookhart(2011)가 제안하고 있는 성취기준 평가를 반영한 기준과 유사하다고 볼 수 있다. 본 연구에서는 성취평가제의 특징인, ‘계획 단계에서부터 성취기준에 부합한 평가계획을 수립한다.’, ‘성취기준에 맞게 학생의 성취수준을 판별할 수 있도록 채점 기준을 만든다.’, ‘성취수준 및 평가기준에 따라 학생의 성취정도를 평가할 수 있어야 한다.’를 반영한 역량 진술문을 개발하고자 하였다.

과정 중심 평가는 2015 개정 교육과정 총론(교육부, 2015)의 ‘교육과정 구성의 중점’ 중, 학생평가와 관련된 항목에서 학생평가를 ‘라. 학습의 과정을 중시하는 평가를 강화하여 학생이 자신의 학습을 성찰하도록 하고, 평가 결과를 활용하여 교수·학습의 질을 개선한다.’라고 기술하고 있다. 과정을 중시하는 평가의 등장 배경을 살펴보면, 학교의 학생평가가 종합적인 평가 성격이 강해 수업한 내용을 평가하는 것보다 평가를 위해 수업이 진행되는 결과 중심 평가의 문제점을 해결하기 위한 것이었다. 이와 같은 결과 중심적인 학생평가의 문제점을 해결하기 위해 과정을 중시하는 평가를 강조한 것으로 볼 수 있다. 즉 과정 중심 평가는 학습 결과에 대한 평가(Assessment of learning)를 학습을 위한 평가(Assessment for learning), 학습으로서의 평가(Assessment as learning)로 평가 패러다임을 확장한 것이다. 나아가 학생평가의 결과를 ‘학생이 자신의 학습을 성찰하도록 하고,’ ‘평가 결과를 활용하여 교수학습 질을 개선’한다고 하는 의미를 통해 교육과정-교수·학습-교육 평가의 연계가 강조되고 있음을 알 수 있다. 이와 같은 분석을 통해, 과정 중심 평가를 “교육과정 성취기준에 기반한 평가계획에 따라 교수·학습 과정에서 학생의 변화와 성장에 대한 자료를 다각도로 수집하여 적절한 피드백을 제공하는 평가(김유향 외, 2017 pp. 2-4; 임은영, 2017)”로 개념화할 수 있다.

위에서 제시된 과정 중심 평가는 Brookhart(2011)가 새롭게 포함시킨 형성평가를 포함하는

것처럼 보인다. 그러나 과정 중심 평가는 단순히 형성평가와 동일한 평가로 보기 어렵다. 한국의 과정 중심 평가는 학생부의 기록과 연계되고 있기 때문이다. 교육부와 한국교육과정평가원의 과정을 중시하는 수행평가 어떻게 할까요 등의 매뉴얼들을 보면, “과정을 중시하는 수행평가에서는 학생의 수행 과정에 대한 상시 관찰 및 누가기록을 바탕으로, 학생의 성취수준의 특성 및 학습활동 참여도 등 특기할 만한 사항을 학교생활기록부 교과학습발달상황의 ‘과목별 세부능력 및 특기사항’란에 기록하도록 권장하고 있다(김유향 외, 2017, p.29). 이처럼 과정 중심 평가는 학생의 과정을 기록하도록 하고 있다. 기록으로 남긴다는 것은 학생평가가 총합평가처럼 작용하고 있다는 것을 의미한다. 즉 과정 중심 평가는 큰 틀에서는 형성평가를 추구하지만, 한국의 학생평가 제도 안에서는 총합평가처럼 작용한다는 것을 의미한다. 과정 중심 평가가 형성평가와 유사한 점은 교수 학습에 활용하고 학생 스스로가 성찰하도록 하는 것이고, 다른 점은 학생의 배움의 과정을 어떻게 기록으로 남기느냐에 대한 것이다. 본 연구에서 개발하고자 하는 학생평가 역량 척도는 형성평가와 관련된 역량을 측정하는 역량 진술문보다는 과정 중심 평가와 관련된 역량 진술문이 필요하다. 이에 본 연구에서는 과정 중심 평가의 가장 큰 특징인 ‘수업과 연계한 평가를 계획한다.’, ‘학생의 인지적 비인지적 측면을 평가한다.’, ‘학생들의 개별 특성에 맞추어 피드백한다.’, ‘결과를 교수학습 개선에 활용한다.’, ‘과목별 세부능력 및 특기사항을 기록한다’ 등과 같은 과정 중심 평가의 특성이 반영되는 역량 진술문을 개발하고자 하였다.

III. 연구방법

1. 연구대상¹⁾

대상 교사 선정은 교육통계샘플링 서비스(<https://kess.kedi.re.kr/index>)에서 초등학교 5%, 중고등학교 10%에 대해 시·도, 설립유형, 지역 크기를 고려한 층화추출법-크기비례법을 사용하여 학교를 추출하였다. 추출된 셀의 학교 수가 1~2개인 경우 응답률을 고려하여 최소 3개 미만의 학교는 제외하였다. 이를 통해 총 초등학교 307개 학교, 중학교 320개 학교, 고등학교 230개 학교에 온라인으로 설문조사를 하였다. 초등학교는 학년별 담임교사 1명 이상과 과목 담당 교사, 중·고등학교는 교과 제한 없이 학년별 교사 1명 이상이 설문에 참여하도록 요청하였다. 설문조사 기간은 2017년 9월 19일부터 29일까지 총 11일 동안 조사되었다. 설문에는 초등학교 623명 중학교,

1) 이 논문은 임은영 외(2017)의 연구를 수행하면서 수집된 자료를 사용한 것임

308명, 고등학교 376명 총 1307명의 교사가 참여하였다. 연구의 교차타당성을 확보하기 위해 요인분석과 실태조사 대상을 달리하여 연구를 진행하였다. 요인분석을 위해서 각 학교 급별로 100명의 교사를 무작위로 추출하여 총 300명의 교사의 응답 결과를 바탕으로 탐색적 및 확인적 요인분석을 하였다. 실태 분석은 나머지 1007명 교사의 응답 결과를 바탕으로 수행되었다. <표 2>는 연구에 참여한 교사의 배경정보이다.

〈표 2〉 연구에 참여한 교사의 배경 정보(N=1,307)

독립 변수	범주	요인분석 (N=300)			기초통계치 (N=1,007)			전체 (N=1,307)		
		초	중	고	초	중	고	초	중	고
	나이	39.19 (9.17)	40.39 (10.16)	42.23 (9.68)	39.09 (8.77)	40.5 (9.19)	40.42 (8.81)	39.11 (8.83)	40.47 (9.50)	40.90 (9.07)
성별	남	29	34	46	144	52	135	173	86	181
	여	68	64	51	353	150	134	421	214	185
	결측치	3	2	3	26	6	7	29	8	10
	전체	100	100	100	523	208	276	623	308	376
교직 경력	5년 미만	18	28	16	119	56	58	137	84	74
	5년 이상~10년 미만	16	15	19	86	26	68	102	41	87
	10년 이상~20년 미만	46	18	32	183	62	80	229	80	112
	20년 이상	20	39	33	135	64	70	155	103	103
	전체	100	100	100	523	208	276	623	308	376

2. 교사의 학생평가 역량 구인 및 역량 진술문

교사의 학생평가 역량에 대한 문헌 조사를 바탕으로 교사의 학생평가 역량에 대한 구인을 ①평가계획 및 방법 선정, ②평가도구 개발, ③평가 시행 및 채점, ④결과 해석 및 활용, ⑤윤리성 및 법적 절차 준수의 5개 구인으로 정하고 교사의 학생평가 역량 진술문을 개발하였다. 기존 연구에서 결과에 대한 의사소통을 독립된 중요 기준으로 제안하고 있지만, 역량 진술문을 개발하는 과정에서 의사소통은 소통의 대상만 다를 뿐 핵심내용은 동일하여, 평가 결과 해석 및 활용에 필요한 역량으로 포함시켰다. <표 3>은 본 연구에서 측정하고자 하는 교사의 학생평가 역량의 구인과 구인에 따라 측정하고자 하는 역량에 대한 설명이다. 역량의 내용에 따라 교사의 학생평가 역량 구인에 따른 교사의 학생평가 역량을 측정한 18개의 역량 진술문이 개발되었다. 이 중 3문항이 성취평가제를, 8문항이 과정 중심 평가의 개념을 반영하도록 개발되었다.

〈표 3〉 교사의 학생평가 역량 구인과 측정하고자 하는 역량

구인	역량
평가계획 및 방법 선정	평가계획 수립 및 학습 목표에 적합한 평가 방법을 선정하는 역량
평가도구 개발	학생의 능력을 평가할 수 있는 평가 도구 채점 기준을 만드는 역량
평가시행 및 채점	평가를 시행하고, 채점 기준에 따라 채점하고, 성적 부여를 하는 역량
결과 해석 및 활용	채점 결과를 분석하고, 결과를 환류하고, 환류를 위한 기록 및 의사소통을 하는 역량
윤리성 및 법적 절차 준수	평가 절차 전반에 대한 윤리성 및 법적 절차 준수 역량

교사의 학생평가 역량 척도의 문항은 초중고 교사 각 1명, 교육과정, 교육사회학, 교육평가 전문가, 교사 전문성 전문가 각 1명, 총 7명이 내용 타당성을 검토하였다. 내용 타당성의 검토 기준은 역량 진술문이 교사의 학생평가 역량을 진술하고 있는지, 교사의 학생평가 역량과 관련 있는지, 학생평가 경향을 반영하고 있는지였다. 이와 같은 검토를 통해 최종 완성된 역량 진술문은 <표 4>와 같다. 최근 학생평가 정책을 반영한 역량 진술문으로는 성취평가제를 반영한 1, 6, 8 역량 진술문과 과정 중심 평가를 반영한 1, 2, 4, 5, 10, 11, 12, 13, 14 역량 진술문이 있다. 척도는 역량 진술문에 대해 자신의 역량의 정도를 5점 리커트 척도로(1-매우 부족, 2-부족, 3-보통, 4-능숙, 5-매우 능숙) 보고하도록 하였다.

〈표 4〉 교사의 학생평가 역량 영역 및 진술문

구인	역량 진술문	학생평가 정책 관련성
평가계획 및 방법 선정	1. 나는 교육과정을 분석하여 성취기준에 부합하는 평가 계획을 수립한다.	성취평가제
	2. 나는 수업과 연계한 평가를 계획한다.	과정 중심 평가
	3. 나는 학습목표에 부합하는 평가방법을 선정한다.	
평가도구 개발	4. 나는 학생 인지적 능력(지식, 기능)을 평가하기 위한 적합한 평가 도구를 만들어 사용한다.	과정 중심 평가
	5. 나는 학생의 비인지적 영역(정의, 태도 등)을 평가하기 위한 적합한 평가 도구를 만들어 사용한다.	
	6. 나는 학생이 도달하고자 하는 성취기준에 맞게 학생의 성취수준을 판별할 수 있도록 채점기준을 작성한다.	성취평가제
평가시행 및 채점	7. 나는 계획한 평가 절차에 따라 평가를 시행한다.	
	8. 나는 성취수준 및 평가기준에 따라 학생의 성취정도를 평가한다.	성취평가제
결과 해석 및 활용	9. 나는 학생평가 결과를 통해 학생의 능력과 특성을 파악한다.	

윤리성 및 법적 절차 준수	10. 나는 학생들의 개별 특성에 맞추어 피드백 한다.	과정 중심 평가
	11. 나는 학생평가 결과를 교수학습 방법 개선에 활용한다.	과정 중심 평가
	12. 나는 학생평가 결과를 바탕으로 과목별 세부능력 및 특기사항을 기록한다.	과정 중심 평가
	13. 나는 학생평가 결과에 대해 학생과 의사소통한다.	과정 중심 평가
	14. 나는 학생평가 결과에 대해 학부모와 의사소통한다.	과정 중심 평가
	15. 나는 학급/학교/교육청/국가수준의 평가 관련 지침 및 규정을 준수하고 있다.	
	16. 나는 학생의 권리를 보호하기 위해 평가도구 및 평가결과의 보안을 유지한다.	
	17. 나는 평가활동 시, 학생의 인격과 권리를 존중하고 보호한다.	
	18. 나는 학생의 특성 및 배경에 상관없이 공정하게 평가한다.	

교사의 학생평가 역량 척도의 구인에 대한 타당성을 검증하기 위해 탐색적 요인분석 및 확인적 요인분석을 실시하였다. 탐색적 요인분석은 SAS 9.4를 사용하였다. 요인추출은 주축분해법을 사용하였고, 요인회전은 사교회전인 harri-kase(0.3)를 설정하여 수행하였다. 요인 수는 스크리도표, 누적 아이겐 벨류, 진술문의 요인 구조를 검토하여 정하였다. 확인적 요인 분석은 탐색적 요인분석의 결과와 연구에서 제시하는 5개 요인에 대해 AMOS 18.0을 사용하여 모형 적합도를 확인하였다.

IV. 연구결과

1. 학생평가 역량 구인에 대한 타당화

<표 5>는 교사 학생평가 역량을 측정하는 18개 역량 진술문에 대한 탐색적 요인분석 결과이다. 스크리ット 플랫과 아이겐벨류 1을 기준으로 요인수를 탐색한 결과, 요인수는 4~5개가 적정하였다. 4개와 5개의 요인수를 지정하여 탐색적 요인 분석을 한 결과 요인이 4개의 경우 문항이 3.0 이상으로 중복 부하되는 문항 수가 5개 요인일 때보다 적어 4개의 요인으로 정하였다. 역량 진술문이 속한 요인의 의미를 살펴보면, 각 요인들은 윤리성 및 법적절차 준수(4문항), 평가계획 및 도구개발(5문항), 결과 환류(4문항), 평가시행(5문항)과 관련된 문항들이 요인에 속하고 있다. 요인의 설명분산의 크기를 살펴보면, 평가시행이 28.74%를 설명하고, 평가계획 및 시행이 25.95%, 윤리성 및 법적 절차가 23.64%, 결과 환류가 21.67%를 설명하고 있다.

〈표 5〉 교사의 학생평가 역량 탐색적 요인분석 결과

문항	윤리성 및 법적절 차 준수	평가계 획 및 도구개 발	결과 환류	평가 시행
16. 나는 학생의 권리를 보호하기 위해 평가도구 및 평가결과의 보안을 유지한다.	<u>0.847</u>	0.098	-0.035	0.056
17. 나는 평가활동 시, 학생의 인격과 권리를 존중하고 보호한다.	<u>0.827</u>	0.041	0.132	0.025
18. 나는 학생의 특성 및 배경에 상관없이 공정하게 평가한다.	<u>0.725</u>	0.052	-0.062	0.224
15. 나는 학급/학교/교육청/국가수준의 평가 관련 지침 및 규정을 준수하고 있다.	<u>0.516</u>	0.080	0.130	0.229
3. 나는 학습목표에 부합하는 평가방법을 선정한다.	0.151	<u>0.786</u>	0.028	-0.006
1. 나는 교육과정을 분석하여 성취기준에 부합하는 평가 계획을 수립한다.	-0.023	<u>0.674</u>	0.089	0.139
2. 나는 수업과 연계한 평가를 계획한다.	0.167	<u>0.661</u>	-0.040	0.161
4. 나는 학생 인지적 능력(지식, 기능)을 평가하기 위한 적합한 평가 도구를 만들어 사용한다.	0.030	<u>0.651</u>	0.051	0.180
5. 나는 학생의 비인지적 영역(정의, 태도 등)을 평가하기 위한 적합한 평가 도구를 만들어 사용한다.	-0.032	<u>0.473</u>	0.385	0.042
14. 나는 학생평가 결과에 대해 학부모와 의사소통한다.	-0.100	-0.050	<u>0.839</u>	0.027
10. 나는 학생들의 개별 특성에 맞추어 피드백 한다.	0.082	0.086	<u>0.699</u>	0.099
13. 나는 학생평가 결과에 대해 학생과 의사소통한다.	0.292	0.193	<u>0.632</u>	-0.109
11. 나는 학생평가 결과를 교수학습 방법 개선에 활용한다.	0.064	0.130	<u>0.436</u>	0.341
12. 나는 학생평가 결과를 바탕으로 과목별 세부능력 및 특기사항을 기록한다.	0.083	-0.007	0.374	<u>0.449</u>
9. 나는 학생평가 결과를 통해 학생의 능력과 특성을 파악한다.	0.173	0.032	0.336	<u>0.446</u>
7. 나는 계획한 평가 절차에 따라 평가를 시행한다.	0.178	0.144	-0.109	<u>0.712</u>
8. 나는 성취수준 및 평가기준에 따라 학생의 성취정도를 평가한다.	0.113	0.095	0.086	<u>0.704</u>
6. 나는 학생이 도달하고자 하는 성취기준에 맞게 학생의 성취수준을 판별할 수 있도록 채점기준을 작성한다.	0.008	0.329	0.101	<u>0.469</u>
설명량	23.64	25.95	21.67	<u>28.74</u>

탐색적 요인분석 결과에 대한 모형의 적합도를 확인하기 위해 확인적 요인 분석을 수행하였다. 또한 교사의 학생평가 역량 척도 개발 시 제시된 5요인과 모형 적합도 비교하기 위해 5요인에 대한 확인적 요인분석도 수행하였다. <표 6>은 교사의 학생평가 역량 척도에 대한 4요인과

5요인의 측정 모형에 대한 적합도 결과이다. 4요인과 5요인 모두 TLI와 CFI는 권장 적합도 지수를 충족하였으나, 5요인의 χ^2 와 RMSEA의 값이 4요인의 값보다 커서 5요인 보다는 4요인 측정 모형이 더 적합함을 알 수 있었다.

〈표 6〉 교사의 학생평가 역량 측정 모형 적합도

적합도	χ^2	df	TLI	CFI	RMSEA
4요인 모델	4.305	2	0.985	0.997	0.062
5요인 모델	36.660	5	0.912	0.971	0.146

〈표 7〉은 교사의 학생평가 역량 척도의 4개의 하위 요인 간 상관 계수로 상관이 0.639~0.831로 하위 요인 간의 상관이 높게 나타났다.

〈표 7〉 교사의 학생평가 역량 하위 요인 간 상관 계수

	윤리성 및 법적절차 준수	평가계획 및 도구개발	결과환류	평가시행
윤리성 및 법적절차 준수	1	0.707**	0.639**	0.804**
평가계획 및 도구개발		1	0.720**	0.831**
결과 환류			1	0.796**
평가시행				1

〈표 8〉은 학생평가 역량 척도 및 요인들 신뢰도를 확인하기 위한 Cronbach's alpha 값이다. 전체 척도의 α 값은 0.965, 윤리성 및 법적절차 준수 0.938, 평가계획 및 도구개발은 0.913, 결과 환류는 0.876, 평가시행은 0.919로 높은 내적 일치도를 보여 교사의 학생평가 역량 척도의 신뢰성이 높음을 알 수 있었다.

〈표 8〉 교사의 학생평가 역량 척도와 하위 요인의 신뢰도 계수

	구분	문항수	Cronbach's alpha
	전체	18	0.965
하위요인	윤리성 및 법적절차 준수	4	0.938
	평가계획 및 도구개발	5	0.913
	결과환류	4	0.876
	평가시행	5	0.919

2. 교사의 학생평가 역량 실태

교사의 학생평가 역량 실태를 조사하기 위해 학교 급, 성별, 경력에 따른 하위 요인의 기초통계치를 살펴보았다. <표 9>는 학교 급별 하위요인의 평균 및 표준편차이다. 중학교, 고등학교의 경우 윤리성 및 법적절차 준수가 가장 높은 평균을 보였고, 평가시행, 평가계획 및 도구개발, 결과환류 순으로 높았다. 초등학교의 경우는 윤리성 및 법적절차 준수의 평균($M=4.28$, $sd=0.73$)이 가장 높고, 평가시행($M=0.48$, $sd=0.74$)이 다음으로 높고, 평가계획 및 도구개발($M=3.96$, $sd=0.73$), 결과환류($M=3.96$, $sd=0.75$)의 평균이 유사하였다. 학교 급을 비교하여 살펴보면, 모든 하위 영역에서 초등학교의 하위요인 평균이 가장 높고, 중학교, 고등학교 순으로 나타났다.

〈표 9〉 학교 급별 하위요인의 평균 표준편차

하위요인	전체 (N=871)		초등학교 (N=452)		중학교 (N=192)		고등학교 (N=227)	
	평균	표준 편차	평균	표준 편차	평균	표준 편차	평균	표준 편차
윤리성 및 법적절차 준수	4.19	0.73	4.28	0.73	4.15	0.72	4.04	0.71
평가계획 및 도구개발	3.88	0.70	3.96	0.73	3.87	0.65	3.72	0.64
결과환류	3.78	0.74	3.96	0.75	3.66	0.70	3.54	0.67
평가시행	3.99	0.71	4.08	0.74	3.96	0.66	3.84	0.66

<표 10>은 성별에 따른 하위 요인 평균 표준편차이다. 전체, 초등학교, 중학교의 경우 모든 하위 영역에서 여교사의 평균이 남교사보다 높았다. 그러나 고등학교의 경우 남교사의 평균이 여교사보다 높게 나타났다.

〈표 10〉 성별 하위요인의 평균 표준편차

하위요인	전체 (N=869)				초등학교 (N=451)				중학교 (N=191)				고등학교 (N=227)				< 표 11 >은 학 교 급 과 경 력 에 따 른 하 위 요 인 들
	남 (N=296)		여 (N=574)		남 (N=132)		여 (N=319)		남 (N=49)		여 (N=142)		남 (N=115)		여 (N=112)		
	평 균	표준 편차	평 균	표준 편차	평 균	표준 편차	평 균	표준 편차	평 균	표준 편차	평 균	표준 편차	평 균	표준 편차	평 균	표준 편차	
윤리성 및 법적절차 준수	4.06	0.75	4.26	0.71	4.06	0.78	4.38	0.69	4.03	0.76	4.19	0.71	4.08	0.70	4.00	0.72	하 위 요 인 들
평가계획 및 도구개발	3.77	0.67	3.93	0.71	3.75	0.74	4.04	0.72	3.81	0.59	3.89	0.68	3.78	0.64	3.67	0.65	
결과환류	3.67	0.70	3.85	0.75	3.73	0.74	4.05	0.73	3.62	0.61	3.69	0.71	3.63	0.69	3.46	0.65	
평가시행	3.85	0.69	4.07	0.71	3.82	0.75	4.19	0.71	3.86	0.61	4.00	0.67	3.88	0.65	3.81	0.67	

의 평균과 표준편차이다. 전반적으로 학교 급에 상관없이 경력이 높을수록 하위요인의 평균이

높아졌다. 그러나 학교 급별 요인별로 다른 경향을 보이기도 한다. 초등학교 교사의 경우 윤리성 및 법적절차 준수의 경우 ‘10년 이상~20년 미만’의 평균($M=4.25$, $sd=0.77$)이 ‘5년 이상~10년 미만’의 평균($M=4.36$, $sd=0.64$)보다 낮았고, 20년 이상에서 평균($M=4.51$, $sd=0.655$)이 다시 높아졌다.

중학교의 경우, 윤리성 및 법적 절차 준수의 경우 경력이 올라가면서 평균이 높아졌다. 평가 도구 개발의 경우 5년 미만이 가장 낮은 평균($M=3.66$, $sd=0.63$)을 보였고, 5년 이상~10년 미만, 10년 이상~20년 미만의 평균이 3.95 가장 높았으면 20년 이상의 평균이 3.92로 5년 이상~20년 미만 경력보다 낮게 나타났다. 평가시행의 경우 5년 미만의 평균이 가장 낮고, 5년 이상~10년 미만, 10년 이상~20년 미만 순으로 평균이 높아지다, 20년 이상에서 10년 이상~20년 미만의 평균보다 낮아졌다. 결과환류의 경우 5년 미만이 가장 낮았고, 5년 이상~10년 미만이, 3.72, 10년 이상~20년 미만이 3.69, 20년 이상이 3.71로 5년 이상의 교사들이 유사한 역량을 보고하였다.

고등학교의 경우 윤리성 및 법적절차 준수와 결과 환류는 경력이 높을수록 평균이 높아지는 경향을 보였지만 평균의 차이가 크지는 않았다. 윤리성 및 법적절차의 경우 5년 미만의 평균이 3.99, 5년 이상~10년 미만이 4.01, 10년 이상이 4.07로 나타났다. 평가도구개발은 다른 하위요인에 비해 5년 미만과 5년 이상의 경력에서 차이를 보이고 있다. 5년 이상~10년 미만의 평균이 3.79로 가장 높았고, 10년 이상~20년 미만이 3.76, 20년 이상이 3.75로 나타났다. 평가시행은 5년 미만의 평균이 3.76으로 가장 낮았고 5년 이상~10년 미만, 10년 이상~20년 미만이 3.85를 유사하였으며, 20년 이상이 3.89로 가장 높은 평균을 보여줬다. 그러나 경력에 따른 평균의 차이가 크지 않았다. 결과환류는 다른 하위요인의 평균에 비해 가장 낮은 평균을 보여주었다. 경력이 올라갈수록 평균이 높게 나타났지만 차이가 크지 않았다. 5년 미만이 3.49, 5년 이상~10년 미만이 3.52, 10년 이상~20년 미만이 3.54, 20년 이상이 3.61로 나타났다. 고등학교의 특징은 경력에 따른 하위 요인들의 평균이 크게 차이를 보이지 않는다는 것이다.

〈표 11〉 경력에 따른 하위요인의 평균 표준편차

학교 급	하위요인	5년 미만		5년 이상~10년 미만		10년 이상~20년 미만		20년 이상	
		평균	표준 편차	평균	표준 편차	평균	표준 편차	평균	표준 편차
전체 (N=870)	윤리성 및 법적절차 준수	3.95	0.75	4.20	0.68	4.20	0.74	4.35	0.69
	평가계획 및 도구개발	3.58	0.66	3.92	0.65	3.95	0.71	3.97	0.70
	결과환류	3.54	0.69	3.74	0.70	3.85	0.75	3.91	0.74
	평가시행	3.74	0.70	3.99	0.69	4.04	0.71	4.12	0.69
초 (N=452)	윤리성 및 법적절차 준수	3.94	0.72	4.36	0.64	4.25	0.77	4.51	0.65
	평가계획 및 도구개발	3.57	0.69	4.00	0.62	4.03	0.76	4.10	0.72

	결과환류	3.58	0.67	3.90	0.69	4.04	0.78	4.15	0.69
	평가시행	3.72	0.72	4.10	0.68	4.11	0.76	4.27	0.69
중 (N=192)	윤리성 및 법적절차 준수	3.94	0.76	4.11	0.68	4.20	0.72	4.29	0.69
	평가계획 및 도구개발	3.66	0.63	3.95	0.67	3.95	0.62	3.92	0.68
	결과환류	3.53	0.74	3.72	0.67	3.69	0.65	3.71	0.74
	평가시행	3.77	0.69	3.97	0.61	4.06	0.65	4.02	0.64
고 (N=227)	윤리성 및 법적절차 준수	3.99	0.82	4.01	0.70	4.07	0.65	4.07	0.71
	평가계획 및 도구개발	3.51	0.63	3.79	0.67	3.76	0.64	3.75	0.63
	결과환류	3.49	0.68	3.52	0.69	3.54	0.63	3.61	0.70
	평가시행	3.76	0.70	3.85	0.73	3.85	0.61	3.89	0.65

〈표 12〉는 교사의 학생평가 역량의 5개의 하위 영역의 각 문항별 평균 및 표준편차이다. 윤리성 및 법적절차 준수는 4개의 역량 진술문으로 구성되어 있고, 진술문의 평균 범위는 4.20~4.28, 초등학교의 평균 범위는 4.15~4.36, 중학교의 평균 범위는 4.08~4.24, 고등학교의 평균 범위는 3.86~4.16이다. 윤리성 및 법적절차에서는 ‘15. 나는 학급/학교/교육청/국가수준의 평가 관련 지침 및 규정을 준수하고 있다’가 가장 낮은 평균을 보였고, ‘18. 나는 학생의 특성 및 배경에 상관없이 공정하게 평가한다.’의 평균이 가장 높았다.

평가계획 및 도구개발은 5개의 문항으로 구성되어 있고, 진술문의 평균 범위는 3.64~4.03, 초등학교 교사의 경우 3.75~4.10, 중학교의 경우 3.64~4.01, 고등학교의 경우 3.42~3.94로 나타났다. 초, 중, 고 모두 ‘5. 나는 학생의 비인지적 영역을 평가하기 위해 적합한 평가 도구를 만들어 사용한다.’가 가장 낮아 교사들이 비인지적 영역의 평가 도구 개발에 어려움이 있음을 추론해 볼 수 있다.

결과환류는 4개의 문항으로 구성되어 있고, 평균의 범위는 3.52~3.91, 초등학교의 경우 3.85~4.02, 중학교의 경우 3.26~3.87, 고등학교의 경우 3.08~3.78로 ‘14. 나는 학생평가 결과에 대해 학부모와 의사소통한다.’의 평균이 가장 낮았다. 평가시행은 5개의 문항으로 구성되어 있고, 평균의 범위는 3.89~4.06, 초등학교의 경우 3.98~4.15, 중학교의 경우 3.84~4.08, 고등학교의 경우 3.75~3.97로, ‘6. 나는 학생이 도달하고자 하는 성취기준에 맞게 학생의 성취수준을 판별할 수 있도록 채점기준을 작성한다.’가 모든 학교 급에서 가장 낮게 나타났다.

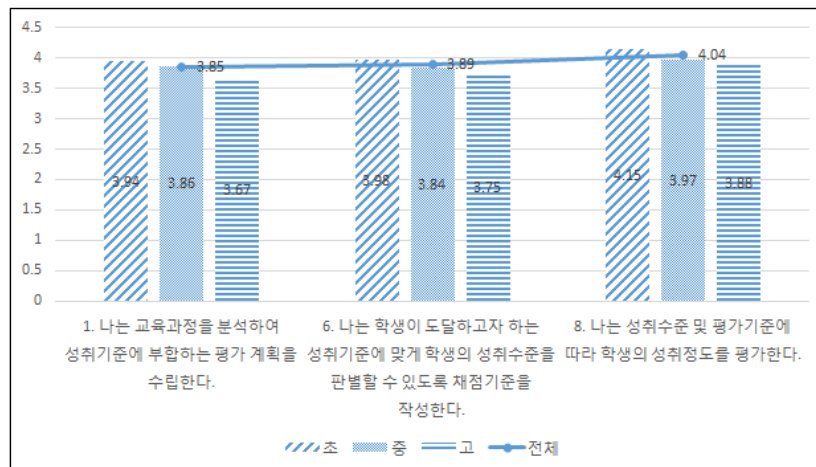
〈표 12〉 하위 영역의 문항별 기초통계치

하위 영역	역량 진술문	전체 (N=871)		초 (N=452)		중 (N=192)		고 (N=227)	
		평균	표준 편차	평균	표준 편차	평균	표준 편차	평균	표준 편차
윤리	16. 나는 학생의 권리를 보호하기 위해 평가도구	4.20	0.79	4.30	0.76	4.15	0.79	4.05	0.80

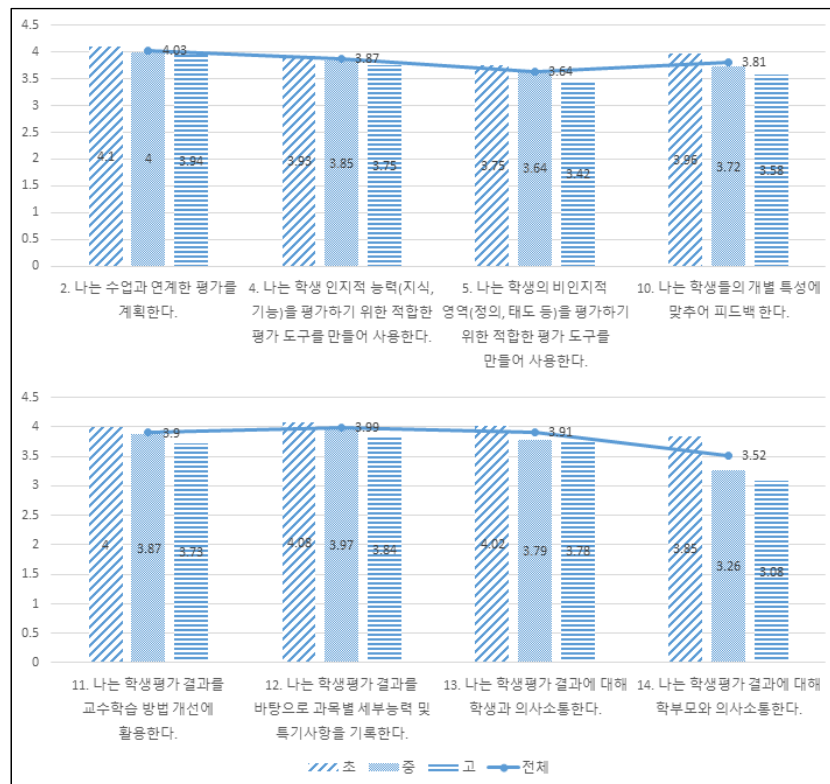
성 및 법적 절차 준수	및 평가결과의 보안을 유지한다.								
	17. 나는 평가활동 시, 학생의 인격과 권리를 존중하고 보호한다.	4.21	0.77	4.31	0.77	4.14	0.78	4.08	0.75
	18. 나는 학생의 특성 및 배경에 상관없이 공정하게 평가한다.	4.28	0.77	4.36	0.76	4.24	0.76	4.16	0.78
	15. 나는 학급/학교/교육청/국가수준의 평가 관련 지침 및 규정을 준수하고 있다.	4.06	0.81	4.15	0.80	4.08	0.77	3.86	0.82
평가 계획 및 도구 개발	3. 나는 학습목표에 부합하는 평가방법을 선정한다.	4.00	0.76	4.07	0.77	4.01	0.73	3.85	0.74
	1. 나는 교육과정을 분석하여 성취기준에 부합하는 평가 계획을 수립한다.	3.85	0.80	3.94	0.81	3.86	0.77	3.67	0.77
	2. 나는 수업과 연계한 평가를 계획한다.	4.03	0.77	4.10	0.79	4.00	0.71	3.94	0.76
	4. 나는 학생 인지적 능력(지식, 기능)을 평가하기 위한 적합한 평가 도구를 만들어 사용한다.	3.87	0.78	3.93	0.81	3.85	0.74	3.75	0.75
	5. 나는 학생의 비인지적 영역(정의, 태도 등)을 평가하기 위한 적합한 평가 도구를 만들어 사용한다.	3.64	0.85	3.75	0.85	3.64	0.79	3.42	0.84
	14. 나는 학생평가 결과에 대해 학부모와 의사소통한다.	3.52	0.97	3.85	0.83	3.26	0.92	3.08	1.02
결과 환류	10. 나는 학생들의 개별 특성에 맞추어 피드백 한다.	3.81	0.85	3.96	0.85	3.72	0.78	3.58	0.84
	13. 나는 학생평가 결과에 대해 학생과 의사소통한다.	3.91	0.82	4.02	0.83	3.79	0.79	3.78	0.77
	11. 나는 학생평가 결과를 교수학습 방법 개선에 활용한다.	3.90	0.79	4.00	0.83	3.87	0.75	3.73	0.73
평가 시행	12. 나는 학생평가 결과를 바탕으로 과목별 세부능력 및 특기사항을 기록한다.	3.99	0.82	4.08	0.81	3.97	0.79	3.84	0.86
	9. 나는 학생평가 결과를 통해 학생의 능력과 특성을 파악한다.	3.98	0.78	4.09	0.79	3.94	0.74	3.78	0.75
	7. 나는 계획한 평가 절차에 따라 평가를 시행한다.	4.06	0.79	4.09	0.83	4.08	0.75	3.97	0.75
	8. 나는 성취수준 및 평가기준에 따라 학생의 성취정도를 평가한다.	4.04	0.77	4.15	0.79	3.97	0.72	3.88	0.76
	6. 나는 학생이 도달하고자 하는 성취기준에 맞게 학생의 성취수준을 판별할 수 있도록 채점기준을 작성한다.	3.89	0.80	3.98	0.82	3.84	0.73	3.75	0.77

3. 최신 학생평가 정책과 관련된 교사의 학생평가 역량

최신 학생평가 정책인 성취평가제와 과정 중심 평가에 대한 교사의 역량 실태를 조사하기 위해 관련 진술문에 대한 평균을 살펴보았다. [그림 1]은 성취평가제 관련 역량 진술문의 학교 급별 평균이다. 성취평가제와 관련한 역량 진술문은 1번, 6번, 8번으로 성취평가제의 기본 특성인, 성취기준에 맞는 평가 계획 수립, 성취기준에 부합하는 학생 성취수준을 판별하는 채점기준 작성, 성취수준 및 평가기준에 따른 학생의 성취정도를 평가하는가에 대한 역량 진술문이다. 성취평가제 관련 역량 진술문의 평균 범위를 살펴보면 전체 교사의 평균 범위는 3.85~4.04, 초등학교 3.94~4.15, 중학교 3.84~3.97, 고등학교 3.67~3.88로 나타났다. 모든 학교 급에서 교육과정을 분석하여 성취기준에 부합하는 평가 계획을 수립하는 것에 대한 역량을 가장 낮게 인식하고 있고, 성취수준 및 평가기준에 따라 학생의 성취정도를 평가하는 것을 가장 높게 인식하고 있었다.



[그림 1] 성취평가제 관련 역량 진술문의 학교 급별 평균



[그림 2] 과정 중심 평가 관련 역량 진술문의 학교 급별 평균

[그림 2]는 과정 중심 평가 관련 역량 진술문의 평균이다. 과정 중심 평가 관련 역량 진술문은 2번, 4번, 5번, 10번, 11번, 12번, 13번, 14번 진술문으로 수업을 연계한 평가, 비인지적, 인지적 평가 도구 개발, 개별 특성에 맞춘 피드백, 결과를 교수학습 방법 개선에 활용, 학생 결과물 과목별 세부능력 및 특기사항에 기록, 학생과 학부모와의 의사소통에 대한 역량 진술문이다. 과정 중심 평가 관련 역량 진술문의 평균을 살펴보면, 전체 교사의 평균 범위는 3.52~4.03, 초등학교 교사의 평균 범위는 3.85~4.10, 중학교의 경우 3.26~4.00, 고등학교의 경우 3.08~3.94로 나타났다. 모든 학교 급에서 수업과 연계한 평가계획이 가장 높은 것으로 나타났고, 학생 평가 결과에 대해 학부모와 의사소통한다는 것이 가장 낮게 나타났다.

V. 결론 및 논의사항

교사의 학생평가 역량 중요성이 부각되는 현 상황에서, 본 연구는 교사의 학생평가 역량을 진단할 수 있는 자기보고식 척도를 개발하고 이를 바탕으로 교사의 학생평가 역량의 실태를 조사하였다. 이를 위해 4개의 요인에 대한 18개의 역량진술문으로 구성된 교사의 학생평가 역량 척도를 개발하였다. 18개의 문항 중, 성취평가제를 반영한 3개의 역량 진술문과 과정 중심 평가를 반영한 8개의 역량 진술문이 포함되어 있어, 최근 학생평가 동향과 관련된 교사의 학생평가 역량 실태를 조사할 수 있었다.

교사의 학생평가 역량 구인은 탐색적 요인분석을 통해 4개의 요인 1)윤리성 및 법적절차 준수(4문항), 2)평가계획 및 도구개발(5문항), 3)결과환류(4문항), 4)평가시행(5문항)으로 구성될 수 있음을 파악하였다. 이에 확인적 요인분석을 통해 연구 개발 단계에서 제시된 5요인과 4요인의 모형 적합도를 비교한 결과, 4요인의 모형 적합도가 더 적절한 것으로 나타나 교사의 학생평가 역량 척도를 4요인으로 구성하였다.

교사의 학생평가 역량의 실태 조사를 위해 전체 척도 및 하위요인, 문항의 기초통계치를 살펴 보았다. 척도 및 하위요인의 평균은 3.50 이상으로 교사들은 자신의 역량 정도를 보통이상이라고 인식하는 것으로 나타났다. 4개의 요인을 학교 급별로 비교해보면, 모든 요인에서 초등학교 교사가 보고하는 학생평가에 대한 역량이 가장 높았고, 중학교, 고등학교 순으로 높았다. 경력에 따른 학생평가 역량을 살펴보면, 초등학교의 경우 경력에 따른 역량의 차이를 보이는 반면 학교 급이 높아질수록 경력에 따른 역량 평균 차이가 작았다. 하위 요인 중에는 윤리성 및 법적절차 준수의 평균이 학교 급과 경력에 상관없이 가장 높게 나타났고, 평가시행, 평가계획 및 도구개발, 결과환류 순으로 보고되었다. 이는 김신영(2002)의 연구에서 유사한 결과로 측정에 사용한 도구가 달라 직접적인 비교에는 어려움이 있지만, 17년이 지난 후에도 교사들의 역량에 대한 순서가 유지되고 있다는 점에서 결과환류, 특히 의사소통에 대한 역량을 높이기 위한 방안 마련이 필요하다는 것을 알 수 있었다. 성별에 따른 하위요인의 차이를 살펴보면 초등학교와 중학교의 경우 모든 요인에서 여교사의 평균이 남교사의 평균보다 높게 나타나지만, 고등학교의 경우 남교사의 평균이 더 높게 나타났다. 이와 같은 학생평가 역량의 실태 조사를 통해 교사들은 자신의 학생평가 역량에 대해 보통이상으로 생각하고 있으며, 경력이 올라갈수록 학생평가 역량에 대해 능숙하다고 인식하고 있음을 알 수 있었다.

최근의 학생평가 정책인 성취평가제와 과정 중심 평가 관련 역량 진술문을 살펴보면, 다른 역량에 비해 평균이 낮은 편이다. 특히 과정 중심 평가의 경우 “14. 나는 학생평가 결과에 대해

학부모와 의사소통한다.”, “5. 나는 학생의 비인지적 영역(정의, 태도 등)을 평가하기 위한 적합한 평가 도구를 만들어 사용한다.”, “10. 나는 학생들의 개별 특성에 맞추어 피드백한다.”, 성취평가제의 경우 “1. 나는 교육과정을 분석하여 성취기준에 부합하는 평가 계획을 수립한다.”의 평균이 낮아 이와 관련된 역량을 높이기 위한 방안 마련이 필요함을 알 수 있었다. 성취평가제의 역량 진술문에 대한 고등학교 교사의 평균이 다른 학교 급에 비해 낮게 나타나, 고교학점제 도입과 맞물려 성취평가제의 중요성이 고등학교에서 더욱 부각되는 시점에서 고등학교 교사의 성취평가제 역량 제고를 위한 방안 마련이 시급함을 알 수 있었다.

본 연구는 교사의 학생평가 역량 구인을 탐색하여, 문항 수가 작은 자기보고식 척도를 개발하였다는 점에 의의를 둘 수 있다. 기존의 교사의 학생평가 역량을 측정하기 위한 척도들은 학생평가에 대한 지식을 측정(Plakw, Impara & Fager, 1993)하거나, 현장 교사의 평가전문성 연구를 AFT, NCME, NEA에서 개발한 35개 문항을 번안하여 사용하거나(김신영, 2002), 60개의 문항으로 구성된 단일 요인의 자기 보고식(송미영, 김경희, 2007) 척도를 개발하였다. 이에 비해 본 연구에서 개발한 교사의 학생평가 척도는 4개의 요인, 18개 문항으로 구성되어 있다. 학생평가 역량 및 전문성 기준들은 상호 독립적이지 않으면 연계되어 있어 범주화하기 쉽지 않았다. 그러나 요인분석을 통해 학생평가의 구인을 4요인으로 확인함으로써, 문항수준 정보뿐만 아니라 교사의 학생평가 역량의 하위 요인에 대한 축적되고 간략한 정보도 알 수 있게 되었다는 점이 본 연구의 의의로 볼 수 있다.

또한 본 연구의 척도는 문항 수가 작은 자기 보고식 척도로 교사에게 신속하게 피드백을 줄 수 있다. 나아가 척도의 역량 진술문들은 학교 현장에서 학생평가 시행에 필요한 핵심적인 역량들이다. 이에 본 연구에서 개발한 척도를 2가지 측면에서 활용할 수 있다. 첫째, 18개의 역량 진술문을 통해 교실에서 이루어지는 교사에게 필요한 학생평가의 역량이 무엇인지 확인할 수 있고, 둘째 척도를 통해 교사 자신의 학생 평가 역량을 진단할 수 있다. 이 척도는 교사 이외에도 시도 교육청, 교육부에서도 활용할 수 있다. 교사의 학생평가 역량을 제고하기 위한 연수 프로그램 개발 시, 연수 대상자에게 필요한 역량을 확인하고 이에 대한 프로그램을 개발하여 교사에게 필요한 연수를 제공할 수 있다. 본 연구에서 결과환류, 특히 의사소통에 대한 역량이 학교 급에 상관없이 교사들은 낮게 인식하고 있었다. 또한 연구 결과 중, 고등학교 교사들은 성취평가제 관련 역량을 낮게 인식하고 있었다. 이와 같은 결과를 바탕으로 교육청이나 교육부는 교사의 학생평가 역량 제고를 위한 연수 개발 시, 결과환류에 대한 연수, 고등학교 교사에게는 성취평가제 역량 제고 연수 개발을 위한 근거 자료로 활용할 수 있다. 이처럼 교사의 학생평가 역량 연수 프로그램 개발 시, 사전 요구 조사용으로 사용할 수 있다. 나아가 교사의 학생평가 역량 관련 연수의 사전, 사후 효과를 측정하기 위한 도구로 사용할 수도 있다.

이와 같은 교사의 학생평가 역량 구인 탐색 및 척도 활용성에 대한 장점에도, 본 연구는 다음과 같은 한계점이 있다. 본 연구에서 개발한 척도는 자기보고식 척도로 교사의 실제적인 역

량을 측정하는 것이 아니어서, 결과 해석에 주의해야 한다. 척도 점수가 높다는 것은 교사의 학생평가 역량이 높다는 것이 아닌 교사가 인식하는 본인의 학생평가 역량이 높다는 것이다. 이에 본 척도는 교사의 학생평가 역량을 평정하는 객관적인 자료로 사용될 수 없다. 객관적인 자료로 사용하기 위해서는 자기보고식이 아닌 교사의 학생평가 역량을 측정할 수 있는 지식, 기술을 평가하는 척도를 개발하고, 개발된 척도와 본 척도의 공인 타당도를 확인한 후 사용할 수 있을 것이다. 이에 후속 연구로 교사의 학생평가 역량을 측정할 수 있는 객관적인 평가를 개발하고, 이를 바탕으로 본 연구 척도의 타당성을 높이기 위한 연구를 제안하고자 한다.

또 다른 한계점은 본 연구가 척도 개발을 위해 설계된 표집을 바탕으로 자료를 수집하지 않아 표본의 대표성이 낮을 수 있다는 것이다. 이와 같은 한계를 극복하고자 척도의 타당성 검증을 위해 각 학교 급에서 100명씩 무선으로 표집하여, 배경 변인의 영향을 줄이고자 하였다.

본 연구는 교사의 학생평가 역량에 대한 학교 급별, 성별, 경력 변인에 대해 기초 분석만 이루어졌다. 배경 변인에 따른 차이가 통계적으로 유의한지 분석하지 않았다. 이는 학교 급별, 성별, 경력 각 배경 변인별 유의도 검증보다는 다양한 배경 변인을 모형에 동시에 포함한 유의도 검증 결과가 심층적인 분석을 위해 더 유용할 것이라고 판단하여, 이번 연구에서는 기초 분석에 기반을 둔 실태조사만 수행하였다. 추후 다양한 배경 변인을 독립변인으로 하위요인을 종속변인으로 하는 MANOVA연구를 수행할 것이다.

마지막으로 본 연구의 교사별 학생평가 역량 척도는 학생평가 역량에 대한 일반적인 역량을 측정하고자 개발한 척도로, 학교 급에 상관없이 또는 교사의 배경 변인에 상관없이 동일 기준으로 학생평가 역량을 확인할 수 있다는 장점이 있다. 그러나 학교 급이나 교과에 대한 특성을 자세히 반영하지 못하고 있어, 학교 급이나 교과에 따른 학생평가 역량을 상세히 분석하는 데는 어려움이 있다. 이를 해결하기 위한 학교 급 및 교과에 부합하는 역량 척도 개발을 후속 연구로 제안하고자 한다.

참고문헌

- 교육부(교육부, 2011.12.14.). 중등학교 학사관리 선진화 방안 발표. 자료출처(검색일 2019.2.26.):
<https://www.moe.go.kr/boardCnts/view.do?boardID=294&boardSeq=35192&lev=0&searchType=null&statusYN=W&page=373&s=moe&m=0503&opType=N>
- 교육부(2015.12.1). 초·등학교 교육과정 총론. 교육부 고시 제2015-80호
- 교육부(정부 24, 2017.11.27.). 고교학점제 추진 방향 및 연구학교 운영계획 발표. 자료출처(검색일 2019.2.24.): <https://www.gov.kr/portal/ntnadmNews/1257935>
- 교육부(대한민국정책브리핑, 2018.11.8). 학생이 꿈꾸는 고교 교육, 고교학점제로 준비한다. 자료출처(검색일, 2019.2.24.):
<http://www.korea.kr/briefing/pressReleaseView.do?newsId=156302901>
- 김경희, 김신영, 김성숙, 지은림, 반재천, 김수동(2006). 교사의 학생평가 전문성 기준 개발. **교육평가연구**, 19(2), 89-112.
- 김신영(2002). 현장교사의 평가전문성 연구. **교육평가연구**, 15(1), 67-85.
- 김재춘, 설현수, 송중현, 신병찬(2005). **교사별 학생평가 방안 연구. 정책연구과제 2005-지정-37**. 교육인적자원부.
- 김지영, 안유민, 홍수진(2014). **문답식으로 알아보는 성취평가제 고등학교 보통 교과**. 한국교육과정평가원(홍보자료 PIM 2014-9).
- 김유향, 노은희, 김인숙, 박종임, 이인화, 김미경, 정연준, 성경희, 김광규, 장근주, 김정효, 유창완(2017). **과정을 중시하는 수행평가 어떻게 할까요?** 한국교육과정평가원(연구자료 ORM 2017-19-2).
- 김희규, 주영호, 엄준용(2016). 중등학교 교사의 생애단계별 직무역량 분석. **교육문제연구**, 29(4), 163-189.
- 박용호, 조대연, 배현경, 이해정(2012). 중등교사의 직무역량 요구분석. **한국교원교육연구**, 29(2), 299-320.
- 박형수(중앙일보, 2018. 5.10). [대입, 넌 누구냐] 수시는 어떻게 대세가 됐다. 자료출처(검색일 2019.1.31.): <https://news.joins.com/article/22609752>
- 서민희, 김희경, 김경희, 박종임, 성경희, 박지현, 김유향, 전경희(2015). **성취평가제 이해 및**

- 실제.** 한국교육과정평가원(연구자료 ORM 2015-109).
- 송미영, 김경희(2007). 교사의 학생평가 전문성 진단도구 개발. **교육과정평가연구**, 10(1), 47-74.
- 조대연(2009). 교사 발달단계별 직무역량 요구분석: 서울초등교사를 대상으로. **한국교원교육연구**, 26(2), 365-385.
- 이인제, 이범홍, 박정, 진재관, 김옥남, 서수현, 김신영(2004). **교사의 학생평가 전문성 신장 모형과 기준**. 한국교육과정평가원 연구보고 RRE 2004-5-2.
- 임은영(2017). 과정 중심 평가의 개념과 의미. **행복한 교육**. 교육부. 자료출처 (검색일, 2019.2.24.)https://happyedu.moe.go.kr/happy/bbs/selectHappyArticleImg.do?bbsId=BB SMSTR_000000000192&nttId=7238
- 임은영, 김유향, 이인화, 심현표, 김성식(2017). 교사의 학생평가 실태 및 현안 진단. **한국교육과정평가원(연구보고 ORM 2017-66-24)**.
- 한혜정, 이광우, 이주연, 김영은, 김현수, 박혜영, 배주경, 임은영, 임해미, 이미숙(2018). **고교학점제 도입을 위한 연구학교 운영 방안**. (연구보고 CRC 2018-4). 한국교육과정평가원.

American Federation of Teachers, National Council on Measurement in Education & National Education Association(1990). *standard for teacher competence in educational assessment of students*.

Brookhart, S., M.(2011). Educational assessment knowledge and skills for teachers. *Educational measurement: Issue and Practice*. 30(1), 3-13.

Plake B., S., Impara, J., C. & Fager, J., J.(1993). Assessment competencies of teachers: a national survey. *Educational measurement: Issue and Practice*. 12, 10-39.

· 논문접수 : 2019.1.3. / 수정본접수 : 2019.2.1. / 게재승인 : 2019.2.18.

ABSTRACT

A Study on teachers' competency of student assessment

Eun-Young Lim

Associate Research Fellow, Korea Institute for Curriculum and Evaluation

The importance of student assessment in the classroom has been increasing due to recent trends and policies on student assessment, and so has the importance of teachers' competencies in student assessment. In this study, a self-report scale was developed to measure teachers' competencies in student assessment. The scale is composed of 18 items including the latest student assessment policies: process-oriented assessment and standard-based achievement assessment system. The survey using the scale was conducted online and a total of 1,307 teachers participated. Among the responses, the responses of 300 teachers' were used to validate the scale using exploratory factor analysis and confirmatory factor analysis, and those of 1007 teachers' were used for descriptive statistics to investigate the current state of teachers' competencies in student assessment. Based on the factor analysis, the scale consists of four factors: administering student assessment ($\alpha = 0.919$), developing assessment ($\alpha = 0.913$), ethics and legal process ($\alpha = 0.938$), and feedback and usage for assessment results ($\alpha = 0.876$). With the developed scale, teachers' competencies in student assessment were analyzed according to the school levels, the years of educational experience, and gender. The results showed that elementary school teachers' competencies were highest in all factors followed by middle school teachers' and then high school teachers'. Also, there were no meaningful gender differences between the school levels in all four factors. According to the years of educational experience, the longer the teachers have served, the higher they think their competencies are. Regarding the recent student assessment policies, the mean of the items related to the process-oriented evaluation was higher than those related to the standard-achievement evaluation.

Key words: teachers' competencies in student assessment, development of scales and validation, the current state of teachers' competencies in student assessment.