

2015년 교육정책네트워크
교육현장지원연구

중등교사의 평가 전문성 제고방안
: 서술형평가 및 수행평가를 중심으로

연구책임자 · 김 석 우(부산대학교)



중등교사 평가 전문성 제고방안 : 서술형 평가 및 수행평가 중심으로

연구책임자 : 김 석 우 (부산대학교)
공동연구진 : 박 상 옥 (동의대학교)
 김 윤 용 (부산국제고등학교)
 장 재 혁 (사직여자고등학교)

연구요약

본 연구에서는 박근혜 정부의 꿈과 끼를 키우는 행복교육 실현의 토대가 되는 교사의 평가전문성 신장 방안을 모색하여, 교사의 평가전문성을 개선시키기 위한 자료 마련 등을 위한 것이다. 서술형 평가 및 수행평가 중심의 중등학교 교사의 평가 전문성 제고 방안 모색은 21세기 교육의 핵심역량인 의사소통능력, 협업능력, 창의력, 비판적 사고능력 등을 길러주는데 일조하게 될 것이다. 이러한 것을 대변한 듯, 현재 학교 현장에서는 서술형 문항의 출제 비율을 50%이상 확보하도록 강조하고 있는 실정이다. 하지만 일선교사들은 여전히 서술형 평가 및 수행평가를 수행 및 평가하는데 어려움을 토로하고 있고, 전반적으로 평가에 관한 지식이 부족한 상태이다(김석우, 2014; 박도순, 2013; 성태제, 2010). 이러한 현 시점에서 학교 현장 교사들에게 학생평가의 수월성을 달성하기 위한 수행도구로서 중등교사의 평가 전문성 제고 방안을 마련하여 제공함으로써 학생평가의 본래의 목적인 학생의 학습과 성장을 지원하는 데 효과적인 정보를 제시하고자 한다.

본 연구의 목적을 달성하기 위하여 선행연구 및 해외사례를 근거로 중등학교 교사의 평가전문성을 규정하였고, 학교현장에서의 실태를 확인하기 위하여 조사연구를 실시하였다. 조사연구의 한계를 극복하기 위하여 장학사, 중학교 및 고등학교 교사를 대상으로 포커스 그룹 인터뷰를 실시하였다. 최종적으로 이를 종합하여 결론을 도출하였고, 다음과 같다.

서술형 평가와 수행평가에 관한 평가 환경은 교육여건(교육과정), 교수학습·방법, 교육평가 세부여건 등으로 고찰해 볼 수 있다.

첫째, 서술형 및 수행 평가에 관한 대부분의 일선 학교교사는 이러한 유형의 평가방식은 학생들의 창의적인 학습과 개념적 지식을 이해하는데 활용 가치가 있지만, 현 교육여건 상 학생들의 학습 부담과 교사의 업무 부담으로 작용함을 피력하고 있다. 이런 점에서 서술형 및 수행 평가에 관한 교사들 간의 협업은 사실상 불가능하고, 명확한 서술형 및 수행평가 기준을 정립할 수 없는 한계점을 지적하고 있다. 가장 시급한 것은 서술형 및 수행 평가에 적합한 즉, 학생을 위한 수업이 이루어질 수 없는 학교 분위기를 언급하고 있다. 무엇보다 현 입시제도에서 서술형 및 수행 평가 본래 취지에 맞는 시험을 준비는 학생들의 공감을 얻지 못하고 있다. 이런 문제점을 해결하기 위해서 정부는 교사가 서술형 평가와 수행평가에 몰입할 수 있는 적합한 평가 환경을 조성해 주어야 한다.

둘째, 일선학교 교사들은 학생들의 수준을 고려하여 평가 기준을 수정하다 보면, 객관적인 기준 자체가 모호해 지는 경향이 있음을 지적하고 있다. 서술형 및 수행 평가는 평가기준이 불분명한 경우가 많고, 이의 비율조정 등 평가에 관한 교사 자율성 여건 형성이 유언하지 못하다. 이런 점에서 서술형 및 수행 평가를 채점하는데 많은 시간과 노력이 필요함을 토로하고 있다. 즉, 협의의 과정(부분점수 기준설정 등)을 거치면서 시간 부족과 업무 부담만을 가중시킨다고 인식하고 있는 것이 일선학교의 현실임을 발견할 수 있었다. 정부는 이

런 점들을 철저히 인지할 필요가 있으며, 교사의 업무부담 경감과 서술형 및 수행 평가 기준(안)을 참고할 수 있는 자료제공이 시급한 실정이고, 단위별로 구체적으로 제공 되어질 필요성이 제공된다. 나아가 서술형 및 수행 평가만을 위한 교수학습·방법 개선을 위한 정책 적 방안 등을 고찰해 보아야 한다.

서술형평가 및 수행평가의 전문성을 향상시키기 위한 시발점은 평가투입이라고 할 수 있다. 이에 대한 논의 및 결론은 다음과 같다.

첫째, 대다수의 교사들이 평가의 목표설정 에 있어서 스스로는 나름의 준거를 가지고 평가의 목표를 설정하는 것으로 인식하고 있으나 그 준거가 명확하지 않다. 이러한 문제로 인하여 문항 개발과 채점기준 마련에 어려움을 느끼는 것으로 판단된다. 다시 말하면, 학교 현장에서 많은 교사들이 경험을 바탕으로 문항개발과 채점기준을 마련한다는 것이다. 또한 교사들은 성취기준을 중시하지 않는다는지, 교육과정에 제시된 성취기준을 무작정 따르려는 경향을 보이고 있다. 교육과정에 제시된 성취기준은 포괄적이고 추상적이기 때문에 학생의 수준에 걸맞게 성취기준을 수정·보완 기회가 마련되어야 할 것이다. 따라서 서술형평가 및 수행평가가 제 기능이 잘 발휘되게 하기 위해서는 교육과정에 제시된 성취기준을 평가 목표에 맞도록 수정·보완하는 것이 필수 선결과제라 하겠다.

둘째, 교사들이 서술형 및 수행평가의 문항 유형을 선정하는 준거가 명확하지 않다. 대다수의 현장 교사들은 ‘자료 제시형’, ‘응답 제한형’ 등과 같은 문항의 형식이나 학생의 응답 범위를 준거로 문항의 유형을 선정하는 것으로 보인다. 그러나 명확한 평가 요소를 정확히 측정할 수 있는 즉, 평가 목표에 부합하는 문항을 선정하는 것이 문항의 유형 선정에 있어서 가장 중요한 준거라 할 수 있다. 이에 본 연구에서는 ‘Bloom의 평가 유형’을 제안하고 있다.

서술형 및 수행평가에 있어서 현재 학교현장에서 교사들이 가장 많은 시간을 투입하고 있는 부분은 평가문항을 제작하고 그 채점기준을 설정하는 일일 것이다. 이 연구에서 서술형 및 수행평가 전문성의 영역에서 평가과정은 평가문항의 제작과 분할점수산출 그리고 채점과정이고 이에 대한 논의와 결론은 다음과 같다.

첫째, ‘교사들이 성취기준 및 이원목적분류표의 내용영역인 지식차원과 행동영역인 인지 단계를 핵심평가요소로 하는 평가문항을 제작하는가?’라는 평가문항 제작의 전문성에 대한 논의를 살펴보면, 성취기준에 대해서는 고등학교 교사들보다 중학교 교사들이 더 잘 이해하고 더 많이 활용하고 있지만 중등학교 교사 전체적으로 평가문항을 제작하는데 일정 수준 이상으로 활용하고 있다고 보인다. 그러나 중학교나 고등학교 모두 이원목적분류표를 문항개발의 핵심도구로 사용하지 않고 있음은 명백하다. 이는 이원목적분류표의 내용영역/지식차원과 행동영역/인지단계를 평가문항의 평가요소로 구체화되지 못하고 있음을 알 수 있다.

둘째, ‘교사들이 평가문항의 채점기준을 설정하는데 채점요소를 구체적이고 명확하게 설정하는가?’라는 채점 기준 설정에 대한 논의를 살펴보면, 기존의 평가요소 분석에서 보듯이 채점기준은 내용영역이 중복되는 형식을 취하고 있어서 기준의 타당성이 떨어진다. 채점기준을 설정하는데 있어서 중학교가 고등학교보다 더 잘하고 있다. 성적 부담이 적고 자유학

기제를 시행하는 중학교는 여유가 있고 대학입시로 인하여 내신성적에 부담이 큰 고등학교의 경우에는 중학교보다 엄격한 채점기준이 적용되고 있다는 의미일 것이다. 중학교와 고등학교의 이러한 차이에도 불구하고 중고등학교 모두 평가문항의 평가요소와 일관되고 정합적인 채점기준설정이 되지 못하고 있다.

셋째, 이상적인 평가문항 제작절차를 현실에 적용하는 것은 노력에 비해 효율이 떨어지는 것으로 나타난다. 현실적으로는 이원목적분류표의 이해와 작성, 서술형 및 수행평가 문항의 제작 그리고 채점기준표의 설정을 중심으로 하는 핵심적 제작절차나 채점절차를 적용하는 것이 비용대비 효과가 크다. 따라서 교사들에게 핵심적 서술형 및 수행평가 문항 제작절차에 관한 연수와 매뉴얼 제공이 기초적으로 제공되어야 할 것이다. 문항제작 및 채점기준 설정의 핵심적 절차과정 중에 성취기준 및 수준과 이원목적분류표에 따른 별 분할점수를 산출하는 것은 개별 평가문항의 난이도와 학생들의 능력이라는 요소를 비교 판단하여 학생의 학업능력을 신장하는 중요한 절차이므로 이에 대한 심도 있는 교사연수가 지속되어야 함은 물론 학생평가(서술형 및 수행평가) 전문성 척도 가운데 이를 포함하여야 할 것이다.

평가결과는 평가목적과 성취기준 및 수준 그리고 이원목적분류표와의 정합성, 성취기준과 평가문항 간의 정합성, 그리고 평가문항과 채점기준과의 일관성 및 정합성에 따른 학생평가결과의 분석과 해석에 대한 고찰이다. 평가결과의 분석은 결과의 양적 분석 및 통계적 분석으로 문항의 정답률, 난이도, 평균을 문항자체를 중심으로 하고 평가결과의 해석은 성취기준 및 수준의 달성정도를 중심으로 개별 학생의 발달정도를 중심으로 한다.

첫째, 평가결과에 대한 교사의 통계적 분석은 나이스에서 제공하는 지필평가에 관한 것이 대부분이었고 서술형 및 수행평가에 관한 통계적 분석은 거의 없었다. 그리고 평가에 대한 전문성(TEP) 전체에 대한 교사의 인식과 평가결과에 대해 분석하는 능력정도를 비교해보았을 때 평가결과의 분석에 대한 능력은 낮게 나타난다. 그리고 평가결과에 대한 분석능력이 평가과정, 즉 평가문항제작과 채점기준설정의 능력보다 현저히 낮다. 중고등학교 교사들이 평가전문성의 다른 영역보다 서술형 및 수행평가의 결과전문성 영역, 즉 평가결과의 통계적 분석 능력을 우선적으로 신장하고, 평가결과를 평가의 목적 및 목표와 연계하는 능력을 제고하도록 교육통계 분석과정을 제공해야 한다. 또한 나이스(NEIC)시스템 상에서 서술형 및 수행평가에 대한 독립적 통계분석 지원하는 것도 동시에 진행되어야 할 것이다.

둘째, 서술형평가 및 수행평가 결과를 이원목적분류표를 기준으로 해석할 수 있는지, 또는 성취기준 및 수준을 기준으로 해석할 수 있는지에 대한 논의에 대하여 입시 중심의 고등학교는 물론 자유학기제를 시행하고 있는 중학교에서조차도 그 해석능력이 현저히 떨어졌다. 이원목적분류표와 성취기준 및 수준에 의한 해석은 학생의 지식 정도와 인지발달의 정도를 파악하고, 학생의 강점과 약점을 파악하여 개별화된 학업향상 처방을 제시할 수 있는 근거가 될 것이다. 따라서 교사들의 평가결과 전문성(ERP)을 제고하기 위해서 성취기준 및 수준과 이원목적분류표에 대한 역량강화를 필수적 과정으로 제공해야 할 것이다. 또한 서술형 및 수행평가를 포함하는 교육(학생)평가전문성 인정제도나 교육(학생)평가 분석사 국가자격증 제도를 도입한다면 이러한 목표를 달성하는데 한층 강력한 동기를 제공할 수 있을 것

이다.

현 일선학교의 서술형 평가와 수행평가에 관한 평가 환류는 교사측면, 학생측면, 교육이해 당사자 측면 등에서 문제점을 고찰할 수 있다.

첫째, 일선 학교 교사는 학생 개인별 서술형 평가와 수행평가 결과를 종합·분석하고 개선방안을 도출하고 있지 않다. 즉, 교사는 서술형 평가와 수행평가 결과를 토대로, 학생들의 재능과 학업성취 정도를 파악하고 학생의 꿈과 끼 실현을 위한 학습계획을 구체적으로 세우고 있지 않으며, 서술형 평가와 수행평가 시행 취지에 부합한 학생들의 강점과 약점을 파악하고 맞춤형 교수학습방법의 개선을 위한 노력이 전혀 이루어지고 있지 않다. 더욱 심각한 것은 서술형 평가 문제를 출제 할 때, 학생 창의성 신장을 도모하기 위하여 출제를 하고 있지만 이에 적합한 효과적인 교육활동을 구상하고 있지 못한 점과 이의 결과를 활용하여, 학생 발달과정을 지속적이고 안정적으로 확인하고 개선방안을 모색하고 있지 않는 점이다. 정부는 이러한 점을 근간으로 서술형 평가와 수행평가의 본래 취지를 위한 평가 환류 측면의 정책적 개선노력이 절실히 필요하다.

둘째, 일선 학교 교사는 서술형 및 수행 평가를 통하여 학생 스스로 성취기준에 제시된 평가내용과 능력(이해, 적용, 분석)을 파악하여 강점과 약점을 고찰하고 있지 못하고, 심지어 교사들은 학생들조차 서술형 및 수행 평가에 대한 학업흥미와 동기를 전혀 유발하지 못하며 오히려 학업부담만 가중시킨다고 인지하고 있다고 밝히고 있다. 나아가 이를 통한 학습 능력 신장에 필요한 정보와 자기 주도 학습이 가능하게 할 수 있는 정보를 담지 못함을 발견할 수 있었다. 정부는 서술형 평가와 수행평가의 본래 취지를 위한 지속적이고 안정적인 학생 성장과정을 도모하기 위한 정책적 대안마련이 시급한 실정이다.

셋째, 일선 학교에서는 서술형 및 수행 평가 자료를 누적적으로 구비되어 있지 못한 실정이다. 즉, 이를 토대로 학생들의 학력 향상, 공교육 정상화와 사교육 경감 등의 교육정책 방향에 구체적으로 활용되고 있지 못한 실정이다. 서술형 및 수행 평가 자료를 토대로, 학생들의 학력신장에 얼마나 기여했는지에 대한 정보를 누적적으로 관리되어질 필요성이 제기된다. 정부는 학력 평가관리 시스템을 적용하고 활용할 수 있는 여건마련과 이를 토대로 교사들의 평가 전문성 향상을 위한 프로그램 개발 등이 시급한 실정이다. 즉, 정부는 학교현장의 안정적이고 지속적인 평가환류체제 구축을 위한 인적, 물적 여건 조성이 필요하다. 특히, 평가 환류에 적합한 맞춤형 교육과정 편성 및 운영, 평가 전문성 향상을 위한 전문가 협의체 구성 및 연수의 필요성 인지, 안정적이고 지속적인 학력 평가관리 및 평가실행 시스템 구축이 필요하다.

본 연구에서 수행한 서술형 평가 및 수행 평가의 평가전문성 제고 방안을 통해 향후 정책 분석 및 개발, 연구에 활용될 수 있는 다양한 정책 제언이 도출되었다. 정책 제언은 본 연구의 전문성 신장 모형을 토대로, 크게 평가 환경, 투입, 과정, 결과, 환류에 근거하여 평가 전문성과 관련한 교육정책 개선 제언 중심으로 구성되었다.

평가 환경은 교사들의 학생 평가 전문성 신장을 위하여 수행되어야 하는 일련의 활동들에 영향을 미치는 전반적인 것을 의미한다. 즉, 교사가 서술형 평가 및 수행 평가의 본질

적인 목표달성을 위해 수행해야 하는 제반적인 활동에 대하여 몰입할 수 있도록 주어져야 하는 여건을 의미하는 것이다. 일선 학교에서 현 평가 환경의 문제점은 서술형 평가와 수행 평가에 적합한 교육여건(교육과정), 교수학습·방법, 교육평가 세부여건 등에서 문제점을 고찰할 수 있는데 이를 개선시키기 위한 제언을 하면 다음과 같다.

첫째, 교육여건을 개선하기 위하여 정부는 우선적으로 교사의 업무 부담을 경감시킬 수 있는 정책적 방안 모색이 절대적으로 필요하다.

둘째, 정부는 서술형 및 수행 평가 취지에 맞는 교육과정이 개편시킬 필요가 있다. 나아가 장기적으로는 이러한 평가를 대입제도 등에 적용시킬 방안 모색도 필요하다.

셋째, 일선 학교 현장에서 교사는 명확한 서술형 및 수행 평가 기준 정립에 어려움을 토로하고 있다. 이를 개선시키기 위해서는 이와 관련한 실질적 실효성을 얻을 수 있는 연수 등의 방안을 적극적으로 모색해야 할 필요성이 제기된다. 나아가 서술형 평가 기준(안)을 참고할 수 있는 다양하고 구체적인 자료제시 방안 모색이 필요하다.

넷째, 일선 학교에서는 서술형 및 수행 평가를 위한 교사 협업이 적극적으로 이루어질 수 있는 학교분위기를 모색할 필요성이 제기된다.

다섯째, 정부는 서술형 및 수행 평가를 위한 교수학습·방법개선을 위한 정책적 방안 모색도 필요하다. 현 정부의 교육정책 즉 “꿈과 끼를 끌어내는 행복교육”에 부합하는 서술형 및 수행 평가를 위한 다양한 교수학습·방법을 창의·인성 프로그램과 연계하여 지속적이고 안정적으로 정착시킬 수 있는 방안 모색이 필요하다. 이를 위하여 서술형 및 수행 평가를 위한 교수학습·방법과 관련한 지속적인 우수사례를 발굴할 수 있는 정책 방안 모색이 절대적으로 필요하다.

여섯째, 서술형 및 수행 평가를 위한 다양한 교수학습·방법을 활용할 수 있는 수업시간 확보와 평가시기 결정 등의 합의점 방안 모색이 필요하다.

일곱째, 교육평가 세부여건 부분에서는 서술형 및 수행 평가 기준을 명확히 정립할 수 있는 안내 책자 제공 및 지속적인 홍보를 통하여 서술형 및 수행 평가 채점 시 교사들 간의 견해 차이를 조율하기 위한 방안 모색이 필요하다. 그리고 서술형 및 수행 평가만을 위한 수업시간 및 시험시간이 확보 방안 마련이 필요하다. 또한 서술형 및 수행 평가의 유연한 평가비율 조정, 시기와 범위문제, 출제와 채점시간에 부담감을 해소시킬 수 있는 방안 모색도 필요할 것이다. 마지막으로 서술형 및 수행 평가에 관한 교사 안내책자뿐만 아니라 학생 답안 작성 방법 등을 담은 학생 교육용 안내 책자 제공도 필요할 것으로 사료된다.

평가 투입은 교사들이 학생 평가를 위한 시작점이라 할 수 있으며, 평가 투입 단계에서는 평가 목표의 설정, 평가 혹은 평가 문항의 유형 선정이 이루어진다. 평가 과정은 투입 단계에서 설정된 평가 목표에 적합한 문항이 개발되어야 하며, 평가 결과는 평가 목표를 고려하여 학생의 성취 수준을 정확히 파악할 수 있어야 함을 의미한다. 평가 투입-과정-결과에서의 문제점은 경험적인 평가 목표 설정과 평가 목표를 고려하지 않은 평가 혹은 평가 문항의 유형을 선정함으로 인하여 평가 목표에 따라 학생의 성취수준을 정확히 측정하지 못하는 것이라 하겠다. 이를 개선하기 위한 제언을 하면 다음과 같다.

첫째, 성취기준과 평가문항 간의 정합성과 일관성에 대한 교사들의 인식과 관심을 강화하기 위해서 이원목적분류표와 평가문항과 관계와 제작원리를 안내하는 것이 교사들의 평가 전문성 제고를 위하여 반드시 필요하다. 그리고 교사들이 평가문항의 제작을 원활하게 할 수 있도록 Bloom의 평가유형에 따른 평가문항의 모듈화를 제안한다.

둘째, 성취기준 및 성취기준 그리고 이원목적분류표에 대한 정확한 이해와 이를 채점기준설정에 활용하는 능력을 신장시키는 위하여 성취기준·수준 및 이원목적분류표에 따른 평가문항 제작과 채점기준의 일관성과 정합성을 신장하는 연수와 매뉴얼이 필요하다. 이를 근본적으로 가능하게하기 위하여 수업과 평가를 통합하는 교수학습과정이 설계되어야 할 것이다.

셋째, 성취기준 및 수준과 이원목적분류표에 따른 분할점수를 산출에 대한 심도 있는 교사연수가 지속되어야 함은 물론 학생평가(서술형 및 수행평가) 전문성 척도 가운데 이를 포함하여야 할 것이다.

넷째, 교사의 통계적 분석 능력을 높일 수 있는 방안 모색이 필요하다.

다섯째, 교사의 역량 강화를 위하여 교육(학생)평가전문성 인정제도나 교육(학생)평가 분석사 국가자격증과 같은 제도적 장치를 마련하는 방안에 대한 모색이 필요하다.

평가 환류는 교사들의 학생 평가를 통하여 수행되어졌던 일련의 활동들을 단계별로 점검 및 관리하는 것을 의미한다. 무엇보다 교사가 서술형 평가 및 수행 평가의 본질적인 목표달성을 위해 수행되어졌던 제반적인 활동에 대하여 시스템적으로 구조화 시키는 것을 의미한다. 일선 학교에서 현 평가 환류의 문제점은 교사측면, 학생측면, 교육이해 당사자 측면 등에서 문제점을 고찰할 수 있는데 이를 개선시키기 위한 제언을 하면 다음과 같다.

첫째, 정부는 교사들이 서술형 평가와 수행평가 결과를 종합·분석하여, 학생들의 강점과 약점을 파악하고 교수학습방법의 개선점을 도모할 수 있는 시스템 구비가 되어야 한다. 이를 통하여, 학생들의 재능과 학업성취 정도를 파악하고 학생의 꿈과 끼 실현을 위한 학습계획을 세울 수 있는 제도적 마련이 시급하다. 이런 제도적 마련은 교사에게 학생의 창의성 신장을 위한 효과적인 교육활동을 구상할 수 있게 하며, 누적적인 자료는 학생 발달과정을 이해할 수 있으며 학생의 지속적이고 안정적인 성장을 도모할 수 있을 것이다.

둘째, 정부는 교사가 학생에게 성취기준에 제시된 평가내용과 능력(이해, 적용, 분석)을 파악하여 강점과 약점을 고찰할 수 있고, 꿈과 끼 실현을 위한 필요영역을 제공해 줄 수 있는 시스템 마련이 시급하다. 이를 위하여 이원목적분류표를 더욱 구체적으로 살펴볼 수 있도록 하면서, 학생 스스로 무엇을 알고, 무엇을 할 수 있는지에 대한 구체적인 분석을 제공할 수 있도록 해야 한다. 이를 활용한다면, 무엇보다 학생의 학습 능력 신장에 필요한 정보를 누적적으로 살펴볼 수 있을 것이다. 나아가 학생입장에서는 스스로 문제점을 파악할 수 있으며, 자기 주도적인 학습이 가능하게 할 수 있을 것이다.

셋째, 정부가 만약 이런 일련의 시스템이 안정적으로 구비시킨다면, 학생들의 학력 향상, 공교육 정상화와 사교육 경감 등에 기여할 수 있을 것이다. 정부가 학력 평가관리 시스템을 적용하여 서술형·수행 평가를 실시하여, 학생들의 학력을 파악하고 분석하는 일련의 절차적

지식을 활용할 수 있으면, 서술형 평가와 수행평가 결과에 대한 신뢰도와 타당도가 높일 수 있고, 무엇보다 교육이해 당사자들의 평가 전문성 신장에 도움이 될 수 있을 것이다. 무엇보다 정부는 안정적이고 지속적인 평가환류체계 구축을 위한 인적, 물적 여건 조성이 절대적으로 필요한 시점에 있다.

차례

I. 서론

1. 연구의 필요성 및 목적	1
2. 연구 내용 및 방법	4

II. 서술형 및 수행 평가의 이해

1. 일반적인 서술형평가의 이해	9
가. 서술형평가의 정의와 범주	9
나. 기존의 선택형 평가와 서술형 평가와의 비교	11
다. 서술형 평가의 설계	13
라. 서술형 평가의 문항 유형	16
마. 서술형 평가의 문항 제작 절차	20
바. 서술형 평가의 문항 채점 절차	30
2. 일반적인 수행평가의 이해	37
가. 수행평가의 정의와 범주	37
나. 선택형 평가와 수행평가와의 비교	40
다. 수행평가의 설계	41
라. 수행평가의 문항 유형	44
마. 수행평가의 문항 제작 절차	57
바. 수행평가의 문항 채점 절차	66
3. Bloom의 서술형 및 수행평가의 이해	69
가. Bloom의 지식 차원의 분류	69
나. Bloom의 인지 단계의 분류	70
다. Bloom의 평가 유형	71
4. 해외 사례 분석	73
가. 미국	73
1) 학생 평가의 이해	73
2) 서술형 평가의 이해	78
3) 수행 평가의 이해	80

4) 시사점	81
나. 핀란드	84
1) 학생 평가의 이해	84
2) 서술형 및 수행평가의 이해	89
3) 시사점	96
Ⅲ. 교사의 서술형 평가 및 수행평가 전문성 규정과 분석 및 제고 방안	
1. 교사의 평가 전문성 규정	98
가. 교사의 평가 전문성 규정 선행연구	98
나. 교사의 평가 전문성 규정	102
2. 교사의 평가 전문성 실태 분석	107
가. 평가 환경 분석	
1) 정량적 분석	107
2) 정성적 분석	108
3) 종합결과	109
4) 평가 환경 개선방안	109
나. 평가 투입 분석	
1) 정량적 분석	111
2) 정성적 분석	113
3) 종합결과	120
4) 우수사례 적용 및 분석	122
5) 평가 투입 전문성 제고 방안	125
다. 평가 과정 분석	
1) 정량적 분석	127
2) 정성적 분석	129
3) 종합결과	138
4) 우수사례 적용 및 분석	140
5) 평가 과정 전문성 제고 방안	144
라. 평가 결과 분석	

1) 정량적 분석.....	147
2) 정성적 분석.....	149
3) 종합결과.....	151
4) 우수사례 적용 및 분석.....	152
5) 평가 결과 전문성 제고 방안.....	152
마. 평가 환류 분석	
1) 정량적 분석.....	155
2) 정성적 분석.....	159
3) 종합결과.....	163
4) 평가 환류 개선방안.....	163
IV. 결론 및 제언.....	164
1. 결론.....	164
2. 제언.....	167
참고문헌.....	172
부록.....	178
◆ 서술형 및 수행평가에 대한 중·고등학교 교사의 인식 설문지	
◆ 서술형 및 수행평가에 대한 중·고등학교 교사의 인터뷰 질문지	
◆ 중등교사 평가 전문성 체크 리스트	

표차례

<표 I -1> 연구내용별 연구방법 및 추진 상황.....	8
<표 I -2> 문항적합성 판정을 위한 교사집단의 특성.....	9
<표 I -3> 설문조사 체계표.....	10
<표 I -4> 설문지 조사 일정.....	11
<표 I -5> 설문지 배포 현황.....	11
<표 II -1> 선택형 평가와 서술형 평가의 비교.....	20
<표 II -2> 수업과 평가의 상관관계.....	25
<표 II -3> 2015학년도 ()학기 ()학년 (서술형)고사 이원목적분류표.....	29
<표 II -4> 2015학년도 ()학기 ()학년 (서술형)고사 채점기준표.....	30
<표 II -5> 서술형 문항카드 A양식.....	31
<표 II -6> 서술형 문항카드 B양식.....	32
<표 II -7> 수행평가와 관련된 용어 정의.....	40
<표 II -8> 선택형 평가와 서술형 평가의 비교.....	43
<표 II -9> 구술시험을 위한 평가 기준표 (예시).....	48
<표 II -10> 구술시험의 진행과정 및 방식.....	49
<표 II -11> 찬·반 토론법을 위한 평가 기준표 (예시).....	51
<표 II -12> 영어 말하기 영역의 수행평가를 위한 채점 기준표의 예.....	53
<표 II -13> ‘읽기’ 능력을 평가하기 위한 평정 척도.....	56
<표 II -14> 자기평가 보고서의 양식 (예시).....	57
<표 II -15> 최종 연구 보고서 평가 기준표 (예시).....	58
<표 II -16> 홍보자료집을 위한 평가 기준표 (예시).....	59
<표 II -17> 수행평가 채점 기준표.....	61
<표 II -18> 2015학년도 ()학기 ()학년 (수행형)고사 이원목적분류표.....	65
<표 II -19> 2015학년도 ()학기 ()학년 (수행형)고사 채점기준표.....	66
<표 II -20> 수행형 문항카드 A양식	67
<표 II -21> 수행형 문항카드 B양식	68
<표 II -20> Bloom의 지식 차원의 분류.....	72
<표 II -21> Bloom의 인지 단계의 분류.....	73
<표 II -22> Bloom의 평가 유형.....	74
<표 II -23> 2015년도 고등학교 졸업자 학력평가 요건(위싱턴 주).....	79
<표 II -24> 평가 항목별 비중.....	83
<표 II -25> 종합평가 지침.....	83
<표 II -26> 한국과 미국의 중등학교 교사의 학생평가 전문성 기준 비교분석.....	85
<표 II -27>언어능력척도(The Language Proficiency Scale).....	91

<표 II-28> 핀란드 중학교 학생 성적표	94
<표 II-29> 수학 평점 8점(good) 취득을 위한 기준	95
<표 II-30>핀란드 헬싱키시 학교 성적 평점 기준(2010-2011).....	96
<표 II-31> 핀란드 중학교 생활기록부	97
<표 III-1> 교사의 평가전문성 주요 영역 및 항목	102
<표 III-2> 교사 평가전문성 신장 구성 영역 및 항목 내용 정리	105
<표 III-3> 평가 환경 전체 문항에 대한 기술통계 및 학교 급에 따른 차이 검정	108
<표 III-4> 평가 환경 전체 문항에 대한 기술통계 및 담당 교과목에 따른 차이 검정	108
<표 III-5> 평가 환경 인터뷰 내용	109
<표 III-6> 평가 환경 개선방안	110
<표 III-7> 평가전문성 전체 및 평가 투입 전문성	112
<표 III-8> 평가투입 전문성에 대한 학교 급별 차이	112
<표 III-9> 평가투입 전문성에 대한 과목별 차이 검증	113
<표 III-10> 평가투입 전문성에 대한 문항별 분석	113
<표 III-11> 평가과정 전문성(EPP)에 대한 학교 급별 차이	127
<표 III-12> 평가과정 전문성(EPP)에 대한 과목별 차이 사후검증	128
<표 III-13> ‘평가도구개발’에 대한 문항별 분석	128
<표 III-14> ‘채점기준 설정’에 대한 문항별 분석	129
<표 III-15> Bloom의 이원목적분류표에 따른 평가문항 제작	144
<표 III-16> Bloom의 이원목적분류표에 따른 채점기준 설정	144
<표 III-17> 평가결과 전문성(ERP)에 대한 학교 급별 차이	146
<표 III-18> 평가결과 전문성(ERP)에 대한 과목별 차이 사후검증	147
<표 III-19> 평가결과 전문성(ERP)에 대한 과목별 차이 사후검증	147
<표 III-20> ‘평가결과 분석 및 해석’에 대한 문항별 분석	147
<표 III-21> Bloom의 이원목적분류표에 따른 평가문항 해석	152
<표 III-22> 평가환류 전체 문항에 대한 기술통계 및 학교 급에 따른 차이 검정	153
<표 III-23> 평가환류 개별 문항에 대한 기술통계 및 학교 급에 따른 차이 검정	153
<표 III-24> 평가환류 전체 문항에 대한 기술통계 및 담당 교과목에 따른 차이 검정	154
<표 III-25> 평가환류 개별 문항에 대한 기술통계 및 담당 교과목에 따른 차이 검정	154
<표 III-26> 평가환류 전체 문항에 대한 기술통계 및 교직 경력에 따른 차이 검정	155
<표 III-27> 평가환류 전체 문항에 대한 기술통계 및 교직 경력에 따른 차이 검정	156
<표 IV-1> 시기별 정책제언	183

그림차례

[그림 II-1] 응답 제한형의 예시	20
[그림 II-2] 분량 제한형의 예시	21
[그림 II-3] 내용 범위 제한형의 예시	21
[그림 II-4] 서술양식 제한형의 예시	22
[그림 II-5] 응답 자유형의 예시	23
[그림 II-6] 응답 자유형에서 응답 제한형으로 변형한 예시	23
[그림 II-7] 단독 과제형의 예시	24
[그림 II-8] 자료 제시형의 예시	24
[그림 II-9] 일상적 수행평가 및 서술형 평가문항 제작절차	26
[그림 II-10] 이상적 서술형 평가문항 제작절차	27
[그림 II-11] 핵심적 서술형문항 제작절차	28
[그림 II-12] 서술형 평가 문항의 채점기준에 대한 예시	36
[그림 II-13] 일상적 채점절차	38
[그림 II-13] 이상적 채점절차	38
[그림 II-13] 핵심적 채점절차	39
[그림 II-14] 선택형 문항의 평가와 수행평가의 연속적 개념	42
[그림 II-15] 일상적 수행평가 및 수행형 평가 문항 제작 절차	61
[그림 II-16] 이상적 수행형(구술) 평가문항 제작절차	63
[그림 II-17] 핵심적 수행형(구술) 문항 제작절차	64
[그림 II-18] 수행평가 문항의 채점기준에 대한 예시	71
[그림 III-1] 중등교사 평가 전문성 신장 영역	107
[그림 III-2] 서술형 및 수행평가 문항 제작 흐름도	123
[그림 III-3] 국어과 서술형 평가 예시	126
[그림 III-4] 서술형 평가 문항카드	140
[그림 III-5] 평가문항	141
[그림 III-6] 채점 기준표 및 유의사항	142
[그림 IV-1] 성취기준과 평가결과의 분석 및 해석	152

I. 서론

1. 연구의 필요성 및 목적

박근혜 정부의 교육정책 방향은 꿈과 끼를 키우는 행복교육이다. 행복교육의 의미는 학교 교육을 통하여 배움을 즐기는 것이다. 즉, 많이 알게 하는 교육보다 좋아하게 하는 교육, 더 나아가 배움을 즐길 수 있게 하는 교육을 지향하는 것이다. 꿈과 끼를 키우는 것은 학생 각자의 타고난 소질과 끼를 찾아내어 키워주고 그 꿈을 이루어 가도록 하는 것을 의미한다. 이를 달성하기 위해서는 지식(교과) 중심교육에서 핵심 역량 중심 교육으로 전환이 필요하며, 획일적 지식 전달교육에서 학교 여건, 학생 수준 고려 즉 성취기준을 근간으로 수업 변화 등의 교사 역량 강화가 무엇보다도 중요하다. 현재, 지식의 전수에 집중하는 교육현실은 학생들의 지필식 평가 방식을 통하여 학업성취도 향상의 수월성은 이루었으나, 학생들은 학업에 대한 흥미상실과 창의성을 억압하는 부작용을 초래하는 결과를 가져오게 되었다. 또한 과도한 경쟁과 입시 중심의 교육은 학생들의 스트레스를 해소할 여유를 주지 못하였고, 이로 인해 학교폭력, 집단 따돌림 등의 비인간적인 행동으로 표출되고 자살로 이어지는 등 심각한 부작용을 초래하고 있는 실정이다(곽영순, 2015).

다행히 최근, 이런 현실을 극복하기 위한 하나의 대안으로 기존의 획일화된 교육에서 벗어나 자신의 가치와 능력을 재발견하고, 자신이 갖고 있는 꿈과 끼 탐색을 통하여 진정한 자아실현을 위한 행복교육으로 교육목표의 초점이 바뀌어가고 있다. 이러한 움직임은 교육과정 개정에서 더욱 공고히 하고 있다. 최근 2009 개정 교육과정에서는 전통적인 교육과정 즉 3R(읽기, 쓰기, 수리)을 중시하는 내용에서 지식의 이해와 축적을 넘어 의사소통 능력, 창의력, 협동심, 문제해결력 등 교육의 결과로 나타날 학생들의 역량 변화에 초점을 맞추고 있다. 전통적인 교육과정은 학교에서의 활동 결과를 통해 학생에게 어떤 역량이 함양되었는지에 관심을 갖기보다는 교과 수업의 내용을 얼마나 암기하고 계산하고 글로 표현할 수 있는지에 초점을 두었다(성태제, 2010). 이에 따라 학교 수업은 교육과정에 충실한 수업보다는 교과서의 내용을 전달하는 방식으로 이루어지게 되었다. 이러한 전통적 수업은 결국 학생들의 꿈과 끼를 키우고 잠재력을 드러내게 하는 데에는 한계가 있었다. 특히, 현대 지식 사회에서 강조되는 의사소통 능력, 공동체 정신, 문제해결력 등의 핵심역량을 키우는 것에 대해서는 교육의 목표와 내용을 통해서 혹은 평가를 통해서도 교육의 핵심으로 자리매김하지 못하였다.

2014년 9월 24일 「2015 문·이과 통합형 교육과정의 총론 주요사항 발표」에서 주

요 개정 방향의 하나로써 “미래사회가 요구하는 역량 함양이 가능한 교과 교육과정을 개발한다.”고 제시하였다. 그러면서 미래사회가 요구하는 역량은 핵심원리 중심으로 학습 내용을 감축하고 교수·학습 및 평가방법을 개선하여 학생의 학습 부담을 줄이고 배움의 즐거움을 느낄 수 있도록 한다는 것이다. 세계적인 추세와 2009 개정 교육과정의 총론 수정 고시를 한 2012년 교육과정 총론이나 2015 개정 교육과정 총론의 기본 방향에서 핵심역량 위주로 교육과정을 개편하여 교육의 큰 방향을 제시하였다는 것은 21세기 교육의 핵심은 핵심역량인 의사소통능력, 협업능력, 창의력, 비판적 사고능력 등을 길러주는데 있는 것이다. 2009 개정 교육과정에서 제시하고 있는 학생들의 역량 신장을 위해서는 무엇보다 수업 방법의 개선, 평가 방법의 혁신 등 교육 전반에 걸친 변화가 필요하며 무엇보다 교사의 평가전문성 신장이 선결되어야 한다.

교사가 일상적으로 시행하고 있는 학생 평가는 교육평가의 일환이다(문영주, 2013). 보편적으로 교육평가는 목적 달성을 지향하는 학습자의 향상정도를 파악하기 위한 것이다. 따라서 교육평가의 과정은 본질적으로 교육과정 및 수업의 프로그램에 의하여 교육목표가 실제로 어느 정도나 실현되었는지를 밝히는 과정이다. 또한 평가의 과정은 학습자의 행동변화 및 학습과정에 관한 정보를 수집하고 이용하여 교육적 의사결정을 내리는데 도움을 주거나 혹은 의사결정을 하는 과정이다. 따라서 교육평가는 교육목표를 올바르게 설정하였는지, 목표 실현을 위한 교육의 계획과 과정은 적절한지, 그리고 궁극적으로 교육 목표가 제대로 성취되었는지를 확인·판단하는 일련의 과정이다. 이처럼 교육평가의 궁극적인 목적은 학습과 교육과정에 최대한 도움을 주어 학습의 효과를 극대화시키려는 데 있다. 그러므로 교육평가는 학생의 학업성취도 확인, 교육과정의 평가와 개선, 교육정책 관련 의사결정을 위한 기반 제공 등의 실제적인 목적을 가지고 시행하게 된다.

따라서 교사는 평가의 궁극적 목표 추구 즉, 성취기준에 비추어 제대로 학습했는지 측정할 수 있는 역량을 갖춰야 하며, 수업은 학생이 성취기준에 도달하는데 초점을 두고, 교육 내용과 교수 학습 방법 그리고 평가가 정합성을 이루는 교육을 실현해야 한다. 이를 위해서, 획일화 되어 있는 선택형 지필평가에서 서술, 논술, 구두, 듣기평가, 프로젝트평가, 포트폴리오 등 다양한 평가 방식 활용이 요구된다. 이러한 요구로, 우리나라의 학교교육에서는 1990년대 중반부터 새로운 학생평가 방식으로써 수행평가 개념을 도입하기 시작하였다. 서술형 평가 및 수행평가는 단편적인 지식만을 암기하도록 조장하는 기존의 학력평가 방식을 지양하고 학생의 창의성이나 문제해결력 등 고등 사고 기능을 파악하고 개별적인 학습 신장을 도모하기 위한 것이다(백순근, 2000). 즉, 평가자가 학습자들의 수행 과정 및 결과를 직접 관찰하고, 그 관찰결과를 전문적으로 판단하는 평가방법이라 할 수 있다. 이러한 평가방법의 도입은 교육개혁이나 교육의 재구조화, 교육목표의 변화와 그에 따른 평가체제의 변화 요구 등과 부합되어 신교육체제로서의 평가의 방향이나 위상에 많은 변화를 가져오는 등 긍정적인 측면이 있었다. 하지만, 서술형 평가 및 수행평가의 실제적 활용의 토대가 될 수 있는 필요성, 가치판단을 내리기 위한 타당한 기준 마련, 공정한 채점 방법과 채점에 소요되는 비용의 문제, 평정자의 오류 가능성,

채점자간 또는 수행 과제 간 내적일치도 결여 등 부정적인 측면도 나타나고 있다.

서술형 평가 및 수행평가는 선다형 평가에 비해 더 많은 시간이 요구되며, 각 평가에 포함되는 과제의 수가 더 적기 때문에 각 과제들이 학생들의 점수에 차지하는 상대적인 비중이 매우 커서 학생들에게 상당한 부담을 주고 있다(김신영, 2014). 그리고 서술형 평가 및 수행평가가 가지고 있는 큰 문제점 중의 하나는 신뢰도에 관한 문제이다(강호수 외, 2014; 한국교육과정평가원, 2005). 서술형 평가 및 수행평가는 학습자의 반응을 평가하기 위해서 관찰이나 교사의 판단에 크게 의존할 수밖에 없고, 또 표준화가 어렵기 때문에 신뢰도가 높지 못하다는 문제를 안고 있다. 더욱이 학생 중심 참여 수업 및 협력 학습의 과정과 결과가 평가에 충분히 반영되도록 학생 평가 방법 개선이 요구되고 있지만, 평가제도 개선에 대한 시대적 요구가 있음에도 불구하고, 일선교사들은 학교는 국가가 학교에 부여한 자율권의 범위와 그 내용을 분명히 파악하고 있지 못하는 실정이고(김지선 외 2014; 김경희 외, 2006; 윤정일, 신호정, 2006), 심지어, 교육정책의 방향을 올바르게 이해하고 있지 못하고(박정, 2013; 손준중, 2010), 학교에 주어진 자율권을 최대한 활용하고 있지 못하고 있다(박도순, 2013; 고정화, 2010). 그리고 학교 교육과정 개발 과정에 교사들의 참여가 저조(성태제, 2010) 즉, 교사 스스로가 학교 교육과정 편성의 주체라는 인식이 미약한 실정이다(성태제, 2010; 백순근, 2010).

뿐만 아니라 학교현장에서의 바람직한 서술형 평가 및 수행평가의 실시를 위한 세 가지 현안 과제로 교사, 학생, 학부모, 교육당국의 인식에 대한 문제와 학교 서술형 평가 및 수행평가 실시여부의 자율권과 실시에 따른 각종 지침이나 자료 및 교사의 전문성 향상의 기회, 평가 결과의 내신 반영 절차와 같은 행정상의 문제 등이 있다. 그리고 열악한 교육환경, 교육과정, 교과내용의 부적절성의 문제, 학생의 과도한 학습 부담, 불완전한 서술형 평가 및 수행평가 도구 및 결과의 활용 등의 시행상의 문제를 들고 있다(김석우, 2014). 교사의 전문성과 관련해서는 일선 교사들은 스스로 자신의 학생평가 전문성이 낮다고 생각하지 않지만(백순근, 2010), 교육 전문가와 학생평가 전문가들은 일선 교사들의 평가 전문성이 낮다고 인식하고 있다(이인제 외 2004). 실제적으로 교사들은 업무 시간의 1/3-1/2의 시간을 평가와 관련된 활동에 할애하고 있으며, 이러한 업무의 중요한 부분으로 평가 관련 능력이 필요하지만(박도순, 2013) 교사들은 평가와 관련된 전문적인 훈련을 받은 경험이 없거나 적으며, 평가활동을 수행하는 데 준비가 덜 되어 있다고 스스로 느끼기도 한다(한국교육과정평가원, 2005).

본 연구에서는 박근혜 정부의 꿈과 끼를 키우는 행복교육 실현의 토대가 되는 교사의 평가전문성 신장 방안을 모색하여, 교사의 평가전문성을 개선시키기 위한 프로그램 개발 등을 위한 자료 마련 등을 위한 것이다. 서술형 평가 및 수행평가 중심의 중등학교 교사의 평가전문성 제고 방안 모색은 21세기 교육의 핵심역량인 의사소통능력, 협업능력, 창의력, 비판적 사고능력 등을 길러주는데 일조하게 될 것이다. 이러한 것을 대변한 듯, 현재 학교 현장에서는 서술형 문항의 출제 비율을 50%이상 확보하도록 강조하고 있는 실정이다. 하지만 일선교사들은 여전히 서술형 평가 및 수행평가를 수행 및 평가하는데 어려움을 토로하고 있

고, 전반적으로 평가에 관한 지식이 부족한 상태이다(김석우, 2014; 박도순, 2013; 성태제, 2010). 이러한 현 시점에서 학교 현장 교사들에게 학생평가의 수월성을 달성하기 위한 수행도구로서 중등교사의 평가 전문성 제고 방안을 마련하여 제공함으로써 학생평가의 본래의 목적인 학생의 학습과 성장을 지원하는 데 효과적인 정보를 제시하고자 한다. 이러한 목적을 달성하기 위하여 설정한 본 연구의 연구문제는 다음과 같다.

첫째, 서술형 평가와 수행평가에 대한 학교 현장의 실태는 어떠한가?

둘째, 서술형 평가와 수행평가에 대한 해외 사례는 어떠한가?

셋째, 서술형 평가와 수행평가에 대한 중등학교 교사의 평가 전문성 제고 방안은 무엇인가?

2. 연구 내용 및 방법

본 연구에서 수행한 연구내용 및 연구방법은 <표 I-1>과 같다.

<표 I-1> 연구내용별 연구방법 및 추진 상황

연구내용	하위영역	연구방법
1. 서술형 평가와 수행평가 배경 및 교사평가 전문성 고찰	1-1. 수행평가 및 서술형 평가 배경 고찰	문헌연구
	1-2. 현재 교육정책 기조와의 연계성 고찰	문헌연구, 서술형 평가 및 수행평가 관련 각 교육청 발간 매뉴얼, 교육과학기술부 정책자료
	1-3. 교사평가 전문성 고찰	문헌연구, 서술형 평가 및 수행평가 관련 연구보고서 교육자료, 각종 연수 자료 및 연구보고서, 학교교육계획서
2. 서술형 평가와 수행평가를 시행하는 해외사례 분석	2-1. 미국 사례	문헌연구, 교사 심층면담
	2-2. 핀란드 사례	문헌연구
3. 서술형 평가와 수행평가에 대한 교사들의 인식 조사	3-1. 서술형 평가와 수행평가 요구·실태 조사	설문조사, 심층면담
	3-2. 서술형 평가와 수행평가 관련 전문성 조사	설문조사, 심층면담
4. 서술형 평가와 수행평가 우수 사례 발굴 및 교사 체험 소개	4-1. 서술형 평가와 수행평가 우수사례	심층면담
5. 평가 전문성 제고 방안 마련	5-1. 평가 전문성 체크리스트	문헌연구, 설문조사 결과, 심층면담 결과 통합
	5-2. 평가 전문성 제고 방안	문헌연구, 설문조사 결과, 심층면담 결과 통합

가. 연구 내용별 추진 과정

1) 서술형 평가와 수행평가 배경 및 교사평가 전문성 고찰

먼저 서술형 평가와 수행평가 배경 및 교사평가 전문성 고찰에서는 수행평가 및 서술형 평가 배경 고찰, 현재 교육정책 기조와의 연계성 고찰, 교사평가 전문성 고찰 등의 하위영역으로 나눌 수 있다. 수행평가 및 서술형 평가 배경 고찰에는 관련문헌, 서술형 평가 및 수행평가 관련 연구보고서, 각종 연수 자료 및 연구보고서 등의 자료를 바탕으로 하였다. 이러한 자료를 바탕으로 수행평가 및 서술형 평가의 정의와 범주, 문항비교, 문항 설계, 구조 및 유형, 문항제작절차, 문항 채점 절차 등의 내용을 고찰하였다. 현재 교육정책 기조와의 연계성 고찰 하위영역에서는 관련문헌, 서술형 평가 및 수행평가 관련 연구보고서, 각 교육청 발간 서술형 평가 매뉴얼, 교육과학기술부 정책자료 등의 자료를 바탕으로 하였다. 교사평가 전문성 고찰 하위영역에서는 주로 교사평가 전문성 관련문헌을 바탕으로 평가모형 및 교사 평가전문성 영역, 항목 및 항목 내용을 추출하였다. 이는 이후 연구(서술형 평가와 수행평가에 대한 교사들의 인식 조사)의 근거로 작용한다.

2) 서술형 평가와 수행평가에 대한 교사들의 인식 조사

가) 설문지 개발 및 조사

먼저 서술형 평가와 수행평가에 대한 교사들의 인식 조사를 위해 설문지를 개발하였다. 이 설문지는 서술형 평가 및 수행평가 관련 요구·실태 및 전문성의 내용으로 구성되어 있다. 그런데 서술형 평가와 수행평가 관련 전문성에 대한 진단의 초점을 교사의 실천에 둘 것인가, 알고 있는 그 자체에 둘 것인가의 문제제에 대한 심층적인 논의 과정 있었다. 이 과정에서 서술형 평가와 수행평가 관련 전문성을 능력의 개념으로 정의하였기 때문에 능력에서 우선 알고 있는 지식이 우선되어야 하기 때문에 서술형 평가 및 수행평가 관련 모든 설문문항을 “~ 한다”로 통일하였다. 서술형 평가 및 수행평가 관련 설문지를 구성하기 위해 교사의 평가전문성 규정 선행연구를 바탕으로 157(요구·실태 38문항, 전문성 119문항)개의 1차 예비문항을 작성하였다. 먼저 연구위원들이 협의를 거쳐 내용적 유사성에 기초하여 문항을 통합하고, 의미가 모호한 문항들을 제외시켜 67(요구·실태 17문항, 전문성 50문항)개의 2차 예비문항을 작성하였다. 특히, 1차 예비문항 중 서술형 평가 및 수행평가와 거리가 있는 일반적 평가전문성 문항(69개)을 많이 제외시켰다. 2차 예비문항에 대한 문항적합성 검증을 위해 중등교사 중에서 20명(중학교교사 국어, 영어, 수학, 사회, 과학 각 2명, 고등학교교사 국어, 영어, 수학, 사회, 과학 각 2명)의 검토위원을 선정하였다.

<표 I -2> 문항적합성 판정을 위한 교사집단의 특성

구분		사례수(%)	계(%)
성별	남	12	20(100)
	여	8	
학교 급	중학교	10	20(100)
	고등학교	10	
과목	국어	4	20(100)
	영어	4	
	수학	4	
	사회	4	
	과학	4	

이렇게 구성된 67개의 예비문항에 대하여 서술형 평가 및 수행평가 관련 요구·실태 및 전문성의 내용으로 적합한가, 적합하지 않으면 그 이유는 무엇인가를 문항적합성 검토위원으로 하여금 검토하게 하였다. 그 결과 서술형 평가 및 수행평가 관련 요구·실태 및 전문성의 내용으로 부적합한 것으로 판정된 문항(20명 중 4명 이상이 부적합하다고 반응한 문항)은 31개였다. 이렇게 31개 문항을 제외한 36개 문항에 자유기술 문항 1개, 인구통계학적 문항 3개를 포함하여 최종 40개의 설문문항을 확정하였다. 구체적 설문조사 체계표는 <표 I -3>과 같다.

<표 I -3> 설문조사 체계표

영역	문항 내용	설문번호
배경변인	학교 급	1
	담당 교과	2
	교육 경력	3
서술형 평가 및 수행평가 요구·실태	서술형 및 수행평가 관련 채점 결과 이의(異議) 유무	1
	서술형 및 수행평가 관련 연구 또는 교육 우선순위	2
	서술형 및 수행평가 관련 전문성 향상을 위한 과제	3
	서술형 및 수행평가 관련 연수 실적	4
	서술형 및 수행평가 실시의 어려운 점	5
	서술형 및 수행평가 관련 개인적 견해 기술	6
서술형 평가 및 수행평가 전문성	각급 학교의 서술형 평가 분위기	1
	서술형 및 수행평가 관련 교육과정 이해	2
	서술형 및 수행평가 관련 교사 협의회	3
	서술형 및 수행평가 관련 시간적 여유	4
	성취기준에 따른 수업진행	5
	성취기준 코드 이해	6
	성취기준에 따른 서술형 평가 관련 문항 선택	7
	서술형 및 수행평가 관련 편파성 최소화	8
	성취기준 성취수준 고려	9
	교수학습 개선 정보 고려	10
	학습능력 정보고려	11

	문항개발 절차에 근거한 서술형 평가 관련 문항개발	12
	이원목적분류표에 근거한 서술형 평가 관련 문항개발	13
	서술형 및 수행평가 관련 난이도 예측	14
	서술형 및 수행평가 관련 변별도 예측	15
	서술형 및 수행평가 관련 기준점(분할점수) 산출	16
	서술형 및 수행평가 관련 타당도와 신뢰도 검토	17
	성취기준 성취수준에 따른 채점기준 설정	18
	이원목적분류표에 근거한 서술형 평가	19
	서술형 및 수행평가 관련 채점기준 안내	20
	채점기준 수정·보완	21
	서술형 및 수행평가 관련 결과 통계분석	22
	평가목적에 따른 서술형 평가 관련 결과해석	23
	학생의 강·약점 해석	24
	성취기준 성취수준에 근거한 결과해석	25
	서술형 및 수행평가 결과 관련 수업계획 수정·개선	26
	학생들의 약점에 대한 개선안 마련	27
	성취기준 성취수준의 정교화	28
	학생 개인의 성장정도 확인	29
	학급 및 학년 수준에서의 활용	30
	학교교육 계획 수립의 기초자료로 활용	31

<표 I -4> 설문지 조사 일정

단계	내용	일시
설문지 개발	설문지 초안 완료	06.24
	설문지 협의 및 수정	06.29
	설문지 타당화	07.01
	설문지 최종확정	07.02
설문지 인쇄 및 준비		07.03 - 07.05
예비 설문조사 실시		07.06 - 07.24
본 설문조사 실시		08.24 - 09.11
자료 처리 및 최종 산출물		07.25 - 09.18

이렇게 구성된 ‘서술형 및 수행평가 대한 중고등학교 교사들의 인식 설문지’는 <표 I -4>의 일정과 <표 I -5>의 설문지 배포 현황에 따라 예비 조사 및 본 조사는 실시하였다.

<표 I -5> 설문지 배포 현황

조사단계	지역	학교급	배포수
예비조사 및 본조사	부산	고등학교	8×25=200
		중학교	10×20=200
본조사	경인	고등학교	3×25=75

		중학교	4×20=80
"	충청	고등학교	3×25=75
		중학교	4×20=80
"	전라	고등학교	3×25=75
		중학교	4×20=80
"	경상	고등학교	3×25=75
		중학교	4×20=80
"	강원	고등학교	3×25=75
		중학교	4×20=80
계			1,175

나) 심층면담

중등학교 서술형 평가와 수행평가에 대한 학교 현황 파악, 실행 형태, 애로사항, 모범사례 등을 파악할 목적으로 서술형 평가와 수행평가의 계획, 실시, 채점, 결과활용, 학생평가 교육청 지침, 학생평가 학교관리규정, 학교분위기, 기타 등의 내용을 면담을 통해 살펴보았다. 면담 대상은 중학교 교사 6명, 고등학교 교사 6명으로 설정하였다. 면담 시 주요 내용은 서술형 평가와 수행평가에 대한 요구사항, 문제점, 정책과 관련된 개선점 등으로 중등학교 교사의 평가 전문성을 신장하는데 필요한 시사점을 도출하였다.

다) 서술형 평가와 수행평가 시행하는 해외 사례 분석

서술형 평가와 수행평가가 일찍부터 발달한 해외사례를 분석함으로써, 서술형 평가와 수행평가 시행지침 및 지원방안 마련에 시사점을 얻고자 하였다. 본 연구에서는 문헌 분석을 통하여 미국, 핀란드의 사례를 바탕으로 시사점을 도출하였다.

라) 서술형 평가와 수행평가 시행 지침 개발 및 지원방안 마련

서술형 평가와 수행평가 시행 지침 개발 및 지원방안 마련에서는 앞서 진행된 서술형 평가와 수행평가 이론적 배경, 서술형 평가와 수행평가 관련 해외사례 분석, 서술형 평가와 수행평가 실태 분석, 우수 사례 적용 및 분석 등을 종합하였고, 이를 바탕으로 중등학교 교사의 평가 전문성 제고 방안을 마련하였다. 특히, 평가 전문성 체크리스트의 개발 보급은 현장 교사들에게 전문성을 신장시키는데 많은 도움이 될 것으로 기대된다.

Ⅱ. 서술형 평가와 수행 평가의 이해

1. 일반적인 서술형 평가의 이해

가. 서술형 평가의 정의와 범주

‘서술형 평가’ 시행의 근본 취지는 지식 중심 평가에서 사고력 중심 평가로 전환하자는 데 있다. 서술형 평가의 목적은 특정한 지식을 맞게 알고 있는지를 주로 평가하고자 하는 선택형, 단답형 또는 완성형 평가와는 달리, 정답의 결과보다는 정답에 도달하는 과정에 중점을 두고, 학생의 문제 해결 능력, 종합적이고 고차적인 사고 능력 및 창의력을 평가하는 데 있다. 이런 서술형 평가를 올바르게 활용하기 위해서는 서술형 문항의 개념을 이해하고, 문항 유형에 대하여 알 필요가 있다.

문항 유형은 분류 기준에 따라 다양하게 분류될 수 있는데 일반적으로 사용되는 문항 유형은 응답자의 반응 형태에 따른 선택형(selection type)과 서답형(supply type)이 있다. 선택형은 응답에 필요한 정보를 제공하고 그중에서 정답이라고 판단되는 것을 선택하게 하는 문항 형태로 진위형, 배합형(연결형), 선다형이 있다. 서답형은 응답자가 직접 답이라고 생각하는 지식이나 의견 등을 구성하여 작성하도록 하는 문항 형태로 단답형, 완성형, 서술형이 있다. 서술형 문항 유형은 응답자가 스스로가 답안 작성하도록 하는 것으로 단답형이나 완성형보다는 비교적 긴 답안을 요구한다. 한 문장부터 여러 문장이나 페이지까지 글로 피험자가 구성한 반응을 요구하고, 채점자가 이 반응을 읽고 답의 정확성과 질을 주관적으로 판단한다(박도순 외, 2007). 서술형 문항은 학생의 생각이나 의견을 직접 서술하도록 하기 때문에 학생의 창의성, 문제 해결력, 비판력, 판단력, 통합력, 정보 수집력 및 분석력 등 고등 사고 능력을 쉽게 평가할 수 있다. 최근의 문항 유형에 대한 연구는 문항 유형을 상호배타적으로 독립된 실체로 보기보다는 각 유형들은 다른 특성을 갖고 있지만 유사성도 갖고 있으므로 연속선(continuum)으로 이해하려 한다. Johanson & Motlomelo(1998)는 문항 유형을 대화-구술형, 수행형, 프로젝트형, 에세이, 단답형, 선다형, 진위형의 8가지로 나누고 반응의 자유도와 응답의 제한 정도에 따라 연속선상에서의 각 문항 유형의 위치를 제시한 바 있다.

그러나 서술형 문항이 논술형 문항과 혼용되어 사용되고, 특히 대학입시에서의 논술고사의 문항들을 논술형 문항이라고 지칭하면서 많은 교육 관련자들이 서술형 문항은 교과 내용요소적인 지식을 묻는 문항이 아니라 논리적이고 창의적 또는 비판적인 사고 능력을 중심으로 표현 방식의 논리성과 체계성을 재는 검사로만 인지되는 경향이 있다. 서술형

과 논술형은 혼용되어 사용되기도 하고, 논술형이 서술형의 하위 유형 문항 형태로 사용되기도 하고, 논술형 중에서 정답이 있고, 문장으로 서술하게 하는 방식을 지칭하기도 한다(서울시 교육청, 2010; 경기도 교육청, 2010). 서술형 문항은 사용하는 사람에 따라 다른 개념으로 사용되어 용어상의 혼란이 있고, 학교 현장에서는 임의로 해석하여 서답형 문항과 혼용하여 사용되기도 한다. 일부 교과영역에서는 서술형 문항을 학생이 서술해야 하는 분량이 상대적으로 많지 않고, 또 채점을 할 때 서술된 내용의 깊이와 넓이에 관심이 있고, 논술형 문항은 학생이 서술해야 할 분량이 상대적으로 많고 또 채점을 할 때 서술된 내용이 깊이와 넓이뿐만 아니라 글을 조직하고 구성하는 표현 능력이나 논리적인 일관성 등에도 관심이 있는 경우로 구분하는 경향도 있다. 따라서 서술형 평가의 추진 목적과 학교 현장의 편의성을 고려하여 서술형 평가에서 활용되는 서술형 문항에 대한 조작적 정의가 필요할 것이다.

서술형 문항의 조작적 정의를 내리기 위해 먼저 서술형 평가의 특징을 살펴보면 다음과 같다.

① 서술형 문항은 고등 사고 능력을 측정하기 위한 것으로 가능한 한 분석, 활용, 추론, 문제 해결 등 상위 수준의 행동 요소를 측정하여야 한다.

서술형 문항의 인지적 행동 영역은 교과별로 차이가 있으며, 학생들이 학습 내용을 배우고 나서 학생들이 갖게 되기를 원하는 능력 또는 심리적 속성이다. 피험자가 서술형 문항을 해결할 때는, 답에 이르는 과정을 문장으로 설득력 있게 제시해야 하며, 과목에 따라서는 문장으로뿐만 아니라 수식이나 그림 또는 도표를 이용하여 명확하게 이해하도록 표현해야 한다. 이러한 훈련을 거듭함으로써 학생은 문제 해결 능력을 함양하고 창의력을 계발할 수 있을 것이다.

② 서술형 문항은 ‘응답의 다양성’을 전제로 하고 있다.

서술형 문항은 본질적으로 채점자의 ‘주관성’을 내포하고, 답안에는 본질적으로 ‘다양성’을 내재하고 있다. 서술형 평가에서는 선택형 평가 문항과 같은 수준의 객관성을 요구할 수는 없다. 다만 평가 문항의 완성도를 높이고, 채점 기준의 전문성과 객관성을 높여 평가 결과를 학생과 교사 모두가 수용할 수 있도록 만드는 것이다.

③ 서술형 문항은 서답형 문항의 한 종류로 단답형, 논술형과는 다음과 같은 점에서 구분된다.

단답형(완성형)은 추측 요인을 통제하기 위해 선택형 문항의 답지를 제공하지 않은 유형의 문항으로 한두 단어나 구로 답을 하며, 논술형은 응답자의 관점과 생각을 표현하게 해 진술된 내용의 진위 여부가 아니라 사고의 깊이와 적절성, 표현방식의 논리성과 일관성 등에 비중을 둔다는 특징이 있다.

④ 서술형 문항은 피험자가 구성한 반응이 문장 이상일 것을 요구하고, 문항의 완성도는 문항의 ‘답안 작성의 방향과 방법 제시의 명료성’에 달려 있다.

평가 요소가 불명확하게 제시되면 학생은 어떻게 응답해야 할지 혼란을 겪게 돼 학생들의 반응 폭이 확장되고 답안 내용도 다양해져서 채점 과정에서 어려움을 겪을 가능성

이 커진다. 따라서 문항에서 요구하는 답안의 내용 영역을 명확히 규정하고 제한해야 하며, 이에 따라 문항 구성 요소들을 구조화시킬 필요가 있다. 반응의 자유도를 일정 부분 허용하되 답의 표현은 ‘이곳으로’라는 방향, 즉, “~ 과 관련하여 설명하십시오”, “~을 이용하여 작성하십시오.”, “~개념을 적용하여 설명하십시오.” 등으로 내용 범위를 한정하여 지시해 주는 문항이 되어야 한다. 문항의 발문에 구체적으로 무엇무엇에 대해서 어떻게 서술해야 하는지를 분명히 제시하는 것이 좋다. 그러나 평가 요소의 범위가 막연하게 제시되거나 과도한 분량의 내용 서술을 요구하는 경우 평가의 신뢰도가 저하될 우려가 있다. 평가 요소는 명확하게 제시되어야 하고, 학생들이 작성할 답안의 내용 범위, 분량 등은 적절하게 제시되어야 한다. 한편 학생들의 문항 접근도를 높이기 위해서는 그림, 지도, 그래프 등을 적절하게 사용하는 것도 바람직하다. 학생들은 제시된 평가 요소에 대해 개인적인 경험과 지식을 인출하고, 문제 장면에 독특한 의미를 부여하여 문제를 이해하고 답안을 작성한다. 평가를 하는 교사들은 이런 점 때문에 현실적으로 곤란을 많이 겪는다. 이런 문제를 해결하기 위해서는 우선 문항의 평가 요소와 발문이 명확한가부터 점검해야 한다. 평가 요소나 발문이 다른 관점에서 볼 때 다른 방향으로 해석될 여지는 없는가, 평가 요소나 발문에 쓰인 단어가 다의적이지는 않은가, 단어와 단어의 연결에서 미묘한 의미 차이를 드러내고 있지 않은가 등을 판단해야 한다(서울시 교육청, 2010).

이런 서술형 평가의 추진 목적과 문항 특성을 고려하고, 학교 현장에서 큰 혼란 없이 사용될 수 있는 서술형 문항에 대한 조작적 정의를 하면 다음과 같다.

서술형 문항은 고등 사고 능력을 평가하기 위한 것으로 한두 단어나 구로 간단히 답할 수 있는 완성형, 단답형을 제외하되, 학생이 서술해야 하는 분량이 상대적으로 많지 않고, 또 채점을 할 때 서술된 내용의 깊이와 넓이에 관심이 있고, 문항에 대한 해결책(다양한 인정 답안이 있을 수 있는)을 문장으로 서술하게 하는 방식의 문항을 의미한다.

나. 기존의 선택형 평가와 서술형 평가와의 비교

서술형 평가는 선택형 평가에 비해 수험생의 성취수준을 정확하게 측정할 수 있는 평가 유형이다. 기존의 선택형 평가와 서술형 평가의 특징을 비교하면 <표 II-1>과 같다.

<표 II-1> 선택형 평가와 서술형 평가의 비교

구 분	선택형 평가	서술형 평가
측정하는 학습결과 의 목표분류	•목표 분류상으로 보아 지식, 이해력, 적용 력 및 분석력 수준에서는 적당하지만, 종합 력이나 평가력과 같은 학습 결과를 측정하 는 데에는 부적당하다.	•지식이나 이해력을 측정하는데는 비효과적 이지만, 종합력이나 평가력 수준의 학습결 과를 측정하는 데에는 적합하다.
내용표집	•문항 수가 많기 때문에 다루게 되는 범위 가 넓으며, 따라서 교과내용을 골고루 잘 대표할 수 있게 표집할 수 있다. 따라서 내 용타당도가 높다.	•문항 수가 비교적 적기 때문에 다루게 되 는 범위가 좁아지며, 따라서 교과내용의 대 표적 표집이 어렵다. 따라서 내용타당도가 낮다.
문항제작 과 출제 소요시간	•훌륭한 문항을 제작하기가 어려울 뿐만아 니라 시간도 많이 걸린다.	•훌륭한 문항을 제작하기가 어렵기는 하지 만, 선택형만큼 어렵지는 않다. 출제 소요시 간도 적다.
채점과 소요시간	•채점이 객관적이며, 간단하고, 신뢰도가 높 다. 채점시 소요시간은 적게 걸린다.	•채점이 주관적이며, 어렵고, 신뢰도가 낮다. 채점시 소요시간은 많이 걸린다.
반응의 자유도	•좁음	•넓음
필요한 경우	•평가받을 학생 수가 많거나 동일한 평가 문항을 다시 사용할 가능성이 있을 때 •보다 신뢰성 있는 평가문항을 원할 때 •평가의 공정성과 객관성 및 다른 불필요 한 요인이 점수에 영향을 미칠 가능성을 가 능한 한 제거하고자 할 때 •평가문항보다도 평가실시 후에 그 결과를 시급히 보고해야 하는 경우 등	•평가대상 학생 수가 적고 또한 같은 평가 문항을 다시 사용할 필요가 없을 때 •학생들의 문장구성력이나 표현능력을 아울 려 평가해 보고자 할 때 •학업성취도보다는 학생들의 태도나 의견에 더 많은 관심이 있을 때 •채점자가 답안지를 비판적으로 채점하는 데에 더 자신이 있다고 생각되는 경우
장점	•채점의 객관성이 유지됨으로써 검사의 신 뢰도를 높여준다. •선택형에서는 비교적 많은 수의 문제를 출 제할 수 있기 때문에 출제내용에 있어 포괄 성을 유지할 수 있다. •각각의 문항들에 대한 정답과 오답이 명백 하고 정답과 오답에 대한 분포를 얻을 수 있기 때문에 문항분석을 통하여 문항의양호 도를 검증할 수 있을 뿐만아니라 각종통계 적 분석이 용이하다.	•반응의 자유도가 크다. •고등 정신기능을 측정할 수 있다. •전체성을 강조한다. •문항작성이 용이하다. •인성 진단자료를 얻는다. •학생들의 학습 태도 개선에 효과적이다

단점	<ul style="list-style-type: none"> •단순한 지식의 기억력 측정에 치우칠 가능성이 크다. •주어진 일정 수의 답지에서 정답을 고르기 때문에 추측에 의한 정답의 가능성을 배제하기 어렵다. •피험자로 하여금 나름대로의 창의적인 답을 할 수 없도록 되어 있으므로, 자기표현과 창의력을 발현할 기회가 봉쇄된다. 	<ul style="list-style-type: none"> •채점의 신뢰도와 객관도가 부족하다. 즉, 채점하는 사람에 따라 채점결과가 다를 뿐만 아니라, 같은 사람이라도 채점결과가 다를 수 있다. •대표적인 표집이 부족하다. 즉, 서술형 평가에서는 아무래도 문제를 많이 출제할 수 없기 때문에 검사의 포괄성을 만족시켜 주지 못 한다. •채점의 비능률성으로 채점하는데 시간과 노력의 낭비가 많다. •검사문항을 잘못 제작하면 나온 답을 비교할 수 있는 근거를 마련하기 힘든 경우가 많다.
유형	<ul style="list-style-type: none"> •진위형 : 진위형, 수정형, 군집형 •배합형 : 단순배합형, 복합배합형, 분류배합형, 관계분석형, 공변관계형 •선다형 : 최선답형, 정답형, 다답형, 불완전문장형, 부정형, 불완전답지형, 합답형 	<ul style="list-style-type: none"> •단답형 : 질문형, 진술형 •완성형 : 자유완성형, 제한완성형 •논문형 : 제한반응형, 확대반응형

다. 서술형 평가의 설계

서술형 평가는 문제해결 과정에 대하여 학생 스스로 깊이 사고하도록 하여 학생 자신의 발전된 사고를 논리적으로 기술하도록 하는 평가로 다음과 같은 특성을 가진다. 첫째, 학생에게 스스로 사고하고, 자신의 아이디어를 표현할 기회를 제공한다. 둘째, 학생에게 단 하나의 답을 고르는 대신에 학생 자신의 문제해결을 논리적으로 기술할 것을 요구한다. 셋째, 다지 선다형으로는 거의 불가능한 문제해결의 이해를 깊이 설명할 수 있게 한다. 넷째, 학생이 여러 가지 방법으로 문제를 해결하는 것을 격려한다.

이러한 서술형 평가의 장점을 극대화하기 위하여 다음과 같은 설계 원리를 따라야한다.

① 평가하고자 하는 학생 집단의 특성을 고려하여야 한다.

학생 집단의 특성은 서술형 평가에서 고려되어야 할 중요한 요소이다. 선다형 문항과 비교하여 서술형 문항은 반응의 자유도를 최대한으로 허용하고 있기 때문에 학생이 어떤 특성을 가진 집단인지를 분명히 파악하지 않으면 반응이 지나치게 다양해질 가능성이 있다. 학생 집단에 따라 그 집단을 평가하는 데 좋은 문항이 될 수도 있고 그렇지 않을 수도 있다는 점에 유의할 필요가 있다. 학생 집단의 특성을 고려하는 데 있어서 특히 주의하여야 할 사항은 다음과 같다.

먼저, 문항 내용이나 지시문 등의 어휘 수준이 학생의 수준에 적절하여야 한다. 이는 문항의 내용이나 지시문 내용이 너무 어려운 수준으로 제시될 경우 학생이 질문의 요지를 파악할 수 없어 문제를 해결하는 과정을 기술할 기회를 잃어버릴 가능성도 있기 때문이다. 특히, 지시문의 내용이 학생에게 어렵게 인식될 경우에는 독해력을 요구하게 되므로 평가하고자 하는 수학 내용에 대한 학생의 인지 능력 추정이 부정확해질 수 있다.

다음으로, 문항 수를 학생의 수준에 적절하게 조절함으로써 각 문항에 충실히 답할 수 있도록 응답 시간을 배려하여야 한다. 서술형 평가에서는 다수의 문항을 출제하기는 어렵지만, 학생의 학습 내용을 고르게 측정할 수 없을 정도의 한 문항이나 두 문항 정도로 한정해서도 곤란하다. 그러므로 제한된 시간 내에 학생이 학습한 내용을 고르게 평가할 수 있도록 계획하여 출제할 필요가 있다. 또한 학생에 따라 답하는 시간이 다를 수 있기 때문에 문제 해결 배당 시간을 고려할 때 문제 풀이 과정을 기술하는 속도가 느린 학생도 배려할 필요가 있다. 특히 선다형 문항과 서술형 문항을 혼용하여 검사지를 구성하는 경우, 서술형 평가 문항에 대한 배당 시간을 충분히 고려함으로써 학생의 시험 불안 및 혼란을 경감시킬 수 있도록 하여야 할 것이다.

② 단순 암기 위주의 지식보다는 고등 사고 능력을 측정하도록 한다.

서술형 평가 문항의 가장 큰 장점은 선다형 평가 문항으로 측정하는 데 한계가 있는 사고력, 추리력, 종합력, 비판력, 분석력, 응용력, 표현력, 창의력과 같은 사고 능력을 상대적으로 더 쉽게 측정할 수 있다는 데 있다. 이러한 장점을 최대한 살릴 수 있도록 문항을 구성하여야 하고, 선다형 문항으로 측정할 수 있는 기능을 서술형 평가 문항을 사용하여 측정할 필요는 없을 것이다.

③ 학습 결과를 측정할 수 있도록 질문을 구체화한다.

서술형 평가 문항의 경우, 선다형 평가 문항과 비교하여 문두의 길이가 길고 복잡한 편이다. 서술형 평가 문항에 대하여 학생이 어떻게 응답해야 할지 혼란을 겪고 다양한 반응을 하게 되고, 경우에 따라서는 학생이 평가 의도를 파악하는 데 어려움을 겪어 실질적인 평가가 되지 못할 수도 있다. 그러므로 문항이 요구하는 영역을 규정하고 제한하여 잘 구체화함으로써, 학생 반응의 자유도를 최대한으로 허용하면서 답의 표현에 대한 방향을 지시해 주는 문항이 되어야 한다.

④ 채점 기준과 문항 점수를 미리 제시한다.

채점 기준을 미리 제시하면 학생이 어떤 내용을 평가하는지를 분명히 알게 됨으로써 분명한 방향성을 지니고 답안을 작성하는 데 도움을 줄 뿐만 아니라 준다. 또한 해당 문항의 점수에 대한 정보가 제공되면 학생이 문항의 점수를 고려하여 문항에 응답하는 전략을 세울 수 있다. 채점 기준과 문항 점수를 미리 제시하기 위해서는 먼저 서술형 평가 문항을 출제하는 과정에서 문항 개발과 함께 모범 답안 및 예시 답안을 작성할 필요가 있다. 이러한 모범 및 예시 답안 작성을 통하여 문항의 오류 가능성을 배제할 수 있을 뿐만 아니라 채점 기준을 명확하게 할 수 있다.

⑤ 쉬운 문항부터 어려운 문항의 순으로 배열한다.

서술형 평가 문항은 학생 자신의 사고 과정을 드러내야 하므로 배점이 크고 문항의 난이도가 선다형에 비해 높아 평가를 받는 학생에게 부담을 주게 된다. 첫 문항부터 지나치게 어렵거나 시간이 오래 걸린다면 학생이 좌절하게 되거나 자신의 능력을 제대로 표현할 수 없게 될 수 있다. 따라서 쉬운 문항에서 어려운 문항의 순서로 배열하는 것이 학생의 동기를 유발하는 데 도움이 될 뿐만 아니라 학생의 적극적인 반응을 어느 정도까지 유도할 수 있고, 모든 학생의 수준을 다소 수월하게 파악할 수 있다.

⑥ 문항 간의 점수 차이가 크지 않도록 설정한다.

문항 간의 점수 비중에서 차이가 크게 되면 특정 내용이 전체 점수에 영향을 주게 되는 가능성을 배제할 수 없다. 특히, '사고력과 표현력'을 측정하는 문항을 출제할 경우 문항 간의 비중을 크게 달리하는 것은 합리화되기도 어렵다.

⑦ 여러 개의 문항 중에서 선택하여 응답하도록 하지 않는다.

3개 또는 5개의 서술형 평가 문항을 제시하고 학생으로 하여금 1~2문항을 선택하여 문제를 해결하게 하는 방식은 바람직하지 않다. 이러한 문항 구성 방식은 학생의 부담을 덜어 준다는 입장에서 긍정적으로 보일 수도 있으나, 학생의 입장뿐만 아니라 출제와 채점 과정의 측면에서도 몇 가지 문제가 발생할 수 있다. 먼저, 제시되는 문항의 내용이 다를 수 있기 때문에 학습내용 측정에 대한 타당도가 떨어지게 된다. 각기 다른 문항을 선택해서 답을 하였을 경우 평가하는 학생의 능력이 서로 다르기 때문이다. 둘째, 채점에 소요되는 노력과 시간이 증가한다. 제시하는 모든 문항에 대한 채점 기준과 모범 답안 작성, 유사 답안 추가 등이 필요하므로 문항 개발에 필요로 하는 노력과 시간이 더 증가하게 된다. 셋째, 학생의 입장에서 여러 개의 문항 중에서 선택하여 문제를 해결하는 것에 익숙해지면 원하는 내용만을 학습하는 습관을 가질 수 있기 때문에 학습 방법 및 태도적인 측면에서 바람직하지 않다.

라. 서술형 평가의 문항 유형

서술형 문항의 유형 분류는 분류 기준에 따라 다르고, 각 문항 유형의 이름도 다양하게 불리고 있다. 일반적으로 서술형 문항은 학생의 반응 허용 정도가 어떠한가에 따라 응답 제한형과 응답 자유형(확대 반응형)으로 구분되기도 하고, 문항에 답안 작성과 관련한 자료나 정보를 포함시켜 제시하는가의 여부에 따라 단독 과제형과 자료 제시형으로 구분되기도 한다.

1)

2) 1) 반응의 허용 정도에 따른 분류


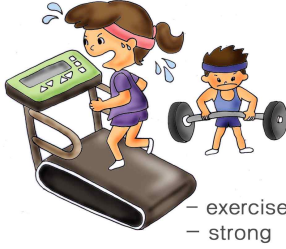

가) 응답 제한형

학생들의 반응 허용 정도가 상당히 제한된 것으로, 반응의 내용과 형식을 구체적으로 지시하여 제한된 방향으로 답을 구성하도록 하는 문항 형태이다. ‘지식’ 수준을 측정하는 것이 아니라 ‘이해’ 수준 이상의 행동을 측정하고, 그 답은 학생의 말로 표현되어야 한다.

(1) 응답 제한형의 주요 특징

- ① 문장 표현력을 제외한 학생 성취도를 측정하는 데 주로 사용한다.
- ② 일반적이고 포괄적인 주제를 주고 학생의 의견을 쓰게 하는 것이 아니라, 구체적인 문제를 주고 학생이 적당한 정보를 재생하고 적절한 방법으로 조직하여 합당한 결론에 도달하도록 한다.
- ③ 단답형 문항보다 학생 반응의 자유도가 덜 제한되며, 응답 자유형 문항에서보다는 반응의 자유도가 더 제한된다.

메신저로 대화하고 있는 친구에게 운동을 추천하려고 합니다. 다음 그림에 제시된 운동 중 한 가지를 선택하고, 주어진 단어나 어구를 이용하여 친구에게 보내는 메시지를 작성하시오. (2-3 문장)

(1)	(2)	(3)
 <p>– walking – healthy</p>	 <p>– exercise – strong</p>	 <p>– climbing – fresh air</p>

[그림 II-1] 응답 제한형의 예시

(2) 응답 제한 방식에 따른 세부 유형

응답 제한형은 응답을 제한하는 방식에 따라 분량 제한형, 내용 범위 제한형, 서술 양식 제한형으로 분류된다. 하지만, 분량과 내용 범위 혹은 서술 양식 등을 조합하여 응답을 제한하는 문항의 사례도 많다.

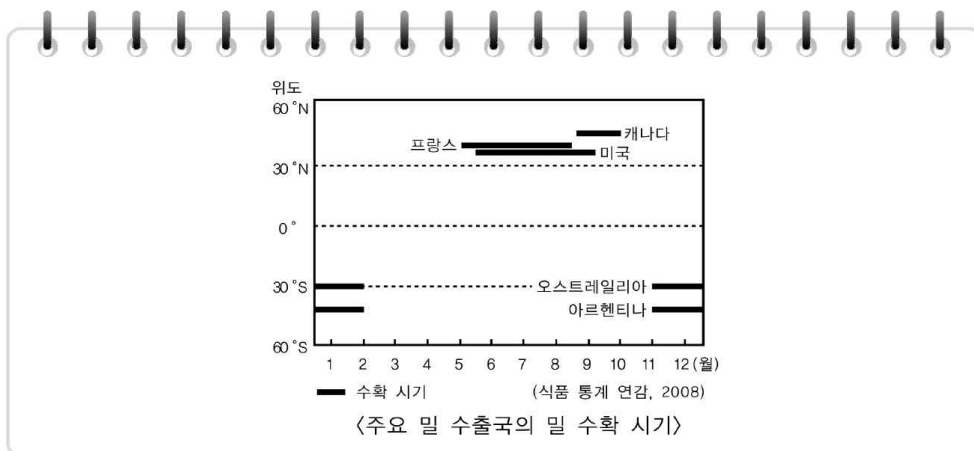
- ① 분량 제한형: 진술 요소의 수, 답안의 길이(문장 수, 글자 수, 또는 행 수) 등에서 물리적 제한이 가해져 있는 문항 유형이다.

자신이 생활신조로 삼고 있는 좌우명을 2가지 쓰고, 그 이유를 10자 이내로 설명하시오. (단, 좌우명은 구체적이고 일상생활 속에서 실천 가능한 것이어야 한다)

[그림 II-2] 분량 제한형의 예시

- ② 내용 범위 제한형: 응답할 내용 범위에 제한을 가하는 문항 유형

[서술형 3] 자료는 주요 밀 수출국의 밀 수확 시기를 나타낸 것이다. 물음에 답하시오. [8점]



- 3-1. 오스트레일리아의 밀 수확 시기가 미국의 밀 수확 시기와 다른 이유를 위도 상의 위치와 연관지어 서술하시오. [4점]

- 3-2. 오스트레일리아가 미국과 프랑스에 비해 밀의 수출에서 유리한 점을 서술하시오. [4점]

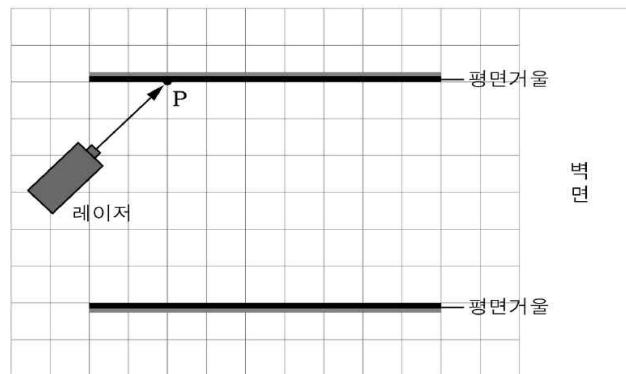
[그림 II-3] 내용 범위 제한형의 예시

③ 서술 양식 제한형: 응답할 서술 양식에 제한이 가해진 문항 유형



문항예시 9

그림과 같이 수평면에 세워져 있는 평면거울의 P점을 향해 레이저를 비추었다.



이 레이저가 P점에서부터 벽면까지 진행한 경로를 그리시오. [7점]

[그림 II-4] 서술양식 제한형의 예시

나) 응답 자유형(확대 반응형)

응답의 내용 범위와 서술 양식에 큰 제한이 가해지지 않고 답안의 길이에 대한 제한도 명시되어 있지 않는 것으로, 답안의 최대 길이가 학생 개개인의 능력 수준과 시험 시간에 의해서만 간접적으로 제한을 받는 문항 형태이다.

(1) 응답 자유형의 주요 특징

- ① 학생들이 조직화된 방법으로 자신의 생각을 ‘응답 제한형’보다 자유롭게 표현하도록 요구한다.
- ② 학생에게 있어 응답 제한형보다 더 도전적인 문항 유형으로서, 중요한 교육 목표의 성취 정도를 보다 직접적으로 파악하는 데 도움이 되고 학생의 개인적 특성을 나타내게 하는 데 유용하게 활용될 수 있다.

요즘 청소년 문화에서 새로이 부상하고 있는 특징 중의 하나는 연예인을 추종하며, 옷이나 좋아하는 물건, 사회에 대한 관점 등 생활의 거의 모든 측면에서 그 연예인을 모방하는 것이다. 이 현상에 대하여 자신의 생각을 쓰시오.

[그림 II-5] 응답 자유형의 예시

이런 응답 자유형 문항은 얼마나 다양하고, 깊이 있게, 그리고 논리적으로 글을 쓰는가에 대한 일반 쓰기 능력 평가에는 적합할 수 있으나 학교의 평가 상황에서 채점의 공정성, 객관성, 용이성을 고려할 때 적용하기 쉽지 않은 형태이다. 학생들이 어떤 방향으로 글을 써야 할지 명료히 해주는 것이 학교의 평가 상황에서는 더 적절할 것이다. 위의 예를 좀 더 구조화하여 응답 제한형으로 변형할 수 있다.

요즘 청소년 문화에서 새로이 부상하고 있는 특징 중의 하나는 연예인을 추종하며, 옷이나 좋아하는 물건, 사회에 대한 관점 등 생활의 거의 모든 측면에서 그 연예인을 모방하는 것이다. 이 현상이 지니는 긍정적, 부정적 측면을 3가지씩 기술하고, 부정적 측면을 개선하기 위한 해결 방안을 3가지 제시하시오.

[그림 II-6] 응답 자유형에서 응답 제한형으로 변형한 예시

(2) 자료 제시 여부에 따른 분류

① 단독 과제형

자료나 정보를 제시하지 않고 특정 질문에 응답하도록 하는 문항 형태이다. 흔히 논술형이라고 할 때 가장 널리 쓰이는 것으로, “응답 자유형”과 밀접히 관련이 있다.

남녀 간의 임금 격차와 업종에 따른 임금 격차의 관련성을 논하시오.

[그림 II-7] 단독 과제형의 예시

② 자료 제시형

문항 속에 학생이 읽을 자료를 제시해 주고 그것을 바탕으로 해서 응답하도록 하는 문항 형태이다. 보통 응답 자유형보다는 “응답 제한형”과 밀접히 관련되어 있다.

[자료 제시형 예시]

37. <보기>를 참고하여 밑줄 친 부분에 나타난 소재의 상징적 의미를 정리해 본 것이다. ㉠에 들어갈 상징적 의미를 서술하시오.

<보 기>

자료1.

이 작품은 15년째 같은 강의 노트로 강의하는 ‘학자’를 조롱하는 ‘도적’과의 대화가 희화적으로 묘사된다.

자료2.

‘무대’에 대한 설명에는 학자 주변의 사물들에 대한 언급이 있다. 이 사물들은 학자의 사회적 지위, 학문적 태도, 생활상 등을 상징적으로 드러내는 장치들이다.

사물	상징적 의미
책상은 비교적 크지만 퇴색함	사회적으로 존경받는 위치이지만 실제로는 그에 걸맞지 않음을 상징한다.
책상에 낡은 책으로 절반 밖에 채워지지 않음	학자는 [㉠]

[그림 II-8] 자료 제시형의 예시

마. 서술형 평가의 문항 제작 절차

1) 서술형 평가의 제작

학교현장에서 서술형 평가(오래된 미래 평가, old future assessment)를 시행하기 위해서 교사들은 학기 초에 계획된 교과 진도와 시험출제 범위 그리고 서술형 평가의 비율을 기준으로 서술형 문항을 만들어 간다. 교사들은 서술형 평가의 근본적인 문제, 즉 왜 서술형 평가를 해야 하는가? 서술형 평가의 목적은 무엇인가? 교사는 서술형시험에서 무엇을 묻고자 하는가? 학생들은 이 서술형시험을 통해서 무엇을 획득할 수 있는가? 교사는 학생의 어떤 능력을 끌어내기 위해서 어떤 서술형문항을 만들려고 하는가? 학생은 이

서술형문항에서 무엇을 어떻게 반응할 것인가? 서술형 평가의 목적과 목표에 이르기 위해서 교사는 수업과 관련하여 문제 장면을 어떻게 선정하고 어떤 문항제작절차를 따라야 하는가? 등 수업과 평가의 목적이나 목표에 대해서는 질문하지 않는다. 일선 학교현장에서 교사들이 가장 많은 시간과 노력을 들여야 할 평가는 내용 없이 형식적으로 ‘평가를 위한 평가’로 진행되고 있다.

서술형 평가문항의 제작은 수업의 목적과 목표 그리고 평가의 목적과 목표를 확인하고 그 전제위에서 이루어져야 한다. 목표 없는 목적은 공허하며 목소리만 크고 실속이 없고, 목적 없는 목표는 분주하기만 할뿐 어디로 가야할지 몰라 헤매는 것과 같다. 평가의 목적은 수업의 목적과 같아야 하고 수업의 목표는 평가의 목표와 같아야 할 것이다.

<표 II-2> 수업과 평가의 상관관계

수업 \ 평가	목적	목표
	일치	불일치
목적		
목표		

평가와 수업의 목표가 일치하는데 목적이 일치하지 않아서 목적 불일치가 발생하는 경우에 수행형(구술) 및 서술형 평가는 타당성을 상실하게 된다. 반면에 평가와 수업이 목적은 일치하는데 목표가 일치하지 않아서 목표 불일치가 일어나는 경우에는 서술형 평가는 신뢰성을 상실하게 된다.

2) 서술형 평가문항의 제작절차

타당도와 신뢰도를 갖춘 서술형 평가문항을 제작하려면 그 절차는 어떠한가?

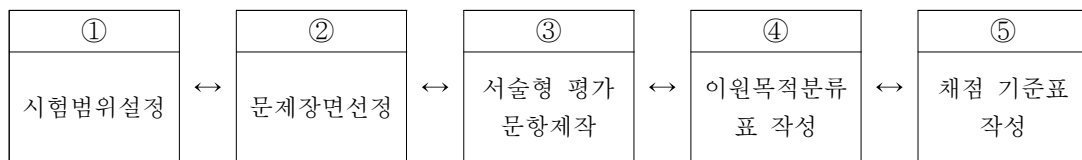
가) 일상적 제작절차

교육현장에서는 일반적인 학교상황에서 서술형 평가문항 제작절차는 대체로 다음 순서로 진행된다.

- ① 수업진도 계획에 따라 주어진 교과서 진도를 시험범위로 설정한다. 이때 출제교사가 2명 이상이면 교사 수에 따라서 문항과 배점을 평균하여 출제한다.
- ② 시험범위 중 교사 개인이 가르친 것 중에서 중요하다고 생각하는 것을 문제 장면으로 상정한다. 이때 여러 명의 교사가 출제를 하면 문항이 중복되는 것을 피하기 위해서 시험범위를 여러 영역을 쪼개어 균등하게 나눈다. 평가문항의 유형에 대해 특별히 규정하지 않으면 같은 유형이 많이 출제되기도 한다. 예를 들면 영어 서술형 평가의 경우, 출제하기 용이하고 채점도 쉽고 평가문항에 대한 이의제기도 발생하지 않는 어법과 빈칸 완성하기가 대표적이다. 따라서 서술형 평가의 타당도를 높이

기 위해서는 여러 명의 교사들이 나누어 출제하는 경우 서술형 출제문항의 유형에 대한 협의가 반드시 필요하다.

- ③ 나눈 교과서 영역을 가지고 기존의 문제지의 질문형식에 따라서 서술형 평가문항을 작성한다. 대부분의 교사들이 전국연합 모의고사나 대학수학능력시험의 질문형식 또는 교과서 출판사의 서술형 평가문제지의 질문형식에 따라서 교과서의 지문을 가지고 서술형 평가 문항을 제작한다. 예를 들면 다음과 같은 수행형(구술) 및 서술형 평가문항의 지시어를 사용하여 문항을 작성한다. 찾아라, 고쳐라, 빈칸을 채워라, 설명하라, 해석하라, 재조직하라, 논하라, 예시하라, 정의를 내려라, 비교하라, 제시하라, 개요를 써라, 요약하라, 분류하라, 항목을 들어라, 서술하라, 종합하라, 관련시켜 설명하라, 분석하라, 원인과 결과를 들어라, 평가하라, 완성하라, 등이다(양길석, 2006). 따라서 서술형 평가문항의 내용영역과 행동영역에 대한 논의나 협의는 없다.
- ④ 이원목적분류표의 작성은 평가문항이 제작된 후에 이루어진다. 따라서 내용영역, 성취기준, 행동영역, 난이도, 그리고 배점은 마지막에 형식적으로 이루어진다. 따라서 서술형 평가가 실시된 후에 평가결과가 학생들의 학습을 위한 정보제공자로서의 역할이나 수업의 개선을 위한 자료로서 환류되지 못하는 것은 너무도 당연하다.
- ⑤ 서술형 평가문항을 제작한 후에 채점 기준표가 작성된다.



[그림 II-9] 일상적 수행평가 및 서술형 평가문항 제작절차

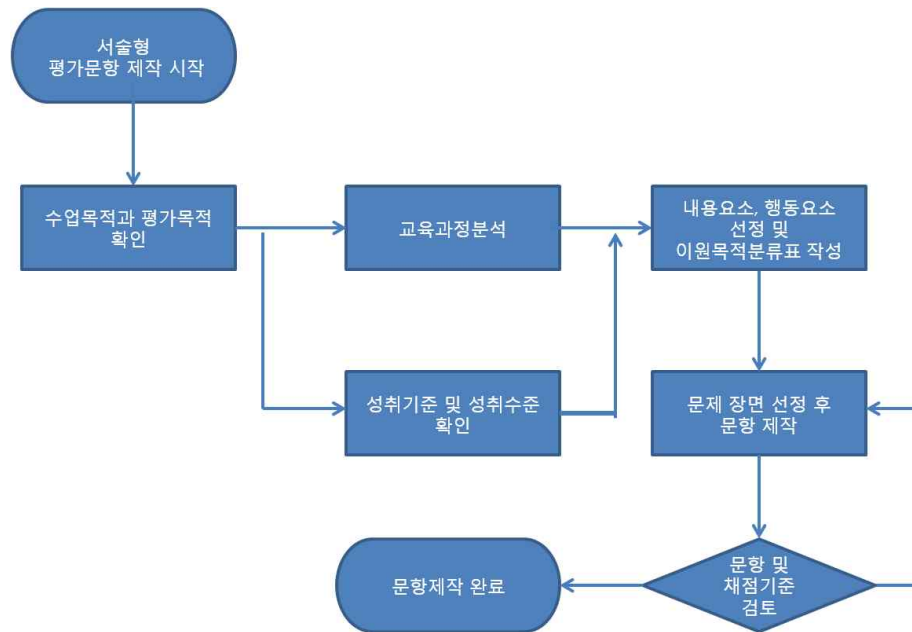
나) 이상적 제작절차

- ① 수업 및 평가 목적 확인: 일반적으로 서술형 평가의 목적은 교수학습의 개선과 학생능력의 향상이다.
- ② 교육과정 분석: 각 교과목의 중·소단위와 내용에 대한 확인과정이다. 각 교과목의 교육 목표와 내용을 분석한다.
- ③ 성취기준 확인과 성취수준 설정: 성취기준을 확인하고 성취기준에 따른 상·중·하 성취수준을 설정하여 문제의 난이도를 가늠하는 과정이다. 이원목적분류표를 작성하는 기초적인 과정이기도 하고 평가문항 출제과정에서 가장 중요한 과정중 하나이다. 교육과정에서 목표와 내용이 잘 조직되었다 하더라도 그 목표와 내용을 ‘평가’ 하기에 적합하도록 진술하는 것이다(김종수 외, 2011). 교육과정의 목표가 수업에서 학생들이 배워야 할 목표가 ‘학습목표’로 진술된다면 학생들이 도달해야 할 평가목표가 ‘성취기준’이며 그 정도가 ‘성취수준’이다. 따라서 양질의 평가문항을 만들기 위해서는 피험자가 도달해야 할 목표와 내용을 구체적으로 진술하여야 하며, 각각 교

육목표에 대하여 최소성취기준을 제시하는 것도 필요하다.

- ④ 핵심성취기준 선정: 성취기준 가운데에서 기본적인 중요한 내용영역과 행동영역에 대한 기준을 선정하는 과정이다. 국가성취기준에 제시된 모든 교육과정 내용은 평가요소가 된다. 그런데 국가성취기준에 제시된 모든 교육과정 내용이 양적으로 많을 뿐만 아니라 중요도에서 차이가 있고 다양하기 때문에 성취기준 가운데에서 일부만 평가요소로 선정할 필요가 있다. 이러한 이유로 평가문항을 제작하는 장면에서는 ‘핵심성취기준’은 최저기준으로서의 ‘핵심 교육목표’이고 학습자의 기본적인 능력을 평가하는 기준이 된다.
- ⑤ 이원목적분류표 작성: 이원목적분류표는 1956년 Bloom이 발표한 교육목표분류학에서 제시된 것으로 측정하고자 하는 인지적 영역과 학습내용의 상호관계를 각각 세로줄과 가로줄로 나타낸 표이다. 행동과 내용의 이원분류표를 작성할 때에는 흔히 행동목표를 종렬에 열거하고 내용은 횡렬에 열거한다(강신천, 2014). 일반적으로 학교에서 성취수준을 평가하기 위해서는 목표의 상세화 및 성취 수준의 결정이 선행되어야 한다. 학교 현장에서는 출제자 스스로 평가 목표의 진술, 평가 기준인 성취 수준을 결정해야 한다. 따라서 출제자는 학습내용을 중심으로 평가목표 및 평가 문항을 먼저 설계하고, 그러한 작업을 정리하기 위해서 이원분류표를 작성해야 할 것이다.
- ⑥ 문제 장면의 선정 후 문항 초안 작성: 초안 작성에 중요한 것은 문두를 작성하는 것인데 이는 앞서 이원목적분류표 작성에서 대부분이 이루어진다.
- ⑦ 문항 검토 및 문항 수정 보완: 동 교과목의 협의가 중요하며 교사들 간의 문항에 대한 평가관을 일치시키는 과정으로 제3자적 관점과 학생의 관점에서 이루어져야 한다. 문항을 수정하고 보완하는 과정에서 채점기준도 수정·보완된다. 이때 교사들 간의 상호작용과 논의의 깊이와 폭이 평가문항의 수준을 높이며 교사들의 평가전문성을 향상시키는데 상당한 기여를 한다.
- ⑧ 최종 문항 완성

이와 같은 절차를 순서도로 그려보면 다음과 같다.



[그림 II-10] 이상적 서술형 평가문항 제작절차

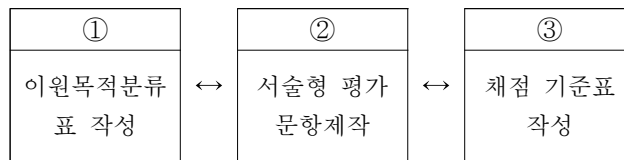
다) 핵심적 서술형문항 제작절차

평가자의 입장이 아닌 순수 문항 제작자로서의 교사입장에서 요구하는 핵심적 제작절차는 이원목적분류표의 내용을 얼마나 정교하게 분석하고 종합하는가와 각 문항에 대한 채점기준표를 완성도 높게 만들어내느냐에 달려있을 것이다. 즉 이원목적분류표를 작성하는 과정이 서술형문항을 제작하는 절차에서 가장 핵심적인 요소라고 할 수 있다.

- ① 이원목적분류표의 내용영역을 확인한다. 이는 교육과정 내용을 분석하고 평가할 지식을 확인하고 선정하는 과정이다. 내용영역은 중소단원의 중심내용이고 단원명이기도 하다. 예를 들어 실용영어1의 교육과정 내용은 “④-1. 주변의 인물, 사물, 사건에 대하여 간단히 쓴다.”이다. 여기서 평가할 내용은 주변의 인물, 사물, 사건이 된다.
- ② 이원목적분류표의 성취기준을 확인한다. 내용영역이 좀 더 구체화되어 평가자와 학습자에게 제시되는 것이 성취기준이다. 이때의 성취기준은 세 가지로 제시되어 있는데 그 가운데에서 하나를 선택한다. 예를 들면 “영고21441-3. 주변의 사건에 대하여 원인, 과정, 결과 등을 간단히 쓸 수 있다.”를 선택할 수 있다. 성취기준에는 평가 지식과 함께 인지적 처리영역인 행동영역이 함께 기술되어 있다. 이때 평가문항 제작자는 지식영역을 압축하여 평가할 내용을 하나로 결정해야 한다.
- ③ 이원목적분류표의 인지적 처리영역인 행동영역을 확인하고 행동영역의 하위요소(기억, 이해, 적용, 분석, 평가, 창의) 가운데에서 하나를 확정한다. 이제 평가문항의 문제유형이 결정된다.
- ④ 이원목적분류표의 난이도를 성취수준을 고려하여 상, 중, 하 가운데에서 결정한다. 사이 가장 어려운 문항이다.
- ⑤ 이원목적분류표의 배점을 결정한다. 배점은 일반적으로 난이도가 높은 문항이 배점

도 높다. 그러나 난이도와 함께 문항이 교육과정에서 차지하는 중요도를 고려하는 것이 필요하다. 핵심적이고 중요한 성취기준이면 쉬운 문항이라도 높은 배점을 해야 할 것이다.

- ⑥ 교과서 등 시험범위의 지문을 이용하여 문항을 출제하고 서술형 채점기준표를 작성한다. 채점기준표는 학교와 학과에 따라서 다양한 양식을 사용하고 있다. 채점기준표에 포함되어야 할 기본적인 내용은 문항번호, 예시답안 또는 모범답안, 채점기준, 그리고 배점은 필수 요소이며 그밖에 난이도, 평가영역, 문항내용 등을 포함하기도 한다.
- ⑦ 문항출제와 채점기준표 작성이 끝나면 평가문항을 검토하고 수정하여 완성하게 된다. 수행형(구술) 및 서술형 평가문항을 제작 후에 문항카드의 작성하기도 하나 앞서 작성한 채점기준표로 문항카드를 대신하는 것이 더욱 실용적이다. 일반적으로 문항카드에 기록되는 모든 요소들이 이원목적분류표와 채점기준표에 포함되기 때문이다.



[그림 II-11] 핵심적 서술형문항 제작절차

<표 II-3> 2015학년도 ()학기 ()학년 (서술형)고사 이원목적분류표

문항	내용영역	성취기준	행동영역			난이도			배점	정답
			기억	이해	적용	상	중	하		
선택1										
선택2										
선택3										
선택4										
선택5										
서술1										
서술2										
서술3										

<표 II-4> 2015학년도 ()학기 ()학년 (서술형)고사 채점기준표

문항 번호	평가과정	평가문항	예시답안 및 채점기준					배점	
1			예시 답안						
			채점 기준	성취수준 성취기준	상	중	하	배점합	
				내용영역					
				행동영역					
전체									
2			예시 답안						
			채점 기준	성취수준 성취기준	상	중	하	배점합	
				내용영역					
				행동영역					
전체									

<표 II-5> 서술형 문항카드 A양식

<div style="text-align: center;"> ○○년도 ○학기 ○○학년 서술형 문항카드 A양식 </div>						<div style="border: 1px solid black; padding: 2px;"> 문항번호 </div>	
과 목 명		출제일	년 월 일	출제자	(인)		
출제영역	내 용 영 역		행 동 영 역				
성취기준			성취수준		난이도		
<문항내용>							
<모범답안 및 채점기준>							
출제 근거	도 서 명		저 자	발행처	발행년도	쪽수	
문항분석	응시인원	난이도 (실제정답률)	변별도	답지반응률(%)			
	특이사항 및 총평						

<표 II-6> 서술형 문항카드 B양식

○○년도 ○학기 ○○학년

서술형 문항카드 B양식

과목명		제작일	년 월 일	출제자	㉠
-----	--	-----	-------	-----	---

□ 문항 번호. _____

출제 영역		평가목표	배점	난이도	예상 소요시간
내용 영역	행동 영역				

문항 내용

모범 답안(예시)

채점 기준

출제 근거 : 도서명, 저자, 발행처, 발행년도, 쪽수 등 표기

특이사항 및 총평 : 문항의 수정 및 검토 과정, 시험의 실시 과정 등에서 나타난 특이사항이나 시험 실시 후 학생의 반응 자료나 결과를 분석하여 해당 문항에 대한 총평을 기입함.

⇒ 향후 동일한 출제 영역에서 문항을 출제할 때 구체적인 도움을 받을 수 있고, 이 과정을 거치면서 자연스럽게 문항 제작의 전문성이 신장될 수 있음.

마. 서술형 평가의 문항 채점 절차

서술형 평가는 객관식 선다형 평가와 비교할 때 채점자의 주관적인 판단이 개입될 가능성이 높다. 일반적으로 주관적인 판단은 검사 결과의 신뢰도를 떨어뜨린다고 생각하기 쉬우나 이것은 근거가 약한 주장일 수 있다. 채점 전문가가 행하는 주관식 서술형 평가가 객관식 선다형 평가보다 상황맥락적인 평가가 가능하며 기계적 처리를 하는 선다형보다 타당도는 물론이고 신뢰도가 동시에 높을 수 있다. 다만 채점자들이 타당하고 신뢰로운 채점을 하기 위해서 자세하고 명확한 채점 기준을 정하고, 이것을 제대로 적용할 수 있는 전문적 평가소양을 개발하고 향상시키는 것이 선행되어야 할 것이다.

1) 채점기준

서술형 평가 문항의 채점기준은 서술형 평가문항의 유형과 종류에 따라서 다양하여 하나의 기준으로 일반하기 어려운 점은 있지만 핵심적인 채점기준은 다음과 같다.

- ① 채점기준은 피험자 집단의 능력을 고려하여 설정하여야 한다. 채점기준을 정할 때, 피험자 집단의 능력에 대한 고려와 한 집단 내에서도 여러 그룹에 대한 고려가 선행됨으로써 평가문항은 피험자의 능력을 더욱 정교하게 측정할 수 있고 피험자인 학생이 더욱 신뢰하는 평가문항을 제작할 수 있다. 서술형 평가문항은 피험자의 반응도를 폭넓게 허용하기 때문에 피험자 집단이나 그룹의 능력에 따라 다양한 반응이 나타나며 그 한계를 분명히 파악하지 않으면 채점기준이 지나치게 엄격하거나 지나치게 느슨해서 평가문항의 질이 떨어지고 평가문항의 신뢰성이 저하될 우려가 있다. 예를 들면, “다음 글을 읽고 3~4 단어이내로 제목을 쓰시오.” 라는 문항에서 하위능력의 피험자는 주로 사실적 표현을 중심으로 하는 제목을 사용할 것이고 상위능력의 피험자는 상상력을 사용하여 비유적 표현을 중심으로 하는 제목을 답안으로 작성할 것이다. 이럴 경우 능력집단에 따른 채점기준을 고려하지 않는다면 이 문항은 피험자의 능력을 변별하는 기능이 약화되어 문항의 질이 저하될 수 있다.
- ② 채점기준은 성취기준과 성취수준에 근거하여야 한다. 채점기준은 평가문항의 유형이나 종류에 따라 다양하기는 하지만 기본적으로 성취기준 및 성취수준, 즉 학습목표와 그 목표에 학습자가 도달한 정도에 따라서 설정되어야 한다. 성취기준과 성취수준은 선택형평가에서와 마찬가지로 서술형 평가에서도 핵심적인 준거이다. 채점기준이 이를 벗어나서 설정될 때 그 평가문항자체가 타당도를 상실하게 되고 피험자의 능력을 잘못 측정하는 문항이 된다. 예를 들면, 성취기준은 “영중9312-1. 일상생활이나 친숙한 일반적 주제에 관한 짧은 글을 듣고 따라 읽으면서 주요 어구의 의미를 파악할 수 있다.”인데 채점기준은 어구의 전체의 의미가 아닌 일부인 단어나 전치사의 의미를 기준으로 할 경우가 성취기준과 채점기준이 정합성을 상실하게 된다.
- ③ 채점기준은 이원목적분류표에 따라 명확히 설정해야 한다. 채점기준은 이원목적분

류표의 내용영역과 행동영역을 상세화하는 수준에서 설정되어야 한다. 학교현장에서 이원목적분류표는 모든 평가문항 핵심요소의 총합이다. 내용영역, 성취기준, 행동영역, 난이도, 그리고 배점을 한 눈에 볼 수 있다. 채점기준은 내용영역은 성취기준 및 행동영역의 관계에서 난이도를 설정하고 난이도를 반영하는 배점이 이루어져야 할 것이다. 특히 피험자의 능력을 변별하기 위해서는 행동영역을 더 세분화하고 그에 따른 채점기준이 상세화될 필요가 있다. 예를 들면, 현재 대부분의 중등학교 교사들이 행동영역을 지식, 이해 적용의 3단계로 나누고 있다. 이를 6단계인 기억, 이해, 적용, 분석, 평가, 그리고 창의로 세분화 하고(강신천, 2014) 그에 따라 채점기준을 상세화하게 되면 학생들이 고등사고능력을 개발하는데 더 많은 도움이 될 것이다.

- ④ 채점기준은 피험자의 학습기준이 될 수 있어야 한다. 채점기준은 피험자의 학습을 위한 현실적 기준이 되어야 한다. 학습자들은 서술형 평가를 준비할 때 서술형문항의 채점기준에 따라서 평가문항을 이해하고 해결하려는 경향이 강하고 채점기준이 지시하는 방향으로 학습자들의 능력이 길러지게 된다. 따라서 채점기준은 지엽적이고 사소한 사항이 아닌 서술형 평가의 목적에 맞닿아있는 핵심적인 이고 본질적인 사항을 학습하도록 설정되어야 한다. 예를 들어, 채점기준이 문법이나 단어의 뜻 등 단편적 지식을 중심으로 설정되어있으면 학생들은 그에 따른 능력을 배양하게 된다. 반면에 사물을 묘사하고 생각을 표현하는 것을 중심으로 채점기준이 설정되어있으면 학생들은 표현적이고 창의적인 능력을 키울 것이다.
- ⑤ 채점기준은 피험자의 학습동기를 유발시킬 수 있어야 한다. 채점기준이 피험자들이 학습을 지속으로 이어가는 외적 동기로서 작용하도록 채점기준을 마련해야 한다. 채점기준이 피험자의 입장에서 이해할 수 있도록 명확히 설명되어야 한다. 채점기준에 대한 자세하고도 친절한 설명은 피험자로 하여금 서술형 평가에 대한 두려움과 불신을 제거한다. 뿐만 아니라 채점기준이 학습을 위한 외적 동기로 긍정적인 역할을 할 것이다.
- ⑥ 채점기준은 집단 내의 피험자들 모두에게 공정해야 한다. 채점기준은 평가문항과 서와 같이 집단 내 피험자들에게 사회문화적으로 편파적이어서는 안 된다. 평가문항이 공정하다고 해도 채점기준이 편파적이라면 피험자들의 능력은 왜곡되거나 거짓이 될 것이기 분명하다. 채점자가 채점기준을 충분히 숙지하고 있다고 하더라도 채점시간이 길어짐에 따라서 채점기준의 해석에 대해 변화가 생겨서 기준의 폭이 더 넓어지거나 좁아지는 현상이 생길 수 있다. 또 반별로 채점을 할 경우에도 채점기준의 해석에 대해 변화가 생겨서 반별로 채점기준이 미세하게 달라지는 경우도 있다. 이러한 경우를 고려하여, 채점기준이 형식적으로나 외적으로는 차이가 없을지라도 채점자의 신체적 심리적 정신적 상태에 따른 편파성을 최소화하려는 제도적 장치가 필요하다.
- ⑦ 채점기준은 서로 중복되거나 누락되는 부분이 없어야 한다. 채점기준을 설정할 때

기준들이 겹치거나 기준 간에 빠뜨리는 부분이 없어야 한다. 채점기준은 해당 평가 문항의 모든 영역을 포괄해야 하며 동시에 소영역 간들은 상호배타적이어야 한다. 이 원칙을 위반하게 되면 채점을 할 수 없는 답안이 생기게 되거나 한 답안이 동시에 두 영역에 속하게 되는 현상이 생기게 되어 채점에 오류가 발생한다. 예를 들어, 아래 평가문항의 채점기준표(교육과학기술부, 2012. 2009 개정 교육과정에 따른 성취기준·성취수준 - 고등학교 역사)를 보면 ‘생명과 신체를 보호’를 서술한 경우가 빠져서 포괄성에 위배되고, 피험자가 ‘농민과 노비의 생명과 신체를 보호’ 라는 ‘의미’로 서술한 경우는 중첩되는 현상이 발생하여 배타성을 위배하게 된다.

【한국사 - 25】

학년군	고등학교 1~3학년군	단원	(1) 우리 역사의 형성과 고대 국가의 발전
성취기준	국1212. 고조선과 초기 철기 시대에 등장한 여러 나라의 사회 모습과 풍속을 설명할 수 있다.		
문항 유형	서답형		
출제 의도 및 평가 내용	고조선 8조의 법 중 알려져 있는 3개 조목의 내용을 통해 당시 사회의 모습을 파악할 수 있는지를 평가한다.		

평가문항

다음 자료를 읽고 물음에 답하시오.[5점]

고조선에서는 백성들에게 금하는 법 8조가 있었다. 그것은 대개 ㉠ 사람을 죽인 자는 즉시 죽이고, ㉡ 남에게 상처를 입힌 자는 곡식으로 갚는다, ㉢ 도둑질을 한 자는 노비로 삼는다, 용서받고자 하는 자는 ㉣ 한 사람마다 50만 전을 내야 한다, 비록 용서를 받아 보통 백성이 되어도 풍속에 역시 그들은 부끄러움을 씻지 못하여 혼인을 하고자 해도 짝을 구할 수 없다. 이러해서 백성은 도둑질을 하지 않아 대문을 닫고 사는 일이 없었다. ㉤ 여자는 모두 정조를 지키고 신용이 있어 음란하고 편벽된 짓을 하지 않았다. 농민은 대나무 그릇에 음식을 담아 먹고, 도시에서는 관리나 장사꾼을 본받아 술잔 같은 그릇에 음식을 담아 먹는다.

- 한서 -

- (1) ㉠~㉤ 중에서 고조선 사회가 노동력을 중시했음을 알 수 있는 근거로 옳은 것만을 모두 골라 표기하시오.[2점]

정답: _____

(2) (1)의 이유를 당시 중요한 생산 자원과 연결시켜 서술할 때 빈칸에 들어갈 내용을 20자 내외로 쓰시오.[3점]

정답: 고조선 사회는 노동력을 중시하여 당시 중요한 생산 자원이었던 _____ 하려고 노력하였다.

예시 답안

(1) ㉠, ㉡

(2) 농민과 노비의 생명과 신체를 보호

채점 기준

* 배점: 아래의 1, 2항목별로 각각 채점하여 합산한다.

채점 기준		배점
(1)	㉠, ㉡을 모두 표기한 경우	2점
	㉠, ㉡ 중에서 한 가지만 맞게 표기한 경우	1점
	㉠, ㉡을 모두 표기하지 못한 경우	0점
(2)	‘농민과 노비의 생명과 신체를 보호’ 구절을 모두 서술한 경우	3점
	‘농민의 생명과 신체를 보호’라는 의미로 서술한 경우	2점
	‘노비의 생명과 신체를 보호’라는 의미로 서술한 경우	
	‘생명과 신체를 보호’라는 구절을 서술하지 못한 경우	0점

[그림 II-12] 서술형 평가 문항의 채점기준에 대한 예시

2) 채점방법

서술형 채점방법은 종합적 방법과 분석적 방법, 가점 방법과 감점 방법, 부분점수 인정방법과 미인정 방법, 그리고 이를 혼합한 채점방법으로 구분할 수 있다.

- ① 종합적 방법과 분석적 방법: 하나의 시험 전체나 하나의 문항 전체를 인상적 직관적으로 3등급이나 5등급 등으로 채점하는 방식이다. 이때는 채점의 공정성과 객관성을 확보하기 위하여 기준 시험 또는 기준 문항(anchor papers/items)이 제공되거나 성취기준이 제시된다. 성취기준평가에서 많이 사용하는 방식이며, 학생의 반응이 다양하게 나타날 수 있는 보다 복합적인 서술형 평가에 사용되는 방식이다. 종합적 채

점방법은 피험자가 표현한 응답을 전체적 인상에 근거하여 평가하기 때문에 ‘종합적으로 어떠하다’고 타당한 판단을 할 수 있고, 채점이 짧은 시간 내에 이루어질 수 있는 장점이 있지만, 근본적으로는 요소별 분석을 하지 않기 때문에 피험자에게 세부적 환류를 하기는 어려우며 채점자의 평가전문성에 더 많은 영향을 받을 개연성이 있다(김수동 외, 2005).

반면에 분석적 방법은 하나의 시험 또는 하나의 문항을 각 요소별로 나누고 요소별로 채점하는 방식이다. 이 방법은 정답이나 모범 답안이 구체적으로 제시되어 채점의 객관성과 신뢰성이 민감할 때 더 많이 사용된다. 예를 들어 영어말하기 평가에서 주관적 채점의 개연성을 최소화하여 공정하고 신뢰성 있는 평가를 진행하기 위해 분석적 방식(Analytical Scoring)을 활용하여 이해력, 유창성, 정확성, 발음의 4가지 영역을 측정하는 경우이다. 분석적 채점방법은 구성 요소들을 따로따로 고려하여 각 구성요소별로 채점한 후 모두 합산하는 방식이기 때문에 피험자 응답의 질과 피험자의 능력에 대해서 보다 세부적으로 환류할 수 있다(김수동 외, 2005). 즉, 분석적 채점 결과는 개별 피험자들에게 이후에 해당 영역의 능력을 향상시키는데 어떤 부분에 어떻게 노력을 해야 하는지에 대한 정보를 제공하기 쉽다. 종합적 채점방법과 비교해서 구체적 채점 기준이 적용되기 때문에, 채점자의 평가전문성이나 소양에 덜 의존하게 된다. 그러나 분석적 채점방법 또한 채점자들이 각 평가요소들이 어떤 것인지 분명히 구분하고 채점기준을 정하고 다양한 채점기준에 알맞게 채점해야 하기 때문에, 채점자의 평가소양과 평가전문성이 종합적 채점방법보다 더 요구된다고 할 수 없다.

- ② 가점 방법과 감점 방법: 가점 방법은 평가 문항에서 점수를 받을 수 있는 요소를 제시하고 그 요소에 가중점수를 더하는 방식으로 채점하는 방법이다. 긍정적 채점방식으로 학생들의 적극적인 도전을 유도하는 채점유형이다. 발달적 교육관을 기반으로 하는 채점방식이다. 반면에 감점 방법은 감점하고자 하는 평가요소를 제시하고 최고점에서 차감하는 방식의 채점유형이다. 이것은 방어적 채점방식으로 실수를 줄이고 완벽을 기하도록 하는데 유용한 방식이다. 채점에 시시비비를 분명히 하며 이의 제기를 미연에 방지하는 효과가 있다. 선별적 교육관에 따른 채점방식에 주로 사용된다.
- ③ 부분점수 인정방법과 미인정방법: 부분점수 인정방법은 평가 문항의 중심적인 요소와 비 중심적 요소가 결합된 평가문항이나 중심적 요소로 이루어진 평가문항의 경우에도 그 중심요소의 어느 정도까지를 인정하는가에 따라서 부분점수를 부여하는 채점 방식이고 부분점수 미인정방법은 평가문항의 중심요소와 비중심요소를 구분하지 않고 일괄적으로 채점하여 부분점수를 인정하지 않는 채점 방식이다. 채점에 의해 평가문항의 질이 훼손되는 것을 방지하려면 부분점수 인정방법이 바람직한 방식이며 피험자의 능력을 변별하는데도 우수할 뿐만 아니라 사고력과 분석력을 향상시키는데도 유익할 것이다. 부분점수 인정방법을 사례나열방식이라고도 한다.

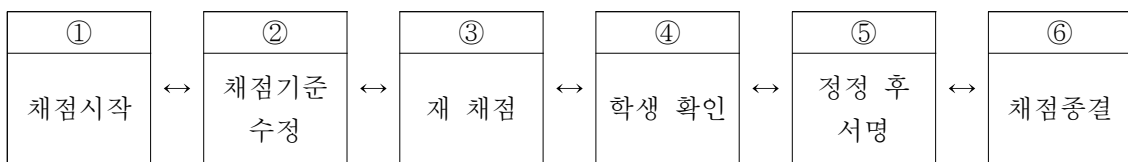
- ④ 혼합방식: 서술형 평가에서 위의 방식을 혼합하여 사용하기도 한다. 예를 들어, 영어 말하기 경시대회에서 참가자는 원고를 제출하고 경시대회 당일 발표를 한다. 이때 발표는 종합적 방법과 가점 방법으로 평가를 하고 원고는 분석적 방법과 감점 방법을 사용할 수 있다. 또는 채점자가 여러 명이려면 두셋은 종합적 방법으로 나머지는 분석적 방법으로 평가한 후에 합산하는 방식으로 채점을 할 수도 있다.

3) 채점절차

가) 일상적 채점절차

일상적으로 학교에서 교사들이 행하는 채점절차는, 예를 들어 한 학년 9개 학급(각반 30명)의 사회를 담당하는 교사를 생각해 보면, 10문항의 서술형 평가를 출제했다면,

- ① 자신이 이미 만들어 놓은 모범답안과 채점기준에 따라 반별로 채점을 시작할 것이다. 한 반을 채점하다보면 생각지 못한 응답이 나오기도 하고 기준을 적용하기 애매한 응답이 나오기도 한다.
- ② 애매한 응답의 경우, 자신의 판단에 따라서 채점기준을 수정하거나 인정답안의 범위를 넓히거나 하여 해결하기도 하고 판단하기가 정말 힘들고 애매한 경우에는 학년은 다르지만 동과의 다른 교사에게 의견을 물어서 해결하기도 한다.
- ③ 채점기준이나 인정답안이 수정되면 앞서 채점했던 피험자의 시험지를 모두 다시 검토하면서 채점한다.
- ④ 채점이 끝나면 학생들의 성적을 입력하고 그 결과를 출력하여 학생들에게 확인하는 절차를 거친다.
- ⑤ 확인과정에서 학생들의 이견이나 주장이 있으면 이를 혼자서 검토하여 처리하거나 동과 교사에게 의견을 물어서 처리한다.
- ⑥ 학생들의 확인 절차가 끝나면 수정입력한 후 출력하여 학생들의 서명을 거치면 채점은 종결된다.



[그림 II-13] 일상적 채점절차

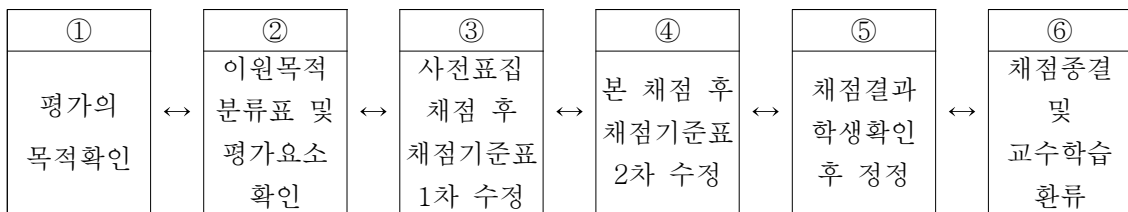
나) 이상적 채점절차

이상의 일상적 채점과정과는 달리 이상적 채점절차는 대체로 다음과 같을 것이다.

- ① 평가 목적 및 목표의 확인(김중수 외, 2011)
- ② 이원목적분류표 확인
- ③ 문항 평가요소 확인
- ④ 사전 표집 채점

- ⑤ 사전 표집채점 후 문제협의 - 채점기준표 1차 수정
- ⑥ 본 채점 실시
- ⑦ 본 채점 후 문제협의 - 채점기준표 2차 수정
- ⑧ 채점결과 학생확인
- ⑨ 채점종결
- ⑩ 채점종결 후 교수·학습 환류

이상적 채점절차에서 중요한 절차는 채점교사들 간의 채점기준에 대한 논의와 소통을 통한 채점기준의 수정과정과 채점과정에서의 학생들과의 소통과 그 결과에 대한 환류이다.



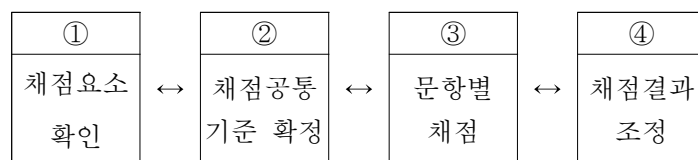
[그림 II-13] 이상적 채점절차

(3) 핵심적 채점절차

그러나 평가자나 학습자의 입장이 아닌 순수 채점자의 입장에서 핵심적 과정은 다음과 같을 것이다.

- ① 채점자는 채점기준표에서 채점기준을 평가요소별로 정확히 이해하고 숙지해야 할 것이다.
- ② 채점자가 여러 명일 경우, 각 채점자들이 일관된 채점기준을 유지하기 위해서 채점기준에 대한 논의를 통해서 개별적 의문과 관점의 차이를 해소하고 공통의 관점을 확고해야 할 것이다.
- ③ 채점자들이 몇몇 평가문항에 대해서 개별적으로 채점을 실시한다.
- ④ 채점결과를 비교하여 차이가 있는 경우 논의를 통해 해석과 적용차이를 해결한다.

이와 같은 채점과정을 통하여 채점자들이 채점 기준을 타당하고 일관성 있게 적용하여 채점할 수 있을 때 서술형 평가의 채점 결과가 신뢰도와 타당도를 가지며, 피험자들의 능력에 대한 의미있는 정보를 환류할 수 있을 것이다.



[그림 II-13] 핵심적 채점절차

2. 일반적인 수행평가의 이해

가. 수행평가의 정의와 범주

수행평가는 지필검사를 대체할 전혀 새로운 평가방법이 아니라 심동적 영역의 행동 특성을 평가하기 위하여 사용되던 평가방법을 인지적 영역 행동 특성을 평가하기 위하여 인지적 행동 특성에 맞게 평가 방법을 적용한 것이다. 그러므로 인지적 영역에서의 수행평가는 기존의 선택형 검사에 근거한 평가방법을 보완하는 측면에서 해석하는 것이 옳다. 남명호(1995)도 수행평가는 관련 있는 과제에 대해 학생들의 실제적인 수행(performance)을 검사하고 판단하는 평가로서 최근에 처음 사용되는 평가방법이 아니라고 하였으며, 오래 전부터 실시되었던 예체능계의 실기평가와 과학실험평가, 그리고 의과대학의 실습 등이 대표적인 수행평가의 예라고 하였다.

성태제(1999)는 수행평가와 혼용되어 사용되고 있는, 참평가, 포트폴리오, 직접평가, 대안적 평가에 대한 용어를 정리하여 <표 II-3>과 같이 설명하고 있다.

<표 II-7> 수행평가와 관련된 용어 정의

용어	정의
수행평가	지식이나 기능, 혹은 기술의 수행정도를 측정하는 평가
참평가	실제 상황에서 수행정도를 측정하는 평가
포트폴리오	개인의 작업이나 작품을 모아둔 자료집이나 서류철
직접평가	표출되는 행위를 직접 관찰에 의하여 실시하는 평가
대안적 평가	기존의 어떤 평가 방법을 대치할 수 있는 평가

수행평가(performance assessment)는 배운 내용이나 지식, 그리고 습득한 기술이나 기능을 행위로 나타내는 정도를 측정하여 판단하는 평가방법을 말한다. 참평가(authentic evaluation)는 실제 생활에서 실시되는 수행평가로써 평가의 목적에 따라 고안되지 않는 특징을 지니고 있다. 그러므로 Meyer(1992)는 교육현장에서 의미 있는 평가방법이 되기 위해서는 수행평가와 참평가는 구분되어야 한다고 주장한다. 정의에 의하면 참평가는 모두 수행평가가 될 수 있으나 모든 수행평가가 참평가가 될 수 없다. 포트폴리오(portfolio)는 작업 결과나 작품 혹은 어떤 수행의 결과를 모아 놓은 자료집이나 서류철을 보고 평가하는 방법으로 자료집에 넣을 작업 결과를 학생이 스스로 선택한다. 그러므로 학생이 일차적으로 평가의 주체가 될 수 있으며 오랜 시간의 수행결과를 평가할 수 있으므로 평가대상의 수행 정도뿐 아니라 발달과정까지 평가할 수 있는 장점을 지닌다. 직접평가(direct assessment)는 간접평가와 반대되는 개념에서 쓰인 용어이다. 인간의 능력은 잠재되어 있기 때문에 잠재되어 있는 속성을 직접 측정하기가 불가능하므로 검사라는 도구를 사용하여 간접측

정을 실시하게 된다. 그러나 수행평가는 검사에 의한 간접 측정이라기보다는 어떤 행위를 보면서 평가하기 때문에 직접평가라는 용어를 사용한다. 대안적 평가(alternative assessment)는 평가방법의 특성을 염두에 두고 사용된 용어가 아니며 기존의 어떤 평가방법을 대체할 수 있는 평가방법을 말한다.

〈표 II-7〉에 설명된 용어를 볼 때 의미상으론 대안적 평가는 수행평가와 동의어가 아니다. 수행평가, 참평가, 포트폴리오, 그리고 직접평가가 동의어이나 이 모두를 포함하여 광의의 수행평가라고도 한다. 그러므로 주목하여야 할 것은 수행평가의 정의가 〈표 3〉에서 설명한 협의의 수행평가의 정의가 있고 광의의 수행평가에 대한 정의가 있음을 알 수 있다. 수행평가를 이해함에 있어 본래 수행평가에 대한 정의를 기초로 수행평가에 대한 범위를 넓히는 것이 수행평가에 대한 오해를 줄일 수 있다.

수행평가가 관심의 대상이 되면서 모든 평가 방법을 수행평가라 하는 경향은 우리나라만의 경향이 아니라 미국도 1990년 초반에 언어영역에서 종전에 사용하던 에세이 형태의 문항도 그리고 수학에서 사용하던 계산형 문제도 수행평가 방법이라 하는 경향이 있었다. 우리나라에서는 선택형 문항이외의 모든 평가 방법(교육부, 1998), 또는 단답형 문항 수준 보다 높은 형태의 문항에 의한 평가를 수행평가(백순근, 1998)라하고 있다. Phopham(1995)도 수많은 교육학자들이 매우 다양한 평가방법을 수행평가라 하고 있으며 단답형 그리고 논술형까지 수행평가 방법이라고 주장한다고 지적하였고 피험자가 답을 구성하는 형태의 평가 방법이면 모두 수행평가와 동등화시키는 경향이 있다고 하였다. 그러나 다른 학자들은 수행평가로 분류하기 위해서는 측정의 절차를 중요시하여 보다 엄격한 기준을 설정하였다. 수행평가 제창자 혹은 옹호자들은 진정한 수행평가가 되기 위해서는 다음 세 가지 기준을 충족시켜야 한다고 주장한다(Popham, 1995).

- ① 다양한 기준: 학생의 수행은 두 가지 이상의 기준에 의하여 평가
- ② 사전에 규명된 질적 기준: 수행평가의 질을 평가하기 이전에 평가기준이 명료하게 제시되어야 함
- ③ 주관적 평가: 점수화하기 보다는 평가자의 주관적 판단에 의존

우리나라에서도 수행평가에 대한 혼돈은 평가방법을 임의적으로 확산하고 그리고 인위적으로 양분하려 하며 종전의 선다형 검사에 의한 평가방법은 나쁜 평가방법으로 간주하려는 경향 때문이다. 이런 이유로 연구자는 교육평가 전문가로서 인지적 행동특성을 평가하는 선택형 문항의 평가와 수행평가를 구조화의 정도에 따라 연속선상의 개념으로 보고 [그림 II-14]와 같은 특징을 설명한다.

진위형 문항에 의한 평가로부터 참평가에 이르는 연속선상에 있는 두 평가방법은 강조점이 다르다. 진위형의 선택형 문항에 의한 평가는 무엇을 알고 있는지 여부를 판단하는 데 초점이 있으나 참평가는 인위적 상황보다는 생활에서 알고 있는 것을 할 수 있느냐에 관심을 둔다. 그러므로

선택형 문항에 의한 평가는 고정된 형태의 평가방법에 의존하나 수행평가는 개방된 형태의 평가방법을 사용하며 결과뿐 아니라 진행과정까지 평가한다. 선택형 문항에 의한 평가는 일반적으로 정답과 정도의 차이를 나타내므로 평가도구의 타당도와 신뢰도를 객관적으로 추정할 수 있는 장점이 있다. 그러나 수행평가는 평가항목이 구조화되어 있지 않으므로 내용타당도와 준거타당도에 의존하여 평가도구의 타당도를 검증하고 관찰에 의하여 평가가 이루어지므로 관찰자인 채점자내, 그리고 채점자간 신뢰도에 의하여 평가 결과의 신뢰도를 검증한다.

등위판정, 서열화 등의 행정적 기능이 강조될 때 측정의 오차와 관련된 신뢰도를 높여야 하므로 선택형 문항에 의한 평가방법이 사용된다. 반면, 행정적 기능보다는 교수적 기능이 강조되고 채점자내와 채점자간 신뢰도에 보다 자유스러우면 수행평가를 실시할 수 있다. 만약 상호경쟁이 치열하여 평가 결과의 공정성과 신뢰성이 절대적이라면 수행평가 결과를 사용하는 데는 한계점이 있다.

선택형 문항에 의한 평가 (구조화)							수행평가 (비구조화)		
진위형	선다형	배합형	괄호형	단답형	논술형	구술시험	수행평가	포트폴리오	참평가
인지							인지 / 정의 / 심동		
앎							행함		
학습결과							학습진행 / 결과		
고정형 평가							개방형 평가		
이분적 평가(하나의 정답)							다분적 평가(다양한 정답)		
타당도 - 내용타당도							타당도 - 내용타당도		
- 구인타당도							- 준거타당도		
- 준거타당도									
신뢰도 - 재검사신뢰도							신뢰도 - 채점자간 신뢰도		
- 동형검사신뢰도							- 채점자내 신뢰도		
- 내적일관성신뢰도									
분석적 접근(analytical approach)							총체적 접근(holistic approach)		
인위적 상황							실제적 상황		
일회적 평가							지속적 평가		
정적평가							동적 평가		
행정적 기능							교수적 기능		

[그림 II-14] 선택형 문항의 평가와 수행평가의 연속적 개념

연속선상에 있는 진위형 문항에 의한 평가와 참평가 사이에서 어떤 평가방법을 수행평가로 볼

수 있는 기준은 논술에 의한 평가 형태 이상의 수준이라 말할 수 있으나 논술에 의한 평가 즉 단순한 논술형 문항을 모두 수행평가라 하지 않는다. 평가방법의 유형으로 수행평가 여부를 판단하는 확실적 기준은 없으나 무엇을 평가하느냐 즉 알고 있는 지식의 정도를 행위로 나타내거나, 행위로 나타내지 못하더라도 어떻게 할 것이라는 내용(how to do)을 서술할 수 있는지를 수행평가의 판단 기준으로 삼는 것이 바람직하다.

수행평가는 습득한 지식, 기능이나 기술을 실제 상황이나 인위적 평가 상황에서 얼마나 잘 수행하는지(doing, performing), 최소한 어떻게 수행할 것인지(how to do, how to perform)를 관찰, 면접 등의 다양한 방법을 통하여 종합적으로 판단하는 평가방법으로서 지식이나 기능에 의한 정답여부나 결과물에만 관심이 있는 것이 아니라 수행과정과 결과를 종합적으로 평가하는 방법이다(성태제, 1999).

나. 기존의 선택형 평가와 수행평가와의 비교

수행평가는 선택형 평가에 비해 수험생의 성취수준을 정확하게 측정할 수 있는 평가 유형이다. 기존의 선택형 평가와 서술형 평가의 특징을 비교하면 <표 II-8>과 같다.

<표 II-8> 선택형 평가와 서술형 평가의 비교

구분	선택형 평가	수행평가
진리관	절대주의적인 진리관	상대주의적인 진리관
철학적인 근거	합리론 경험론 등	구성주의 현상학 해석학 인류학 등
시대적 상황	산업화 시대 소품종 대량 생산	정보화 시대 다품종 소량 생산
학습관	직선적·위계적·연속성 과정 추상적·객관적 상황 중시 학습자의 기억·재생산 중시	인지구조의 계속적 변화 구체적·주관적 상황 중시 학습자의 이해·성장 중시
평가체제	상대평가 양적평가 선발형평가	절대평가 질적평가 충고형평가
평가목적	선발·분류·배치 한 줄 세우기	지도·조언·개선 여러 줄 세우기

평가내용	선언적 지식 (내용적 지식) 학습의 결과 중시 학문적 지능의 구성요소	절차적 지식 (방법적 지식) 학습의 결과 및 과정도 중시 실천적 지능의 구성요소
평가방법	선택형 평가 위주 표준화 검사 중시 대규모 평가 중시 일회적·부분적인 평가 객관성·일관성·공정성 강조	수행평가 위주 개별 교사에 의한 평가 중시 소규모 평가 중시 지속적·종합적인 평가 전문성·타당도·적합성 강조
평가시기	학습활동이 종료되는 시점 교수·학습과 평가활동 분리	학습활동의 모든 과정 교수·학습과 평가활동 통합
교사의 역할	지식의 전달자	학습의 안내자·촉진자

다. 수행평가의 설계

수행평가를 실시하기 위하여 평가자는 다음과 같은 질문들에 대하여 깊이 생각해야 되며, 이에 대한 해답을 제시할 수 있어야 한다.

첫째, 평가하려는 능력은 무엇인가?

둘째, 평가하려는 능력을 측정하는데 적절한 평가방법이 수행평가인가?

셋째, 사용될 수행평가 방법이 신뢰롭고 타당하며, 사용하기에 편리한가?

넷째, 수행평가 방법이 특정집단의 학생들에게 이익이나 불이익을 주지 않고, 모든 학생들에게 공정하게 적용되는가?

수행평가에 필요한 일반적 실시 절차를 크게 분류하면 평가목적, 성취행동의 평가형태, 평가방법의 설계, 평가결과의 채점방법의 4단계가 된다(Stiggins, 1994). 각각의 세분화된 단계는 수행평가를 위해서 필수적이다.

① 평가이유를 명확히 밝힌다.

평가란 그 자체가 목적이 아니라 설정된 목표를 확인하기 위한 하나의 수단 혹은 도구에 불과하다. 따라서 수행평가는 평가를 하는 이유와 평가 결과의 사용처를 분명히 밝힘으로써 출발해야 한다.

㉠ 평가목적을 구체적으로 서술한다.

평가목적은 명확하게 규명함으로써 그 목적에 적절한 평가내용과 형태가 결정될 수 있다. 교육현장에서는 새로운 학습상태에서 학생들을 능력에 따라서 반편성을 하거나 능력별 집단편성을 하기 위한 배치고사(placement test)를 실시하고, 선수 학습능력의 측정이나 새로 전개하게 될 교과목표와 관련하여 도입단계의 학업 수행능력을 알아보기 위해서 진단평가(diagnostic evaluation)를 실시한다. 그리고 수업과정과 관련하여 개선점을 찾기 위한 제반 정보를 얻기 위해서는 형성평가(formative evaluation)를 실

시하고, 학습단원이 끝났거나 학기, 학년이 끝날 무렵에 의도했던 학습목표의 달성여부와 학생들의 학업성취능력을 측정하여 성적의 판정이나 자격부여 등을 목적으로 총괄평가(summative evaluation)을 실시한다.

㉠ 평가결과를 활용할 사람을 구체적으로 서술한다.

평가결과를 활용할 사람으로는 교사, 평가실시자, 학생, 학부모, 교육청, 자격부여기관 등이 될 수 있다.

㉡ 평가결과의 용도를 구체적으로 서술한다.

평가결과에 대한 가치부여 혹은 해석방법에 따라 규준지향평가(norm referenced evaluation)와 준거지향평가(criterion referenced evaluation)로 분류된다.

㉢ 평가대상을 구체적으로 서술한다.

수행평가를 실시할 평가대상의 수, 평가대상 학년, 평가대상의 특징 등을 구체적으로 기술한다. 평가대상의 특징으로는 우수 학생집단, 열등 학생집단 등이 될 수 있다.

② 평가할 성취행동을 명확히 밝힌다.

㉠ 평가내용이나 기술을 구체적으로 서술한다. 평가할 성취행동의 내용이나 기술 등을 자세하게 기술한다.

㉡ 평가할 성취행동의 유형을 선정한다.

평가할 구체적인 성취행동의 유형을 기술한다. 수행평가는 실제로 성취행동이 일어날 때 그 행동을 관찰하거나 성취행동의 결과물을 보고 이루어진다. 따라서 언제, 어떤 과제물을 평가하는 것이 학생의 능력을 가장 정확하게 평가하는지를 파악하는 것이 매우 중요하다.

㉢ 평가기준(performance criteria)을 열거한다.

평가기준을 결정하는 것은 수행평가의 질을 좌우할 만큼 중요하다. 그러므로 적절한 평가기준을 기술함으로써 수행평가 중 관찰한 성취행동이나 과제물에 등급을 정확하게 매길 수 있어야 한다.

③ 평가방법을 설계한다.

㉠ 평가방법의 형태를 결정한다.

평가를 어떤 방법으로 할 것인가를 결정하여 기술한다. 성취행동을 가장 적절하게 평가할 수 있는 방법을 결정해야 하며, 경우에 따라서는 두 가지 이상의 방법을 모두 사용하여 평가할 수 있다. 주의할 점은 모든 학생들에게 자신의 능력을 기회가 공정하게 주어질 수 있는 방법을 선정해야 한다는 것이다.

㉡ 평가시행의 공고여부를 결정한다.

학생들에게 대한 평가의 공고여부를 결정한다. 미리 학생들에게 공고하고 평가를 하는 경우와 그렇지 않은 경우로 구분하여 피험자의 동기유발과 시험불안 정도를 비교할 수 있다.

㉢ 평가자료 수집의 양을 결정한다.

표집할 평가 자료의 수를 기록한다. 평가 자료의 수집에는 첫째, 한 번의 성취행동에서

한 개의 자료를 수집하고 둘째, 한 번의 성취행동에서 여러 개의 자료를 수집하고 셋째, 여러 번의 성취행동을 통하여 여러 개의 표본을 수집하는 것이다. 평가자료 수집방법은 평가결과의 중요성, 표집의 대표성, 평가에 소요되는 시간과 비용에 따라 결정된다. 따라서 수행평가를 실시할 때에는 피험자 수와 각 피험자로부터 표집된 평가 자료의 양을 고려해서 채점할 수 있는 시감과 비용이 충분한지를 염두에 두어야 한다.

④ 채점계획을 수립한다.

수행평가의 마지막 단계는 성취행동을 점검하고 그 평가결과를 필요한 사람에게 알려주는 채점 기록방법을 결정하는 것이다. 채점계획에서는 어떤 형태의 점수가 필요하며, 누가 채점을 하며, 어떻게 채점 자료를 기록할 것인가를 결정해야 한다.

㉠ 채점방법을 결정한다.

채점방법은 총괄적 채점(holistic rating)과 분석적 채점(analytic rating)으로 구분되며, 이는 평가결과의 용도에 따라 결정된다. 예를 들어, 평가의 목적이 선발이나 배치, 순위를 필요로 하는 경우에는 일반적으로 총괄적 채점이 적용되는 반면에 개인 및 집단의 진단이나 최소한의 숙련 정도를 파악하는 경우에는 주로 분석적 채점이 사용된다. 그러나 실제로는 총괄적 채점과 분석적 채점이 함께 사용되는 경우가 많이 있다.

㉡ 채점자를 결정한다.

채점자는 교사나 자격이 있는 전문가가 될 수 있다. 혹은 피험자도 채점자가 되어서 자신의 수행결과나 다른 피험자의 수행결과를 채점할 수 있다. 교사나 자격이 있는 전문가가 채점을 하는 경우에는 채점시 전문적, 기술적 지식, 경험을 필요로 하거나 중요하고 경쟁이 심한 평가에서 공정하게 채점을 해야 할 때이다. 피험자가 채점을 할 경우는 피험자가 평가기준을 적용할 수 있으며, 평가기준표를 보고 채점과정을 배울 수 있거나 채점결과가 채점자간에 약간의 차이가 있더라도 평가결과에 큰 영향을 미치지 않을 때이다.

수행평가 결과를 채점할 때에는 채점자간의 차이(variability)와 편파(bias)가 수행평가의 문제점으로 지적될 수 있다. 이를 해결하기 위해서는 정확한 채점기준표의 작성과 채점자들의 훈련이 절대적으로 필요하다. 채점자들의 개인차로 인하여 채점기준표를 다르게 이해하고 적용할 수 있으므로 채점훈련을 통하여 채점결과가 일치하도록 하는 것이 요구된다.

그리고 채점자를 두 사람이상으로 하는 방법도 채점결과를 신뢰롭게 하기 위해서 매우 중요하다. 만약 두 사람이 동일한 문항을 따로 채점하여 그 결과의 차이가 정한 범위를 넘었을 경우에는 다시 채점해야 하며, 이런 상황이 자주 발생하면 채점자를 바꾸는 문제도 고려해야 한다. 또한 채점자들은 모든 평가문항을 채점하지 말고 몇 개의 문항만을 전담하여 채점하는 것이 여러 측면에서 효과적이다. 동일한 문항만을 계속 채점하면 채점속도가 빨라지며 채점결과는 정확해 질 것이다.

㉢ 평가결과 기록방법을 명확히 밝힌다.

수행평가의 마지막 과정은 평가결과를 표시하는 방법을 정하는 것이다. 평가결과는 여러 방법으로 기록될 수 있다. 예컨대 성취행동의 평가결과를 점수로 기록하거나, 성취행동에서 어느 항목의 해당여부를 체크리스트에 표시하거나, 성취행동의 숙련정도를 척도로 표시할 수 있다. 그리고 평소에 관찰한 사항을 기록하거나 포트폴리오로 보관할 수 있다. 피험자가 많으면 성취행동을 점수로 표시하거나 체크리스트에 표시하는 방법이 가장 적절하다.

라. 수행평가의 문항 유형

수행평가 방법으로는 서술형(주관식) 검사, 논술형 검사, 구술시험, 토론법, 실기시험, 실험·실습법, 면접법, 관찰법, 자기평가보고서법, 동료평가보고서법, 연구보고서법, 포트폴리오(portfolio)법 등이 있다. 이러한 평가 방법들은 수행평가의 본질을 구현하기 위한 하나의 도구들이며, 수행평가가 전제하고 있는 중요한 특성들을 평가 상황에서 제대로 구현될 때 비로소 수행평가가 시행되고 있다고 할 수 있다. 따라서 수행평가의 여러 가지 방법에 대한 문항의 유형을 제시하면 다음과 같다.

1) 서술형 및 논술형 검사

서술형 평가 혹은 논술형 평가란 영어의 에세이 테스트(essay test)를 번역한 것으로, 흔히 ‘주관식 평가’라고 하기도 한다. 이 평가 방식은 학생으로 하여금 출제자가 제시한 답을 ‘선택’하도록 하는 평가 방식이 아니라 학생이 답이라고 생각하는 지식이나 의견 등을 직접 ‘서술’하도록 하는 평가 방식이다. 서술형 및 논술형 평가의 가장 큰 특징은, 학생의 생각이나 의견을 직접 서술하도록 하기 때문에 학생의 창의성, 문제 해결력, 비판력, 판단력, 통합력, 정보 수집력 및 분석력 등 고등 사고기능을 쉽게 평가할 수 있다는 것이다.

평가 방식을 세밀하게 분류할 때는 서술형 평가와 논술형 평가를 서로 구별하기도 한다. 즉, 논술형 평가도 일종의 서술형 평가이기는 하지만, 개인 나름의 생각이나 주장을 창의적이고 논리적이면서도 설득력있게 조직하여 작성해야 함을 강조한다는 점에서 일반 서술형 평가와 구별하기도 한다. 정도의 문제이기는 하지만, 대체로 서술형 평가라고 할 경우에는 학생이 서술해야 하는 분량이 많지 않고 또 채점을 할 때 서술된 내용의 깊이와 넓이에만 관심이 있다고 한다면, 논술형 평가라고 할 경우에는 학생이 서술해야 할 분량이 상대적으로 많고 또 채점을 할 때 서술된 내용의 깊이와 넓이뿐만 아니라 글을 조직하고 구성하는 표현 능력이나 논리적인 일관성 등에도 관심이 있다. 그러나, 학교 교육 현장에서 교사들이 문항을 제작할 때 서술형 평가 문항과 논술형 평가 문항을 서로 구분하기 힘든 경우가 많이 있고 또 구분한다는 것이 특별한 의미가 없을 경우도 많이 있다. 따라서 이 글에서는 서술형 평가와 논술형 평가를 특별한 경우를 제외하고는 서로 구별하지 않고 ‘서술형 및 논술형 평가’와 같이 묶어서 사용하고자 한다.

서술형 및 논술형 문항의 유형은 학생들이 서술할 내용을 제한하는 정도에 따라 크게 응답 제한형과 응답 자유형으로 나눌 수 있다(한국교육과정평가원, 1999). 이에 대한 자세한 예시는 ‘서술형 평가 유형’에서 이미 소개하였다.

2) 구술시험

구술시험은 종이와 붓이 발명되기 전부터 시행되어 오던 가장 오래된 수행평가의 한 형태로서, 학생으로 하여금 특정 교육 내용이나 주제에 대해서 자신의 의견이나 생각을 발표하

도록 하여 학생의 준비도, 이해력, 표현력, 판단력, 의사소통능력 등을 직접 평가하기 위한 방법이다. 구술시험은 학생들이 대학이나 대학원을 졸업하기 위해서 졸업시험이나 논문심사의 한 과정으로 거쳐야 하는 관문처럼 생각하는 시험의 한 형태이지만, 초·중등학교에서도 쉽게 사용할 수 있다. 예컨대 중학교 국사 시간에 ‘신라가 삼국을 통일할 수 있었던 이유’라는 주제에 대해서 학생들에게 발표 준비를 하도록 한 다음, 개별적으로 약 5분간 발표하도록 하여 평가할 수 있다. <표 II-9>에 제시된 각 평가요소별 가중치는 평가 상황에 따라 달라질 수도 있을 것이며, 앞에서 예시적으로 제시된 서술형 및 논술형 문항의 【채점 기준표】와 같이 각 평가 요소의 상·중·하 수준에 대해서 그 내용을 구체적으로 진술함으로써 보다 타당하고 공정한 평가를 유도할 수도 있을 것이다.

구술시험은 위와 같이 주제나 질문을 사전에 미리 알려주는 형식을 취할 수도 있지만, 특별한 내용영역만 알려 주고 난 다음, 구술시험을 시행할 때 평가자가 그 내용영역과 관련된 주제나 질문을 제시하고 학생이 답변하는 형식을 취할 수도 있다. 예컨대, 중학교 국사 시간에 학생들에게 ‘삼국시대의 문화’에 대해서 공부하도록 한 다음, 구술시험을 칠 때 평가자가 학생에게 ‘백제 문화의 특징을 고구려 문화나 신라 문화의 특징과 비교하여 설명해 보도록 요구할 수 있다.

<표 II-9> 구술시험을 위한 평가 기준표 (예시)

평가 요소*	점수**			비고
	상	중	하	
준비도 (구술시험을 위해 관련 자료 등 준비를 제대로 하였는가?)	2	1	0	
이해력 (질의한 내용을 제대로 이해하고 발표하는가?)	2	1	0	
조직력 (발표할 내용을 제대로 조직하여 체계적으로 발표하는가?)	2	1	0	
표현력 (자신의 의견을 제대로 표현하는가?)	2	1	0	
판단력 (주어진 시간 안에 적절히 발표하는가?)	2	1	0	
의사소통 능력 (다른 사람들의 시선을 끌면서 설득력 있게 발표하는가?)	2	1	0	
발표 태도 (다른 사람들 앞에서의 발표 태도는 바람직한가?)	2	1	0	
계				

* 평가요소 간 가중치를 다르게 줄 수도 있음.

** 상·중·하 수준에 대해 그 내용을 구체적으로 진술하는 것이 바람직함.

참고로, 프랑스의 경우 대학입시에서 자국어인 프랑스어 과목에 대한 구술시험을 치르는

데, 그 내용과 방법을 소개하면 다음과 같다 (정구향 외, 1998). 프랑스어 과목에 대한 구술 시험은 20~30분에 걸쳐 시행되며 1부와 2부로 구성된다. 1부는 응시생이 시험관에게 자신이 공부한 문헌 목록을 제출하고 이에 관련된 문제에 대해 답하는 형식으로 이루어지고, 2부는 시험관이 제시한 글을 응시생이 읽고 설명함으로써 최종 점수를 받게 된다. 이 경우 구술시험에서는 말하는 내용 뿐 아니라 표현과 발음 및 태도도 평가의 대상이 된다(<표 II-10> 참조).

<표 II-10> 구술시험의 진행과정 및 방식

	시험관	학생(응시자)
시 험 전	① 구술시험 준비를 한다.	① 20분의 답안준비 시간이 주어지며, 답안 골격과 주요사항 및 목록을 메모하는 정도로 한다. ② 시험관의 편의를 위하여 텍스트의 부분(복사물)을 가져가서 개별적으로 전하도록 한다.
시 험 과 정	① 10분씩 2회에 걸쳐 시행된다. 간단한 답도 좋지 않지만, 시간을 넘기면 중단시킨다. ② 답의 특성 부분에 관한 보충설명(상세화나 근거 제시)을 요구할 수 있으며, 방법적 독서와 관련한 질문에서 종합적 질문으로 넘어가는 과정에서 보완적 성격의 질문을 할 수 있다. ③ 응시자의 답변이 일단 끝난 후에 간단한 토론을 벌일 수 있다.	① 무엇을 어떻게 읽어야 하는지에 초점을 두는, 방법적 독서에 관한 질문에서 텍스트를 이미 읽고 있는 경우, 독서기준·근거·참고·인용을 밝히고 결론을 유도한다. ② 텍스트 및 자료를 모르는 경우, 아는 바를 바탕으로 융통성을 발휘하여 간단한 기준을 도입하고 일반적 문제를 중심으로 풀어 나가도록 한다. ③ 종합적 질문에서는 분석·인용하고자 하는 자료를 분명하게 이용하고 상호 관련시키도록 한다.
비 고	① 텍스트의 내용을 요약 제시하거나 다른 말로 반복하는 답안은 용납되지 않는다. ② 작품을 읽은 소감을 피력하는 식의 답안은 논지를 벗어난 것으로 처리한다.	① 답변 시 너무 자신 있거나 꾸미거나 암송하는 말투는 피하여 자연스럽게 말하도록 한다. ② 말을 하다가 멈추지 말고 반드시 끝을 맺도록 한다.

예시된 프랑스어 구술시험에서는 각 문항별 구체적인 채점기준이 별도로 주어지지 않고, 시험관이 ‘관례에 따른’ 채점 기준을 적용한다. 프랑스의 대학 입시를 위한 구술시험에서 사용된 문항들을 예시하면 다음과 같다.

- ① 루소의 ‘고백록’을 읽고 진실을 보는 관점과 이를 통한 자서전의 교훈성에 대해 생각하는 바를 말하라.
- ② 보들레르의 ‘악의 꽃’에 나타난 가을의 의미와 인간의 가을이라는 계절을 통한 시간의 흐름의 지각에 대해 생각하는 바를 말하라.
- ③ 까뮈의 ‘이방인’을 읽고 행복의 요소와 행복의 상대성을 설명하고 인간의 영원한 행복

의 추구에 대해 생각하는 바를 말하라.

④ 샤르트르의 ‘말’을 읽고 나쁜 교육이라는 개념이 성립하는지에 대해 생각하는 바를 말하라.

⑤ 파스칼의 ‘팡세’를 읽고 인간과 시간의 관계에 대해 생각하는 바를 말하라.

3) 토론법

토론법이란, 교수·학습 활동과 평가 활동을 통합적으로 수행하는 대표적인 평가방법으로, 특정 주제에 대해 학생들이 서로 토론하는 것을 보고 평가하는 것이다. 수행평가에서는 특히 찬·반 토론법을 많이 사용하는데, 찬·반 토론법이란 사회적으로나 개인적으로 서로 다른 의견을 제시할 수 있는 토론 주제(예컨대 대통령 중심제와 내각 책임제의 권력구조, 자동차 10부제 시행 여부 등)를 가지고, 개인별로 혹은 소집단별로 찬·반 토론을 하도록 한 다음, 찬성과 반대의견을 토론하기 위해 사전에 준비한 자료의 다양성이나 충실성, 그리고 토론 내용의 충실성과 논리성, 반대 의견을 존중하는 태도, 토론 진행 방법 등을 총체적으로 평가하는 방법이다. 교사는 이러한 찬·반 토론의 과정을 자세히 관찰함으로써 토론 진행과정에서 지도력을 발휘하여 토론을 이끌어 가는 사람, 당당하게 자기주장을 피력하는 사람, 남의 의견을 차분히 듣고 모두의 의견을 집약하는 능력을 발휘하는 사람, 또는 상대방에게 의견을 자유스럽게 제시하도록 한 후 결론은 자기 의견대로 끌고 가는 사람 등등 학생들의 여러 가지 유형의 성격도 파악할 수 있다.

이러한 찬·반 토론법은 논술형 검사와 구술시험을 통해 얻을 수 있는 정보를 모두 얻을 수 있는 장점이 있는 반면, 학생 수가 많을 경우 개별 학생들이 충분히 발언할 기회를 갖지 못하는 단점도 있다. 찬·반 토론법을 시행할 때에도 구술시험과 비슷하게 <표 II-11>과 같은 평가 기준표를 이용하여 특정 학생의 준비도, 이해력, 조직력, 표현력, 판단력, 의사소통 능력, 토론 태도 등을 평가할 수 있다.

<표 II-11> 찬·반 토론법을 위한 평가 기준표 (예시)

평가 요소*	점수**			비 고
	상	중	하	
준비도 (찬·반 토론을 위해 관련 자료 등 준비를 제대로 하였는가?)	2	1	0	
이해력 (토론할 내용을 제대로 이해하고 의견을 발표하는가?)	2	1	0	
조직력 (토론할 내용을 제대로 조직하여 체계적으로 발표하는가?)	2	1	0	
표현력 (자신의 의견을 제대로 표현하는가?)	2	1	0	
판단력 (상대편이 발표한 내용의 핵심을 제대로 파악하고 대응하는가?)	2	1	0	
의사소통 능력 (다른 사람들의 시선을 끌면서 설득력 있게 발표하는가?)	2	1	0	
토론 태도 (상대방의 의견을 존중하면서 토론을 진행하는가?)	2	1	0	
계				

* 평가요소간 가중치를 다르게 줄 수도 있음.

** 상·중·하 수준에 대해 그 내용을 구체적으로 진술하는 것이 바람직함.

참고로 토론법을 사용한 고등학교 공통사회 과목의 평가도구를 예시하면 다음과 같다(김정호 외, 1998).

① 국토 및 지역개발에서 개발과 환경 보전간의 갈등적 입장에 대해 입장을 정리한 후 토론하기

4) 실기시험

수행평가에서 말하는 실기시험은 종래에 말하던 실기시험과 그 강조점에 있어서 다소 차이가 있다. 종래의 실기시험과 수행평가에서 말하는 실기시험의 차이는 시험을 상황이 통제 혹은 강요된 상황인가, 아니면 자연스러운 상황인가의 차이에 기인한다. 종래의 실기시험에서는 평가가 이루어지는 상황이 통제되거나 강요되는 경우가 대부분이었다. 예컨대 종래의 농구 실기시험의 상황을 보면, 공을 5개 주고 슛을 하게 하여 골인된 공의 개수와 슛을 하는 자세를 보고 평가하거나, 두 사람이 나와서 서로 여러 번 패스하게 한 후 이를 보고 평가하는 경우가 대부분이었다. 그러나 수행평가의 실기시험에서는 가능한 한 학급 또는 학교 대항 농구 시합에서(통제되거나 강요되는 상황이 아니라 자연스러운 상황에서) 학생이 실제로 하는 것을 여러 번 관찰하여 실제 수행 능력을 평가한 것이다.

수행평가를 위한 실기시험에서 교수·학습활동과 평가활동은 가능한 한 분리되지 않는다. 종래에도 초등학교 과학시간에 ‘알코올 램프에 불을 켜고 끄기’를 제대로 할 줄 아는지를 평가하기 위해, 종종 실기시험이 실시되었다. 그러나 대부분의 경우 그것은 학생들을 서열화하기 위한 선발형 평가방안으로 사용되지는 않았다. 즉 별도의 시험시간을 정해놓고 평가

의 공정성을 높인다는 구실로 학생들이 다른 동료가 어떻게 하는지 보고 배울 수 없도록 칸막이를 해 놓고 그 안으로 한 학생씩 불러서 알코올 램프에 불을 켜고 끄도록 하였다. 그리고 시험을 치른 학생은 아직 시험을 치르지 않은 학생들과 서로 이야기하지 못하도록 하였다. 이에 반해, 수행평가를 위한 실기시험에서는 교수·학습활동과 상호 통합적으로 진행되는 것을 바람직하다고 생각하기 때문에, 별도의 시험시간을 정하지 않고 교사가 ‘알코올 램프에 불을 켜고 끄기’를 학생들에게 가르치는 수업시간에 개별 학생들을 관찰하면서 지도·조언해주는 과정이 곧 평가활동이 된다. 또한 잘 못하는 학생이 잘하는 학생을 보고 배울 수 있도록 학생들 간의 상호작용을 적극적으로 장려하기도 한다.

자연스러운 상황 하에서의 실기시험은 음악, 미술, 체육, 실과(기술·가정)분야에서 많이 이용되는 방법이나 요즈음은 읽기, 듣기, 쓰기, 말하기 등 언어 분야에서도 많이 활용되고 있다. 예컨대 직접 학생들에게 말하거나 읽거나 듣거나 쓰게 한 후에 보통 2명 이상의 교사나 채점자가 평가하는 방식이 있을 수 있다. 그 한 가지 예로 영어의 ‘말하기’ 영역에 대한 수행평가를 실시하는 방법을 제시하면 다음과 같다.

미국 버클리(Berkeley) 대학에서는 영어를 모국어로 사용하지 않는 학생들의 영어 말하기 영역에 대한 수행평가를 하기 위해 채점 기준표를 사용하고 있는데 <표 II-12>와 같다. <표 II-12>에서는 영어의 ‘말하기’라는 영역을 분석하여 그 하위 내용으로 ‘발음’, ‘유창성’, ‘문법’, ‘어휘’, ‘조직’, ‘듣고 이해하기’, 그리고 ‘질의 응답’으로 나누고 있다. 그 후에 각 하위 내용에 대한 성취 수준을 5단계(0~4점)로 나누고 있다. 이 때 ‘말하기’ 영역의 전체적인 성취수준은 각 하위 내용의 성취수준의 점수를 합한 것의 평균이 된다. 학교 당국에서는 이러한 채점 기준표를 제작함에 있어서 가능하면 모든 사람들이 합의할 수 있을 만큼 객관적이면서도 타당하게 되도록 노력했을 뿐만 아니라, 채점의 공정성을 유지하기 위하여 채점자 훈련을 여러 번 시행한 후 실제 평가에 활용하고 있다.

또한 학교 당국은 평가 결과를 보고할 때, 총점뿐만 아니라 각 하위 영역에서 개별 학생이 성취한 평정 점수를 채점 기준표와 함께 제공함으로써, 교사는 물론 개별 학생도 각 하위 영역에서의 성취수준에 대한 자세한 정보를 알 수 있도록 하고 있다. 교사는 이러한 정보를 특정 학생에 대한 ‘말하기’ 영역에서의 강·약점을 구체적으로 파악하고, 해당 학생에게 적절한 교육적 처방을 개별적으로 내리는 데 활용하고 있으며, 학생은 이러한 정보를 자기 자신을 제대로 인식하고 차후의 학습활동을 하는 데 활용하고 있다. 예컨대 A라는 학생은 7개의 하위 영역 중 ‘발음’과 ‘유창성’에서는 4점을 받았고, ‘문법’이나 ‘어휘’에서는 0점을 받았고, 나머지 3개의 하위 영역에서는 각각 1점을 받아 총점 11점이 된 반면, B라는 학생은 ‘문법’이나 ‘어휘’에서는 모두 4점을 받았고, ‘발음’과 ‘유창성’에서 0점을 받았고, 나머지 3개의 하위 영역에서 각각 1점을 받아 총점 11점을 받았다고 하자. 이 A학생과 B학생의 경우 총점은 11점(평균은 1.57로 2수준과 3수준의 사이)으로 동일하지만, A학생을 지도할 때는 ‘문법’이나 ‘어휘’의 지도에 초점을 두어야 할 것이며, 반면에 B학생을 지도할 때는 ‘발음’과 ‘유창성’의 지도에 초점을 두어야 할 것이다. 이와 같이 수행평가를 위한 영역별 채점 기준표는 비록 동일한 총점을 얻은 A, B 두 학생일지라도 그들 각각에게 서로 다른 적절한 교육적 처방을 내리는 데 적극적으로 활용될 수 있다.

<표 II-12> 영어 말하기 영역의 수행평가를 위한 채점 기준표의 예

수준 소영역	1수준(0점)	2수준(1점)	3수준(2점)	4수준(3점)	5수준(4점)
발음	무슨 말인지 이해가 불가능하다.	틀리게 발음하는 경우가 잦다; 영어를 모국어로 사용하지 않는 사람들과 접촉이 잦은 모국어 사용자만이 이해할 수 있다.	가끔 틀리게 발음하는 경우도 있으나 노력하면 이해할 수 있을 정도이다.	엑센트가 어색하기는 하지만 알아듣는 데 무리가 없다; 모국어 사용자와 거의 비슷하게 발음한다.	잘못 발음하는 경우가 거의 없다.
유창성	너무 더듬거려 대화가 불가능하다.	일상적인 표현을 빠고는 부자연스럽게 말한다.	유창하지는 않으나 자신 있게 말한다; 주저하며 말한다; 약간의 바뀌 쓰기를 한다; 문체가 고르지 못하다.	어는 정도 유창하게 말한다; 거의 더듬지 않는다; 바뀌 쓰기나 완곡한 표현을 쉽게 한다.	상당히 유창하게 말한다; 별로 힘들이지 않고, 부드럽게 말한다.
문법	전혀 문법에 따르지 않는다.	자주 문법상의 실수를 한다; 영어를 모국어로 사용하지 않는 사람들과 접촉이 잦은 모국어 사용자만이 이해할 수 있다.	단순한 문장으로 의미전달이 정확하다; 복잡한 문법의 사용은 피하거나 잘못 표현된다.	모든 기본구조들을 사용하여 말한다; 복잡한 구조도 사용한다; 종종 틀린 유형을 사용하기도 하지만 의미전달은 정확하다.	아주 드물게 문법 상 실수를 한다; 잘못된 유형을 사용하지 않는다; 어려운 수준의 문장구조를 사용한다.
어휘	단순한 말조차도 부적절한 어휘를 사용한다.	특정한 일을 표현하기에 필요한 어휘가 부족하다.	특정한 일에 대해 완곡하고 간단하게 표현하기에 충분한 어휘를 사용한다.	특정한 일에 대해 적당한 어휘를 사용한다; 문장 상 중요하지 않은 어휘를 가끔 틀리게 사용한다.	광범위하고 적절한 어휘를 사용한다; 세세한 어휘를 사용하여 구체적인 일을 표현한다; 정확한 의사전달을 위해 어휘를 부연 설명한다.
조직	설명을 이해하기가 불가능하다.	전체적인 설명구조가 불명확하다; 사고의 과정을 이해하는데 어려움이 있다.	단순하고 명확하게 설명한다; 전환상의 어색함이 있다; 부연하는 예가 부족하다; 꼭 필요한 부연설명이 부족함이 있다.	충분히 자세하고 명확하게 설명하나, 때때로 주제에서 벗어나기도 한다.	논리적 설명이 완벽하고 명확하며 잘 조직되어 있고, 명확한 표현을 사용하여 말한다; 부연 설명이 충분하다.

듣고 이해하기	단순한 일대일 상황에서도 듣고 이해하는 데 문제가 있다.	설명이나 반복에도 불구하고 오해의 소지가 있다.	친숙한 주제로 일대일 의사소통을 할 때는 별 문제가 없다; 잘못된 부연설명이 필요하다.	대부분의 이야기를 이해할 수 있다; 때때로 설명이나 반복이 필요하다.	행정가나 대학생의 이야기를 완전히 이해한다.
질의 응답	질문을 이해하지 못한다; 오해를 해명하기 위한 아무런 시도도 하지 않는다.	종종 질문을 이해하지 못한다.; 비논리적으로 응답한다.	응답하는 데 어려움이 있다; 불명확하게 진술된 질문은 잘 다루지 못한다.	명확하게 응답한다; 반응할 때 약간 지체하기도 한다; 불명확하게 진술된 질문에 대해서 혼돈을 일으킨다.	질문을 명확하게 이해한다; 오해가 생기면 쉽게 해명한다; 지체없이 응답한다.

5) 실험·실습법

실험·실습법은 자연과학 분야에서 많이 사용하는 것으로, 학생들로 하여금 직접 실험·실습을 하게 하고 그에 대한 과정이나 결과에 대한 보고서를 쓰게 하여 제출하게 하고, 제출된 보고서와 함께 담당 교사가 학생들이 실험·실습을 하는 과정을 관찰했던 것은 모두 고려하여 평가하는 방법을 말한다. 이 때 개인 단위로 실험·실습을 하게 할 수도 있고 팀을 구성하여 공동 작업으로 하게 할 수도 있다. 이러한 방법은 실험·실습을 위한 기자재의 조작 능력이나 실험·실습 태도뿐만 아니라, 지식을 적용하는 능력이나 문제해결 과정에 대해서 포괄적이면서도 종합적으로 평가할 수 있다. 참고로, 실험·실습법을 사용한 고등학교 공통과학 과목의 평가도구를 예시하면 다음과 같다 (이양락 외, 1998).

① 빗면 위에서 역학 수레를 움직이게 한 다음 이의 운동을 시간기록계를 이용하여 기록하고, 이의 결과를 분석하여 물체의 속력과 가속도를 구해 보자.

6) 면접법

면접법이란 평가자와 학생이 서로 대화를 통해서 얻고자 하는 자료나 정보를 수집하여 평가하는 방법이다. 즉 평가자가 학생과 직접 대면하여 평가자가 질문하고 학생이 대답하는 과정을 통해 지필식 시험이나 서류만으로는 알 수 없는 사항들을 알아보고 그것을 평가하는 방법이다. 앞에서 설명한 구술시험도 평가자와 학생이 대면하여 질문하고 대답하는 과정을 거치나, 구술시험은 주로 인지적인 영역을 중심으로 한 학업성취도를 평가하기 위한 방법이라고 한다면, 면접법은 주로 정의적인 영역이나 신체적인 영역에 대한 것을 평가하기 위한 방법이라고 할 수 있다. 면접하는 방법은 한 명의 평가자와 한 명의 학생이 일 대 일로 하는 면접, 여러 명의 평가자와 한 명의 학생이 다수 대 일로 하는 면접, 반대로 일 대 다수가 하는 면접, 마지막으로 다수 대 다수가 하는 면접 등이 있다. 면접법의 장점으로 보자 심도 깊은 정보를 얻을 수 있다는 점, 사전에 예상할 수 없었던 정보나 자료를 얻을 수 있다는 점, 진행상 융통성을 발휘할 수 있다는 점 등을 들 수 있다. 특히 최근에는 여러

직장에서 신입사원을 선발할 때, 열린 면접이라고 하여 평가자와 응시자가 함께 어울려 식사도 하고 운동도 하고 회의도 하면서 자연스러운 상태에서 응시자의 사회성이나 창의력, 조직 적응력 등을 평가하기도 한다.

면접법의 한 예로 서울에 있는 ○○대학교 사범대학 신입생 선발을 위한 면접시험에서 사용한 질문 문항의 예는 다음과 같다.

- ① 자신의 장·단점에 대해 간단히 말해 보시오.
- ② 사범대학을 지원한 이유를 간단히 말해 보시오.
- ③ 특별히 잘하는 과목이나 특별히 흥미를 가지고 있는 과목은 무엇입니까?
- ④ 대학에 입학한다면, 어떤 동아리 활동에 참여하고 싶습니까? 그 이유는 무엇입니까?
- ⑤ 졸업 후, 교사가 된다면 어떠한 일을 가장 먼저 하고 싶습니까?
- ⑥ 취미나 특기가 무엇입니까?
- ⑦ 최근에 감명 깊게 읽은 책이 있습니까? 그 내용은 무엇이고 특별히 어떤 감명을 받았습니까?
- ⑧ 우리 나라 교육문제 중에 가장 심각한 것이 무엇이라고 생각하십니까? 그 문제를 어떻게 해결하는 것이 바람직하다고 생각하십니까?

7) 관찰법

관찰은 학생을 이해하고 평가하기 위한 가장 보편적인 방법 중의 하나이다. 교사들은 늘 학생들을 접하고 있으며 개별 학생 단위로나 집단 단위로나 항상 관찰하게 된다. 예컨대 학생들 간의 사회적 관계 구조를 파악하기 위해 한 집단 내에서 개인 간 또는 소집단간의 역동적 관계를 집중적으로 관찰할 수 있다. 특히 나이가 아주 어리거나 지적 능력이 지나치게 낮은 학생들을 대상으로 평가하기 위해서는 평가하기 위한 상황을 의도적으로 마련할 수 없는 경우가 많기 때문에 인위적이 아닌 자연적인 상황에서의 관찰법을 자주 사용하게 된다. 객관적이고 정확한 관찰을 위해서는 관찰 대상을 있는 그대로 기술하는 일화기록법이나, 체크리스트(check list)나 평정 척도 등을 이용하기도 하고, 경우에 따라서는 비디오(video) 녹화를 한 후 분석하기도 한다. 참고로 국어 수업 시간에 수시로 학생들에게 ‘읽기’를 시킨 후에 개별 학생의 읽기 능력을 자연스러운 관찰을 통하여 평가하기 위해 <표 II-13>과 같은 평정 척도를 사용할 수 있다 (박인기, 1995).

<표 II-13> ‘읽기’ 능력을 평가하기 위한 평정 척도

문항	매우 아니다	아니다	보통 이다	대체로 그렇다	매우 그렇다
1. 글의 형식과 성격을 반영하며 읽고 있는가?	1	2	3	4	5
2. 글의 내용을 잘 통제·조정하며 읽고 있는가	1	2	3	4	5
3. 글을 표현 전달하는 연출적 기능이 잘 발휘되고 있는가?	1	2	3	4	5
4. 문장 간의 연결, 단락간의 연결이 자연스러운가?	1	2	3	4	5
5. 글의 총체적인 분위기를 반영하며 읽고 있는가?	1	2	3	4	5
6. 낭독 과정의 적절한 곳에서 적절한 호흡을 구사하고 있는가?	1	2	3	4	5
7. 발음과 억양은 정확한가?	1	2	3	4	5
8. 청중들의 반응은 좋은가?	1	2	3	4	5
9. 바깥으로 나타나는 유창력은 좋은가?	1	2	3	4	5
10. 낭독 후 내용 요약은 잘하는가?	1	2	3	4	5

8) 자기평가 및 동료평가 보고서법

자기평가(self-evaluation) 보고서법이란 개별 학생 스스로가 특정 주제나 교수·학습 영역에 대하여 학습 과정이나 학습 결과에 대한 자세한 자기평가 보고서를 작성·제출하도록 한 다음, 그것을 이용하여 교사가 평가하는 것이다. 예컨대 학생들에게 어떤 교육프로그램을 이수한 후, 그 프로그램을 이수하면서 자기 스스로의 학습과정이나 학습 결과에 대해 자기평가 보고서를 작성·제출하도록 하여 그것을 교사가 평가하는 것이다. 이러한 자기평가 보고서는 학습자로 하여금 자신의 학습 준비도, 학습 동기, 성실성, 만족도, 다른 학습자들과의 관계, 성취수준 등에 대해 스스로 생각하고 반성할 수 있는 기회를 제공할 뿐만 아니라, 교사로 하여금 교사가 시행한 해당 학습자에 대한 관찰이나 수시로 시행한 평가가 타당하였는지를 비교·분석해 볼 수 있는 기회를 제공하게 된다. 이와 유사한 것으로 동료평가(peer-evaluation) 보고서법이 있는데, 이는 동료 학생들이 상대방을 서로 평가하도록 하여 동료평가 보고서를 작성·제출하도록 한 다음, 그것을 이용하여 교사가 평가하는 것이다.

특히 학생 수가 많아서 담당 교사 혼자 힘으로 모든 학생들을 제대로 평가하기 어렵다고 판단될 때, 자기평가 보고서와 동료평가 보고서를 이용하여 학생을 평가하도록 한 다음 그 결과를 성적에 반영하는 방법도 있다. 예컨대, 체육 실기시험을 치를 때, 동료 학생으로 구성된 5명의 평가위원들에게 학생들의 자세나 동작을 평가하도록 한 다음 가장 높은 점수와 가장 낮은 점수를 제외한 나머지 3명의 평가위원의 점수를 학생의 성적에 반영한 경

우이다. 이때, 최고점과 최하점을 뺀 3명의 점수를 평균한 것과 교사 자신이 직접 평가한 점수를 각각 50:50 혹은 30:70 등으로 하여 학생의 최종 성적으로 사용한다면 교사의 주관성을 배제할 수 있을 뿐만 아니라 성적처리 방식에 대해서 학생들이 매우 공정하다고 인식할 것이다. 참고로, 특정한 주제에 대해서 연구를 수행하도록 한 후, 개별 학생들이 작성하도록 만든 자기평가 보고서의 양식을 예시로 제시하면 <표 II-14>와 같다.

<표 II-14> 자기평가 보고서의 양식 (예시)

평가요소*	평가 관점	점수			비고
		그렇다	보통이다	아니다	
준비도 및 성실성	선정한 연구 주제에 대한 보고서를 작성하기 위하여 충분한 시간과 노력을 부여하였는가?	2	1	0	
학습 동기	이 연구를 수행하는 것이 나 자신에게 꼭 필요한 것이었는가?	2	1	0	
성취감 및 만족도	이 연구를 통해서 많은 것을 배우게 되었으며, 연구 결과에 대해 스스로 만족스러운가?	2	1	0	
다른 학생과의 관계	이 연구를 수행하는 과정에서 다른 학생들과 상호 토론을 하거나 협력 학습을 하였는가?	2	1	0	
자료 수집	관련되는 자료나 정보를 만족할 만큼 다양하고 충분하게 수집하였는가?	2	1	0	
자료 분석 및 종합	수집한 자료나 정보를 분석하거나 종합한 것에 대해 스스로 만족하는가?	2	1	0	
보고서 작성 방법	보고서 작성법에 맞도록 보고서를 작성하였는가?	2	1	0	
계					

* 평가요소 간 가중치를 다르게 줄 수도 있음.

9) 연구 보고서법

연구 보고서법이란 개별 과목과 관련되거나 범교과적인 연구 주제 중에서 학생의 능력이나 흥미에 적합한 주제를 선택하여, 그 주제에 대해서 자기 나름대로 자료를 수집하고 분석·

종합하여 연구 보고서를 작성·제출하도록 하여 평가하는 방법을 말한다. 이 때 연구는 그 주제나 범위에 따라 개인적으로 할 수도 있고, 관심 있는 학생들이 함께 모여서 소집단별로 할 수도 있다. 이러한 연구 보고서법을 흔히 프로젝트(project)법이라고도 하며, 학생들은 연구를 수행하고 보고서를 작성하는 과정에서 연구하는 방법, 각종 정보를 수집하는 방법, 다양한 자료를 종합하고 분석하는 방법, 보고서 작성법 등을 익히게 될 것이다. 연구 보고서 발표회나 학생들간 연구 보고서의 상호 교환을 통해서도 많은 것을 배우게 될 것이다. 학생들의 연구 보고서를 평가할 때 사용할 수 있는 평가기준을 예시하면 <표 II-15>과 같다 (백순근 편, 1998).

<표 II-15> 최종 연구 보고서 평가 기준표 (예시)

평가 요소*	평가 관점	점수**		
		상	중	하
연구 영역 및 내용	보고서에 진술한 연구내용이 해당 국가의 여행에 얼마나 필요한 것인가?	2	1	0
자료 수집 능력	해당 국가에 관한 자료나 정보를 얼마나 다양하게 수집하였는가?	2	1	0
자료 분석 및 종합 능력	수집한 자료나 정보를 분석하거나 종합하는 능력은 어느 정도인가?	2	1	0
보고서 작성 방법	보고서 작성법에 맞도록 보고서를 작성하였는가?	2	1	0
계				

* 평가요소 간 가중치를 다르게 줄 수도 있다.

** 상·중·하 수준에 대해 그 내용을 구체적으로 진술하는 것이 바람직함.

10) 포트폴리오법

포트폴리오(portfolio)법이란 자신이 쓰거나 만든 작품을 지속적이면서도 체계적으로 모아둔 개인별 작품집 혹은 서류철을 이용한 평가 방법이라 할 수 있다 (백순근, 1995b). 예컨대 이것은 어떤 화가 지망생이 유명한 화가에게 지속적으로 지도를 받으면서 자신의 작품을 그린 순서대로 차곡차곡 모아 뒀으로써 자기 자신의 변화·발전 과정을 스스로 파악할 수 있도록 하고, 그 작품집을 이용하여 자기의 스승뿐만 아니라 다른 사람들에게도 쉽게 평가를 받을 수 있게 하는 것과 유사한 것이다. 그림뿐만 아니라 사진이나 시, 소설 등도 작품집을 이용하여 평가될 수 있으며, 글짓기한 것, 과목 과제물이나 연구 보고서나 실험·실습의 결과 보고서 등을 정리한 자료집을 이용하여 평가할 수도 있다. 포트폴리오를 통해 학생들은 자기 자신의 변화 과정을 알 수 있고, 자신의 강점이나 약점, 성실성 여부, 잠재 가능성 등을 스스로 인식할 수 있으며, 교사들은 학생들의 과거와 현재의 상태를 쉽게 파악할 수 있을 뿐만 아니라 앞으로의 발전 방향에 대한 조언을 쉽게 할 수 있다. 이 평가 방법은 단편적인 영역에 대해 일회적으로 전체적이면서도 지속적으로 평가하는 것을 강조하는 것으로 수행평가의 대표적인 방법 중의 하나로 각광받고 있다.

예를 들어 고등학교 1학년의 공통사회과 ‘각 지역의 생활’이라는 단원을 가르친 후 학생들에게 ‘각 지역의 생활에 관련되는 자료나 사진을 모으고, 정리하고, 편집하여 각 지역의 생활에 대한 홍보 자료집(즉 포트폴리오)을 만들어 제출하게 한 후, 그 결과에 대하여 <표 II-16>과 같은 평가 기준을 이용하여 평가할 수 있다.

<표 II-16> 홍보자료집을 위한 평가 기준표 (예시)

평가 요소*	점수**			비고
	상	중	하	
내용 구성 (홍보할 만한 내용으로 구성되어 있는가?)	2	1	0	
자료 수집 능력 (수집한 자료나 사진이 다양하고 충분한가?)	2	1	0	
자료 분석 및 종합 능력 (수집한 자료나 사진을 적절히 분석하고 종합하였는가?)	2	1	0	
조직력 (홍보물의 조직 및 편집은 제대로 되었는가?)	2	1	0	
호응도 (다른 사람들의 시선을 끌면서 설득력 있도록 제작되었는가?)	2	1	0	
계				

* 평가요소간 가중치를 다르게 줄 수도 있음.

** 상·중·하 수준에 대해 그 내용을 구체적으로 진술하는 것이 바람직함.

11) 기타

이미 지적한 바 있듯이, 창의력이나 문제 해결력 등 고등 사고기능의 신장을 중시하고, 교수·학습의 과정을 개선하는 것을 중시하며, 개별 학생에게 지도·조언하고 충고하기 위한 목적으로 사용되기만 한다면 어떠한 평가방법도 수행평가 방법에 포함될 수 있다. 특히 수행평가에서는 교수·학습 활동과 평가활동이 상호 통합적으로 진행되는 것을 강조하기 때문에 다양한 교수·학습 방법들이 곧 수행평가를 위한 평가방법이 될 수 있다. 예컨대, 도덕과나 사회과 등에서 많이 사용하는 역할놀이(role playing)는 교수·학습 활동을 위한 중요한 방법이 됨과 동시에 모여서 공동으로 수행하는 직소우(jigsaw) 등도 좋은 수행평가 방법이 된다. 이 밖에도 열린교육에서 교수·학습 방법으로 널리 사용되고 있는 신문활용교육(NIE : newspapers in education), 현장조사, 작품감상, 만들기, 전시회, 발표대회, 협력학습, 개념도(conceptual map) 만들기 등도 수행평가를 위한 좋은 평가 방법이 될 수 있는 것들이다.

참고로, 고등학교 공통과학 과목의 에너지 단원에서 개념도(개념들 간의 상호 관련성을 그림으로 나타내는 것으로, 구체적인 하위 개념들이 그 보다 일반적인 상위 개념에 통합되거나 포섭될 때 의미 있는 개념 학습이 효과적으로 이루어진다는 Ausubel의 이론에 근거하고 있음) 만들기 등을 이용하여 평가하기 위한 평가문항을 예시하면 다음과 같다 (이양락 외, 1998).

① 다음 개념 카드는 에너지 단원에서 나온 개념들을 나타낸 어휘들이다. 이것들을 가위로 잘라서 백지에 붙이고, 이를 선으로 연결하여 개념도를 작성하라. 이때 각 개념간의 관계가 명확히 표시될 수 있도록 선 위에 적절한 연결 어휘를 써넣어라.

마. 수행평가의 문항 제작 절차

수행평가에는 여러 가지 방법이 있다. 서술형 평가를 수행평가에서 열외로 하더라도 구술시험, 면접, 토론, 보고서, 실험실습, 포트폴리오, 등이 있다. 수행평가의 본질을 가장 잘 구현하고 있고 동시에 평가방법으로서도 널리 사용되고 있는 구술평가의 문항제작을 다루고자 한다. 구술시험은 종이와 붓이 발명되기 전부터 시행되어 오던 가장 오래된 수행평가의 한 형태이다(백순근 1999). 이 장에서는 구술평가를 질의응답, 구두발표, 면접, 그리고 토론을 포함하는 것으로 정의한다. 이 오래된 미래형 평가인 수행평가(구술)를 시행하기 위해서 교사들은 학기 초에 계획된 교과 진도와 시험출제 범위 그리고 수행평가의 비율을 기준으로 수행평가를 시행한다.

1) 수행형(구술) 평가문항의 제작절차

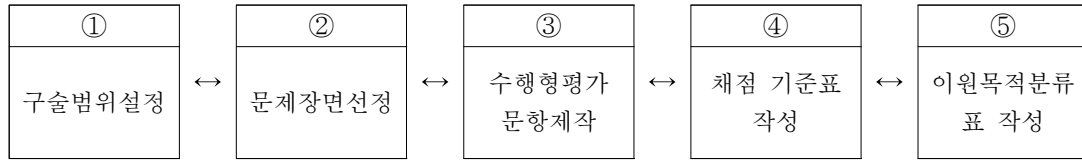
타당도와 신뢰도를 갖춘 수행형 평가문항을 제작하려면 그 절차는 어떠한가?

가) 일상적 수행형(구술) 문항 제작절차([그림 II-15] 참고)

교육현장에서는 일반적인 학교상황에서 수행형(구술) 평가문항 제작절차는 대체로 다음과 같다.

- ① 수행형(구술) 평가문항은 수업진도 계획에 따라 주어진 교과서 진도에 따른 시험범위로 설정한다.
- ② 수행형(구술) 평가문항은 시험범위 중 교사 개인이 가르친 것 중에서 중요하다고 생각하는 것을 문제 장면으로 상정한다.
- ③ 학생들에게 제시될 수행평가의 질문문항들을 제작한다. 이때 문항은 성취기준과는 별개로 대체로 글의 소재에 따라 결정된다. 질문문항은 교사개인이 중요하다고 생각하는 것주를 이루며 대체로는 표준적 절차 없이 만들어 진다.
- ④ 수행형(구술) 평가문항을 제작과 함께 채점 기준표가 작성된다. 일반적으로 수행형(구술) 평가에서는 서술형 평가에 비해서 평가문항보다는 오히려 채점 기준표가 평가에 더 중요한 역할을 한다. 왜냐하면 평가요소가 성취기준에 따르기 보다는 준비도, 표현력, 발표태도, 유창성, 논리력 등의 요소가 평가대상이 되기 때문이다(<표 II-17> 참조).
- ⑤ 이원목적분류표의 작성은 평가문항이 제작된 후에 이루어진다. 따라서 내용영역, 성취기준, 행동영역, 난이도, 그리고 배점은 마지막에 형식적으로 이루어진

다. 문항제작절차적인 측면에서 보면 이원목적분류표를 작성하는 순서가 가장 뒤에 위치한다([그림 II-15] 참고)



[그림 II-15] 일상적 수행평가 및 수행형 평가 문항 제작 절차

<표 II-17> 수행평가 채점 기준표

평가 요소	점수			비고
	상	중	하	
준비도				
이해력				
조직력				
표현력				
의사소통능력				
발표 태도				
계				

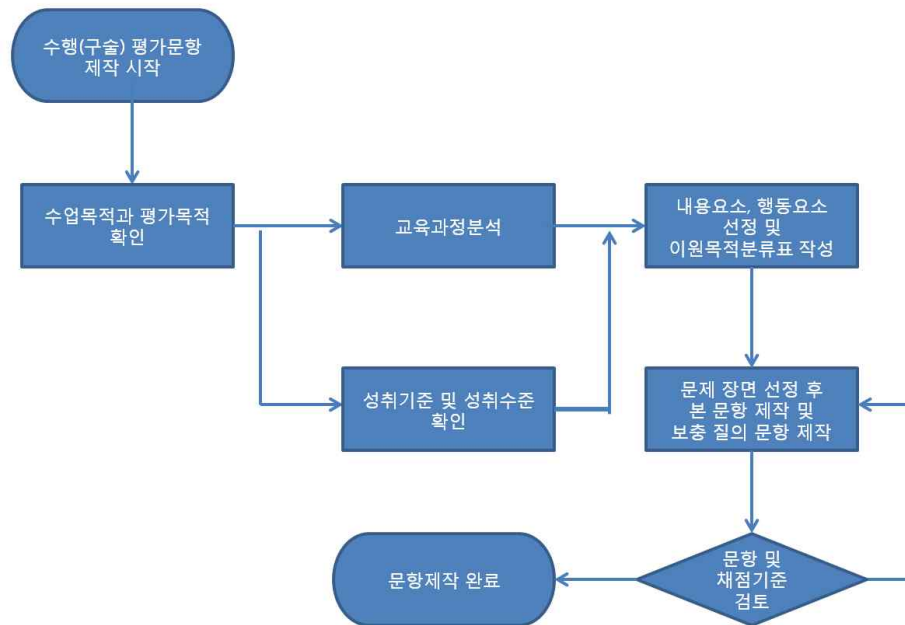
나) 이상적 수행형(구술) 문항 제작절차([그림 II-16] 참고)

- ① 수업 및 평가 목적 확인: 일반적으로 수행형(구술)평가의 목적도 서술형 평가와 같이 교수학습의 개선과 학생능력의 향상이다.
- ② 교육과정 분석: 각 교과목의 중·소단위와 내용에 대한 확인과정이다. 각 교과목의 교육목표와 내용을 분석하여 수행형(구술) 평가문항을 제작하기 위한 방향을 설정한다.
- ③ 성취기준 확인과 성취수준 설정: 성취기준을 확인하고 성취기준에 따른 상·중·하 성취수준을 설정하여 문제의 난이도를 가늠하는 과정이다. 수행형 평가문항을 제작하기 위해서, 특히 구술평가를 위한 성취수준의 타당도가 학생들의 능력수준과 어울리도록 설정되어야 한다. 이 과정은 이원목적분류표를 작성하는 기초적인 과정이기도 하고 평가문항 출제과정에서 가장 중요한 과정중 하나이다. 교육과정에서 목표와 내용이 잘 조직되었다 하더라도 그 목표와 내용을 ‘평가’하기에 적합하도록 진술하는 것이다(김중수 외, 2011). 교육과정의 목표가 수업에서 학생들이 배워야 할 목표가 ‘학습목표’로 진술된다면 학생들이 도달해야 할 평가목표가 ‘성취기준’이며 그 정도가 ‘성취수준’이다. 따라서 양질의 수행형(구술) 평가문항을 만들기 위해서는 피험자가 도달해야 할 목표와 내용을 구체적으로 진술하여야 하며, 각각 교육목표에 대하여 최소성취기준을 제시하는 것도

필요하다.

- ④ 핵심성취기준 선정: 수행형(구술) 평가의 특징은 일회성과 휘발성이 강하다. 따라서 성취기준 가운데에서 기본적인 중요한 내용영역과 행동영역에 대한 기준을 선정하는 핵심성취기준의 선정이 중요한 과정이다. 국가성취기준에 제시된 모든 교육과정 내용은 평가요소가 된다. 그런데 국가성취기준에 제시된 모든 교육과정 내용이 양적으로 많을 뿐만 아니라 중요도에서 차이가 있고 다양하기 때문에 성취기준 가운데에서 일부만 평가요소로 선정할 필요가 있다. 이러한 이유로 평가문항을 제작하는 장면에서는 ‘핵심성취기준’은 최저기준으로서의 ‘핵심 교육목표’이고 학습자의 기본적 능력을 평가하는 기준이 된다.
- ⑤ 이원목적분류표 작성: 이원목적분류표는 1956년 Bloom이 발표한 교육목표분류학에서 제시된 것으로 측정하고자 하는 인지적 영역과 학습내용의 상호관계를 각각 세로줄과 가로줄로 나타낸 표이다. 행동과 내용의 이원분류표를 작성할 때에는 흔히 행동목표를 종렬에 열거하고 내용은 횡렬에 열거한다(강신천, 2014). 일반적으로 학교에서 성취수준을 평가하기 위해서는 목표의 상세화 및 성취 수준의 결정이 선행되어야 한다. 학교 현장에서는 출제자 스스로 평가 목표의 진술, 평가 기준인 성취수준을 결정해야 한다. 따라서 출제자는 학습내용을 중심으로 평가목표 및 평가 문항을 먼저 설계하고, 그러한 작업을 정리하기 위해서 이원분류표를 작성해야 할 것이다.
- ⑥ 문제 장면의 선정 후 문항 초안 작성: 초안 작성에 중요한 것은 문두를 작성하는 것인데 이는 앞서 이원목적분류표 작성에서 대부분이 이루어진다.
- ⑦ 본문항 및 보충질의 문항 제작과 검토: 수행형(구술) 평가문항제작에서는 본문항은 물론 특히 보충질의 문항을 개발하는 것이 필요하다. 동 교과와 협의가 중요하며 교사들 간의 문항에 대한 평가관을 일치시키는 과정으로 제3자적 관점과 학생의 관점에서 이루어져야 한다. 문항을 수정하고 보완하는 과정에서 채점기준도 수정·보완된다. 이때 교사들 간의 상호작용과 논의의 깊이와 폭이 평가문항의 수준을 높이며 교사들의 평가전문성을 향상시키는데 상당한 기여를 한다.
- ⑧ 최종 문항 완성: 최종 문항이 완성되었다 하더라도 수행형(구술) 평가은 유동성과 임기응변적 즉흥성이 상당한 영향을 행사한다는 것을 고려해야 한다.

이상의 수행형(구술) 평가문항을 제작하는 절차는 서술형 평가문항을 제작하는 절차와 별 차이가 없다. 그러나 수행형 평가는 즉시성, 일회성, 유동성이 강함으로 이를 고려한 문항 제작이 이루어져야 한다.



[그림 II-16] 이상적 수행형(구술) 평가문항 제작절차

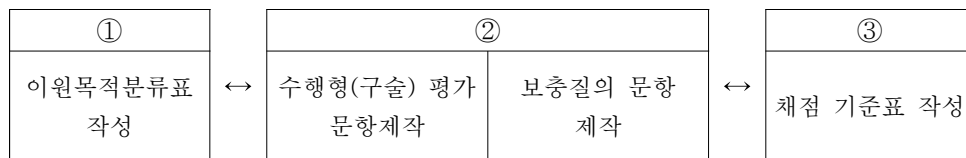
다) 핵심적 수행형(구술) 문항 제작절차

수행형(구술) 평가문항을 제작하는 과정에서 핵심적인 절차를 상정하는 것은 수행평가의 목적과 목표를 효율적으로 달성하기 위해 반드시 필요한 절차이다. 일선에서 평가문항을 제작하고 채점을 하는 교사입장에서 요구되는 핵심적 제작절차는 이원목적분류표의 내용을 얼마나 정교하게 분석하고 종합하는가와 각 문항에 대한 채점 기준표를 완성도 높게 만들어내느냐에 달려있을 것이다. 즉 이원목적분류표와 수행형 채점 기준표를 작성하는 과정이 수행형 평가문항을 제작하는 절차에서 가장 핵심적인 요소라고 할 수 있다.

- ① 이원목적분류표의 내용영역을 확인한다. 이는 교육과정 내용을 분석하고 평가할 지식을 확인하고 선정하는 과정이다. 내용영역은 중소단원의 중심내용이고 단원명이기도 하다. 예를 들어 중학교 영어의 교육과정 내용은 “①-1. 일상생활이나 친숙한 일반적 주제에 관하여 주요 내용을 묻고 답한다.”이다.
- ② 이원목적분류표의 성취기준을 확인한다. 내용영역이 좀 더 구체화되어 평가자와 학습자에게 제시되는 것이 성취기준이다. 이때의 성취기준은 세 가지로 제시되어 있는데 그 가운데에서 하나를 선택한다. 예를 들면 “영중9211-2. 학교생활이나 지역 사회 활동에 관하여 주요 내용을 묻고 답할 수 있다.”를 선택할 수 있다. 성취기준에는 평가지식과 함께 인지적 처리영역인 행동영역이 함께 기술되어 있다. 이때 평가문항 제작자는 지식영역을 압축하여 평가할 내용을 하나로 결정해야 한다.
- ③ 이원목적분류표의 인지적 처리영역인 행동영역을 확인하고 행동영역의 하위요소(기억, 이해, 적용, 분석, 평가, 창의) 가운데에서 하나를 확정한다. 이에 따라서 평가문항의 문제유형이 결정된다.
- ④ 이원목적분류표의 난이도를 성취수준을 고려하여 상, 중, 하 가운데에서 결정한다.

상이 가장 어려운 문항이다.

- ⑤ 이원목적분류표의 배점을 결정한다. 배점은 일반적으로 난이도가 높은 문항이 배점도 높다. 그러나 난이도와 함께 문항이 교육과정에서 차지하는 중요도를 고려하는 것이 필요하다. 핵심적이고 중요한 성취기준이면 쉬운 문항이라도 높은 배점을 해야 할 것이다.
- ⑥ 교과서 등 시험범위의 지문을 이용하여 문항을 출제하고 수행형(구술) 채점기준표를 작성한다. 채점기준표는 학교와 학과에 따라서 다양한 양식을 사용하고 있다. 채점기준표에 포함되어야 할 기본적인 내용은 문항번호, 예시답안 또는 모범답안, 채점기준, 그리고 배점은 필수 요소이며 그밖에 난이도, 평가영역, 문항내용 등을 포함하기도 한다. 특히 수행형(구술) 평가문항의 채점 기준에는 평가과정과 방법을 명확히 해야 한다.
- ⑦ 문항출제와 채점기준표 작성이 끝나면 평가문항을 검토하고 수정하여 완성하게 된다. 수행형(구술) 및 수행형 평가문항을 제작 후에 문항카드의 작성하기도 하나 앞서 작성한 채점기준표로 문항카드를 대신하는 것이 더욱 실용적이다. 일반적으로 문항카드에 기록되는 모든 요소들이 이원목적분류표와 채점기준표에 포함되기 때문이다.



[그림 II-17] 핵심적 수행형(구술) 문항 제작절차

<표 II-18> 2015학년도 ()학기 ()학년 (수행형)고사 이원목적분류표

문항	내용영역	성취기준	행동영역			난이도			배점	정답
			기억	이해	적용	상	중	하		
선택1										
선택2										
선택3										
선택4										
선택5										
수행1										
수행2										
수행3										

<표 II-19> 2015학년도 ()학기 ()학년 (수행형)고사 채점기준표

문항 번호	평가과정	평가문항	예시답안 및 채점기준					배점	
1	1) 준비시간 2) 발표시간 3) 보충설명 등		예시 답안						
			채점 기준	성취수준 성취기준	상	중	하	배점합	
				내용영역					
				행동영역					
전체									
2			예시 답안						
			채점 기준	성취수준 성취기준	상	중	하	배점합	
				내용영역					
				행동영역					
전체									

〈표 II-20〉 수행형 문항카드 A양식

<div style="text-align: center;"> ○○년도 ○학기 ○○학년 수행형 문항카드 A양식 </div>						<div style="text-align: center;"> 문항번호 </div>	
과 목 명		출제일	년 월 일	출제자	(인)		
출제영역	내 용 영 역		행 동 영 역				
성취기준			성취수준		난이도		
<문항내용>							
<모범답안 및 채점기준>							
출제 근거	도 서 명		저 자	발행처	발행년도	쪽수	
문항분석	난이도 (실제정답률)		소요시간				
	특이사항 및 총평						

<표 II-21> 수행형 문항카드 B양식

○○년도 ○학기 ○○학년

수행형 문항카드 B양식

과목명		제작일	년 월 일	출제자	㉠
-----	--	-----	-------	-----	---

□ 문항 번호. _____

출제 영역		평가목표	배점	난이도	소요시간
내용 영역	행동 영역				

문항 내용

모범 답안(예시)

채점 기준

출제 근거 : 도서명, 저자, 발행처, 발행년도, 쪽수 등 표기

특이사항 및 총평 : 문항의 수정 및 검토 과정, 시험의 실시 과정 등에서 나타난 특이사항이나 시험 실시 후 학생의 반응 자료나 결과를 분석하여 해당 문항에 대한 총평을 기입함.

⇒ 향후 동일한 출제 영역에서 문항을 출제할 때 구체적인 도움을 받을 수 있고, 이 과정을 거치면서 자연스럽게 문항 제작의 전문성이 신장될 수 있음.

바. 수행평가의 문항 채점 절차

수행형 평가는 객관식 선다형 평가와 비교할 때 채점자의 주관적인 판단이 개입될 가능성이 높다. 일반적으로 주관적인 판단은 검사 결과의 신뢰도를 떨어뜨린다고 생각하기 쉬우나 이것은 근거가 약한 주장일 수 있다. 채점 전문가가 행하는 수행형 평가가 객관식 선다형 평가보다 상황맥락적인 평가가 가능하며 기계적 처리를 하는 선다형보다 타당도는 물론이고 신뢰도가 동시에 높을 수 있다. 다만 채점자들이 타당하고 신뢰로운 채점을 하기위해서 자세하고 명확한 채점 기준을 정하고, 이것을 제대로 적용할 수 있는 전문적 평가소양을 개발하고 향상시키는 것이 선행되어야 할 것이다.

1) 채점기준

수행형 평가 문항의 채점기준은 수행형평가문항의 유형과 종류에 따라서 다양하여 하나의 기준으로 일반하기 어려운 점은 있지만 핵심적인 채점기준은 다음과 같다.

- ① 채점기준은 피험자 집단의 수행(구술)능력을 고려하여 설정하여야 한다. 채점기준을 정할 때, 피험자 집단의 능력에 대한 고려와 한 집단 내에서도 여러 그룹에 대한 고려가 선행됨으로써 평가문항은 피험자의 능력을 더욱 정교하게 측정할 수 있고 피험자인 학생이 더욱 신뢰하는 평가문항을 제작할 수 있다. 수행형 평가문항은 피험자의 반응도를 폭넓게 허용하기 때문에 피험자 집단이나 그룹의 능력에 따라 다양한 반응이 나타나며 그 한계를 분명히 파악하지 않으면 채점기준이 지나치게 엄격하거나 지나치게 느슨해서 평가문항의 질이 떨어지고 평가문항의 신뢰성이 저하될 우려가 있다. 예를 들면, “다음 글을 읽고 글쓴이가 주장하는 바를 한 문장으로 말하십시오.”라는 문항에서 하위능력의 학생이 사용하는 어휘와 문장구조와 상위능력의 학생들이 사용하는 어휘와 문장구조가 상이하다. 이러한 점이 채점기준에 고려되지 않으면 이 문항은 피험자의 능력을 변별하는 기능에 있어서 타당도가 낮은 문항이 될 수 있다.
- ② 채점기준은 성취기준과 성취수준에 근거하여야 한다. 채점기준은 평가문항의 유형이나 종류에 따라 다양하기는 하지만 기본적으로 성취기준 및 성취수준, 즉 학습목표와 그 목표에 학습자가 도달한 정도에 따라서 설정되어야 한다. 성취기준과 성취수준은 선택형평가에서와 마찬가지로 수행형(구술) 평가에서도 핵심적인 준거이다. 채점기준이 이를 벗어나서 설정될 때 그 평가문항자체가 타당도를 상실하게 되고 피험자의 능력을 잘못 측정하는 문항이 된다. 예를 들면, 성취기준은 “영중9252-1. 개인생활이나 가정생활에 관한 자신의 느낌이나 의견을 말할 수 있다.”인데 채점기준은 느낌이나 의견이 아닌 개인생활이나 가정생활의 현실을 기준으로 할 경우가 성취기준과 채점기준이 정합성을 상실하게 된다.
- ③ 채점기준은 이원목적분류표에 따라 명확히 설정해야 한다. 채점기준은 이원목적분류

표의 내용영역과 행동영역을 상세화하는 수준에서 설정되어야 한다. 학교현장에서 이원목적분류표는 모든 평가문항 핵심요소의 총합이다. 내용영역, 성취기준, 행동영역, 난이도, 그리고 배점을 한 눈에 볼 수 있다. 채점기준은 내용영역은 성취기준 및 행동영역의 관계에서 난이도를 설정하고 난이도를 반영하는 배점이 이루어져야 할 것이다. 특히 피험자의 능력을 변별하기 위해서는 행동영역을 더 세분화하고 그에 따른 채점기준이 상세화될 필요가 있다. 예를 들면, 현재 대부분의 중등학교 교사들이 행동영역을 지식, 이해 적용의 3단계로 나누고 있다. 이를 6단계인 기억, 이해, 적용, 분석, 평가, 그리고 창의로 세분화 하고 그에 따라 채점기준을 상세화하게 되면 학생들이 고등사고능력을 개발하는데 더 많은 도움이 될 것이다.

- ④ 채점기준은 피험자의 수행(구술)능력향상을 위한 학습기준이 될 수 있어야 한다. 채점기준은 피험자의 학습을 위한 현실적 기준이 되어야 한다. 학습자들은 수행형 평가를 준비할 때 수행형문항의 채점기준에 따라서 평가문항을 이해하고 해결하려는 경향이 강하고 채점기준이 지시하는 방향으로 학습자들의 능력이 길러지게 된다. 따라서 채점기준은 지엽적이고 사소한 사항이 아닌 수행형평가의 목적에 맞닿아있는 핵심적인 이고 본질적인 사항을 학습하도록 설정되어야 한다. 예를 들어, 채점기준이 문법이나 단어의 뜻 등 단편적 지식을 중심으로 설정되어있으면 학생들은 그에 따른 능력을 배양하게 된다. 반면에 사물을 묘사하고 생각을 표현하는 것을 중심으로 채점기준이 설정되어있으면 학생들은 표현적이고 창의적인 능력을 키울 것이다.
- ⑤ 채점기준은 피험자의 수행(구술)능력을 자극할 학습동기가 될 수 있어야 한다. 채점기준이 피험자들이 학습을 지속으로 이어가는 외적 동기로서 작용하도록 채점기준을 마련해야 한다. 채점기준이 피험자의 입장에서 이해할 수 있도록 명확히 설명되어야 한다. 채점기준에 대한 자세하고도 친절한 설명은 피험자로 하여금 수행형(구술) 평가에 대한 두려움과 불신을 제거한다. 뿐만 아니라 채점기준이 학습을 위한 외적 동기로 긍정적인 역할을 할 것이다.
- ⑥ 채점기준은 집단 내의 피험자들 모두에게 공정해야 한다. 채점기준이 형식적으로나 외적으로는 차이가 없을지라도 채점자의 신체적 심리적 정신적 상태에 따른 편파성을 최소화하려는 제도적 장치가 필요하다. 예를 들어, 아래 예시 평가문항의 채점기준표(예시, 【영어 - 4】)를 보면 성취기준이 개인생활과 가정생활을 말하도록 되어 있다. 이 성취기준에 대한 평가문항은 학생의 사회경제적 배경에 따른 영향을 무시할 수 없을 것이다. 따라서 평가문항뿐만 아니라 채점기준에서 이에 대한 고려가 있어야 한다.
- ⑦ 채점기준은 서로 중복되거나 누락되는 부분이 없어야 한다. 채점기준을 설정할 때 기준들이 겹치거나 기준 간에 빠뜨리는 부분이 없어야 한다. 채점기준은 해당 평가문항의 모든 영역을 포괄해야 하며 동시에 소영역 간들은 상호배타적이어야 한다. 이 원칙을 위반하게 되면 채점을 할 수 없는 답안이 생기게 되거나 한 답안이 동시에 두 영역에 속하게 되는 현상이 생기게 되어 채점에 오류가 발생한다.

예시) 【영어 - 4】

학년군	중학교 1~3학년군	영역	말하기
성취기준	영중9252-1. 개인생활이나 가정생활에 관한 자신의 느낌이나 의견을 말할 수 있다.		
문항 유형	수행 평가		
출제 의도 및 평가 내용	개인생활이나 가정생활에서 겪을 수 있는 문제 상황에 대한 글을 읽고, 문제 해결에 도움이 되는 적절한 조언을 말할 수 있는가를 평가한다.		

평가문항

다음 글을 읽고, Mary에게 해줄 적절한 조언을 <조건>에 알맞게 말하십시오.

My name is Mary, and I'm 16 years old. Ever since I was a kid, I've felt that my mother likes my younger sister much more than me. My sister is beautiful, smart and does everything well. I'm just an ordinary girl and I don't like to hear my mom compliment my sister all the time. I'm always worried I'm too plain. What should I do?

*compliment: 칭찬하다

—<조 건>—

- ① 두 가지 이상의 조언을 말할 것.
- ② 다음의 언어형식 중 하나를 사용하여 완전한 문장으로 말할 것.
I think you should...
I think you had better...

예시

답안

- I think you should(had better) talk about your feelings with your mom.
- I think you should(had better) remember you are special, too.
- I think you should(had better) not be so hard on yourself.
- I think you should(had better) not compare yourself with your sister.
- I think you should(had better) be more confident.
- I think you should(had better) not think about it so much.

[그림 II-18] 수행평가 문항의 채점기준에 대한 예시

3. Bloom의 서술형 및 수행평가의 이해

본 연구에서는 학교현장 교사들의 서술형 및 수행평가 전문성 향상을 위한 도구로 Bloom의 개정 이원목적분류표(Anderson & Krathwohl 외, 2001)를 제시한다. 다만 현재 학교에서 교사들이 주로 사용하고 있는 Bloom의 이원목적분류표는 1956년에 발간된 것이고 이 분류법에는 내용이 일차원적으로 제시되고 인지적 영역 안에서 포괄하고 있다. 즉, 이원목적인 내용영역과 행동영역에서 행동영역의 인지적 과정을 중심으로 6단계, 지식, 이해, 적용, 분석, 종합, 평가로 분류하고 있다. 그러나 성취기준 중심의 교육과정 및 수행평가의 설계와 실행, 문제해결에 Bloom의 개념들이 실용적인 가치가 있다(Anderson & Krathwohl 외, 2001). 이러한 가치를 살리고 내용영역을 정교화하기 위해서 Bloom의 공동저자였던 Krathwohl이 Anderson과 함께 이 핸드북을 2001년에 개정하였다. 여기에서 지식 영역을 ‘사실적 지식’, ‘개념적 지식’, ‘절차적 지식’, ‘메타인지 지식’ 등 4차원으로, 인지영역은 ‘기억하다’, ‘이해하다’, ‘적용하다’, ‘분석하다’, ‘평가하다’, ‘창안하다’ 6유목으로 분류한다. 본 연구에서는 학교현장의 편의와 평가의 효율성을 위하여 지식영역은 ‘메타인지 지식’을 ‘절차적 지식’에 포함하여 3차원으로, 인지영역은 ‘기억하다’는 ‘이해하다’에 포함시키고, ‘분석하다’에 ‘평가하다’, ‘창안하다’를 포함시켜 3유목으로 분류하고 이를 바탕으로 평가유형을 9가지로 분류하고 이를 ‘Bloom의 평가 유형’으로 명명한다.

가. Bloom의 지식 차원의 분류

Bloom의 지식 차원의 분류는 <표 II-20>과 같다.

<표 II-20> Bloom의 지식 차원의 분류

주요 유형 및 하위 유형	예시
A. 사실적 지식-교과나 교과외의 문제를 해결하기 위해 숙지해야 할 기본적인 요소	
Aa. 전문용어에 대한 지식	전문용어 음악 부호 주요 자원
Ab. 구체적 사실과 요소에 대한 지식	신뢰로운 정보원
B. 개념적 지식-요소들이 통합적으로 기능하도록 하는 상위	구조 내에서 기본 요소들 사이의 상호관계
Ba. 분류와 유목에 대한 지식	지질학 연대 기업소유 형태 피타고라스 정리 수요와 공급의 법칙
Bb. 원리와 일반화에 대한 지식	진화론
Bc. 이론, 모형, 구조에 대한 지식	의회조직
C. 절차적 지식-어떤 것을 수행하는 방법, 탐구방법, 기능을	활용하기 위한 준거, 알고리즘, 기법, 방법
Ca. 교과에 특수한 기능과 알고리즘에 대한 지식	수채화를 그리는 기능

Cb.	교과에 특수한 기법과 방법에 대한 지식	정수 나눗셈 알고리즘 면접기법 과학적 방법
Cc.	적절한 절차의 사용 시점을 결정하기 위한 증거에 대한 지식	뉴턴의 제2법칙이 포함된 절차의 적용시 점을 결정하기 위한 증거, 사업비용 추정 방법의 실현가능성을 판단하기 위한 증거
D.	메타인지 지식-지식의 인지에 대한 인식 및 지식과 인지	전반에 대한 지식
Da.	전략적 지식	교재 단원의 구조를 파악하기 위한 수단 으로서 개요를 작성하는 지식, 발견법 활 용에 대한 지식
Db.	인지과제에 대한 지식(적절한 맥락적 지식 및 조건 적 지식 포함)	특정 교사가 실시하는 시험유형에 대한 지식, 과제의 인지적 요구에 대한 지식 논문을 비판하는 것은 개인적 강점이지만
Dc.	자기-지식	논문을 작성하는 것은 개인적 약점이라는 지식, 자신의 지식수준에 대한 인식

나. Bloom의 인지 단계의 분류

Bloom의 인지 단계의 분류는 <표 II-21>과 같다.

<표 II-21> Bloom의 인지 단계의 분류

인지과정	유목	관련된 용어	예시
1.	기억하기-장기기억으로부터	관련된 지식을 인출한다.	
1.1	재인하기	확인하기	제시된 자료와 일치하는 지식을 장기기억 속에 넣기(예, 미국사의 주 요사건들의 날짜를 재인한다.)
1.2	회상하기	인출하기	장기기억으로부터 관련된 지식을 인출하기(예, 미국사의 주요사건의 날짜를 회상한다.)
2	이해하다-구두, 문자, 그래픽을 포함한 수업	메시지로부터의 의미를 구성한다.	
2.1	해석하기	명료화하기 바꿔쓰기 표현하기 번역하기	하나의 표현형태(예, 숫자)를 다른 표현형태(예, 단어)로 바꾸기(예, 주 요 연설문이나 서류를 바꿔쓴다.)
2.2	예증하기	예를들기 실증하기	개념이나 원리의 구체적인 예나 범례 찾기(예, 다양한 미술양식의 예 를 든다.)
2.3	분류하기	유목화하기 포섭하기	사물의 특정 유목(예, 개념이나 원리)에 속한다는 것을 결정하기(정신 적으로 정리되지 않은 관찰되거나 기술된 사례를 분류한다.)
2.4	요약하기	추상하기 일반화하기	일반적인 테마나 요점을 요약하기(예, 비디오테이프에 나타난 사상들 에 대한 짧은 요약문을 쓴다.)
2.5	추론하기	결론짓기 외삽하기 내삽하기 예언하기	제시된 정보로부터 논리적인 결론을 도출하기(예, 외국어 학습에서 여 러 사례로부터 문법적 원리를 추론한다.)

2.6	비교하기	대조하기 도식화하기 결합하기	두 개의 아이디어, 대상들 간에 일치점을 탐색하기(예, 역사적 사건들을 현재 상황에 비교한다.)
2.7	설명하기	모델구성하기	인과관계 체제 모델 구성하기(예, 프랑스의 18세기 주요사건들의 원인을 설명한다.)
3.	적용하다-특정한 상황에서 어떤 절차를 사용하거나 시행한다.		
3.1	집행하기	시행하기	어떤 절차를 유사한 과제에 적용하기(예, 하나의 정수를 다른 정수로 나눈다.)
3.2	실행하기	사용하기	어떤 절차를 친숙하지 못한 과제에 적용하기(예, 뉴턴의 제 2법칙을 적절한 상황에 활용한다.)
4.	분석하다-자료를 구성부분으로 나누고, 그 부분들 간의 관계와 부분과 전체 구조나 목적과의 관계가 어떻게 되어 있는가를 결정한다.		
4.1	구별하기	변별하기 식별하기 초점화하기 선정하기	제시된 자료를 관련된 부분과 관련되지 않은 부분으로, 중요한 부분과 중요하지 않은 부분으로 구분하기(예, 수학적 문장제 문제에서 관련된 수와 관련되지 않은 수를 구분한다.)
4.2	조직하기	발견하기 정합성 찾기 통합하기 윤곽그리기 해부하기 구조화하기	요소들이 구조 내에서 어떻게 기능하는가를 결정하기(예, 역사적으로 기술된 증거들을 특정한 역사적 해설에 대한 찬반의 증거로서 구조화한다.)
4.3	귀속하기	해체하기	제시된 자료를 기반으로 하고 있는 관점, 편견, 가치 혹은 의도를 결정한다. (예, 저자의 정치적 관점에 따라 그의 관점을 결정한다.)
5	평가하다-준거나 기준에 따라 판단한다.		
5.1	점검하기	조정하기 참지하기 모니터하기 검사하기	과정이나 산출물 내부의 오류나 모순을 탐지하기, 과정이나 산출물의 내적 일관성 여부를 결정하기; 절차가 진행될 때 그 효과성을 참지하기(예, 과학자들의 결론이 관찰된 데이터로부터 도출되었는지의 여부를 결정한다.)
5.2	비판하기	판단하기	어떤 결과와 외적 기준간의 불일치 여부를 탐지하기, 어떤 결과가 외적 일관성을 가졌는지의 여부를 결정하기, 특정문제에 대한 절차의 적절성을 참지하기(예, 두 가지 방법 중 어느 것이 주어진 문제를 해결하는 최상의 방법인지를 판단한다.)
6.	창안하기-요소들을 일관되거나 기능적인 전체로 형성하기 위해 함께 둔다; 요소들을 새로운 패턴이나 구조로 재조직한다.		
6.1	생성하기	가설세우기	준거에 기반을 둔 대안적 가설을 제안하기(예, 관찰된 현상을 설명하기 위해 가설을 설정한다)
6.2	계획하기	설계하기	어떤 과제를 성취하기 위한 절차를 고안하기(예, 특정한 역사적 토픽에 관한 연구 보고서를 계획한다.)
6.3	산출하기	구성하기	어떤 절차를 만들어 내기(예, 특정한 목적을 위한 거주지를 건설한다.)

다. Bloom의 평가 유형

Bloom의 평가 유형은 <표 II-22>와 같다.

<표 II-22> Bloom의 평가 유형

유형	정의	예시
제1 사실이해	구체적 사실과 현상을 명료화하거나 바꾸어 표현하는 유형	다음 그림을 글로 표현하시오.
제2 사실적용	구체적 사실과 현상을 특정한 상황에 시행하거나 사용하는 유형	재료들을 1m 높이로 쌓으시오.
제3 사실분석	구체적 사실과 현상을 구별하거나 조직, 통합 또는 해체하는 유형	다음 조각들로 아파트를 조립하시오.
제4 개념이해	구체적 사실 및 현상간의 상호관계에 관한 지식을 명료화하거나 바꾸어 표현하는 유형	아래 자료에서 수요와 공급의 법칙을 추론하시오.
제5 개념적용	구체적 사실 및 현상간의 상호관계에 관한 지식을 특정한 상황에 시행하거나 사용하는 유형	진화론을 교실환경에 사용하시오.
제6 개념분석	구체적 사실 및 현상간의 상호관계에 관한 지식을 변별하거나 조직, 통합 또는 해체하는 유형	뉴턴의 제2법칙을 해체하시오.
제7 절차이해	교과의 특수한 기능이나 사용준거에 관한 지식을 명료화하거나 바꾸어 표현하는 유형	나눗셈 알고리즘을 설명하시오.
제8 절차적용	교과의 특수한 기능이나 사용준거에 관한 지식을 특정한 상황에 시행하거나 사용하는 유형	글의 의미를 파악하는 과정을 그림을 이해하는데 사용하시오.
제9 절차분석	교과의 특수한 기능이나 사용준거에 관한 지식을 구별하거나 조직, 통합 또는 해체하는 유형	글의 의미를 파악하는 과정을 조직화하시오.

4. 해외사례 분석

서술형 및 수행평가에 관한 교사 평가 전문성 신장 방안을 모색하기 위하여 미국과 핀란드에서 활용하고 있는 서술형 및 수행평가의 사례와 효율적 운영 방안(운영관리 행정체제 등)을 살펴보았다. 이 두 나라를 선정한 이유는 미국의 경우, 서구의 여러 나라들 중에서 서술형 및 수행평가의 운영과 이의 교사 전문성 신장에 관하여 서구 어느 나라들 보다 활발히 의견을 개진하고 있다. 특히 이 두 나라에서는 단위학교의 교육의 질 개선을 위하여 서술형 및 수행 평가에 관한 교사 평가 전문성을 국가차원에서 적극적으로 적용 및 운영에 노력하고 있다. 즉, 이 두 나라는 서술형 및 수행 평가에 관한 교사 평가 전문성 개선을 위해서 이의 운영 사업에 다양한 프로그램을 개발하여 적용해오고 있으며, 단위학교의 교사 평가 전문성뿐만 아니라 평생교육 및 인적자원 개발에의 활용 등 교사 전문성 신장을 통하여 단위학교의 교육 질 개선을 위해서 매우 활발하게 이루어지고 있다. 따라서 이 두 나라는 우리나라의 서술형 및 수행 평가에 관한 교사 평가 전문성 신장을 위한 검사도구 개발 및 운영을 위한 벤치마킹에 적합하다고 볼 수 있다.

또한 이 두 나라는 행정단위가 연방정부, 주 정부, 자유시, 광역단체, 기초단체 등의 자치제도는 우리나라의 지방교육 자치와 유사한 형태를 보인다. 즉, 이 두 나라의 자치 제도 역시 우리나라와 마찬가지로 일선 학교에 대한 자율성 및 책임성이 확대되고 있다. 특히, 핀란드의 광역단체와 기초단체는 교사 평가 전문성의 효율적인 운영 및 관리를 위한 전담부서가 마련되어 있다. 이를 통해, 일선학교의 교사가 학생들이 높은 교육적 성취를 달성할 수 있도록 돕는 디자인을 개발 및 운영관리하고 있다. 그리고 핀란드의 교사 평가 전문성은 여러 시행착오를 거쳐 현재, 서술형 및 수행 평가 전문성 척도를 개발·시행하고 있고, 호주, 홍콩, 스페인, 싱가포르와 중국 등지에서 이를 벤치마킹하여 개발·시행하고 있다. 무엇보다 이를 통해, 우리나라 교육현장에 보다 적합한 서술형 및 수행 평가에 관한 교사 평가 전문성 신장 방안 도출 및 효율적인 운영 방안을 위한 시스템을 구축하는데 활용하고자 한다.

가. 미국

1) 학생평가 이해

미국 중등학교(secondary schools)의 역할은 다음의 4가지로 정의하고 있다. 첫째, 지식을 습득하여 지성의 능력을 키운다. 둘째, 사회의 생산적인 일과 경제적 책임을 질 수 있도록 구직의 능력을 키워준다. 셋째, 복잡한 사회에 대처하는 사회성을 준비하는 사회적, 시민적 의식을 함양해준다. 넷째, 개인의 재능을 발굴하고 자유 의사표현을 발전시킨다는 등의 이슈가 있다. 미국 중학교(Junior High school; Intermediate School)는 7-9 학년으로, 교육과정은 영어, 수학, 사회, 과학 등 주요 과목 외에 외국어 등의 희망 선택과목을 학습할 수 있도록 구성되어 있다. 중학교부터는 한국의 대학 교과 선택 방식과 유사하게 학생 각자

가 강의 시간마다 학과 교실을 찾아가 수강하게 된다. 또한 학생들은 Home Room 시간을 통해 우리나라의 담임선생님 개념인 Home Room 선생님으로부터 생활지도를 받게 된다. 미국 고등학교(High School)는 9 - 12학년으로, 교과과정은 매우 다양하게 편성되어 있어 있다. 대학진학을 준비하는 학생, 졸업 후 취업을 준비 하는 학생 모두에게 필요한 과목을 제공하고 있다. 그러나 일반적으로 사립 고등학교의 경우는 공립의 포괄적인 교과과정과는 약간의 차이를 보이며 대부분의 사립학교는 대학진학 준비에 중점을 둔 College Preparatory 교과과정을 제공한다.

미국의 중등학교 교육제도는 기본적으로 평준화 정책을 따르고 있지만, 실질적으로는 학생의 재능에 따라 차별화된 교육이 이뤄지고 있다. 상위권 학생은 학력을 더욱 증진시키는 공통적인 프로그램 즉 고급반과 국제학사학위 등을 운영 하고 있다. 고급반은 일종의 선행 학습 프로그램으로 대학 수준의 수업을 제공한다. 이 제도는 SAT를 주관하는 칼리지보드의 승인을 받아야 한다. 국제학사학위는 인재양성을 목적으로 고등학교에서 미리 대학 수준의 수업을 듣고 학점을 인정받는 제도이다. 그리고 영재학교(schools for the gifted and talented)제도도 있다. 마그네틱 스쿨로 학군에 상관없이 학생을 선발하여 전문적으로 훈련, 저소득층이 많은 지역에 설치되어 있다. 반면, 하위권 학생을 위한 제도로는 일정한 학업수준까지 오를 수 있도록 정부(카운티)가 책임지는 제도를 갖추고 있다. 이는 교육관이 같은 부모와 시민단체가 카운티 정부와 협약을 맺고 운영하는 차터 스쿨, 집에서 공부하는 홈스쿨링, 클럽활동 등이 대표적이다. 주목할 점은 창의력을 무엇보다 중시하는 교육과정이 전반적으로 이루어져 있으며, 학기초반 커리큘럼 박람회를 열어 관심 분야를 조사하고 20명 이상의 학생들이 관심을 보인 분야는 새로 수업을 만들 수 있는 구조도 갖춰져 있다. 학생 평가는 방과 후 활동(스포츠 활동 등)과 다양한 활동(프로젝트의 리포트, 프레젠테이션, 시뮬레이션 등)으로 이루어진다. 미국 중등학교의 학생평가에 대하여 구체적으로 살펴보면 다음과 같다.

우리나라와 비교할 때 미국의 성적표의 가장 큰 특징은 준거지향 평가의 원칙에 입각하여 담당교사가 각 교과목이나 영역별 성취수준을 기록하며, 학생들의 상대적인 서열을 나타내는 석차를 기재하지 않는다는 것이다. 미국의 성적표는 교사의 절대적인 평가가 기본으로 되어 있고 평가 대상과 점수를 매기는 법도 학생의 발육 상태, 학습에 대한 적응도, 학교생활에서의 참여, 사회활동 등에 대해 자세하게 적는다. 또 다른 특징은 저학년으로 내려갈수록 행동 발달 상황이나 교과 학습 발달 상황에 대해서 자세히 기록하지만, 고학년으로 올라갈수록 성취수준을 등급화하고 있다는 점이다 (정홍희, 2014, 우리 가족 미국 교육 체험기 중에서).

즉, 이는 상위 학교 급으로 올라 갈수록 교과 학습 발달 상황을 자세히 기록하여 개별 학생에게 필요한 교육적인 처방을 하고 개별 학생의 교육과 관련하여 학부모와의 의사소통 수

단으로 활용할 수 있도록 하고 있다. 학업 성취수준에 관해서 점수 이외에도 공부 태도와 협동심의 수준을 표시하도록 하고 있다.

미국에서는 정답만 달랑 시험지에 적으면 채점자는 많은 점수를 주지 않는다. 정답보다는 왜 그렇게 생각했는지 어떻게 답을 산출했는지 이유와 풀이과정이 아주 중요한 것 같았다. 비록 정답을 정확하게 맞히지 못했더라도 나름대로 논리정연하게 풀이과정을 기술했다면 그 나름대로 의미 있는 것으로 간주하고 좋은 점수를 받을 수 있다. 이는 결과보다는 과정이 중요함을 시사하고 있다(정홍희, 2014, 우리 가족 미국 교육 체험기 중에서).

미국 중등교육에 있어 학생평가는 이런 과정의 중요성에 부합하여 진행된다. 예를 들어, 영어 작문 평가 시험은 창작영역을 제외한 분야의 영어 작문능력과 영어 문체와 수사학에 대한 이해에 대한 평가를 목적으로 3시간 동안 진행된다. 시험의 첫 부분은 한 시간 동안 진행되는데 학생들은 지문을 읽고 나서 지문의 의미, 목적, 구조, 문체, 용어에 대한 문제를 풀게 되며 나머지 두 시간은 설득적·분석적·기술적인 세 가지의 에세이를 작문하는 데 할애된다.

미국에서는 학교숙제가 많고 미국 학생들은 학교숙제에 많은 시간을 할애한다. 학교에서 내주는 숙제는 철저하게 검사되며 학생 성적으로도 연결된다. 학기 초 평가에 대해 안내가 나가는데 이때 시험성적 뿐 아니라 태도, 숙제가 평가에 많은 부분을 차지하게 된다(정홍희, 2014, 우리 가족 미국 교육 체험기 중에서).

뉴욕 주는 교과목 학습 표준을 충족한 것을 증명하기 위하여, 학생 과제 등의 모든 학생 기록에 대하여 학생들이 교과목 학습 표준(적절한 읽기쓰기 및 수학 등) 능력을 갖추기 위해 발전하고 있는지, 그리고 진척이 더더 관심을 기울여야 할 학생들에게 적절한 지원과 교육중재 전략이 전달되고 있는지를 확실히 하기 위해 정기적인 검토를 실시한다. 나아가 학부모들은 자녀의 교과목 학습 표준(적절한 읽기쓰기 및 수학 등) 능력 발전에 관한 정보를 제공받고 진척을 위해 참여하며 수업과 숙제 지원에 참여할 기회를 제공받는다. 교육 전략 및 중재를 위한 전략으로는 학생들의 적절한 교과목 학습 표준(적절한 읽기쓰기 및 수학 등) 능력을 충족하기 위하여 기초 읽기쓰기 표준, 뉴욕 주 수학 학습 표준, 조기아동 읽기쓰기 평가 시스템(ECALS-2), 연장학습-일 및 연장학습-년 수업, 개인지도와 기타 활동 등을 활용하는 것으로부터 이끌어 내고 있다.

미국 중등학교의 학생평가는 학업성적(SAT 점수 등)도 중요하지만 그 외의 클럽활동, 스포츠 활동, 봉사활동, 사회경험도 매우 중

요시 한다. 심지어 이런 것들이 대학진학에 있어서도 지대한 영향을 미친다. 즉, 대학교에서는 공부만 잘하는 학생을 원하지 않는다는 사실이다. 좋은 학교에 가려면 공부를 잘하는 것은 당연한 자격이기에 공부 이외에 무엇을 더 할 수 있는 학생인지 여부를 가린다고 한다. 특히, 미국 중등 교육에서는 운동을 교육의 중요한 부분으로 간주하고, 공부뿐만 아니라 운동이나 봉사 활동 같은 사회 활동도 중요시 여긴다. 또한 미국 중등교육에서 봉사활동은 필수다. 대충 시간만 채우는 그런 식의 봉사활동이 아니라 다른 사람을 위해 자신의 시간과 에너지를 사용하는 습관을 배우는 것을 중요시하고 있다(정홍희, 2014, 우리 가족 미국 교육 체험기 중에서).

학습결과물평가(COE)는 학생이 교사의 지도를 받으며 완성한 일련의 수업 학습물 표본을 평가하는 것이다. 즉, 각 학생의 표준화 포트폴리오는 이 학생의 평가 점수가 학생의 실제 학습성과를 정확히 반영하고 있는지 판가를 하는 중요한 잣대이다. 담임 교사는 각 학생의 진급 포트폴리오의 평가 점수가 학생의 학습 표준 습득 정도를 명확하게 반영하고 있는지 여부를 검토한다. 표준 절차를 활용하여 담임교사는 본인의 최상의 판단 하에 각 학생이 8학년에서 진급을 하기 위한 영어 학습 표준 및 수학학습 표준을 충족하기 위해 레벨 1, 레벨 2 또는 상위 레벨 2에 동일한 성취도를 보였는지 여부를 밝힐 수 있어야 한다. 일반적으로 미국에서 시행되어 온 교육평가는 평가의 결과인 점수가 어떤 목적으로 사용되는지에 따라 크게 네 가지 종류로 분류된다. 즉, 교육이 시작되기 전에 시행하여 학생이 교육이수에 필요한 선행지식과 수행 능력을 갖추고 있는지를 가늠해보는 배치평가(placement assessment), 교육기간 중에 일어나는 학업성취의 진행을 모니터링 하고자 시행하는 형성평가(formative assessment), 교육이수 기간 중 계속해서 학습어려움을 보이는 부분을 파악하고자 시행하는 진단평가(diagnostic assessment), 그리고 교육이수가 완료되는 시점에 시행하여 학생이 교육목표를 달성했는지를 보여주는 누적평가(summative assessment)가 있다. 다음 <표 II-23>은 위의 일련 내용을 종합정리하면서 2015년도 고등학교 졸업자 학력평가 요건(워싱턴 주)을 검토한 것이다.

<표 II-23> 2015년도 고등학교 졸업자 학력평가 요건(워싱턴 주)

영어(ELA)	과학	수학
ELA 요건 충족 방법 (하나를 선택):	과학 요건 충족 방법 (하나를 선택):	수학 요건 충족 방법 (하나를 선택):
독해 & 작문 고등학교능력시험 (HSPE)*	생물학 과정수료시험(EOC) 합격	1종의 수학 과정수료시험(EOC)

독해 & 작문 HSPE에서 2단계 (기초) 점수 취득(특수교육 학생에 한함)+	생물학 EOC에서 2단계(기초) 점수 취득(특수교육 학생에 한함)	수학 EOC 시험 또는 EOC 졸업 시험에서 2단계 (기초) 점수 취득(특수교육 학생에 한함)
워싱턴 주 대안적평가체계 (WAAS) ELA 포트폴리오 또는 워싱턴 주 교육 및 평가 접근 (WA-AIM) 합격(현저한 인지장애가 있는 학생에 한함)	워싱턴 주 대안적평가체계 (WAAS) 과학 포트폴리오 또는 워싱턴 주 교육 및 평가 접근 (WA-AIM) 합격(현저한 인지장애가 있는 학생에 한함, IEP 팀 판정 기준)	워싱턴 주 대안적평가체계 (WAAS) 수학 포트폴리오 또는 워싱턴 주 교육 및 평가 접근 (WA-AIM) 합격(현저한 인지장애가 있는 학생에 한함)
독해 & 작문 평가 시험으로 승인된 AP 시험 중 1종에서 3점 이상, 또는 IB 시험 중 1종에서 4점 이상 취득	과학 평가 시험으로 승인된 AP 시험 중 1종에서 3점 이상, 또는 IB 시험 중 1종에서 4점 이상 취득	수학 평가 시험으로 승인된 AP 시험 중 1종에서 3점 이상, 또는 IB 시험 중 1종에서 4점 이상 취득
SAT나 ACT 독해 또는 SAT나 ACT Plus Writing 작문에서 높은 점수 취득	ACT 과학에서 높은 점수 취득 (주 교육위원회가 정하는 커트라인 점수)	SAT 또는 ACT 수학에서 높은 점수 취득
학습결과물평가(COE) 독해 및 작문 합격	학습결과물평가(COE) 생물학 합격	학습결과물평가(COE) 수학 합격

* HSPE: 고등학교 능력 평가(High School Proficiency Exams)

** COE: 학습결과물평가(Collection of Evidence)

*** EOC: 과정수료시험(End-of-Course Exam)

**** IEP: 개별화 교육 프로그램(Individualized Education Program)

2) 서술형 평가 이해

가) 평가목적

미국 중등학교의 서술형 평가는 2001년에 시작된 No Child Left Behind Act of 2001 (낙오방지법)의 시행에 기인하는데, 채점의 어려움에도 불구하고 복잡한 학습결과를 효과적으로 측정하기 위하여 서술형 평가의 중요성은 더욱 부각되고 있는 실정이다(Atkin, J.M., Black, P., & Coffey, J., 2001; 김재춘 외, 2004). 미국 중등학교에서 학생 평가를 위하여 서술형 평가에 많은 비중을 두는 이유는 글쓰기 능력 평가보다 전인적, 창의적 학생으로의 성장을 도모하기 위한 것이다(Fisher, D. & Frey, N., 2007; 박도순, 2004). 또한 서술형 평가를 원활하게 수행하기 위해서, 무엇보다 종합적인 능력이 필요함에 의견을 같이하고 있다(송인수, 2005; 윤지희, 2005). 이러한 중요성으로 말미암아, 대학입시에서는 학생의 작문 능력(essay)을 종합적으로 살피고 있다. 이의 목적은 첫째, 내용이다. 대학이 요구하는 에세이 주제는 자기 자신에 대한 내용이다. 이는 고교 졸업 때까지 수많은 과정 중 핵심을 글에 투영해내야 하는 것을 의미한다. 대학 측에서는 내용을 보고 해당 학생이 어떻게 살아왔으며 그것을 토대로 향후 어떤 미래를 살아갈 것인지를 파악해낸다. 즉, 글 안에 인생관, 가치관, 행동양식, 잠재의식까지 드러난다는 것이다. 둘째는 글쓰기 능력이다. 글쓰기 능력은 단순히 재주를 말하는 것이 아니다. 전문가들은 글쓰기 능력은 독서량과 일맥상통함을 주장하고 있다(이용관, 2005). 글쓰기 능력은 풍부한 어휘력과 표현력, 문장력, 구성력에 의해 판단되어진다(진동섭, 2004). 셋째, 글쓰기를 잘하면 말하기와 발표력도 좋아진다. 대학과 대학 이후 인생에서는 글쓰기, 말하기, 발표력이 매우 중요하다. 대학에선 논문을 써야 하고 졸업 후에는 사업제안서나 연설을 할 때도 많이 생긴다(김재춘 외, 2004).

에세이 유형은 반응의 자유도가 크기 때문에 학생 자신의 지적 배경에 따라 적절한 자료와 정보를 선택하여 자신의 말로 표현이 가능하고, 문항의 지시 범위 내에서 자유로운 문제 접근이 가능하다. 표현력, 사고력 및 문제해결력 같은 일반적 교육목표와 개념, 원리의 적용과 같은 구체적인 학습목표를 평가하는데 용이하다. 에세이 유형은 좋은 문항을 만드는데 상당한 노력이 요구되어지지만, 선택형 문항에 비하면 시간과 노력이 적게 든다. 에세이 유형은 답안 작성에 허용되는 자율성의 정도에 따라 제한형과 확장형으로 구분되어 진다(Waugh & Gronlund, 2013, p. 131).

나) 문항종류와 채점방식

미국 중등학교에서는 교과 담당 교사가 자신이 지도한 학생의 성취 수준을 절대평가방식으로 기록하며, 학생들의 상대적인 등급이나 석차는 기록하지 않는다. 학생들의 교과학습발달상황은 중학교 단계보다는 고등학교 단계에서 더 자세하게 기록된다(Miller, M. D., Linn, R. L.,

Gronlund, N. E., 2009). 예컨대, 중학교의 교과 성적(subject marks)은 A(매우 우수), B(우수), C(만족) 세 등급으로 분류하여 표기하는데 비하여, 고등학교에서는 중학교보다 더 세분화된 다섯 등급 즉 A(매우우수), B(우수), C(보통), D(미흡), F(매우미흡)로 분류하여 학생의 성적을 표기한다.¹⁾ 무엇보다 미국 중등학교에서는 개별 교사가 평가 문항을 직접 개발하여 교육적 목적으로 활용하고 있으며, 학생들의 학업 성취 정도를 등급으로 평정하여 기록하며, 학생들의 원점수나 상대적인 석차를 산출하지 않고 있다.

서술형 평가의 분석적 영역은 ‘문제이해’, ‘문제해결’, ‘정답제시’로 나타낸다. 이의 각 영역의 상(high), 중(mid), 하(low)로 구분하여 각 영역별로 학생의 강점이나 취약점을 파악하고 있다. 구체적으로 살펴보면, ‘문제이해’의 상(high)은 문제를 완전히 이해하고 있음을 의미하는 것이며, 중(mid)은 문제를 부분적으로 이해하면서 해결전략을 모색하는 것을 의미한다. 하(low)는 문제를 전혀 이해하고 있지 못함을 의미한다. ‘문제해결’의 상(high)은 문제를 이해한 것에 기초하여 정답을 이끌어 가고 있음을 의미한다. 중(mid)은 문제를 이해한 것에 기초하여 부분적으로 해결전략을 제시하고 있음을 의미한다. 하(low)는 문제를 이해한 것과 상관없이 부적절한 해결전략을 제시하고 있음을 의미한다. ‘정답제시’의 상(high)은 논리적인 과정으로 정답을 명확히 제시하고 있음을 의미하는 것이며, 중(mid)은 부분적인 오답을 제시하고 있음을 의미하고, 하(low)는 답을 기재하고 있지 않거나 완전한 오답을 기재하는 경우를 의미한다. 미국 중등학교의 서술형 평가는 각 주마다 상이한 방식을 취하고 있지만 보편적으로 살펴보면 이러한 방식을 활용하고 있다. 이러한 평가방식의 장점은 교사가 학생들에게 즉각적인 피드백을 해 줄 수 있는 것이다. 단점 또한 각 영역의 배점에 대한 논란이 있으며, 부분점수의 차이가 있지만 동일한 총점이 야기되는 문제점이 노출될 수 있다.

서술형 평가 중 논술형(essay) 문제의 항목과 기준을 살펴보면 다음과 같다. 먼저, 논술형 문제의 항목과 기준은 내용의 정합성과 문단 구성 및 논지 전개, 독창성 및 논리성, 전체구조, 표현, 문법, 형식과 분량 등으로 이루어진다. 구체적으로 살펴보면, 내용의 정합성과 문단 구성 및 논지 전개는 제시된 주제에 관하여 지문의 내용을 정확하게 이해하고 작성하는지 평가한다. 독창성 및 논리성은 논지의 전개가 논리적인지, 논지의 전개과정에서 독창적인 해석과 생각이 반영되어 있는지를 평가한다. 전체구조에서는 서론, 본론, 결론의 구성이 유기적으로 치밀하게 전개되는지 평가한다. 각 문단의 구성이 두괄식 혹은 미괄식으로 논지 전개가 뚜렷하고, 문단으로서의 완결된 구조를 갖고 있는지 평가한다. 즉 하나의 문단이 하나의 중점적인 정보나 견해를 표현하며 그 이외의 성분은 주요 논지를 설득력 있고 논리적으로 뒷받침하기 위해 사용되고 있는지를 평가한다. 표현, 문법은 어휘 구사 수준과 표기의 정확함을 평가한다. 그리고 문법의 적합성, 표점 부호의 정확한 사용 등을 평가한다. 형식과 분량은 제한된 분량 내에서 완성된 한편의 논술을 작성하였는지를 평가. 지시사항(일반 지침)의 준수 여부, 분량의 초과 및 미달 여부 등을 감점처리 하여 평가한다.

1) 미국의 중등학교 학생 평가방식은 각 주마다 상이하지만 보편적으로 살펴보면 앞서 설명한 것과 같음

<표 II-24> 평가 항목별 비중

형식과 분량	내용의 정합성과 문단 구성 및 논지 전개	독창적 견해와 논리적 흐름	전체 구조	표현·문법·절차
감점	50%	30%	10%	10%

각 에세이를 전체적으로 A+, A, B+, B, C-D, Fail로 평가하면서, 위의 세부적인 평가항목을 고려하여, 대부분의 수험생들의 총점을 90 ± 6 으로, 즉 96점~84점 사이에서 채점한다. 아주 우수한 에세이에 대해서는 97점~100점으로, 기준 이하의 에세이에 대해서는 79점~60점으로 평가할 수 있다. 96점~84점 범위에 드는 에세이의 채점에 세밀한 주의를 기울여서, 가능한 한 1점 단위의 등위로 상대 평가한다. 종합평가 지침을 살펴보면 다음과 같다.

<표 II-25> 종합평가 지침

종합점수	A+	A	B+	B	C-D	Fail
평가내용	· 모든 평가 항목 충족한 우수한 에세이	· 논제 내용의 정합성이 우수 · 원인분석, 목적 지향의 명확성 · 논리성 충족	· 논제 내용의 정합성이 다소 부족 · 평가항목을 상당히 충족 · 논제에 대한 생각의 심화가 다소 미흡한 에세이	· 논제 내용의 정합성이 상당부분 부족 · 논제에 대한 심화가 다소 미흡한 에세이	· 논제 내용의 정합성이 관련 없음 · 평가항목 기준 이하인 에세이	· 논제 내용의 정합성이 전혀 관련 없음 · 평가항목 기준 이하인 에세이

3) 수행평가 이해

미국 중등학교의 수행평가는 1980년대 후반 교육개혁의 일환으로 본격 도입했다. 83년에 나온 ‘위기에 처한 국가’라는 교육경쟁력 관련 보고서가 계기가 됐다. 이 보고서는 “미국 학생의 학업성취 수준이 국제적으로 매우 낮은 데다 17세 전체 학생의 약 30%가 글을 읽을 수 없다”며 학력 저하를 우려했다. 학업성취도를 끌어올리기 위한 미국의 교육개혁은 크게 세 가지로 요약할 수 있다. 첫째, 학교·교사의 책무성 강화. 둘째, 주 단위의 학업성취도 시험 도입과 이를 통한 성적 향상. 마지막으로 수행평가의 확대다. 미국 중등학교의 수행평가는 연구·조사 보고서나 실험·실습 등 프로젝트형 과제를 내고, 서술·논술형 시험을 출제하고, 다양한 팀 단위 활동 등을 수업에 도입하는 방식으로 수행평가를 했다. 이후 수행평가는 암기 위주 교육에서 탈피해 지식을 적용·융합할 줄 아는 창의적 인재를 육성하는 새로운 교육모델로 각광받았다. 미국 중등학교 수행평가는 1980년대 후반 본격적으로 도입하기 시

작했고, 우리나라에는 1996년에 처음 소개된 뒤 1999년 전국 초·중·고교에서 일제히 시행했다. 각 시·도 교육청은 각 과목별 수행평가의 최소 비율에 대해서만 지침을 제시할 뿐 평가 방식과 시행 시기 등과 관련한 구체적인 사항은 각 학교 재량이다.

그리고 미국 중등학교의 수행평가 영역 중 하나는 학창시절에 교내외에서 참가했던 다양한 활동(Activity)이다. 이러한 활동은 대학진학 혹은 취업을 위하여 중요한 참고 잣대가 되며, 중등학교에서 심층적으로 다양한 활동에 대한 학습과 평가가 이루어지고 있다. 심지어 미국대학에서는 이런 활동만을 평가하는 심사위원회가 조직되어 있다. 즉, 미국대학 입학위원회는 다양한 성격과 문화적인 배경과 학교를 활기치게 이끌어 나갈 수 있는 학생을 원한다. 이런 부분을 표현하는 중요요소로 중등학교에서의 다양한‘Activity’를 중요 참고자료로 활용하고 있다. 구체적인‘Activity’를 살펴보면, Community Service, Vacation봉사(개인 비용으로 방학 때 캄보디아 등을 방문하여 하는 기획봉사), Volunteer, Rock Band, Internship(대학생도 구하기 힘든 인턴쉽을 고등학생이 했다는 것은, 보통 부모님회사나 부모의 추천으로 들어갈 수 있는 회사에서 간단한 일을 했다는 사실을 입증하는 것이고 이런 활동은 좋은 평가를 받기 어려움) 등이 있다. 반면, 좋은 활동의 예로서 Sports, Writing/ Art Clubs, Political Activism(Amnesty, STAND), Research, Math/ Science Competitions, Computer Programming 등을 추천하고, 이런 활동을 장려하고 있다.

4) 시사점

미국 중등학교의 학생평가는 학습자의 학업 수행 과정을 다양한 기법을 동원하여 종합적으로 평가하고 있다. 무엇보다 타당성과 다양성의 입장에서 학생평가가 이루어지고 있는 점에 주목할 필요가 있다. 구체적으로 살펴보면, 첫째, 타당성은 교사가 목표-학습내용-교수-학습활동 등 일련의 과정을 평가에 담고 있다. 즉, 수업 중의 구두질문, 쪽지시험, 학습자의 발표 등이 평가항목에 구체화되어 있다. 이런 점에서 수업 중 학습자들의 자발적인 토의학습을 많이 활용되며, 교사는 학습자의 개인적 특성을 고려하여 응답할 수 있는 발문과 무엇보다 학습자의 장점과 개성을 찾는 것을 주요한 모티브로 한다. 한 예로, 수학과 과학교사들 대부분이 문제를 직접 해결해 주지 않고, 학습자 스스로 해결할 것을 권유한다. 어려운 문제에 대해서는 해결에 필요한 실마리만 제공하고, 학습자의 자율 해결을 권유한다.

둘째, 다양성은 학생평가와 관련한 교수·학습 자료가 풍부하고, 사용절차가 간편하다. 업무시간에는 각종 자료실, 실험, 실습실 등이 개방되어 있으며, 자원봉사자나 보조직원이 자료 활용을 보조해줄 준비를 갖추고 있다. 몇 가지의 주요한 특성을 찾아보면, ①교과 담당 교사들이 수업을 시작할 때, 진단 검사 단계에서 반드시 수업 목표를 다양한 방법으로 주지시킴으로서 학습자들의 학습준비를 자연스럽게 시키고 있다. ②학습지도 내용을 선정하는데 있어서, 다양성과 실용성을 학습자들에게 주지시킨다. 수학과목인 경우, 주변에서 수학 원리가 적용되고 있는 현상을 찾아서 실제적인 문제를 해결하도록 한다. 또한 수학과 과학과목의 교사가 협업을 통한 강의와 개별 학습자의 관심과 학습능력 또는 특성을 고려하여 학습자의 응답을 이끌어내고 있다. ③교사는 학생평가에 있어 학습자의 장점과 독특성, 자율적인 학습능력을 고려한다.

④종합평가나 퀴즈 등의 결과를 피드백해주고, 학습자의 자율적인 확인 또는 상호-교환 확인 과정을 거침으로써, 평가의 공정성과 객관성을 유지하고 있다. ⑤철저한 절대평가를 하고 있다. 종합평가에서 실패한 학습자에게 2~3번의 기회를 더 줌으로써, 학습능력 향상을 위한 자율적인 노력을 조장하고 있다. 학습자간의 상대적인 비교를 거의 하지 않는다. ⑥교수·학습자료의 제작과 활용에 학교, 교사 및 학습자가 함께 참여한다. 학습자에게 부과된 프로젝트를 평가하여 우수한 작품을 선정, 활용하고 있다.

전반적으로 미국의 중등학교 학생평가는 종합평가나 퀴즈 등의 결과를 피드백 해주고, 학습자의 자율적인 확인 또는 상호-교환 확인 과정을 거침으로써, 준거지향평가를 주로 시행함으로써 평가의 타당성이 유지되고 있다. 수행평가에 있어서도 종합평가, 연구과제, 숙제, 협동작업, 퀴즈, 구두질문, 쪽지시험, 학습자의 발표 등 수행평가도 다양한 방식으로 수행됨으로써 그 본래의 취지를 살리고 있다. 이와 같이 뚜렷한 학교의 교육목표 달성을 지향하면서 수업내용과 평가가 일관성을 유지하기 위하여 노력하고 있는 점 등을 면밀히 고찰할 필요가 있는 것으로 사료된다. 주목할 점은 미국 중등학교 수행평가에서는 내가 무엇이 부족하고 어떤 점을 보완해야 하는지 수행평가 결과에 대해 상세하게 알려준다. 이런 일련의 과정들로 대체로 교사의 평가를 신뢰한다. 미국도 수행평가 도입 초기엔 평가과정에서의 공정성과 객관성 문제 등으로 혼란이 많았다(Attali, Y., & Burstein, 2006). 하지만 지금은 많이 안정되었다(Popham, W.J., 2008). 수행평가 채점 기준을 사전에 정확하게 공지하고 평가 후에도 교사가 학부모에게 평가 결과를 상세하게 안내하자 학부모가 평가 결과를 신뢰하게 되었다(Abedi, J., 2007). 교육과정 평가원(2011)은 미국의 수행평가 정착 과정에 대해 “피드백 과정이 중요했다”고 평가했다. 한 예로 글쓰기 수업이라면 한 학기 동안 그 학생의 글쓰기를 지도한 교사가 교과과정 중에 어떤 평가와 조인을 했는지, 그리고 최종 결과물에 대한 평가까지 다 모아 한 권의 책으로 엮어 학부모에게 발송해준다. 학부모는 이 책만 보면 아이에게 무엇이 부족한지, 그리고 집에서 어떻게 도와야 하는지를 파악할 수 있어 앞으로의 공부 계획을 짤 수 있다. 이를 토대로, 한국과 미국의 중등학교 교사의 학생평가 전문성 기준 비교분석하면 다음 <표 II-26>과 같다.

<표 II-26> 한국과 미국의 중등학교 교사의 학생평가 전문성 기준 비교분석

한국	미국
(1) 평가방법의 선정	(1) 수업실제에 적합한 평가방법을 선택하는 전문능력
(2) 평가도구의 개발	(2) 수업실제에 적합한 평가방법을 개발하는 전문능력
(3) 평가실시·채점 및 성적부여	(3) 외부 평가와 교사자작 평가를 실시, 채점, 해석하는 전문능력
(4) 평가결과의 해석·분석·활용·의사소통	(4) 개별 학생, 수업계획, 교육과정 개발, 학교 개선 등과 같은 의사결정을 내릴 때, 평가결과를 활용하는 전문능력
(5) 평가의 윤리성	(5) 학생평가의 타당한 점수 부여 절차를 개발하는 전문능력

	(6) 평가결과를 학생, 학부모, 일반 시민, 그리고 다른 교육자들과 의사소통하는 전문능력
	(7) 평가방법과 평가정보가 비윤리적으로, 불법적으로, 부적절하게 활용되고 있는지를 간파하는 능력

* **한국:** 한국교육과정평가원(2005). 교사의 학생평가 전문성 신장연구 (Ⅱ) (연구보고 RRE 2005-3). 서울: 한국교육과정평가원.

** **미국:** McMillan, J.H.(2007). *Classroom assessment: Principles and practice for effective standards-based instruction*(4th ed.). Boston, MA: Pearson Education, Inc.

나. 핀란드

1) 학생 평가 이해

가) 핀란드의 교육제도

핀란드는 세계가 인정하는 교육 강국이다. 핀란드는 교육 경제력 강화를 위해 끊임없이 노력한다. 2006년부터는 학습 부진아를 위한 계획, ‘Joustava Perusopetus’를 적극적으로 추진하고 있다. 학습능력이 떨어진다는 이유로 학교를 중도 포기하는 일이 없도록 국가가 직접 돕는 정책을 취하고 있다. 한화로 약 90억 원 정도의 재정을 1000명 정도의 학습부진아를 위해 마련했다. 단 한명의 낙오자로 막으려고 노력한다. 이 제도는 의무적으로 시행하며 전혀 새로운 학습 환경을 조성하려는 것이다. 학교는 모든 학생에게 동등한 교육기회를 제공하여 평등교육을 실현하려 한다. 학생평가는 이를 위한 수단의 역할을 한다.

핀란드의 교육기관은 사회복지기관들이 운영하는 어린이집, 취학 직전 1년간의 유치원(preschool), 9년제의 종합학교(comprehensive school), 후기 중등학교, 그리고 대학교로 구성되어 있다. 9년제의 종합학교만이 의무교육에 해당한다. 여기에서 기초교육(Basic Education), 6년의 초등교육과 3년의 전기 중등교육이 이루어진다. 취학 전 1년간의 유치원 교육은 의무교육은 아니지만 대부분의 아이들이 다니고 있으며 유치원의 교육과정은 국가가 마련하여 다음 단계인 종합학교교육과 연계된다. 9년간의 의무교육을 마치면 후기 중등교육(upper secondary education)을 위해 우리의 전문계 고등학교에 해당하는 직업학교나 일반계 고등학교에 해당하는 집나지움에 진학한다.(배준형, 2011). 이것은 3년 과정으로 이 두 학교 간에는 이동이 가능하다. 이를 마치면 대학이나 기술전문대학(Polytechnics)을 진학하게 된다. 두 학교간의 학위는 상호 인정이 되지 않는다. 학사 학위와 박사 학위는 대학에서만 수여가 가능하고 기술전문대학은 기술전문대학 학사학위와 현장경험 후의 기술전문대학 석사학위를 수여할 수 있을 뿐이다(한국교육연구네트워크 총서기획팀, 2010).

나) 핀란드의 중등교육

전기 중등교육은 기초교육의 일부로서 학습자에게 다양한 성장, 학습 그리고 건강한 자존감의 개발을 위한 기회를 제공해야하고 그러므로 학생들은 삶에 필요한 지식과 기술을 획득할 수 있고 더 깊은 학습이 가능하고 민주적 사회를 발전시킬 수 있다. 또한 전기 중등교육은 학습자의 언어적 문화적 정체성과 모국어의 발전을 지원해야 한다. 전기 중등교육의 또 다른 목적은 평생교육의 욕구를 일깨우는 것이다(Finnish National Board of Education, 2004). 7~9학년의 전기 중등학교의 교육목표와 역할은 1~6학년의 초등교육의 목표와 역할과 크게 다르지 않다.

후기 중등교육(upper secondary education)의 역할은 전기 중등교육의 과제들을 지속하면서 광범위하고 전반적인 심화학습의 가능성을 제공해야한다(Finnish National Board of Education, 2003). 후기 중등교육은 사회와 환경이 제시하는 문제들을 해결할 수 있는

능력을, 그리고 문제들을 다른 관점에서 평가할 수 있는 능력을 기르는 것을 목적으로 한다. 학생들은 사회와 미래노동에서 책임 있는 시민으로 활동하도록 지도받는다. 후기 중등학교 수업에서는 학생의 자기 지식의 발전과 성인이 되기 위한 긍정적인 성장을 지원하고 학습자들이 지속적인 자기발전과 평생학습을 지원해야 한다. 후기 중등학교에 배운 기능과 지식은 인정서, 졸업 및 입학자격시험 인정(matriculation examination certificate), 학위증(diploma), 등에 표시된다.

한국의 중학교 과정인 전기 중등교육(lower secondary education)은 7~9학년의 핵심과목은 모국어 및 문학, 외국어, 환경, 보건교육, 역사, 사회(social studies), 수학, 물리학, 화학, 생물학, 가정경제(home economics), 체육, 음악, 그리고 미술과 공예이다. 이에 더하여 선택 교과(목)들이 있다. 전기 중등교육은 1~6학년이 담임을 중심으로 한 반별수업과는 달리 교과별 수업이 중심을 이루면서 더 많은 선택이 가능하고 현장학습기간(workplace guidance period)도 포함된다. 학생들은 개설된 여러 과목 중에서 학부모와 담임교사와 의논하면서 함께 과목을 선택하고 과목의 목표를 설정한다.

한국의 고등학교 과정에 해당하는 후기 중등교육(upper secondary education)은 한국의 대학 과정과 비슷하다. 핀란드의 후기 중등교육 기간인 3년간 학습자가 취득한 학점에 따라서 졸업 여부가 결정된다. 거의 전 과목이 자유 선택이기 때문에 자신의 취향과 적성에 맞는 수업을 듣게 된다. 따라서 학습자가 학교에 있는 시간도 제각각이다. 마지막 학년이 되면 학점을 거의 취득한 상태가 되어 일주일에 두 번 정도밖에 등교하지 않는다. 담임교실에는 수업하는 일이 없고 과목마다 교실을 이동한다. 특히 한 것은 쉬는 기산에는 교실이 잠겨있고 선생님도 교실에 없다. 즉, 쉬는 시간은 원칙적으로 모두 교실 이외의 장소에서 보내야 한다. 쉬는 시간은 10분이고 수업 종이 치면 선생님이 교실 문을 연다. 자유롭게 앉고 출석을 부르면 수업이 시작된다(지스카와 마유, 2009).

다) 핀란드의 학생평가

핀란드는 교육기본법(Basic Education Act 628/1998)’에서 ‘교육의 목적은 첫째, 취학 전 유치원 아동에 대해서는 학습을 위한 능력을 향상하고, 취학 초중등학생들에 대해서는 인간성 향상과 사회의 책임 있는 구성원으로서의 성장을 지원하며 삶에 필요한 지식과 기능을 제공하는 것이며, 둘째, 사회 내 문명과 평등을 향상하고 학습자들이 자신들의 일생동안 스스로를 개발하고 교육에 참여하기 위한 선수요건들을 향상시키는 것이며, 셋째, 전 국가적으로 교육에 있어서의 적절한 형평성을 보장하는 것이다.’고 규정하고 이러한 목적에 따라서 교육평가의 목적은 교육발달을 지원하고 학습 환경을 향상시키는 것으로 규정하고 교육제공자가 자신이 제공하는 교육과 그 영향을 평가하며 외부평가(external assessment)에 참여하는 것으로 한다. 외부평가는 교육부 산하의 교육평가위원회에서 시행하며 그 내용은 정부 시행령으로 규정된다.

(1) 내부 평가

내부 평가(external assessment)로서의 학생평가는 학습자들의 학습을 안내하고 격려하

며 자기평가를 위한 능력을 개발하는 것을 목적으로 하며, 학생의 학습과 과업 그리고 태도는 다양한 방식으로 평가된다(Basic Education Act 628/1998). 기초교육을 위한 국가 핵심 교육과정에서 학생평가는 학업과정기간 평가와 최종 평가로 나뉘어 있다(Finnish National Board of Education, 2004).

학업과정기간의 학생평가는 다양한 증거를 기초로 하며, 학습자의 학습과 학습의 타 영역에서 발전을 심도 있게 다루는 것이다. 이 과정에서 교사의 지속적인 환류(feedback)이 중요한 역할을 한다. 평가를 통해서 교사는 학습자가 자신의 생각과 행동을 인식하도록 안내하고 자신이 배우고 있는 것을 이해하도록 돕는다. 학습자의 발전, 기능과 태도는 교육과정의 목적과 좋은 수행 묘사와 관련하여 평가된다. 평가는 수적 등급으로 표시할 수도 있고 언어적 기술로 나타낼 수도 있다. 등급은 1~10급으로 나타내며 4등급이하의 미이수(fail)로 처리되고 8등급이상이 우수 등급이다. 5등급은 지식과 기능이 적당(adequate), 6등급은 상당(moderate), 7등급은 충분(satisfactory), 8등급은 우수(good), 9등급은 매우 우수(very good), 10등급은 탁월(excellent)하다는 것을 나타낸다. 학습자들은 학년말에 평가보고서를 받는다. 행동, 태도와 관련된 생활기록은 담임교사가 작성하고 성적에 대해서는 교과 담당교사가 전권을 가지고 기록한다. 초등학생은 언어적 기술을 우선하고 수적 등급을 보조로 한다. 중학생은 수적 등급을 우선하고 언어적 기술을 보조적으로 한다. 학년말 보고서 이외에도 중간보고서, 다른 형태의 고지, 평가토론회, 등을 통해서 평가에 대한 환류(feedback)가 이루어진다.

핀란드에는 유급이 있다. 학습과 관련하여 이수해야 할 과목에서 4등급이하를 받게 되면 재이수 또는 유급이 결정된다. 미이수 과목을 다음 학년도에서 이수할 수 있다고 판단이 되면 유급을 면할 수 있다. 반대로 미이수 과목이 없는 경우에도 학생의 학교생활 전반에 대해 적절한지를 판단하고 학부모와의 대화를 통하여 유급이 결정될 수 있다. 일반적으로 한 과목이 미이수인 경우에는 진급을 결정하고 미이수 과목에 대한 보충교육을 받고 두 과목 이상인 경우에는 유급을 하게 된다. 학년별 진급이 아닌 개별적 학업 프로그램에 따라 모듈식으로 교육과정을 이수하는 학생들은 그 모듈의 이수 여부에 따라 진급을 하며 9학년까지 기초교육과정을 완성하면 기초교육인증서를 받는다.

학업과정기간의 학생평가의 구성요소는 과목평가, 작업기능평가, 태도평가, 자기평가로 이루어져있다. 자연과 환경과목을 제외하고는 모든 과목들이 개별적으로 평가된다. 1학년부터 4학년까지 자연과 환경과목은 하나의 모듈로 평가된다. 5~6학년에서는 생물과 지리가 하나의 모듈이고 물리와 화학이 하나의 모듈로 평가된다. 7~9학년에서는 생물, 지리, 물리, 화학 그리고 건강교육이 각각 따로 평가된다. 한국과 달리 핀란드에서는 작업기능(work skills)이 평가된다. 작업기능은 학생들이 학습을 하는 방법과 관련된 기능을 말한다. 그것은 학생들이 자신의 학업을 계획하고, 규제하고, 수행하고, 평가하는 기능을 가리킨다. 학생들이 자신의 일을 책임 있게 하는 방법과 다른 학생과 공동 작업하는데 있어서 협력하고 소통하는 방법을 고려한다. 태도평가(Behavioural assessment)는 학생이 다른 사람과 환경을 고려하고 규칙을 어떻게 준수하는지를 평가한다. 모든 교사는 자신이 가르치는 학생들의 태

도를 평가한다. 학생태도에 대한 목적과 목표는 지역 교육과정에 설정되어 있다. 태도목표를 설정할 때 학교의 교육목적이 고려된다.

마지막으로 자기평가는 학생 자신이 스스로를 평가하는 것이다. 기초 교육과정 중 하나는 학생들이 자기평가능력을 계발하는 것이다. 자기평가능력을 계발하는 목적은 자기 지식의 성장과 학업 기능의 계발을 지원하여 학생의 자존감, 학습자로서의 우호적 자아상, 그리고 몰입감을 강화하는 것이다. 자기평가기능을 계발함으로써 학생들은 자신의 발전과 학습목표를 인식하고 스스로 학습목표를 설정하며 자신의 학습과정을 규제하게 된다. 자기평가기능을 발달시키기 위해서 학생들은 자신의 학습과정을 검토하고 학습과 작업의 기능들을 평가하도록 안내된다. 이것은 학생들이 자신의 작업에 규칙적인 환류를 하게하고 다양한 방식으로 자신의 학습과 수행을 평가하도록 안내받고 격려 받는다.

최종 평가(final assessment)는 학습자가 최종적으로 각각 다른 과목에서 기초교육과정의 목표를 얼마나 잘 달성했는지를 평가하는 것이다. 최종평가는 국가적으로 비교 가능해야 하고 학습자들이 평등하게 취급되어야 한다. 각각의 핵심 과목에서 최종등급은 기초교육의 마지막 단계, 8~9학년에서의 학생 수행에 기초한다. 기초교육의 최종 평가의 범주는 모든 핵심 과목을 준비하도록 되어있다. 학생의 수행은 그 범주와 다양한 증거에 기초하여 평가된다.

만약 기초교육의 마지막 단계 이전에 핵심 과목의 학업이 끝나면 그 학생의 수행이 교과목의 최종평가 범주에 따라 평가하는 것이 의심스럽게 된다. 최종 평가 범주는 8등급에 필요한 지식과 기능 수준으로 정의되어있다. 만약 평균적으로 그 학생이 과목 범주에서 요구하는 수행수준을 입증하면, 즉 몇몇 과목의 수준을 충족하지 못하더라도 다른 과목의 기준을 초과함으로써 보충이 된다면 그 학생은 8등급을 받도록 범주가 작성되어있다. 만약 학생이 어느 정도 그 범주에서 요구하는 수행 수준을 증명할 수 있다면 그 학생은 기초교육이 요구하는 지식과 기능을 적절히 획득한 5등급을 받게 된다. 최종 평가에서 작업기능(work skills))에 대한 평가는 과목 등급에 포함된다.

기초교육의 최종 평가에서 등급으로 표시되는 핵심과목은 모국어와 문학, 제1~2 국어, 수학, 물리, 화학, 생물, 지리, 건강교육, 종교 또는 윤리, 역사, 사회, 음악, 시각예술, 공예, 체육, 그리고 가정경제이다. 모국어나 제2국어 또는 외국어를 변경하게 되면, 해당 학생이 마지막으로 수업했던 과목을 평가한다. 선택과목은 일주일에 두 시간 이상, 일 년 동안 수업을 한 과목은 등급이 매겨진다. 그 이하로 수업한 과목이나 모듈로 된 과목들은 언어적으로 기술된다(Finnish National Board of Education, 2004).

핀란드의 후기 중등학교의 학생평가는 교육과정의 목적을 성취하였는지 그리고 관련 과목의 전반적인 발전에 대해 학생들에게 환류를 제공하는 것이 목적이다. 학생들의 학습에 대한 평가의 역할은 학생들에게 자신의 발전과 학습결과에 대한 환류를, 그리고 후기 중등 학업 동안 학업의 완성에 대해 환류를 제공하는 것이다(Finnish National Board of Education, 2003). 이러한 환류의 목적은 학업 중에 있는 학생들에게 용기를 주고 안내를 하려는 것이다. 게다가 평가는 학생의 보호자, 심화학습 제공자, 작업장 대표자, 그리고 유

사한 집단들에게 정보를 제공한다. 또한 학생의 학습에 대한 평가는 교사와 전체로서의 학교 공동체가 교육의 효과성을 평가하는 것을 돕는다. 등급을 매기는 것은 평가의 한 유형이다. 평가는 학생을 긍정적으로 격려하여 자신을 객관화하고 자신의 작업 방법을 재조정할 것이다.

각 과목이나 과목군에 대한 학생의 지식과 기능은 교사나 교사집단에 의해 평가된다. 후기 중등교육에서도 등급은 4등급 이하는 낙제이고 5등급은 통과, 6등급은 7등급은 만족, 8등급은 좋음, 9등급은 매우 좋음, 9등급은 탁월함을 나타낸다. 실패한 학생은 과정을 완성하기 위해서 ‘별도의 시험’을 치를 기회를 가진다. 학생이 필요한 수의 과목을 완성하면 인증서(a leaving certificate)이 수여된다. 여기에서 시험이란 단어의 사용은 일종의 내적 평가 유형의 지표이다. 이 단계에서 학교는 졸업 및 대입자격시험(Matriculation Examination, ME)을 충족할 평가를 학생들에게 준비시키고 있을 지도 모른다.

(2) 외부 평가

외부평가(external assessment)는 교육부 산하의 교육평가위원회(Educational Evaluation Council)에서 시행하며 그 내용은 정부 시행령으로 규정된다. 핀란드의 후기 중등과정(한국의 고등학교과정)에서는 일반계와 직업계가 뚜렷이 구분되어 있고, 대학에 진학하기 위해서는 국가가 주관하는 졸업 및 대입자격시험(ME)을 통과해야 한다. 일반계 후기 중등학교와 직업계 후기 중등학교간의 이동은 가능하며 직업계열 학생이 대학을 진학하고자 하는 경우에도 졸업 및 대입자격시험(ME)에 통과해야 한다. 대학입학사정을 위한 전형 요소는 학교등급, ME, 전공별 본고사로 구성되어 있다. 전형요소별 비율은 학교등급이 20%, ME가 40-45%, 전공별 본고사가 35-40%로 이다. 전공별 본고사의 영향력이 상대적으로 크게 작용하기도 한다.

핀란드의 졸업 및 대입자격 시험(ME)은 후기 중등교육의 말에 행하여지는 외부평가이다. 이 시험은 12학년에 있는 학생들을 대상으로 1년에 두 번 봄, 가을에 핀란드의 모든 고등학교에서 동시에 시행된다. 이 시험은 국가 주관으로 1년에 2번 실시되며 별도의 대입자격시험위원회(Matriculation Examination Board)에 의해 관리된다. 학생은 연속 3학기 이내에 정해진 모든 시험을 통과해야 하며 기간 내에 통과하지 못할 경우 처음부터 재시험을 봐야 한다. 개인에 따라 원하는 경우 한 학기에 전체 시험을 모두 통과할 수 있다. 시험과목은 최소 4과목이며, 필수와 필수선택으로 구분된다. 모국어는 필수이며, 나머지는 학생이 4개영역 중 3개 과목 선택할 수 있다. 시험과목은 모국어, 제2국어(핀란드어, 스웨덴어), 외국어(영, 독, 불, 서반아, 러, 이탈리아어), 수학, 과학 & 인문학(10개 과목) 등으로 구분되며 원하는 경우 필수과목 외의 선택과목에 응시할 수 있다. 각 과목에 응시하기 위한 자격을 갖추기 위해서는 고교 교육과정이 요구하는 관련 교과목을 이수해야 한다. 필수과목 중 수학, 제 2국어, 그리고 외국어는 2단계의 난이도(basic, advanced)로 구성되는데, 필수과목 4개 중에서 적어도 한 과목은 고급(advanced) 단계로 시험을 봐야 한다.

시험방식은 모국어, 핀란드어, 스웨덴어는 문맥 기능(textural skills)과 에세이, 제2국어

와 외국어는 등기, 독해, 쓰기형식이고 수학은 10개 문항 그리고 인문자연과학은 6~8개 문항이 출제된다. 평가결과는 7등급 상대평가를 원칙으로 하며, 등급간 비율은 최우수 5%(L) - 15%(E) - 20%(M) - 24%(C) - 20%(B) - 11%(A) - 탈락 5%(I)로 정해져 있으나 등급의 상대적 비율은 시험과목과 시험시기에 따라 조금씩 달라진다(정일용·김든, 2011). 필수 과목에서 탈락한 경우 해당 과목에 대해 두 번의 재시험 기회가 제공된다.

2) 서술형 및 수행평가 이해

전문가로서 핀란드의 교사들은 학생을 평가할 때 학생의 교실작업, 프로젝트, 교사가 만든 시험, 그리고 포트폴리오를 활용한다. 수행평가와 서술형평가가 교실수업과 함께 이루어진다. 핀란드에서 의무교육인 종합학교 학생들은 표준화된 국가주도의 전국적인 평가를 하지 않는다. 학생평가를 어떻게 어떤 방법으로 할 것인가는 전적으로 교사의 자율에 달려있다. 국가는 다만 각 과목별로 성취기준과 동시에 어떻게 평가할 것인가 그 방향과 기준을 제시하고 있을 뿐이다. 그 기준을 모국어, 외국어, 수학, 자연, 사회과목을 중심으로 살펴보면 다음과 같다.

가) 모국어와 문학

모국어와 문학과정에서는 과목의 다양한 영역에 초점을 맞추면서 과목평가를 서술형, 구술형 그리고 참여활동에 의해 등급을 결정한다. 교사와 또래집단으로 부터의 개인적 환류는 학생의 언어소통과 서술기능의 발달과 관련하여 중요하다. 평가와 수업은 학생이 자신을 화자(speaker), 독자(reader), 그리고 저자(writer)로서의 긍정적인 현실적인 상을 만들 수 있도록 자기평가기능(self-assessment skills)을 발전시키도록 이루어져야 한다. 수행평가로서의 자기평가는 학생이 그 교육과정의 목적과 내용을 인식하는 것이 기본이며 따라서 그 과정 자체에서 학생이 자신의 발전을 살펴볼 수 있도록 허용된다. 최종 평가는 개별 학생이 성취한 독자로서, 언어사용자로서의 성숙도를 고려하게 될 것이다.

나) 외국어

외국어 과목의 평가는 개별 외국어의 우선적 특성에 따라서 모든 영역(듣기, 말하기, 읽기, 쓰기)에서 언어역량을 고려하여야 한다. 외국어 과목에 대해서는 A1.1~C1.1에 이르는 10단계의 언어역량 척도를 제시하고 있다. 간략한 척도는 아래 <표 II-27>과 같다.

<표 II-27>언어능력척도(The Language Proficiency Scale)

	L i s t e n i n g Comprehension	S p e a k i n g	R e a d i n g Comprehension	W r i t i n g
A1.1	First stage of elementary proficiency			
A1.2	Developing elementary proficiency			
A1.3	Functional elementary proficiency			

A2.1	First stage of basic proficiency
A2.2	Developing basic proficiency
B1.1	Functional basic proficiency
B1.2	Fluent basic proficiency
B2.1	First stage of independent proficiency
B2.2	Functional independent proficiency
C1.1	First stage of fluent proficiency

출처: National core curriculum for upper secondary schools 2003

다) 수학

수학수업에서의 평가는 해결책을 제시·발표하는 능력을 발달시켜야하고, 수학적 개면을 형성하는 과정에서 학생들을 지원하고, 서술형 발표 기능을 평가하고 자신의 작업을 평가하도록 학생들을 가르쳐야 한다. 수학에서는 해결책을 제시·발표하고 수학적 개념을 형성하는 과정을 평가하며 글쓰기 발표 기능을 평가하도록 하고 있다. 기능과 지식에 대한 평가는 계산기능, 방법선택, 자세하고 논리적인 정당화에 초점을 맞춘다.

라) 자연과학(생물, 지리, 물리, 화학)

생물과목의 평가는 생물학의 핵심적 개념을 숙달하고 이용하며 생물학 지식을 적용하는 능력을 평가하는데 중점이 있다. 평가는 자연과학의 법칙과 원인관계, 즉 상호의존적 의미에 대한 통찰과 전체로서의 인식을 이해하는데 주목한다. 기능 평가는 과학적 작업에 대한 태도, 단체 행동 그리고 생물학적 정보를 습득하는 능력 그리고 정보를 비판적으로 평가하는 능력에 주목한다. 생물학의 다른 영역에서 입증된 관계가 평가의 부분으로 고려될 수도 있다.

지리과목의 평가는 지식과 기능에 대해 지리적 사고의 발달에 주목한다. 평가의 목표는 기본 개념에 대한 능통함, 지리적 진술과 주장을 정당화하는 능력, 그리고 지역적 관계를 관찰하는 능력을 포함한다. 또한 평가는 지리적 지식내용을 해석하고 평가하는 능력과 다른 상황에서 지리적 지식을 적용하는 능력에 주의를 기울인다. 그 기능들은 협동기술은 물론이고 지도 해석과 다른 지리적 기능과 같은 지리적 정보를 분석하고 처리하며 발표하는 기능을 포함한다.

물리과목의 평가는 국가핵심교육과정에 설정된 것과 같은 구체적 물리 기능을 성취했는가 또는 특히 수학적 모형을 사용하는 응용 기술을 성취했는가에 주목한다. 평가의 목표는 학생의 정보처리기능, 실험기능 그리고 물리적 문제해결 과정을 분석하는 능력과 같이 그들의 학업을 지원하는 다른 기능들을 포괄한다.

화학과목의 평가는 화학적 지식과 기능을 이해하고 적용하는 능력에 주목한다. 화학 지식을 이해하고 적용하는 능력에 주목한다. 게다가 정보습득기능과 정보처리기능의 발달에 주목해야한다. 이것은 관찰하고 계획하고 실험과 측정하는 것, 장비와 시약을 안전하게 사용하는 것, 결과를 구술과 서술로 발표하는 것, 결과를 분석하고 모형화하고 평가하는 것, 결론을 도출하고 그 결과를 적용하는 것을 포함한다. 화학에 사용된 평가 방법은 과정시험, 참여활동, 실험작업, 작업보고서, 프로젝트, 발표 또는 연구논문을 포괄한다. 게다가 개념적 방법론적 지식과 기능을 지속적으로 관찰한다.

마) 사회과학(사회, 역사, 윤리)

사회과목의 평가는 교육과정의 문제와 핵심개념에 능통함, 지식을 구조화하여 발표하는 능력, 그리고 다른 형식으로 제공된 수적 지리적 정보와 사회현상을 비판적으로 평가하고 해석하는 능력에 주의·집중한다. 이러한 능력이 다양한 시험과제나 다른 대안적 평가방법에 의해서 평가될 것이다.

역사과목은 역사에 특징적인 기능과 사고유형에 기초하여, 교육과정의 핵심내용의 능통함에 기초하여 평가된다. 지식에 기초하여 전체적 구조를 세우는 능력에 특별히 주목하고, 필요한 정보와 필요하지 않은 정보를 구분하고, 일시적 관계와 원인적 관계를 이해하며 비판적으로 역사의 현상과 정보근원을 평가하는 능력을 평가한다. 과정평가(course assessment)는 다양한 방법을 차용할 것이다. 시험대신에 학습과제, 연구논문, 그리고 다른 대안적 평가방법을 사용하는 것이 가능하다.

윤리과목의 평가는 사실적 지식의 흡수와 철학적 이해의 발달에 더하여 다양하고 기술적이고 창의적으로 생활철학의 주제를 탐험하고 표현하는 능력에 초점을 둔다. 생활철학의 문제는 종종 개인적이다. 그러나 이에 대한 성찰은 지적인 미덕, 즉 비판적 사고, 논리, 일관성과 시스템을 기초로 한다. 생활철학, 가치론, 신념에 대한 판단과 묘사와 관련하여 평가는 다른 관점 그리고 다른 사람의 사고방식에 대한 배려와 관용을 가치 있게 여긴다.

바) 평가의 실제

핀란드 전기 중등학교에서는 교과 담당 교사가 자신이 지도한 학생의 성취 수준을 절대평가 방식으로 기록하며 석차는 기록하지 않는다. 핀란드 교실에서는 한국의 중간고사나 기말고사와 같이 정기고사라는 것이 없고 교사들의 재량으로 수업과 함께 평가가 수시로 이루어진다. 이 과정에서 서술형 평가와 구술시험, 토론평가, 과제물, 조별활동 등 다양한 수행평가가 활용된다(김순남, 2013). 전기 중등학교 2학년 수학교사의 평가실제를 비교해 보면, 학생이 과제(homework)를 하는 횟수는 한국이 핀란드보다 아주 적다. 과제에 대한 환류는 완료점검을 확인하는 환류는 한국이 많이 하나 과제와 관련하여 토론하는 환류는 핀란드가 월등히 높다. 평가 유형별 활용도를 살펴보면, 핀란드는 진행과제와 교실평가를 한국보다 아주 많이 활용한다. 외부 평가를 하는 빈도는 한국이 핀란드보다 월등히 많다. 평가문항 유형별 출제빈도를 살펴보면, 한국은 기억이나 탐색문항 출제빈도가 상대적으로 높았고, 핀란드는 설명/증명 문항의 출제빈

도가 높았다(손원숙 · 신이나 · 배주현, 2014). 핀란드 수학과학 교사연합회가 출제한 연례 과학 시험을 보면 한 시험지 안에 선택형과 수행형, 서술형이 함께 조화를 이루고 있다. 예를 들면 1번 A~H 까지 8문항은 선택형이다. 4번 A~D는 모눈종이를 이용한 수행형이고 7번 A~C는 서술형이다(Jari Lavonen, 2007). 그 시험 문항을 제시하면 다음 [그림 II-19]와 같다:

- | |
|------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| <p>1. Change right alternatives,</p> <p>A. Particle radiation is</p> <ol style="list-style-type: none"> 1) α-radiation 2) β-radiation 3) γ-radiation <p>B. Consequence of the rotation of the earth around the axis is</p> <ol style="list-style-type: none"> 1) a day 2) a month 3) a year <p>C. Joule is</p> <ol style="list-style-type: none"> 1) the unit of work 2) the unit of power 3) the unit of energy <p>D. Absolute zero is equal to</p> <ol style="list-style-type: none"> 1) 0°C 2) 0°F 3) 0°K <p>E. Boiling point of water will increase if</p> <ol style="list-style-type: none"> 1) pressure increase 2) salt is dissolved to the water 3) adjusting the knob of the hotplate is turned <p>F. The gravity on the moon is smaller than on the earth. On the moon the weight of an astronaut</p> <ol style="list-style-type: none"> 1) the weight of an astronaut increases 2) the weight of an astronaut becomes smaller 3) the mass becomes smaller <p>G. The effect doubles when</p> <ol style="list-style-type: none"> 1) the work doubles 2) the time doubles 3) the time is halved <p>H. The voice is</p> <ol style="list-style-type: none"> 1) a longitudinal wave motion 2) a transversal wave motion 3) both wave motion and particles |
|------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|

4. A certain material was warmed 60 minutes. In the investigation, the results given below were obtained:

A. Present the measurement results in the grid(time-temperature)

A full page of blank graph paper with a uniform grid of small squares. The grid consists of 20 columns and 20 rows, creating a total of 400 small square units. The lines are thin and black, set against a white background. There are no margins or additional markings on the page.

D. What is the form of the material when it has been warmed for 13 minutes?

〈표 II-28〉 핀란드 중학교 학생 성적표

- 94 -

듣고 이해하기 (Listening comprehension)	X				
말하기(Speaking)	X				
읽고 이해하기 (Reading comprehension)	X				
쓰기(Writing)	X				
수학(9학년(Mathematics(grade: 9)					
문제해결 능력 (Problem solving skills)	X				
연산(Arithmetic)	X				
기하 및 측정 (Geometry and easurement)	X				

* 출처: Ressu 종합학교 8학년에 재학 중인 학생 성적 성적표를 영어로 번역한 것임(김순남, 2012)

위의 <표 II-28>의 핀란드 중학교 학생 성적표에서 보듯이 핀란드 중학교 과정에서 학생평가 기록 양식과 기록방법 역시 학교 및 교사마다 다르다. 즉 기록양식이나 기록방법은 각 학교 및 교사들이 다양한 방식으로 만들어 활용하고 있다. 단지 교사들은 국가핵심교육과정에서 정해진 성취기준을 근거로 하여 해당 영역들에 대한 학생평가를 실시한다. 예를 들면, 헬싱키시의 수학과목의 대체적인 8점 취득기준과 평점기준은 다음 <표 II-29> ~ <표 II-31>과 같다(National Board of Education, 2004:170).

<표 II-29> 수학 평점 8점(good) 취득을 위한 기준

사고 기법 및 방법	<ul style="list-style-type: none"> - 학생은 다양한 사건들 사이의 규칙과 관계를 알 수 있다. - 학생은 대화에서 그리고, 또는, 그래서, 아니오, 존재하지 않는 등의 논리적 요소들을 사용할 수 있다. - 학생은 간단한 명제의 참을 증명할 수 있다. - 간단한 문제를 수학적 개념을 사용하여 구성하고, 문제를 규정하고 풀며 해결할 수 있다. - 수학적 문제 해결에서 분류, 분석할 수 있다. - 다양한 표현 방법을 활용하여 대안적 해결 방법을 체계적으로 나타낼 수 있다.
수와 계산	<ul style="list-style-type: none"> - 수학적 문제해결을 위해 가능한 결과를 예측하고 문제해결 계획을 세울 수 있다. - 필요한 경우 제곱근을 사용하여 문제를 풀 수 있다. - 일상생활에서의 문제를 해결하기 위하여 비율, 퍼센트 등 다양한 계산 방법을 활용할 수 있다.
대수	<ul style="list-style-type: none"> - 1차 방정식을 풀 수 있다. - 간단한 대수 표현으로 축약할 수 있다. - 고급 계산을 수행할 수 있다. - 일상의 문제해결과 관련하여 간단한 방정식을 구성하여 활용할 수 있다. - 간단한 문제해결을 위해 몇 개의 방정식을 함께 사용할 수 있다. - 결과의 논리를 평가하고 해결의 다양한 측면을 검토할 수 있다.

함수	<ul style="list-style-type: none"> - 좌표 시스템에서 좌표를 결정할 수 있다. - 주어진 규칙에 따라 수의 짝으로 표를 만들 수 있다. - 선형 함수에서 영점을 찾을 수 있다. - 주어진 규칙에 따라 수 계열을 알 수 있고 수 계열에서 규칙을 찾아낼 수 있다. - 선형 방정식에서 직선과 곡선의 의미를 이해한다.
----	---------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------

<표 II-30>핀란드 헬싱키시 학교 성적 평점 기준(2010-2011)

학점	성취수준	성취 내용
10	excellent	학생이 모든 학습목표를 성취한다. 학생은 주어진 정보와 자신의 숙련도를 잘 분석하고 평가한다. 학생은 다른 물건들의 관계를 잘 이해하며 어떤 상황에서도 물건들을 적절히 이용할 줄 안다. 학생은 창의적이고 통찰력을 항상 지니며, 과제를 훌륭하게 잘 수행한다. 학생은 적극적이고 완벽하게 자기에게 주어진 과제를 수행한다. 학생은 책임감이 강하며 협조적이다.
9	very good	학생은 모든 학습목표를 성취한다. 학생은 주어진 정보와 자신의 숙련도를 잘 분석하며 평가한다. 학생은 어떠한 상황에서도 학습목표를 잘 적용할 줄 안다. 학생은 창의적이고 통찰력을 지니며 과제를 훌륭하게 수행한다. 학생은 적극적이고 완벽하게 자기에게 주어진 과제를 수행한다. 학생은 책임감이 강하며 협조적이다.
8	good	학생은 필수적인 학습목표를 성취한다. 학생은 대체적으로 주어진 정보와 자신의 숙련도를 잘 분석하며 평가한다. 학생은 다양한 학습상황에서 학습목표를 적용할 줄 안다. 학생은 어느 한 교과과목에서 부족함을 다른 교과를 통해 채운다. 학생은 가끔 창의력과 통찰력을 발휘한다. 학생은 적극적이며 주어진 과제를 책임지고 완수한다. 학생은 책임감이 있고 협조적이다.
7	satisfactory	학생은 대부분의 중요한 학습목표를 성취한다. 학생은 때때로 주어진 정보와 자신의 숙련도를 잘 분석하며 평가한다. 학생은 기본적인 학습상황에서 학습목표를 적용할 줄 안다. 학생은 대체적으로 적극적이며 주어진 과제를 책임지고 완수한다.
6	mediocre	학생은 약간의 학습목표를 성취한다. 학생은 가끔 학습 목표를 잘 이해하지 못한다. 학생은 자기 수준에 맞는 정보를 해석하고 사용할 줄 모른다. 학생은 보통 학습상황에서 학습목표를 적용하는데 도움이 필요하다. 학생은 상당히 적극적인 편이다.
5	poor	학생은 목표를 이루기 위해 노력은 하나 대부분의 목표가 수행되지 않는다. 학생은 주어진 과제에 좀처럼 참여하지 않으며 적극적이지 않다.
4	fail	학생은 어떠한 학습목표도 성취하지 못한다. 학생은 적극적이지 않으며 학습목표에 참여하지 못한다.

[자료출처] Helsinki Department of Education (2010). p. 5.

<표 II-31> 핀란드 중학교 생활기록부

학생 이름(Name of the student)		생일 (Date of birth) 23.10.1996			
학교 명(Name of the school) 레수 종합학교(Ressu Comprehensive School)		학기 시작일(Starting date in class) 16.08.2010			
과업수행 기술(working skills)					
	규칙적 수행 (Regularly)	대체로 수행 (Mostly)	가끔씩 수행(With variability)	때때로 수행 (Sometimes)	전혀 수행 안함 (Never)
지침이나 절차에 따름 (Follows the instructions and order)	X				
계획에 따름 (Follows the schedules)	X				
책임감 있음(Is responsible)	X				
집중하여 과업 수행 (Works with concentration)	X				
독립적인 과업 수행 (Studies independently)	X				
성실한 과업 수행 (Conscientiousness in studies)	X				
사회적 관계 기술(Social Skills)					
양호한 협력 기술 (Good co-operative skills)	X				
양호한 학급 규율 준수 (Good class discipline)	X				
친절함(Is friendly)	X				

* 출처: Ressu 종합학교 8학년에 재학 중인 학생의 생활기록부를 영어로 번역한 것임(김순남, 2012)

3) 시사점

핀란드에서 9년의 의무·기초 교육과정의 학생평가는 학교와 교사에 따라 다른 수행평가와 서술형 중심의 절대평가이다. 고등학교 진학을 위해서 마지막 학년에 언어와 수학 등 주요과목은 4점~10점의 평점이 매겨지고 그 평점의 평균점수로 진학할 학교가 결정된다. 고등학교 과정인 후기 중등학교 간에는 학력 격차가 거의 없고 무학년제로 운영되기 때문에 학생들은 자신이 선택한 교과목에서 학점을 취득하고 1년에 두 번 봄, 가을에 실시하는 Matriculation Examination에 응시하여 지정된 4과목을 합격하면 졸업이 된다. 이 시험문항은 모두 서술형이다(후쿠타 세이지, 2009). 핀란드의 학생평가로부터 교사의 평가전문성에 대하여 우리가 얻을 수 있는 시사점은 다음과 같다.

가) 형평성과 차별

핀란드는 학생평가를 통해서 학생들이 현재 얼마나 다른지를 확인하고 그 차이에 따라 학생들을 차별적으로 교육하고자 한다. 85년 수준별 반편성 정책을 폐지했다. 못하는 학생, 느린 학생을 차별적으로 지원·지도함으로써 교육에 있어서 형평성을 유지하고 낙오되는 학생이 없도록 교육복지를 실현하고자 한다. 기초 교육과정에서 학생들의 차이에 따른 교육적 지원은 학생평가에도 적용이 된다. 수행평가와 서술형 평가를 통해서 학생 개개인에게 적합한 맞춤형 평가가 가능하게 된다. 교육 초기에 학습결손을 제거함으로써 학생의 학업능력과 자신감을 향상하는 것이 가능하게 할 뿐만 아니라 장기적으로는 시민으로써의 통합과 경제 발전의 기초가 된다.

나) 성취기준과 타당도

핀란드는 국가의 과목별 성취기준에 따른 절대평가를 시행하고 있고 개인 간 변별보다는 학생 개인이 성취기준을 어느 수준까지 도달했는지를 중요하게 여겨진다. 기초학교에서 상대평가를 금지함으로써 학생들이 자신의 적성을 찾게 하고 깊이 사고할 수 있는 여유를 제공하게 된다. 즉 분석력, 종합평가력, 창의력, 등 고등사고력을 기를 수 있는 터전이 된다. 성취기준에 따른 평가를 함으로써 각자의 기준에 도달하는 것이 중요하게 되고 학생들이 서로를 도울 수 있는 협동학습을 가능하게 한다. 학생의 발달상황을 이해하고 지원하는 평가 방식이다(이윤미, 2010).

다) 교사의 전문성과 신뢰

핀란드는 교사의 전문성에 의하여 교육과정을 실현한다. 교육당국은 큰 범위의 기준을 제시할 뿐, 그것을 강제하는 감시와 장학 장치를 90년대 중반에 폐지했다. 수업과 직결되는 간단한 업무를 제외하고는 모든 행정 업무를 맡지 않는다. 행정 업무는 교장, 교감, 행정직원이 모두 한다. 교사는 교육과정을 자율적으로 결정하고 교과서를 독점적으로 선택할 권한을 가지며 학생평가의 전문적 권한과 예산편성에 강한 발언권을 가지고 있다. 핀란드의 교사는 의사나 대학교수와 같은 사회적 지위를 가지고 있으며 교육 전문가로서의 전문성 향상에 매진한다. 핀란드는 25여년에 걸친 장기적인 교육개혁에서 교사를 우대하고 신뢰하는 과정이 끊임없이 진행되어 왔다는 것이다. 오늘날 핀란드 교육의 성과는 교육에서 피그말리온 효과가 연쇄적으로 일어난 결과라고 보여진다. 국가 사회가 교사를 신뢰하고 교사를 우대함에 따라서 교사는 학생을 존중하고 신뢰하는 선순환이 이루어지고 학생과 교사의 상호간 신뢰가 핀란드를 세계최고의 교육 강국으로, 경제 부국으로 자리매김하게 만든 것이다.

Ⅲ. 교사의 서술형 평가 및 수행평가 전문성 규정과 분석 및 제고 방안

1. 교사의 평가전문성 규정

가. 교사의 평가전문성 규정 선행 연구

<표 Ⅲ-1>은 교사의 평가전문성 규정에 관한 주요 국내 선행연구 결과를 연구자, 연구명, 주요특성, 주요 결과를 중심으로 정리한 것이다. 먼저, 연구자에 따라 학생평가 전문성의 개념에 대하여 다양한 의견을 제시하고 있다. 구체적으로 살펴보면, 학생평가 전문성이란 다양한 경로를 통해 수집된 정보에 근거하여 적절한 의사결정을 내릴 수 있는 평가 능력이다(김신영, 2002; McMillan, 2004). 그리고 학생평가 전문성은 교수학습과정에서 이루어지는 학생평가 활용 능력, 평가도구를 개발하고 활용할 수 있는 능력, 그 결과의 질에 대하여 이해할 수 있는 능력, 평가결과를 산출하고 활용할 수 있는 방법 등에 일련의 평가 활동을 의미한다(송인발, 박주호, 2014; 이인제 외, 2004; AFT, MCME, & NEA, 1990). 한국교육과정평가원(2005)에서는 교사의 학생평가 전문성을 ‘학생의 학습과 성취에 관한 평가 정보를 수집하고 해석하여 활용할 수 있는 능력’으로 정의 하고, 학교 교육의 내실화와 타당한 학생평가를 위한 교사의 학생평가 전문성 제고를 위한 다음 다섯 가지 방안을 제시하였다. 첫째, 평가방법의 선정 둘째, 평가도구의 개발 셋째, 평가의 실시·채점·성적 부여 넷째, 평가 결과 분석·해석·활용 다섯째, 평가의 윤리성 등이다. 구체적으로 살펴보면, 먼저 교사들이 겪는 어려움 중 하나가 평가방법의 선정임(한국교육과정평가원, 2005)을 판단하고, 학습내용을 평가하는 방법에 대한 합의의 필요성과 교사와 교과별로 공유된 이해와 합의에 도달할 수 있는 평가방법과 학생평가 경험에 대한 다양한 정보의 제공 등이 이루어져야 함을 주장하고 있다. 현재 일선교사들은 학습 내용에 관한 학습목표 및 평가 실시 목적의 타당성 정도에서 평가방법의 선정 기준이 되고 있는 것으로 판단된다(강호수 외, 2014; 김신영, 2014; 김지선 외, 2014; 고정화, 2010). 또한 교사별 혹은 교과별 특성에 따라 기존의 평가방법과 새로운 평가방법을 적용하는 선호도가 다소 차이를 보이는 것으로 판단된다(손준중, 2010; 김경희 외, 2006; 윤정일, 신효정, 2006). 무엇보다 평가 항목에 대한 기준과 서술형, 단답형 선호 정도, 조별 활동 반영 정도 등의 그 설정에 대한 이견도 있는 것으로 파악된다(김지선 외, 2014). 이러한 문제점을 해결하기 위해서 무엇보다 평가방법의 선정 기준에의

평가전문성 신장이 필요하다(한국교육과정평가원, 2005). 둘째, 학생평가를 위해 교사가 평가방법을 선정하였다면, 교사는 평가의 목적과 내용에 적합한 평가도구를 개발하거나 선택하여 사용해야 하고, 평가도구의 질을 스스로 점검하고 개선해야 할 필요성이 제기된다(한국교육과정평가원, 2005). 일반적으로 진단평가, 형성평가, 수행평가 및 총괄평가에 따라 평가문항을 개발하고, 선택형, 서답형 등 평가도구 개발에 필요한 기본 원리와 절차를 이해하고 변형하여 적용해야 한다(김지선 외, 2014; 김신영, 2014; 송인발, 박주호, 2014). 또한 기본적으로 평가문항은 교과서에서 산출하고, 평가도구개발을 실시해야 한다(AFT, MCME, & NEA, 1990). 이때 단계별 학습 진행상황과 핵심내용 및 원리의 이해 정도를 항목별로 채점하는 방법과 풀이과정이나 부분점수 반영방법을 활용하는 것도 좋은 예가 될 수 있다(박정, 2013; 이인제 외, 2004; AFT, MCME, & NEA, 1990). 셋째, 평가의 목적과 내용에 적합한 평가도구를 개발하거나 선택한 이후 교사는 학생평가를 실시하기 위한 적절한 시기를 결정해야 한다(한국교육과정평가원, 2005). 평가 실시를 위한 시간 확보는 학교 운영 일정 또는 학사 일정을 통해 확인해야 한다(강호수 외, 2014; 김지선 외, 2014). 즉, 교사는 학습 진도와 학사 일정을 고려해 평가 실시의 횟수를 판단하고 확보할 수 있음을 의미한다. 평가 실시 일정과 횟수가 결정되었다면 채점기준과 관련된 타당도와 신뢰도를 확보해야 할 필요성이 제기된다(백순근, 2006; 손준중, 2010). 이것은 채점기준에 대한 일관성을 유지하고 결과에 대한 타당도와 신뢰도를 높이기 위한 방법으로 교사 간 채점기준의 사전 공개 및 논의, 문항별 채점 기준 세분화, 채점기준의 부분점수 적용 등을 사용할 수 있음을 의미하는 것이다(송인발, 박주호, 2014). 또한 과거 채점기준을 근거로 채점을 실시했을 때 나타날 수 있는 장·단점을 정리하고 문제가 제기된 기준은 최대한 배제하는 방향으로 채점을 진행하는 방법도 좋은 대안이 될 수 있음을 조언하고 있다(한국교육과정평가원, 2005). 이러한 방법을 적용해 채점기준에 대한 타당도와 신뢰도가 확보된다면 평가목적에 부합하는 성적을 부여할 수 있으며, 도출된 결과는 학생의 학습내용 이해와 앞으로 진행해야 할 학습의 방향성 설정에 많은 도움이 될 수 있음을 언급하고 있다. 넷째, 교사는 평가 결과를 정확하고 타당하게 분석 및 해석해야 한다(한국교육과정평가원, 2005). 즉, 교사는 평가 결과를 수업 및 학생에 대한 교육적 의사 결정에 활용해야 함을 의미하는 것이다. 이러한 의견에 대하여 국내·외 많은 연구자들은 공감하고 있다(강호수 외, 2014; 김지선 외, 2014; 손준중, 2010; 김경희 외, 2006; 윤정일, 신효정, 2006; 김수동 외, 2005; 이인제 외, 2004; 원효현, 2002; AFT, MCME, & NEA, 1990). 특히, 교사는 학생, 학부모와 평가 결과에 대해 정확하게 의사소통해야 하고(이인제 외, 2004; AFT, MCME, & NEA, 1990), 평가 결과 분석과 해석에 대한 정확한 이해를 위해 기초적인 교육 통계 개념과 문항분석 방법을 숙지(김지선 외, 2014)하고 있어야 함을 주장하고 있다. 특히, 규준(norm)이나 준거(criteria)에 근거하여 평가결과를 적절하게 해석하여 학습기회 제공 및 학습환경 개선에 활용해야 함에 의견을 같이하고 있다(강호수 외, 2014). 이외에 학생의 수준, 발달, 성장 및 성취 정도(손준중, 2010; 김경희 외, 2006) 그리고 학습계획 및 진로 등(김지선

외, 2014)을 결정하는 데 도움을 줄 수 있는 도출된 결과에 대한 분석(강호수 외, 2014)과 해석(손준중, 2010)이 필수적이다. 이러한 의미는 학습자 중심 학습 활동의 가장 대표적인 활동 중 하나가 평가 이후 학생에게 제공하는 피드백임을 강조하고 있는 것이다. 학습에 대한 피드백 제공은 학생의 지적 수준을 진단하고 학생의 이해수준을 고려한 학습 방향을 설정할 때 유용한 방법이기 때문이다. 마지막으로 교사는 학생의 인격을 존중하여 모든 평가활동 시 윤리적·법적 책임을 준수해야 함을 지적하고 있다(한국교육과정평가원, 2005). 즉, 교사는 학생의 특성 및 배경을 고려하여 공정하게 평가를 실시(박도순, 2013)하고, 평가의 적절성(백순근, 2000)을 판단해야 함을 의미하는 것이다. 평가활동 시 교사는 채점 및 성적에 영향을 미치지 않도록 평가도구 보안을 철저히 유지해야 한다(김명숙, 2003). 또한 국가수준의 평가 관련 지침 및 규정에 관한 규정과 법적 책임에 대해 숙지(성태제, 2010; 백순근, 2000)하고 평가활동의 전 과정을 합리적이고 공정하게 진행(성태제, 2010)해야 함을 강조하고 있다.

전반적으로 교사의 평가전문성 규정에 관한 주요 국내 선행연구는 미미한 실정이다. 하지만 교사의 평가전문성은 학습과 관련한 학생의 다양한 교육 활동을 교육적 가치관을 바탕으로 평가하는 것이기 때문에 중요하게 고려되어야 함에는 의견을 같이하고 있다(강호수 외, 2014; 김신영, 2014; 김지선 외, 2014; 고정화, 2010; 성태제, 2010; 한국교육과정평가원, 2005; 김명숙, 2003). 또한 이의 효율적이고 안정적인 정착을 위해서는 다음과 같은 요건이 충족되어야 함을 제시하고 있다. 교사는 교육 활동을 통해 학생의 인지 능력과 개별 특성을 파악해야 한다(김신영, 2007). 또한 인지 능력과 개별 특성을 파악하기 위해 다양한 측정 도구와 방법을 활용해 학생평가를 실시한다(곽영순, 2015; 윤정일, 신효정, 2006). 그리고 교사는 일련의 과정을 통해 수집한 학생 정보를 바탕으로 학생에게 맞는 교육의 방향성과 적절한 교수활동, 학습자 자기 이해 등을 고려하여 앞으로 진행될 교육적 처치와 학습 진도 계획을 수립해야 한다(강호수 외, 2014; 김신영, 2014; 한국교육과정평가원, 2005). 이러한 일련의 과정은 앞으로 진행해야 할 교육적 방향성 설정과 지도 방법, 교육 활동개선에 대한 결정에 영향력을 미치기 때문에 교사의 평가전문성 신장을 위한 선결요건이 되는 것이다(한국교육과정평가원, 2005). 무엇보다 교사는 학생평가 실시 전에 정확성, 신뢰성 및 타당성이 담보된 평가 기준을 숙지하고 준수해야 함을 강조하고 있다(박도순, 2013).

<표 III-1> 교사의 평가전문성 주요 영역 및 항목

항목	평가 환경				평가 투입			평가 과정		평가 결과		평가 환류	
	학교 분위기	교육여건 및 교사 자율성	교육방법 및 프로그램	교육평가 환경(평가 비율, 채점시간 등)	목표 설정	평가방 법 및 유형 선정	평가 윤리성	평가도 구 개발	채점기 준 정립	평가결과 분석	평가결과 해석	평가결과 공개 및 피드백	평가결과 활용(학생의 변화, 교육활동에 대한 만족도)
곽영순(2015)		○				○	○	○		○			○
강호수 외 (2014)					○			○	○	○			
김신영(2014)		○	○	○	○			○	○	○			
김지선 외 (2014)		○	○	○	○			○	○	○		○	
송인발, 박주호 (2014)	○					○	○	○		○	○		○
박정(2013)					○	○	○	○	○	○	○		
고정화(2010)					○	○	○	○	○				
권순달(2010)					○	○	○	○	○				
손준중(2010)					○	○	○	○	○	○	○		
권순달(2010)													
김신영(2007)						○	○	○		○			○
김경희 외(2006)						○	○	○		○	○		○
윤경일, 신효정(2006)	○	○	○	○									
김수동 외(2005)					○	○	○	○	○				
이인제 외(2004)					○	○	○	○		○			○
원효현(2002)					○	○	○	○	○			○	
AFT, MCME, & NEA (1990)						○	○	○		○			○

나. 교사의 평가전문성 규정

본 연구는 선행연구에서 제시하고 있는 교사의 평가전문성 규정에 관한 요소와 교육현장의 실태 분석을 통한 교사의 평가전문성 저해 요소를 파악하여 <표 III-2>와 같이 교사의 평가전문성 주요 영역 및 항목을 제시하였다. 구체적으로 살펴보면, 교사의 평가전문성 주요 영역은 평가 환경, 평가 투입, 평가 과정, 평가 결과, 평가 환류 등의 5개 영역으로 구성하였다. 평가 환경은 학교분위기, 교육여건 및 교사 자율성, 교육방법 및 프로그램, 교육평가 환경(평가비율, 채점시간 등)의 4개 항목으로 구성하였다. 평가 투입은 목표 설정 평가방법 및 유형 선정, 평가 윤리성 등의 3개 항목으로 구성하였고, 평가 과정은 평가도구 개발, 채점기준 정립 등의 2개 항목, 평가 결과는 평가결과 분석과 평가결과 해석 등의 2개 항목으로 구성하였다. 마지막으로 평가 환류는 평가결과 공개 및 피드백, 평가결과 활용(학생의 변화, 교육활동에 대한 만족도) 등의 2개 항목으로 구성하였다.

교사의 평가전문성 신장을 위한 5개 영역과 13개 항목을 구체적으로 살펴보면 다음과 같다. 먼저, 최근 교육평가는 집단표준에 근거한 측정·평가에서 성취기준에 근거한 측정·평가로의 전환을 강조하고 있다. 하지만 일선학교의 평가환경은 상이하다. 무엇보다 수행 및 서술형 평가에 관한 학교분위기는 학교마다 상이하고(강진구, 2013; 김선영, 2014), 심지어 2009 개정교육과정의 기본 개념과 성취평가제도를 인지하고 있지 못하는 교사가 다반사다(김소라, 2015; 김윤자, 2015). 그리고 성취기준 정립과 이에 따른 다양한 교육방법과 수업설계 미적용 등이 현재의 전반적인 교육방법 및 교육 프로그램의 현 주소이다(김진아, 2013; 박원택, 2014; 박혜인, 2015). 몇몇 학교에서는 서술형문항 검토를 위한 교사협의회를 개최하고 있었지만, 그렇지 못한 학교가 다반사였고, 심지어 평가문항에 대한 논의조차 이루어지고 있지 못하는 학교가 많았다. 일선교사들은 무엇보다 서술형 평가를 할 시간적 여유와 채점시간이 충분하지 않음을 토로하고 있다(성명진, 2015; 양정훈, 2014; 이수희, 2014). 교사의 평가전문성 신장을 위해서는 무엇보다 이에 알맞은 평가환경 구비가 선행되어야 하며, 본 연구에서는 평가환경 영역에 학교분위기, 교육여건 및 교사 자율성, 교육방법 및 프로그램, 교육평가 환경(평가비율, 채점시간 등)의 4개 항목으로 구성하였다.

평가 투입 영역에는 목표 설정 평가방법 및 유형 선정, 평가 윤리성 등의 3개 항목으로 구성하였다. 학생평가는 교육 활동을 평가하는 것이다. 즉, 학생평가는 가르치고 배우는 일에 대한 평가로서 평가의 대상이 학생이 아니라 교육 활동이라는 것이다. 지금까지 학교의 교육평가는 교육 활동에 대한 평가보다는 학생들의 학업성적을 평가하고 점수 매기는 일에만 많은 관심을 두어 왔던 것이 사실이다. 이런 까닭에 교육평가를 시험, 성적, 석차로 생각하는 경향이 높다(이윤승, 2015; 이지현, 2013). 그러나 학생평가는 곧 교육 활동에 대한 평가라는 의미에서 볼 때, 학생평가는 이제 교육에 영향을 주는 모든 요인들에 대한 평가로 그 범위가 넓어지게 된다. 그 의미는 학생평가가 단순히 학생들의 학업성적을 점수화하는 일만이 아니라, 교육목표에 대한 타당성 점검으로부터 시작하여 교육과정이나 수업활동의 적절성 판단 그리고 교육의 제반여건이나 학교 내의 인간관계에 이르기까지 학교교육의 모든 측면이 바로 학생평가와 연관되어 있는 것이다. 따라서 서술형 평가의 성취기준 설정, 성취기준에 따라 서술형 평가의 비율 설정과 더불어 교육과정에 따라 성취기

준을 만들 수 있는 역량을 기본적으로 구비되어야 한다. 나아가 성취기준에 따른 명확한 서술형 평가유형 마련과 선택 나아가 성취수준에 따른 서술형 평가유형이 준비되어있어야 하겠다. 그리고 학생평가는 교육적 가치의 판단과 제시이다(장선자, 2014; 전보람, 2013). 평가이기 위해서는 시험의 결과로 나온 점수에 대해 어떤 형식이로든 가치 판단이 이루어져야 한다. 잘했다든가 못했다든가, 어떤 것이 좋다든가 나쁘다든가, 또는 무엇이 만족스럽다든가 혹은 부족하다든가 등의 판단이 주어져야 그것을 평가라고 할 수 있을 것이다. 진정한 평가라면 교사는 마땅히 학생의 점수 결과에 대해 가치를 판단해야 하고, 그 판단결과에 대해 어떤 방식으로든 지도나 처방을 해주어야 하는 것이다. 이렇듯 앞서 많은 연구자들은 교사 평가전문성 신장에 중요하게 고려되어야 하는 것이 평가 윤리성임을 재차 강조하고 있는 것이다.

평가 과정은 평가도구 개발, 채점기준 정립 등의 2개 항목, 평가 결과는 평가결과 분석과 평가결과 해석 등의 2개 항목으로 구성하였다. 과거 학교 현장의 평가는 주로 학생의 인지적 측면의 학업성취 측면을 중심으로 실시되었다(정연주, 2013). 그로 인해 암기식 위주의 학습 풍토와 평가제도가 나타났으며, 이로 인해 상대적으로 교육이 추구하는 근본적인 배움의 즐거움, 심층적 지식 습득, 자아성찰과 같은 측면이 상대적으로 위축되었다. 그러나 이제는 학생평가 활동 시 학업성취 방법 및 도구와 학습자 특성에 관한 중요성이 점차 강조되고 있다(정홍명, 2014). 특히, 학생평가에서 다양한 활동의 학업성취 이외에 학습에 영향을 미칠 수 있는 정의적 측면을 고려한 학습자의 개인차를 반영한 평가 활동의 중요성이 부각되고 있다(한은주, 2015). 즉, 학교 현장에서 학생평가는 학습자 개인뿐 아니라 학생을 둘러싼 다양한 요인들이 반영되고 있는 것이다. 따라서 서술형 평가의 목적과 내용에 맞는 문항 개발 특히, 이원목적분류표의 내용영역과 행동영역을 이해하고 그 기준에 따라 문항을 개발할 수 있는 역량이 선행되어야 한다. 나아가 서술형 평가문항의 난이도와 정답율을 구하고 예측할 수 있는 역량과 능력집단별로 정답율을 구하고 예측할 수 있는 역량 등이 요구된다. 이러한 평가도구 개발뿐만 아니라 채점기준 정립에 관한 역량 또한 요구된다. 성취기준과 성취수준에 따른 채점기준을 설정할 수 있는 역량, 내용영역, 행동영역에 따른 채점기준을 설정할 수 있는 역량 나아가 학생에게 채점기준을 정확하게 안내하고 학생의 반응을 참조하여 채점기준을 수정·보완하면서 채점의 명확성을 검토할 수 있는 역량이 평가과정에서 필수적인 역량이 되는 것이다.

그리고 학생평가는 교육 활동을 변화시킨다(정홍명, 2014; 정연주, 2013). 학생평가가 교육에 도움 주기 위해서는 평가의 결과가 다시 교육 활동과 연결되어야 한다. 다시 말해서 평가의 결과는 학생들에게 학습활동에 대한 정보를 줄 뿐 아니라, 교사의 교수활동을 수정·보완해 나가는 기능을 해야 한다. 그리고 학생평가는 학습자의 정의적 측면을 고려한 평가이다(전보람, 2013). 이는 학습 결과를 진단·확인하고 그에 관한 치료와 처방을 제공해주는 기능이 수반되어야 함을 의미하는 것이다. 나아가 학습자가 지닌 장점과 약점을 확인할 수 있고, 장점을 더욱 강화시키는 심화 학습 방안과 약점을 보완해 주는 보충학습 방안을 강구할 수 있는 기준을 제공할 것이다. 무엇보다 학생평가의 목적은 교육과정의 목표와 내용 및 학습 지도방법 등의 개선과 평가 그 자체의 개선을 위한 자료를 제공해 주기 위한 것이다. 따라서 평가결과 영역에서는 서술형 평가 결과를 양적·질적으로 분석할 수 있는 역량이 요구되며, 서술형 평가에 대한 다양한 통계자료를 산출하면서

서술형 평가의 내용과 성취기준에 따른 적절한 평가결과를 해석할 수 있는 역량이 요구된다.

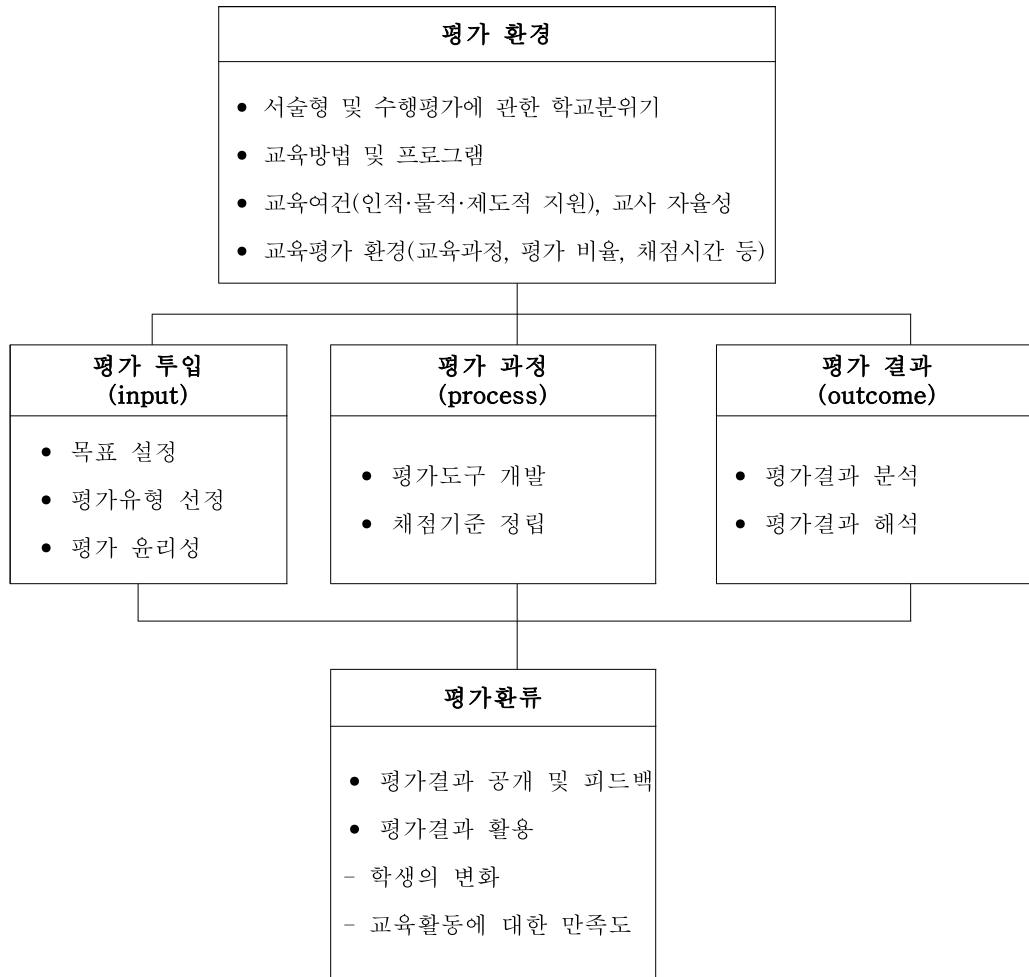
마지막으로 평가 환류는 평가결과 공개 및 피드백, 평가결과 활용(학생의 변화, 교육활동에 대한 만족도) 등의 2개 항목으로 구성하였다. 이 단계에서는 평가를 통해서 교육목표 설정의 타당성을 비롯해서 교육내용이 과연 학생에게 적합한 수준의 것이었는지, 학생들의 학습경험은 알맞게 선정되었는지, 교육내용과 경험의 조직(단원계획)은 적절하였는지, 그리고 학습지도는 제대로 진행되었는지를 확인·진단하고 또 평가 계획과 도구의 선정 및 평가 과정이 적절하였는지의 여부도 확인해서 다음의 교육과정 계획을 위한 기초 처방을 제시해 줄 것이다. 이 과정에서는 일정 기간의 학습도달에 관한 평가나 성적 등 위에 관한 평가보다 교수·학습 활동의 수정에 유용한 정보를 얻을 수 있는 피드백을 강조하고 있다. 따라서 평가환류 영역에서는 서술형 평가시험과 채점기준 공시와 더불어 결과분석 자료 공시, 평가결과에 기초하여 수업계획을 수정·변경과 학생들의 장단점에 대한 개선안을 제공할 수 있어야 한다. 즉 평가결과는 성취기준, 성취수준에 따라 활용될 수 있으며, 이러한 전반적인 결과는 학생수준, 학급수준, 학년수준, 학교수준 나아가 시·도교육청수준 등에서 활용할 수 있어야 하겠다.

<표 III-2> 교사 평가전문성 신장 구성 영역 및 항목 내용 정리

영역	항목	항목 내용
평가환경	• 수행 및 서술형 평가에 관한 학교분위기	· 교사들이 서술형 및 수행평가에 우호적이다. · 학생들이 서술형 및 수행평가를 신뢰한다. · 학부모들이 서술형 및 수행평가를 신뢰한다.
	• 교육과정	· 2009 개정교육과정의 기본 개념을 설명할 수 있다. · 성취평가제에 대해 설명할 수 있다.
	• 교육방법 및 프로그램	· 성취기준을 명확히 정립할 수 있다. · 성취기준에 따라 다양한 교육방법을 적용할 수 있다. · 성취기준에 따라 다양한 수업설계를 적용할 수 있다.
	• 교육여건(인적·물적·제도적 지원), 교사 자율성	· 서술형 및 수행평가 문항 검토를 위한 교사협의회를 개최한다. · 교사들이 평가문항에 대한 논의를 자유로이 한다. · 서술형 및 수행평가를 하는 교사에게 다양한 인센티브를 제공한다.
	• 교육평가 환경(평가 비율, 채점시간 등)	· 서술형 및 수행평가를 지필평가와 따로 실시한다. · 서술형 및 수행평가를 할 시간적 여유가 충분하다. · 서술형 및 수행평가를 채점할 시간이 따로 마련되어있다.
평가투입	• 목표 설정	· 서술형 및 수행평가의 성취기준을 설정한다. · 성취기준에 따라 서술형 및 수행평가의 비율을 설정한다. · 교육과정에 따라 성취기준을 만들 수 있다. · 성취기준을 수정·보완한다. · 성취기준의 코드를 이해하고 활용한다.
	• 평가유형 선정	· 성취기준에 따른 명확한 서술형 및 수행평가 유형이 마련되어있다. · 성취기준에 따라 서술형 및 수행평가유형을 적절하게 선택한다. · 성취수준에 따른 서술형 및 수행평가유형이 준비되어있다.
	• 평가 윤리성	· 서술형 및 수행평가에 따른 성적 차별성을 제거하려고 노력한다. · 서술형 및 수행평가에 따른 경제적 차별성을 제거하려고 노력한다. · 서술형 및 수행평가에 따른 문화적 차별성을 제거하려고 노력한다.

영역	항목	항목 내용
평가과정	<ul style="list-style-type: none"> 평가도구 개발 	<ul style="list-style-type: none"> 서술형 및 수행평가의 목적과 내용에 맞는 문항을 개발한다. 문항개발의 절차를 이해하고 그에 따라 문항을 개발한다. 이원목적분류표의 내용영역과 행동영역을 이해하고 그 기준에 따라 문항을 개발한다. 서술형 및 수행평가 문항의 정답율을 구하고 예측한다. 능력집단별로 정답율을 구하고 예측한다. 성취수준별 기준점을 산출한다. 서술형 및 수행평가 도구의 타당도 및 신뢰도를 검토한다.
	<ul style="list-style-type: none"> 채점기준 정립 	<ul style="list-style-type: none"> 성취기준에 따른 채점기준을 설정한다. 성취수준에 따른 채점기준을 설정한다. 내용영역에 따른 채점기준을 설정한다. 행동영역에 따른 채점기준을 설정한다. 학생에게 채점기준을 정확하게 안내한다. 학생의 반응을 참조하여 채점기준을 수정·보완한다. 채점의 명확성을 검토한다.
평가결과	<ul style="list-style-type: none"> 평가결과 분석 	<ul style="list-style-type: none"> 서술형 및 수행평가 결과를 양적으로 분석한다. 서술형 및 수행평가 결과를 질적으로 분석한다. 서술형 및 수행평가에 대한 다양한 통계자료를 산출한다.
	<ul style="list-style-type: none"> 평가결과 해석 	<ul style="list-style-type: none"> 서술형 및 수행평가의 내용에 따른 평가결과를 해석한다. 서술형 및 수행평가 결과에 따라 학생의 장단점을 해석한다. 성취기준에 근거하여 평가결과를 적절하게 해석한다.
평가환류	<ul style="list-style-type: none"> 평가결과 공개 및 피드백 	<ul style="list-style-type: none"> 서술형 및 수행평가의 채점기준을 공시한다.(중간, 기말만 공시) 서술형 및 수행평가의 결과분석 자료를 공시한다. 평가결과에 기초하여 수업계획을 수정·변경한다. 평가결과에 기초하여 학생들의 장단점에 대한 개선안을 제공한다.
	<ul style="list-style-type: none"> 평가결과 활용 	<ul style="list-style-type: none"> 평가결과를 성취기준에 따라 활용한다. 평가결과를 성취수준에 따라 활용한다. 서술형 및 수행평가 결과를 학생수준에서 활용한다. 서술형 및 수행평가 결과를 학급수준에서 활용한다. 서술형 및 수행평가 결과를 학년수준에서 활용한다. 서술형 및 수행평가를 학교수준에서 활용한다. 서술형 및 수행평가를 시도교육청수준에서 활용한다.

중등교사 평가전문성 신장을 위한 영역은 평가환경, 평가투입, 평가과정, 평가결과, 평가환류 등 평가시스템 전반에 대한 포괄적 평가모형을 통해 구성된다. 영역별 구체적인 내용은 다음 [그림 Ⅲ-1]과 같다.



[그림 Ⅲ-1] 중등교사 평가 전문성 신장 영역

2. 교사의 평가 전문성 실태 분석

가. 평가환경 분석

1) 정량적 분석

가) 학교 급

평가 환경 전체 문항에 대하여 학교 급에 따른 기술통계 및 차이 검정을 실시하였다. 그 결과 학교 급에 따라 교사의 평가환경유에 대한 평가 전문성 지수에 통계적인 차이가 나타났으며, 중학교가 고등학교에 비하여 서술형 및 수행평가를 위한 평가 환경이 좋은 것으로 나타났다.

〈표 III-3〉 평가 환경 전체 문항에 대한 기술통계 및 학교 급에 따른 차이 검정

	학교 급	평균	표준편차	N	F
평가환경	중학교	3.6994	.59781	489	42.368***
	고등학교	3.2571	.65085	531	
	전체	3.4691	.66308	1,020	

***p<.001

나) 담당 교과목

평가 환경 전체 문항에 대하여 담당 교과목에 따른 기술통계 및 차이 검정을 실시하였다. 그 결과, 담당 교과목에 따라 평가환경에 대한 지수는 통계적인 차이가 없었다.

〈표 III-4〉 평가 환경 전체 문항에 대한 기술통계 및 담당 교과목에 따른 차이 검정

	교과목	평균	표준편차	N	F
평가 환경	국어	3.4315	.65937	186	1.773
	영어(외국어)	3.4519	.60979	234	
	수학	3.7000	.68109	165	
	과학과	3.3500	.70891	150	
	사회과	3.4336	.65804	192	
	기타	3.4435	.66052	93	
	합계	3.4691	.66308	1,020	

2) 정성적 분석

〈표 Ⅲ-5〉은 평가 환경에 관한 인터뷰 내용을 정리한 것이다.

〈표 Ⅲ-5〉 평가 환경 인터뷰 내용

“늘 해오던 것이라 당연하게 여기지만 서술형과 수행평가 채점에 대한 부담은 늘 가지고 있다. 교육부에서 제시하는 성취기준과 평가방법이 일관적이지 않고 모호하여 서술형 평가에 대한 자신감이 결여되어 있다. 그리고 성적산출에 있어서 민원이 생길 확률이 선다형 평가에 비해 많으므로 부담이 크다.”

“학생들이 글을 쓰는데 익숙하지 않음으로 인해 생기는 감점에 대해 민감하게 반응한다. 서술형 평가의 수준과 방법이 학교마다 교사마다 다르기 때문에 신뢰도가 떨어진다. 최소한 학년별, 학교 급별 평가문항이 일관성을 가지도록 교육부의 가이드라인 배무와 그의 홍보가 필요하다.”

“서술형 평가가 대부분 총점에서 차지하는 비중이 선다형보다 낮고 점수가 비슷하여 학생들, 교사들 모두 서술형 평가에 대해 크게 인지하고 있지 않다. 즉, 수업방식도 따로 설정하지 않는다. 서술형 평가의 비중을 높이고 서술형 평가의 양식도 세부적으로 기준을 정해 실패와 성공만 산출한다면 당연히 수업방식도 바뀔 것이다.”

“교사들은 자신이 출제한 문제에 대하여 자부심을 느끼고, 자신이 출제한 문제에 대해 장단점을 얘기하는 것을 좋아하지 않는다. 그래서 본인이 출제한 문항과 타인이 출제한 문항을 모아서 시험을 내는 경향이 많아서 각자 알아서 하는 편이다. 자유로운 논의를 위해서는 문제출제 방식의 변화가 필요하다.”

“서술형 평가는 시간과 노력이 많이 필요한 영역이라 상시로 실시하기가 어려운 부분이 많다. 서술형 평가를 지필평가와 같이 실시하는 것이 왜 잘못된 것이고 왜 보완해야 하는지 잘 모르겠다. 평가를 수시로 하는 것은 학생들에게 부담이 되고 교사도 짧은 일과시간 안에 이루어 내기가 힘이 든다. 교사입장에서 한과목이지만 학생들은 10과목 전후가 되는데 이를 수시로 친다는 것은 이해가 안 된다.”

“집중이수제로 단위는 늘었지만 여전히 가르쳐야 할 분량이 많다. 또 문항출제를 위해서는 주말을 할애하거나 일과이후에 남아서 해야 한다. 교무실이라는 곳은 학생들의 출입이 잦은 곳이라 보안 유지가 어렵다. 채점을 할 때도 세 번 검사까지 해야 하는 상황이라 같은 답안을 세 번이나 본다는 것도 쉽지 않다. 보완방법은 명확한 성취기준을 제시하고 실패와 성공 여부만 확인하여 점수를 준다면 채점의 노력을 줄일 수 있다.”

3) 종합결과

평가 환경은 교사들의 학생 평가 전문성 신장을 위하여 수행되어야 하는 일련의 활동들에 영

향을 미치는 전반적인 것과 무엇보다 교사가 서술형 평가 및 수행 평가의 본질적인 목표달성을 위해 수행해야 하는 제반적인 활동에 대하여 몰입할 수 있도록 주어져야 하는 여건 등을 의미한다.

평가 환경의 문제점은 서술형 평가와 수행평가에 적합한 교육여건(교육과정), 교수학습·방법, 교육평가 세부여건 등에서 문제점을 고찰할 수 있다.

첫째, 교육여건의 측면에서 서술형 및 수행 평가는 학생들의 창의적인 학습과 수학적 개념을 이해하는데 활용가치는 있지만, 현 교육여건 상 학생들의 학습 부담과 교사의 업무 부담으로 작용하고 있다. 교사들 간의 협업을 통한 서술형 평가의 기준을 정립할 수 있는 학교 분위기가 미약하고, 서술형 및 수행 평가에 적합한 즉, 학생을 위한 수업이 이루어질 수 없는 학교 분위기이다. 현 입시제도에서는 서술형 및 수행 평가 본래 취지에 맞는 시험을 출제(논리력 향상 등)하기보다는 학생들이 적을 수 있는 시험을 출제하고 있다.

둘째, 교수학습·방법의 측면에서 서술형 및 수행 평가만을 위한 교수학습·방법을 고찰하기란 불가능하다.

셋째, 교육평가의 측면에서 학생들의 수준을 고려하여 평가 기준을 수정하다 보면, 객관적인 기준 자체가 모호해 지는 경향이 있다. 이로 인하여 서술형 및 수행 평가는 평가기준이 불분명한 경우가 많고, 이의 비율조정 등 평가에 관한 교사 자율성 여건 형성이 유연하지 못하다. 또한 서술형 및 수행 평가를 채점하는데 시간과 노력이 많이 필요하다. 즉, 협의의 과정(부분점수 기준설정 등)을 거치면서 시간이 부족하면서 업무 부담으로 여겨지는 것이 현실이다.

4) 개선방안

학교 현장의 서술형 및 수행평가를 위한 평가환경에서의 문제점을 토대로 개선방안을 제시하면 <표 III-6>과 같다.

<표 III-6> 평가 환경 개선방안

평가 환경 문제점 영역	개선 방안
교육여건(교육과정)	<ul style="list-style-type: none"> 교사의 업무 부담을 경감시킬 수 있는 정책적 방안 모색 필요 서술형 및 수행 평가 취지에 맞은 교육여건(교육과정 등) 개선 방안 모색 필요 서술형 및 수행 평가 기준을 명확히 정립할 수 있는 방안 모색 필요(연수 등) 서술형 및 수행 평가를 위한 교사 협력이 이루어질 수 있는 학교분위기 모색 필요 서술형 평가 기준(안)을 참고할 수 있는 다양하고 구체적인 자료제시 방안 모색 필요

교수학습·방법	<ul style="list-style-type: none"> • 서술형 및 수행 평가를 위한 교수학습·방법개선을 위한 정책적 방안 모색 필요 • 서술형 및 수행 평가를 위한 다양한 교수학습·방법을 창의·인성 프로그램과 연계하여 지속적이고 안정적으로 정착시킬 수 있는 방안 모색 필요(지속적인 우수사례를 발견할 수 있는 정책적 방안 모색) • 서술형 및 수행 평가를 위한 다양한 교수학습·방법을 활용할 수 있는 수업시간 확보와 평가시기 결정 등의 합의점 방안 모색 필요
교육평가 세부여건	<ul style="list-style-type: none"> • 서술형 및 수행 평가 기준을 명확히 정립할 수 있는 안내 책자 제공 및 지속적인 홍보 방안 모색 필요 • 서술형 및 수행 평가 채점 시 교사들 간의 견해 차이를 조율하기 위한 방안 모색 필요 • 서술형 및 수행 평가만을 위한 수업시간 및 시험시간이 확보 방안 마련이 필요 • 서술형 및 수행 평가의 유연한 평가비율 조정, 시기와 범위문제, 출제와 채점시간에 부담감해소 방안 모색 필요 • 서술형 및 수행 평가에 관한 교사 안내책자뿐만 아니라 학생 답안 작성 방법 등을 담은 학생 교육용 안내 책자 제공이 필요

나. 평가투입 분석

1) 정량적 분석

가) 전반적 분석

중등학교 교사들은 평가투입 전문성(M=3.67)과 평가전문성 전체(M=3.65)를 비교해 볼 때 거의 같게 인식하고 있다. 중등학교 교사들이 인식하는 평가투입 전문성의 하위요인별 전문성을 살펴보면 평가 윤리성 전문성(M=3.73), 평가유형선정 전문성(M=3.67), 목표설정 전문성(M=3.64)의 순으로 나타났다.

<표 III-7> 평가전문성 전체 및 평가 투입 전문성

		N	평균	표준편차
평가전문성 전체		1,020	3.65	.56
평가투입 전문성	목표설정	1,020	3.64	.74
	평가유형선정	1,020	3.67	.63
	평가윤리성	1,020	3.73	.68
	전체	1,020	3.67	.75

나) 학교 급별 비교

중학교 교사와 고등학교 교사가 인식하는 평가투입 전문성을 비교해보면 <표 III-10>에서와 같이 중학교 교사가 고등학교 교사보다 평가투입 전문성을 더 높게 인식하고 있는 것으로 나타났다($F=70.821$, $p<.05$). 중학교 교사들이 고등학교 교사들보다 평가투입 전문성을 더 높게 인식하는 것은 자유학기제를 실시하는 중학교에서 서술평가 및 수행평가에 대한 관심과 실행이 정착된 것으로 풀이된다. 한편 고등학교에서는 선다형 지필고사로 이루어지는 수학능력시험에 초점을 맞추어지기 때문에 중학교만큼 서술 및 수행평가에 대한 관심이 높지 않다고 풀이된다.

<표 III-8> 평가투입 전문성에 대한 학교 급별 차이

		N	평균	표준편차	F
평가 전문성 전체	중학교	489	3.8915	.47496	70.821*
	고등학교	531	3.4283	.53497	
	합계	489	3.6504	.55687	
평가투입 전문성	중학교	531	4.0353	.57929	92.928*
	고등학교	489	3.3376	.74436	
	합계	531	3.6721	.75489	

*p<.05

다) 과목별 분석

아래 <표 III-11>는 평가투입 전문성 중 평가문항의 유형 선택 및 성취기준과 성취수준의 재설정을 중심으로 하는 평가투입 전문성에 대하여 중등학교 교사들의 과목별로 유의미한 차이(Wilks' λ =.912, $p>.05$)가 없는 것으로 나타났다.

<표 III-9> 평가투입 전문성에 대한 과목별 차이 검증

과목	N	평균	표준편차	Wilks' λ	F
국어	186	3.5766	.67385	.912	1.227
영어	234	3.5962	.72213		
수학	165	3.8364	.81681		
과학	150	3.5800	.81822		
사회	192	3.7266	.76858		

라) 문항별 분석

아래 <표 III-12>는 평가투입 전문성에 대한 문항별 분석이다. 목표설정과 관련하여 6번 문항 '성취기준 코드를 이해하고 활용한다.'은 중학교 교사가 고등학교 교사에 비해 이해 및 활용이 높게 인식하는 것으로 나타났다($t=8.629$, $p<.001$). 그러나 9번 문항 '성취기준과 성취수준을 재설정한다.'라는 문항에 대해서 중학교 교사와 고등학교 교사간에 차이가 없는 것으로 나타났다. 평가유형의 선정을 묻는 문항인 7번 문항 즉 '성취기준에 따라 서술형 평가문항의 유형을 적절하게 선택한다.'라는 질문에 대해 중학교 교사가 고등학교 교사에 비해 유형선택을 더 적절하게 선택하는 것으로 나타났다($t=6.589$, $p<.001$). 마지막으로 평가 윤리성 전문성인 8번 문항 '서술평가에 따른 성별, 가정적 요인에 따른 편파성을 최소화하려고 노력한다.'라는 질문에 대해 중학교 교사가 고등학교 교사에 비해 더 많은 노력을 기울이는 것으로 나타났다($t=5.417$, $p<.001$).

<표 III-10> 평가투입 전문성에 대한 문항별 분석

문항번호	학교 급	N	평균	표준편차	평균차(d)	t
6번	중학교	489	4.05	.760	.78	8.629***
	고등학교	531	3.27	.892		
7번	중학교	489	3.98	.780	.59	6.589***
	고등학교	531	3.39	.850		
8번	중학교	489	4.02	.864	.56	5.417***
	고등학교	531	3.46	1.027		
9번	중학교	489	3.70	.869	.11	1.193
	고등학교	531	3.59	.771		

***p<.001

2) 정성적 분석

가)목표설정

(1) 성취기준 수정·보완

“선생님께서 몸담고 있는 학교 또는 교과 교사들이 학생의 수준에 따라 성취기준을 수정·보완하는 수준은 어느 정도라고 생각하십니까?”라는 질문에 대하여 중학교 교사들이 학생의 수준에 따라 성취기준을 수정·보완하는 수준이 대체로 낮게(44.4%) 나타났다. 특히 수학교사는 교육부가 제시하고 있는 성취기준을 전혀 수정·보완하지 않고 그대로 따른다고 진술하였다.

고등학교 교사들이 학생의 수준에 따라 성취기준을 수정·보완하는 수준이 보통(50.0%)으로 나타났다. 특히 고등학교 국어교사와 영어교사는 교육부가 제시하고 있는 성취기준을 전혀 수정·보완하지 않고 그대로 따른다고 진술하였다. 이 질문에 대한 중등교사들의 심층면접 내용은 다음과 같다.

학교급	중학교	고등학교
항목		
성취기준 수정·보완	<p>“기존의 성취기준 및 성취수준이 잘 설정되어있어서 수정·보완의 필요를 느끼지 않으며, 또 성취기준을 그대로 활용 하는 것이 아니라, 어차피 수업에서는 좀 더 학생수준에 맞는 구체적인 활동을 하지만, 그 활동을 새롭게 용어로 고쳐 만들기는 번거롭고 어려우므로 성취기준 및 성취수준을 그대로 사용하거나 약간만 수정·보완하여 사용하고....</p> <p>나아가 성취기준에 대한 다양한 예시 자료를 제시하여 준다면,</p>	<p>“학생의 수준에 따라 성취기준을 수정·보완하는 것은 효과적인 수업의 첫걸음이라 생각하기 때문에 대체로 성취기준을 높이 설정하여 수업을 진행합니다”(사회교사)</p> <p>“학생들의 수준에 맞추어 성취기준을 수정·보완하고 있고....”(과학교사)</p> <p>“성취기준을 학생의 수준에 따라 수정·보완하지 않는다. 일반적으로 교사들이 성취기준을 변경해야 한다고 생각지 않을 뿐 아니라 수정·보완의 필요성을 느</p>

	<p>교사들이 필요하면 골라 쓸 수 있도록 하면 될 것...”(국어교사)</p> <p>“우리학교는 다른 학교에 비해 학생들의 학습능력이 조금 떨어지는 편이라, 교사들이 학생들의 수준과 수업의 내용들을 고려하여 성취기준을 선택하고 조합하여 반영합니다.”(영어교사)</p> <p>“교육부 제시 성취 기준에 따를 뿐이고...”(수학교사)</p> <p>“성취기준에 따라 학생의 성취정도에 대한 정보 및 학습개선을 위한 피드백을 적용하는 수준입니다.”(사회교사)</p>	<p>끼지 못하고..</p> <p>학생의 학업능력을 정교하게 평가하기 위하여 성취기준을 분할·결합하는 과정이나 새롭게 성취기준을 제작하는 과정을 제공할 필요가 있다는 생각....”(영어교사)</p> <p>“성취기준 자체를 중요시 하지 않기 때문에 수정 보완하지 않고...</p> <p>개선책으로는 입시제도 개편 , 성취기준에 따른 출제를 학기 시작 전에 시행, 다양한 방식으로 교과협의회 대신 짝 티칭을 활성화해야...”(국어교사)</p>
--	-------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------	-------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------

(2) 성취기준 재구성

중학교 교사들의 성취기준 재구성에 대한 설문조사 응답 결과는 3.70/5로 74.0% 정도가 성취기준을 재구성하는 것으로 나타났다. 설문조사 분석 결과에 비추어 볼 때 심층면접 교사들 중 수학 교사는 성취기준을 수정·보완하지 않고 그 기준에 따르려는 것으로 해석된다.

고등학교 교사들의 성취기준 재구성에 대한 설문조사 응답 결과는 3.59/5로 71.8% 정도가 성취기준을 재구성하는 것으로 나타났다. 설문조사 분석 결과에 비추어 볼 때 심층면접 교사들은 성취기준을 재구성하는 정도가 보통이고(56.6%) 특히 고등학교 국어, 영어 교사는 성취기준을 수정·보완하지 않고 그 기준에 따르려는 것으로 진술하였다. 이 질문에 대한 중등교사들의 심층면접 내용은 다음과 같다.

학교급 항목	중학교	고등학교
성취기준 재구성	<p>“기존의 성취기준이 잘 설정되어 있어서 수정·보완의 필요를 느끼지 않으며, 번거롭고 어려우므로 있는 성취기준의 수준을 그대로 사용하거나 약간만 수정하고.... 성취기준에 대한 다양한 예시 자료를 제시하여 주어, 교사들이 필요하면 골라 쓸 수 있도록 하면 된다”(국어교사)</p> <p>“성취기준은 절대적 진리가 아니기에 학교 수업 현장에서 능동적으로 수정·보완되어야 한다. 배움의 제반적인 환경에 따라</p>	<p>“현실적으로 고등사고능력을 평가하기 위해 개발된 성취 기준에 없는 새로운 성취기준을 재구성하기도...”(사회교사)</p> <p>“학생들의 수준에 맞는 성취기준의 재구성은 반드시 필요하지만, 실제로 재구성한 교과내용을 가지고 수업을 하는 경우는 거의 없다. 예를 들면 분수의 덧셈과 뺄셈을 모르는 학생들에게 미분과 적분을 가르치는 상황이 되는 경우가 많고.... 학생들 수준에 맞는 성취기준의 재구성에</p>

	<p>역동적으로 수정·보완된 그 학교만의 성취기준이야말로 진정한 성취기준이라고 생각된다. 동 교과 교사와 협의를 통해 수업에 대한 학생들의 피드백을 살펴보고 성취기준의 중요한 요소들을 보다 쉽고 효율적으로 성취할 수 있는 방법들을 논의해야 한다.”(영어교사)</p> <p>“교육부 제시 성취 기준에 따라서 시행할 뿐이고....”(수학교사)</p> <p>“성취기준에 따라 학생의 성취 정도에 대한 정보 수집 및 학습 개선을 위한 피드백을 적용하는 수준이고...”(사회교사)</p>	<p>많은 시간과 노력을 투자해야 하지만 특별한 인센티브도 없는 상황에서 교사의 양심에만 호소하는 현실이 안타깝기도 하고....”(과학교사)</p> <p>“성취기준을 학생의 수준에 따라 수정·보완하지 않는다. 일반적으로 교사들이 성취기준을 변경해야 한다고 생각지 않을 뿐 아니라 수정·보완의 필요성을 느끼지 못하고....”(영어교사)</p> <p>“재구성하지 않고, 성취기준에 대한 인식도 없고...</p> <p>개선책으로는 재교육기간 확보, 방학 중 소집교육, 성취기준 자체를 자율적으로 설정할 수 있게 하는 제도적 장치가 마련되어야...”(국어교사)</p>
--	-----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------	---------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------

(3) 평가유형 선정

성취기준에 부합하는 평가 유형 선택에 대하여 중학교 교사들은 대체로 높은 응답 반응(77.8%)을 보이고 있다. 성취기준에 부합하는 평가 유형 선택에 대하여 고등학교 교사들은 대체로 보통 응답 반응(56.7%)을 보이고 있다. 이 질문에 대한 중등교사들의 심층면접 내용은 다음과 같다.

학교급 항목	중학교	고등학교
평가유형 선정	<p>“성취기준에 따라 어떤 평가유형이 어울릴지를 미리 결정하기에 주로 수행평가에 어울리는 단원은 수행평가로 실시하고 나머지를 모두 지필평가 대상으로 삼으며, 지필평가의 문제 출제 시에 서술형 평가에 어울리는 문제를 먼저 만들고, 그 외의 나머지는 지필평가로 문제를 만들고..”(국어교사)</p> <p>“성취기준과 교과서의 학습요소들을 살펴보고 평가해야 할 요소에 대해 동 교과 교사끼리 논</p>	<p>“단순 지식, 이해 정도를 평가할 때는 지필평가를 실시하고, 주로 고등사고능력을 평가할 때는 서술형 평가를 실시하고, 수행과정을 평가하고자 할 때에는 수행평가를 실시하고...”(사회교사)</p> <p>“일단 성적관리규정에서 규정되어 있는 형식만은 철저히 지키고 있고...(과학교사)</p> <p>“성취기준에 따라 평가유형을 설정할 필요가 없고... 세 가지 평가 유형은 대부분이 성취기준에 적용 가능하고.. 성취기준</p>

	<p>의한다.”(영어교사)</p> <p>“교육부 권장 수준 준수할 뿐이고...”(수학교사)</p> <p>“지필평가는 단순 지식, 이해 정도를 평가하고, 사고력, 문제해결력, 창의력 등 고등 사고 기능은 서술, 논술형 평가를 하며, 수행과정은 연구·실험보고서, 프로젝트, 역할극, 협력학습, 포트폴리오 등 과정 중심의 다양한 수행평가 도구를 활용합니다.”(사회교사)</p>	<p>의 상세화가 필요하다고....”(영어교사)</p> <p>“관습적인 평가유형을 그대로 시행하기에 개선책으로 평가유형 모형개발, 교육청 단위의 모형 제시, 업무부담 경감 등이 필요하고...”(국어교사)</p>
--	-----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------	-----------------------------------------------------------------------------------------------------------------------

(4) 서술형 평가 및 수행평가 비율 설정

중등학교 교사들은 대체로 서술형 평가 및 수행평가의 비율을 성취기준에 따라 설정하는 것으로 반응을 보였다(72.2%, 70%). 이 질문에 대한 중등교사들의 심층면접 내용은 다음과 같다.

학교급 항목	중학교	고등학교
서술형 평가 및 수행평가 비율설정	<p>“성취기준에 따라 어떤 평가가 어울릴지를 결정하는데 주로 수행평가에 어울리는 단원은 수행평가로 실시하고 나머지를 모두 지필평가 대상으로 삼으며, 지필평가의 문제 출제 시에 서술형 평가에 어울리는 문제를 먼저 만들고, 그 외의 나머지는 지필평가로 문제를 만들어...”(국어교사)</p> <p>“수행평가 비율은 교육청의 권고를 많이 따르는 편이며, 서술형 평가는 올해 혁신학교를 도입하며 배움 중심의 수업을 제대로 해보자는 취지에서 그 연장선상에 있는 평가를 위해 서술형 평가를 반영 비율을 늘이려고 노력하고 있습니다.”(영어교사)</p> <p>“교육부 권장 수준 준수할 뿐이고...”(수학교사)</p> <p>“과목별로 성취기준 도달 여부</p>	<p>“단순 지식, 이해 정도를 평가할 때는 지필평가를 실시하고, 주로 고등사고능력을 평가할 때는 서술형 평가를 실시하고, 수행과정을 평가하고자 할 때에는 수행평가를 실시하고...”(사회교사)</p> <p>“과학 교과의 평가 비율 중에서 서술형 평가 및 수행평가의 비율은 30~40% 정도의 비율로 실시되고...”(과학교사)</p> <p>“성취기준에 따라 서술형 평가 및 수행평가의 비율을 설정할 필요를 느끼지 못하고...</p> <p>성취기준과 수준을 더 상세화하고 그에 따라 서술형 문항 및 수행형 문항의 비율이나 점수를 적정화해야...”(영어교사)</p> <p>“핵심 성취기준에 더 많은 비율을 배분하고 있고...”(국어교사)</p>

	를 평가할 수 있는 서술형 평가 및 수행평가의 비율을 성취기준에 따라 설정하고 있다”(사회교사)	
--	-------------------------------------------------------	--

(5) 성취수준 코드체계 이해·활용

중등학교 교사들의 성취기준 코드체계 이해 및 활용에 대한 설문조사 응답 결과는 4.05/5로 81.0% 정도가 성취기준 코드체계를 이해하고 활용하는 것으로 나타났다. 설문조사 분석 결과에 비추어 볼 때 심층면접 교사들은 성취기준 코드체계를 이해는 하나 활용 수준이 낮은 것(52.2%)으로 해석된다. 이 질문에 대한 중등교사들의 심층면접 내용은 다음과 같다.

학교급 항목	중학교	고등학교
성취수준 코드체계 이해·활용	<p>“대부분 자신들의 과목에 대한 성취기준 코드 체계를 이해하는 것이 아니라, 성취기준의 내용을 보고 교과서의 내용과 어떻게 일치하는지 확인하고...”(국어교사)</p> <p>“교사들은 코드체계를 인식하기 보다는 성취기준의 내용에 더 관심을 기울이는 편입니다. 성취기준의 코드체계 도입에 대한 필요성과 그 실체가 어떠한지 연수가 요구되는 ...”(영어교사)</p> <p>“성취기준 코드체계에 대한 기초연수는 평가계에서 하기는 하나 구체적인 이해 잘 되지 않고...”(수학교사)</p> <p>“체계적이고 반복적인 연수 필요한데.. 특히 전달연수가 정착되어야 할 것으로....”(사회교사)</p>	<p>“어떤 의미인지 몰랐기 때문, 일부 교사(연구부장, 평가담당 교사)만 연수를 받았고...</p> <p>성취기준 코드체계를 이해할 수 있도록 연수강화, 성취평가제 연수강화, 성취평가제 중요성 인식 제고가 필요하고...특히 전달연수가...”(사회교사)</p> <p>“성취기준 코드에 대한 교육은 실시하고 있으며 평가에 반드시 반영하도록 하고 있으나, 기존에 관행적인 평가 습관이 아직도 몸에 배어 있는 교사들이 있어”(과학교사)</p> <p>“성취기준과 코드를 그대로 옮겨 적을 뿐이고...”(영어교사)</p> <p>“코드로만 활용할 뿐이고...”(국어교사)</p>

(6) 성취기준에 적합한 서술형 평가 유형선택

성취기준에 부합하는 서술형 평가 유형 선택에 대한 설문조사 응답 결과는 3.98/5로 79.6% 정도가 성취기준에 부합하는 평가 유형 선택하는 것으로 나타났다. 그러나 심층면접 교사들도 대체로 이보다 낮은 응답 반응(중학교교사71.1%,고등학교교사 70.0%)을 보이고 있다. 이 질문에 대한 중등교사들의 심층면접 내용은 다음과 같다.

학교급	중학교	고등학교
-----	-----	------

항목		
<p>성취기준에 적합한 서술형 평가 유형선택</p>	<p>“수행평가는 성취기준대로 하지만, 서술형 평가는 지필평가를 출제하면서 문제를 서술형으로 내면 적당하다고 판단되는 것을 서술형으로 내기 때문에, 서술형 평가의 기준은 성취기준에 적합성 여부를 먼저 기준으로 삼지 않고...”(국어교사)</p> <p>“성취기준이나 평가요소에 대해 파악이 이뤄진다면이라도 적절한 서술형 평가 문항에 대한 유형을 선택하지는 않는 것 같다. 교육청에서 제공하는 서술형 문제 유형을 참고하기도 하나 일단 유형에 대한 데이터베이스가 약한 편이다.”(영어교사)</p> <p>“교육부 제시 매뉴얼을 참고하여 선택 활용하고...”(수학교사)</p> <p>“중학교 급에서 고등사고능력을 평가하기는 어려우므로 주로 자료를 제시하여 이를 근거로 응답하도록 하는 유형을 선택하고...”(사회교사)</p>	<p>“고등사고능력을 평가하기 위해 서는 서술형 평가가 적합하기 때문에 자료 제시형으로 문항에 글이나 그림, 도표 등의 자료를 제시하고 이를 근거로 응답하도록 하는 문항 이 많고..”(사회교사)</p> <p>“성취기준에 부합하는 서술형 평가 문항을 제작하고 있으나 난이도는 학생 수준을 고려하여 제작하고...”(과학교사)</p> <p>“측정 가능한 성취기준을 적용하지 않는다. 서술형 평가 문항이 단답형 중심이고 다양하지 않습니다. 현재 일반적으로 알려져 있는 서술형 평가문항은 분류를 위한 분류일 뿐 성취기준이나 채점기준을 위한 유형이 아니라고 봅니다.</p> <p>성취기준과 수준과 서술형 평가 문항을 연계하는 평가 유형이 개발되어야 합니다.”(영어교사)</p> <p>“개방형, 발산형 문항을 알고 잘 활용하고 있고....”(국어교사)</p>

(7) 성취기준에 적합한 수행평가 유형선택

학교급 항목	중학교	고등학교
성취기준에 적합한 수행평가 유형선택	<p>“대부분 수행평가의 유형을 결정할 때 성취기준을 중심으로 결정하고...”(국어교사)</p> <p>“수행평가에 적절한 유형을 선택하고 있습니다. 영어과이므로 4영역을 평가하기 위해 듣기는 교육청에서 실시하는 듣기평가로 대체하고 있으며, 말하기는 원어민과 1분간에 주어진 주제에 대해 대화를 주고받는 평가입니다. 쓰기는 교과서의 문법을 이용해 작문해내거나 교과서의 내용을 토대로 자기생각을 기술하는 평가입니다. 읽기는 수행평가로 실시하지 않고 지필평가를 통해 이루어집니다.”(영어교사)</p> <p>“성취기준에 적합한 학습과제 수행 과정 및 결과를 직접 관찰하고 그 결과를 판단하므로 다양한 수행평가 유형 즉 구술, 연구·실험보고서, 프로젝트, 역할극, 협력학습, 포트폴리오, 토론, 자기평가, 동료평가 등을 활용합니다.”(사회교사)</p>	<p>“학생들이 다양한 장면에서 자신의 지식과 기능을 활용할 수 있는 능력이 어느 정도인지를 평가할 수 있는 방법이기 때문에 구술시험, 연구보고서법, 포트폴리오법, 실험실습법 등 다양한 유형을 적용하고...”(사회교사)</p> <p>“성취기준에 적합한 수행평가 유형을 선택하여 문제를 출제하고... 과제보고서, 실험실습등의 방법을 많이 활용하고...”(과학교사)</p> <p>“성취기준을 수행평가에 적용하지 않는다. 태도, 과제감사, 실험, 독후감 등 수행평가는 행동요소가 인지적 영역과 정의적 영역에 혼재해 있고 성취기준도 설정되지 않아서...</p> <p>성취기준 및 성취수준과 수행평가 유형을 연계할 수 있도록 수행평가 유형이 많이 개발되어야 합니다”(영어교사)</p> <p>“국어과 전체에서는 직접 말하기, 프로젝트 보고서 등의 수행평가를 시행하고 있고....”(국어교사)</p>

(8) 성별, 경제적, 문화적 요인에 따른 편파성 최소화 노력

서술형 평가 및 수행평가를 시행함에 있어서 학생의 성별, 경제적, 문화적 요인에 따른 편파성을 최소화하려는 노력에 대한 설문조사 응답 결과는 4.02/5으로 80.4% 정도가 학생의 성별, 경제적, 문화적 요인에 따른 편파성을 최소화하려고 노력하는 것으로 나타났다. 심층면접 교사들은 이보다 훨씬 높은 반응(94.4%)을 보이고 있다.

학교급 항목	중학교	고등학교
성별, 경제적, 문화적 요인에 따른 편파성 최소화 노력	<p>“서술형이나 수행평가를 칠 때 장애학생의 편의를 봐주는 정도이고 나머지 학생들은 일괄적으로 같은 조건에서 시행함으로</p>	<p>“학생평가의 기본원리 중의 하나는 공정성이라 여기기에 학교현장에서 보면 남녀요인 따라 편파성이 있을 수 있는가를</p>

	<p>성별, 경제적, 문화적 요인에 따른 편파성은 거의 없다고...”(국어교사)</p> <p>“서술형 평가나 수행 평가를 실시할 때는 학생들이 가지는 성별, 경제적, 문화적 차이에 대해 최대한 중립적이어야 한다는 것은 당연한 일입니다. 학생들이 가지는 배경에 의해 평가가 영향을 받아서는 안된다고 생각합니다.”(영어교사)</p> <p>“평가 내용에 성별, 경제적, 문화적 요인 반영 요소 없고...”(수학교사)</p> <p>“서술형평 및 수행평가를 시행함에 있어서 객관성, 합리성, 신뢰성 등을 기준으로 학생의 성별, 경제적, 문화적 요인에 따른 편파성을 최소화하려고 노력하고...”(사회교사)</p>	<p>면밀히 조사하고...성별, 경제적, 문화적 요인에 따른 편파성은 거의 없다고 판단되고...”(사회교사)</p> <p>“학업성적관리 규정에 이에 관한 내용이 있으며, 평가 문항 검토 시 점검하고 있기에..”(과학교사)</p> <p>“배운 범위 내에서 평가하기 때문에, 동일 문화권이어서 편파 문항은 없다고 생각하지만.. 편파문항에 대한 이해와 예시 필요하고..(예, 국어문법, 표준어 문제 - 서울경기출신 유리, 부산 경남출신 불리)“(영어교사)</p> <p>“수행, 서술형 평가 시에는 교과서 학습활동 위주의 출제를 지향하여 편파성을 최소화하려고....”(국어교사)</p>
--	--------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------	-------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------

3) 종합결과

평가투입 전문성의 제 1의 하위요인은 목표설정이다. 목표설정 요인의 첫 번째 항목은 ‘학생의 수준에 따른 성취기준의 수정·보완’에 대한 심층면접 결과를 살펴보면 대체로 낮은 수준이다. 심층면접 자료에서 보듯 현장 교사들은 성취기준을 중시하지 않는다는든지, 아예 교육부에서 제시된 성취기준을 무작정 따르려는 경향을 보이고 있다. 예외로 “학생의 수준에 따라 성취기준을 수정·보완하는 것은 효과적인 수업의 첫걸음이라 생각하기 때문에 대체로 성취기준을 수정·보완하고 이를 바탕으로 수업을 진행합니다.”는 교사도 있다. 이는 서술형 평가 및 수행평가 문항의 기능이 잘 발휘되게 하기 위해서는 학생집단의 특성을 고려하여 성취기준을 수정·보완하는 것이 필수 선행결과제라 하겠다.

목표설정 요인의 두 번째 항목은 ‘서술형 평가 및 수행평가를 위하여 성취기준을 재구성하는 정도’이다. 이에 대한 설문조사 응답결과는 높게 반응했으나 심층면접 결과를 살펴보면 대체로 재구성하지 않는 것으로 나타났다. 물론 예외적으로 ‘서술형 평가 및 수행평가가 고등사고능력을 향상시킬 수 있는 과제이어야 한다는 의미에서 적극적으로 성취기준을 재구성한다’라는 교사도 있긴 하지만 대체로 심층면접 교사들은 이를 등한시 하고 있다. 아예 ‘성취기준에 대한 인식이 없다.’라고 응답한 교사도 있다. 바람직한 평가로서 서술형 평가 및 수행평가가 제 기능 즉, 학생의 고등사고능력을 향상시킬 수 있는 평가, 학습결과 뿐만 아니라 학습 과정까지 다룰 수 있는 평가, 수

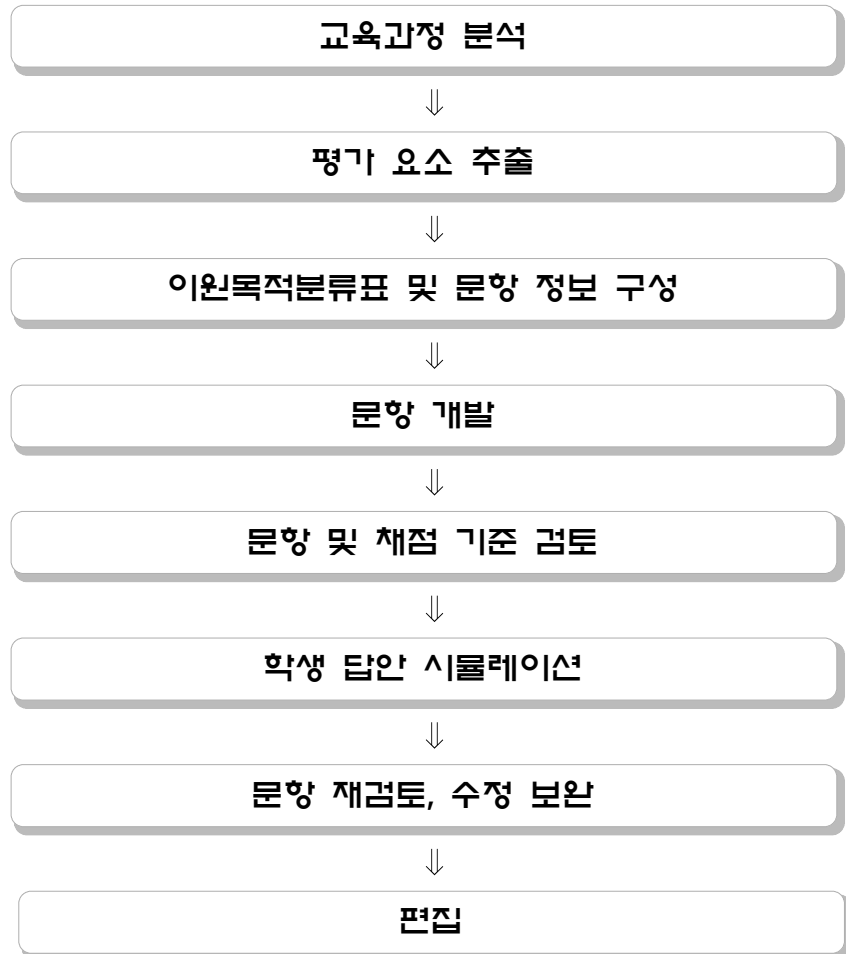
업과 통합된 지속적인 활동이 될 수 있는 평가, 학습 개선을 위한 형성평가가 되기 위해서는 성취 기준을 재구성하는 것은 필수적인 사안이라 하겠다.

평가투입 전문성의 제 2의 하위요인은 평가유형설정이다. 평가유형설정 요인의 첫 번째 항목인 ‘성취기준에 적합한 서술형 평가 유형선택’문항이다. 이에 대한 설문조사 및 심층면접 결과를 살펴보면 설문조사에서는 80%정도가 그렇게 하는 것으로 나타났으나 심층면접에서는 대체로 낮은 응답 반응을 보이고 있다. “서술형 평가는 지필평가를 출제하면서 문제를 서술형으로 내면 적당하다고 판단되는 것을 서술형으로 내기 때문에, 서술형 평가의 기준은 성취기준에 적합성 여부를 먼저 기준으로 삼지 않고...”, “서술형 평가 문항이 단답형 중심이고 다양하지 않다. 현재 일반적으로 알려져 있는 서술형 평가문항은 분류를 위한 분류일 뿐 성취기준이나 채점기준을 위한 유형이 아니다.”라고 부정적으로 응답하고 있다.

평가유형설정 요인의 두 번째 항목인 ‘성취기준에 적합한 수행평가 유형선택’문항이다. 이에 대한 응답은 대체로 높은 반응을 보이고 있다. 그러나 심층면접 결과 중학교 교사들은 수행평가의 취지에 걸맞게 “다양한 수행평가 유형 즉 구술, 연구·실험보고서, 프로젝트, 역할극, 협력학습, 포트폴리오, 토론, 자기평가, 동료평가 등을 활용한다.”라고 응답하고 있다. 그러나 고등학교 교사들은 성취기준에 적합한 수행평가의 유형선택이 이루어지지 않을 뿐만 아니라 대부분 서술형 평가로 대체하고 있는 실정이다.

평가투입 전문성의 제 3의 하위요인은 편파성 최소화 노력이다. 이 항목인 ‘성별, 경제적, 문화적 요인에 따른 편파성 최소화 노력’에 대한 응답 결과를 살펴보면 대체로 높은 수준의 노력을 경주하고 있다고 파악된다. 심층면접 결과를 살펴보면 “서술형 평가나 수행 평가를 실시할 때는 학생들이 가지는 성별, 경제적, 문화적 차이에 대해 최대한 중립적이어야 한다는 것은 당연한 일이다. 따라서 학생들이 가지는 배경에 의해 평가가 영향을 받아서는 안된다고 생각한다.” 라고 응답을 보이고 있다. 또한 “학생평가의 기본원리 중의 하나는 공정성이라 여기기에... 학교현장에서 보면 남녀요인 따라 편파성이 있을 수 있는가를 면밀히 조사하고...성별, 경제적, 문화적 요인에 따른 편파성은 거의 없다고 판단되고...”라는 응답을 보이고 있다.

4) 우수사례 적용 및 분석



[그림 III-2] 서술형 및 수행평가 문항 제작 흐름도

서술형 및 수행평가의 우수사례를 일반적인 관점에서 분석하면 문항 제작 흐름도에 따라 문항을 제작하였으며([그림 1] 참조), 흐름상의 각 과정을 구체적으로 진술하였다. 특히 이원목적 분류표 및 문항 정보 구성이 아주 구체화되어있다. 또한 교육부에서 마련한 성취기준을 서술형 평가 문항제작에 적합하고 자세하게 기술하고 있다. 문항개발과정에서는 문항유형, 문제장면, 구체적 예시, 평가요소, <보기> 자료, <조건> 등을 구체적으로 설명하고 있다. 문항 및 채점 기준 검토과정에서는 동료교사와의 협의를 통한 신뢰도 및 타당도를 확보하는 과정을 거쳤다는 점이다. 문항 곤란도 기준의 명세화, 학생 답안 시뮬레이션, 조건별 분할점수 제시, 채점기준을 내용과 형식으로 나누어 제시하고 있다. 또한 문항검토 및 수정과정을 거침으로써, 즉 서술형 평가 제작과정을 충실하게 따름으로서 서술형 평가의 타당성을 높이고 나아가 교사의 서술형 평가 전문성의 전형을

보여주고 있다.

서술형 및 수행평가의 우수사례를 평가투입 전문성의 관점에서 분석하면 첫째, 평가투입 전문성의 하위요인인 성취기준의 수정·보완을 잘 이루어졌다는 점이다. 즉 교육부에서 개발한 성취기준인 ‘731-1 읽기의 개념을 말할 수 있다.’를 ‘731-1. 읽기가 글의 의미를 파악하는 과정이라는 것을 설명할 수 있다.’로 성취기준을 수정·보완하여 성취기준과 서술형 평가문항 간의 유기적 연관성을 높이고 있다. 둘째, 평가요소(교육과정, 성취기준, 성취수준 등에 근거하여 서술형 평가에서 묻고자 하는, 즉 학생이 응답하기를 요구하는 요소)추출과정에서 ‘개념이해’라는 평가유형을 설정하고 있다. 이는 본 연구에서 제시한 Bloom의 평가 유형(<표 II-22> 참고)에서는 제 4유형에 해당된다. 셋째, 평가투입 전문성의 제 3의 하위요인은 성별, 경제적, 문화적 요인에 따른 편파성 최소화 노력이다. [그림 III-3]의 서술형 평가 문제에서는 편파적인 요소가 전혀 보이지 않는다. 따라서 [그림 2]의 서술형 평가 문제는 평가투입 전문성의 관점에서도 우수사례라고 볼 수 있다.

문항 정보

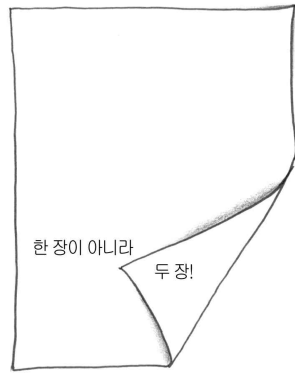
영역	읽기	학년	중1
곤란도	하	배점	4
문항 유형	정리·요약형		
성취 기준	731-1. 읽기가 글의 의미를 파악하는 과정이라는 것을 설명할 수 있다.		
평가 요소	읽기가 의미를 파악하는 과정임을 알기		

문항 01

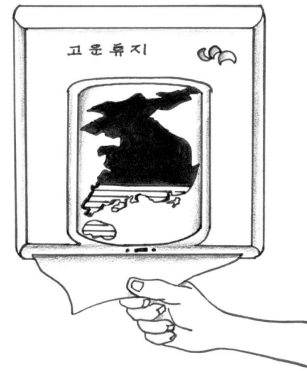
(가)와 (나)를 읽고 <보기>의 밑줄 친 부분에 적절한 말을 쓰시오. 【4점】

(가) 글자만이 읽기의 대상인 것은 아니다. 다른 많은 것들도 읽기의 대상이 된다. 우리의 생활은 읽기의 연속이라고 해도 지나친 말이 아니다. 횡단보도의 신호등이 푸른색으로 바뀌었을 때 우리는 그것이 건너도 좋다는 의미라는 것을 안다. 이때 우리는 신호등을 단순히 바라본 것에 그치지 아니하고, 신호등의 의미를 읽은 것이 된다. 반가운 친구의 얼굴에 눈길이 가는 것은 ‘보는’ 것이지만, 친구의 기분이 어떠한지 얼굴 표정을 살펴 알아내는 것은 ‘읽는’ 것이 된다. 지하철의 노선도도 그냥 보는 사람에게는 복잡한 무늬에 불과하지만, 그것의 의미를 읽을 줄 아는 사람에게는 어떻게 지하철이 운행되는지를 알려 준다. 우리는 인터넷 아이콘도 읽고, 지도도 읽고, 신문, 텔레비전, 영화도 읽는다. 우리는 그것들을 읽을 줄 앎으로써 비로소 우리 생활에 활용하고, 또 거기에 들어 있는 재미를 즐길 수 있다.

(나) 다음은 여러 가지 공익 광고의 일부분이다.



<그림 1>



<그림 2>

<보 기>

질문	<그림1>	<그림2>
무엇을 그린 거지?	종이 한 장이 말려 올라가 있고 “한 장이 아니라 두 장” 이라고 적혀 있네.	휴지를 뽑으니까 지도에서 숲을 의미하는 짙은 부분이 줄어들고 있네.
그림의 의미는 뭘까?	종이에 쓸 수 있는 면이 두 면이라는 의미인 것 같아.	휴지를 사용할수록 우리나라의 숲이 없어진다는 의미인 것 같아.
그림을 그린 이유가 뭘까?	그림을 보는 이에게 어떠한 교훈을 전달하기 위한 의도겠지.	
두 그림에서 공통적으로 읽어내야 할 교훈은 무엇일까?	“ _____ ” 라는 교훈을 알 수 있겠네.	



답안



답안 및 채점 기준

1 기본 답안 및 인정 답안

구분	내 용
기본 답안	<ul style="list-style-type: none"> · 자원을 절약하자. · 절약하는 마음을 갖자. · 종이와 휴지를 아껴 쓰자.
인정 답안	<ul style="list-style-type: none"> · 우리 주변의 물건을 아끼자. · 우리 주변의 물건을 아껴야 한다. · 종이와 휴지를 아껴서 자연을 지키자.

2 채점 기준

채점 요소	채점 기준	배점
내용	두 그림이 주는 공통적 교훈을 적절하게 읽어낸 경우	4
	· ‘휴지를 아껴 쓰자.’, ‘종이를 아껴 쓰자.’ 와 같이 두 그림의 의미를 각각으로 이해한 경우	3
	· 그림의 의미를 ‘자연, 환경’ 등 포괄적으로 해석한 경우	
형식	맞춤법에 어긋난 경우(횃수에 관계없이)	-1

[그림 Ⅲ-3] 국어과 서술형 평가 예시

5) 평가 전문성 제고 방안

목표설정과 관련하여 학생의 학업능력을 정교하게 평가하고 나아가 교사의 평가투입 전문성을 향상시키기 위해서 성취기준을 분할·결합하는 과정이나 새롭게 성취기준을 제작하는 연수 과정이 필요하다. 또한 [그림 Ⅲ-3] 국어과 서술형 평가 예시에서 보듯 교육부에서 개발되어 제시된 성취기준은 포괄적이고 추상적이기 때문에 학생의 수준에 걸맞게 성취기준을 수정·보완 기회가 마련되어야 할 것이다. 또한 서술형 평가의 취지인 학생의 고등사고능력을 향상시킬 수 있는 평가, 학습 결과뿐 만아니라 학습 과정까지 다룰 수 있는 평가, 수업과 통합된 지속적인 활동이 될 수 있는 평가, 학습 개선을 위한 형성평가를 먼저 생각하고 이에 부합하는 서술형 평가의 유형을 선정해야 할 것이다.

다. 평가 과정 분석

1) 정량적 분석

가) 학교별 분석

평가과정의 두 가지 평가전문성 핵심요소는 평가문항의 제작과 평가문항의 채점기준 및 루브릭 작성인데, 아래 <표 III-7>에서 보듯이 평균이 3.5를 넘어 교사들의 ‘평가과정 전문성(EPP)’이 우수하다고 해석될 여지가 있다. 그러나 이 수치는 일반적으로 교사집단이 자기평가와 동료평가에서 보여주는 평균적 수치($M > 4.0$)에 비하여 낮게 나타난다는 점에 유의할 필요가 있다. 이는 전반적으로 중등교사들이 평가과정에 대한 전문성이 떨어진다는 것을 인식하고 있다는 것을 의미한다고 해석할 수 있다.

중학교 교사와 고등학교 교사의 평가과정 전문성을 비교해보면 <표 III-7>에서와 같이 중학교 교사가 고등학교 교사보다 평균이 더 높고 그 차이($d = .4199$, $p < .05$)가 유의미하다. 중학교 교사들이 고등학교 교사들보다 평가과정 전문성이 더 높게 나타나는 것은 1,2학년만 성취평가제를 시행하고 대학입시의 영향력이 강한 고등학교보다 성취평가제가 전 학년 전 교과에서 실행되고 있고 또 자유학기제를 하는 중학교에서 서술과 수행에 대한 관심과 실행이 높기 때문일 것이다. 그리고 전체평가 전문성(TEP)과 평가과정 전문성(EPP)을 비교해보면 전체에 비하여 평가과정이 더 높다. 중학교와 고등학교 두 집단 모두에 같은 현상이 나타난다. 이는 중등학교 교사들이 평가환경, 평가투입, 평가결과 그리고 평가환류 영역보다 평가과정 영역, 즉 서술형 평가나 수행평가와 관련하여 시험문제출제와 채점기준설정을 상시적으로 시행하고 있기 때문인 것으로 풀이된다.

<표 III-11> 평가과정 전문성(EPP)에 대한 학교 급별 차이

		N	평균	표준편차	F
전체평가 전문성(TEP)	중학교	489	3.8915	.47496	70.821*
	고등학교	531	3.4283	.53497	
	합계	340	3.6504	.55687	
평가과정 전문성(EPP)	중학교	489	3.9340	.56302	41.731*
	고등학교	531	3.5141	.62993	
	합계	1,020	3.7154	.63375	

* $p < .05$

나) 과목별 분석

아래 <표 III-13>은 평가문항의 제작과 평가문항의 채점기준 및 루브릭 작성을 중심으로 하는 평가과정 전문성에 대하여 중등학교 교사들의 과목별로 유의미한 차이(Wilks' $\lambda = .937$, $p < .05$)가 있다는 것을 보여준다. 그러나 수학과 과학 교과간의 차이($d = .3377$, $p < .05$)를 제외하면 나머

지 과목은 유의미한 차이를 드러내지 않는다. 즉 수학교과와 평가과정 전문성(EPP) 평균값이 4.0178로 다른 교과보다는 높기는 하지만 교사집단의 자기평가나 동료평가의 평균수치(M>4.0)의 최저수준으로 낮은 것으로 해석할 수 있다. 오히려 과학교과가 실험, 보고서 등 수행평가 및 서술형 평가문항의 제작과 채점기준 및 루브릭 작성에 많은 어려움이 있어서 평가과정 전문성이 더욱 낮게 나타남으로써 수학과 과학교과 간의 전문성 차이가 발생하는 것으로 보아야 할 것 같다.

<표 III-12> 평가과정 전문성(EPP)에 대한 과목별 차이 사후검증

과목(I)	N/ 평균	과목(J)	N	평균	평균차(d)	Wilks' λ	F
수학	165/ 4.0178	국어	186	3.8162	.2015	.937	2.230*
		영어	234	3.7465	.2713		
		과학	150	3.6801	.3377*		
		사회	192	3.7731	.2447		

*p<.05

다) 문항별 분석

아래 <표 III-14>는 평가도구의 개발과 제작에 대한 문항별 분석이다. 문항개발과 관련하여 9번 문항, ‘문항개발 과정에서 학생의 수준을 고려하여 성취기준과 성취수준을 재설정할 수 있다.’라는 설문에 대하여 중학교와 고등학교 간의 차이는 없다. 그러나 13번 문항, ‘이원목적분류표의 내용영역 및 성취기준과 행동영역에 따라서 문항을 개발한다.’라는 설문에서는 중학교 교사들이 고등학교 교사들보다 전문성(d=.36)이 높고 유의미한 차이(t=8.033, p<.05)가 있는 것으로 나타난다. ‘문항개발과정에서 성취수준별 분할점수를 산출한다.’라는 16번 문항에서도 중학교와 고등학교는 유의미한 차이를 보인다. 평가도구의 신뢰도와 타당도를 검토하는 17번 문항에서 중학교가 고등학교보다 평균이 높고(d=.42) 유의미한 차이(t=4.941, p<.05)를 보인다.

<표 III-13> ‘평가도구개발’에 대한 문항별 분석

문항번호	학교 급	N	평균	표준편차	평균차(d)	t
9번	중학교	489	3.70	.869	.13	1.193
	고등학교	531	3.57	.771		
13번	중학교	489	4.30	.730	.36	8.033*
	고등학교	531	3.64	.788		
16번	중학교	489	3.64	.910	.34	3.487*
	고등학교	531	3.30	.915		
17번	중학교	489	3.87	.759	.42	4.941*
	고등학교	531	3.45	.790		

*p<.05

아래 <표 III-14>은 ‘채점기준 설정’에 대한 문항별 분석이다. 18번 문항, ‘성취기준과 성취수준에 따라서 채점 기준을 설정한다.’라는 설문에 대하여 중학교와 고등학교 간의 평균차이($d=.51$)는 유의미하다($t=6.047, p<.05$). ‘이원목적분류표의 내용영역 및 행동영역에 따라서 채점기준을 설정한다.’라는 19번 설문도 평균차이가 .57로 유의하다. 21번 ‘평가에 대한 학생들의 응답 반응을 참조하려 채점기준을 수정·보완한다.’라는 설문도 평균차 .39로 유의하다. 이 세 문항에 대하여 중학교 교사들이 고등학교 교사들보다 전문성이 더 높은 것으로 나타난다. 특히 이원목적분류표에 따라서 채점기준을 설정하는 부분에서는 더욱 그러하다.

<표 III-14> ‘채점기준 설정’에 대한 문항별 분석

문항번호	학교 급	N	평균	표준편차	평균차(d)	t
18번	중학교	489	4.25	.686	.51	6.047*
	고등학교	531	3.74	.840		
19번	중학교	489	4.18	.764	.57	6.554*
	고등학교	531	3.61	.853		
21번	중학교	489	4.08	.801	.39	4.480*
	고등학교	531	3.69	.804		

* $p<.05$

2) 정성적 분석

가) 평가도구 개발·제작

(1) 서술형 평가 전문성

“선생님이 있는 학교 또는 교과 교사들의 ‘서술형 평가 전문성’은 어느 정도라고 생각하십니까?” 또는 “선생님 학교 교사들은 서술형 평가에 대해 전문적 지식과 기능을 가지고 능숙하게 활용하십니까?”라는 질문에 대하여 중학교 교사들은 ‘잘 하고 있다’, ‘전문성이 높다’고 인식하고 있는 반면에 고등학교 교사들은 ‘좀 부족하다’, ‘전문성이 낮다’고 인식하고 있다. 굳이 수치로 표현하자면 6점 기준으로 중학교 교사들은 4.67점, 고등학교교사들은 3.25점 정도이다. 이 질문에 대한 교사들의 면접 내용은 다음과 같다.

학교급	중학교	고등학교
항목		
서술형 평가 전문성	“성취기준과 자신의 수업을 반추하여 그를 반영한 서술형 평가를 출제하고 타당한 평가기준을 중심으로 채점하며, 학생들에게 확인시켜 피드백을 제공한다는 면에서 전문성이 높다고 생각하고	“서술형 평가 전문적 지식과 기능이 부족으로 문항의 목적과 제작 간에 정합성이 떨어지며, 문항제작과 채점기준 작성에 이원목적분류표를 활용하지 못하고 ...” (영어)

	<p>... ” (영어)</p> <p>“전공 교과에 대한 전문성이 있으니까 ... ” (수학)</p> <p>“평가윤리, 평가목적과 목표의 설정, 평가유형과 방식의 이해 및 선정에 대하여 잘 알고 있으므로” (국어)</p> <p>“수행평가 영역에 서술·논술형 평가를 반영하기도 하고 서술형 평가에서 학생들의 답안 기술능력을 감안하여 서술형 평가 도구 개발하고 교과협의회를 통해 출제계획 수립 후 출제하니까 어느 정도 전문성이 있다고 생각하는데...” (사회)</p>	<p>“전문성을 확보할 기회가 없었음” (국어):</p> <p>“서술형 평가 업무(출제 및 채점, 시행)에 대한 과도한 시간과 노력” (사회)</p> <p>“학생들이 서술형 평가에 대한 관심과 응답률이 낮다보니 교사들에게도 중요도가 많이 떨어짐. 학생들의 긍정적인 반응을 이끌어 내기 위해서는 많은 노력이 필요함” (과학)</p>
--	---------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------	------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------

(2) 서술형 문항제작 전문성

“선생님이 있는 학교 또는 교과 교사들의 서술형 '평가문항 제작의 전문성'은 어느 정도라고 생각하십니까?” 또는 “선생님 학교 교사들은 서술형 평가의 제작에 대해 전문적 지식과 기능을 가지고 있고 이를 능숙하게 활용하십니까?”라는 질문에 대하여 중학교 교사들은 ‘잘 하고 있다’, ‘전문성이 있다’고 인식하고 있는 반면에 고등학교 교사들은 ‘좀 부족하다’, ‘전문성이 낮다’고 인식하고 있다. 굳이 수치로 표현하자면 6점 기준으로 중학교 교사들은 4.05점, 고등학교교사들은 4.00점 정도이다. 이 질문에 대한 교사들의 면접 내용은 다음과 같다.

학교급	중학교	고등학교
항목		
서술형 문항제작 전문성	<p>“학생들은 수업시간에 배움을 일으킨 내용을 토대로, 다양한 텍스트 속에서 자신이 배움을 적용하고 문제해결을 해나가고, 이것이 지필이 가지지 못하는 서술형 평가의 진정한 힘이고, 이러한 맥락에서 교사들은 수업시간에 다루어진 내용을 토대로 성취기준을 고려하여 문항을 제작하고 있고, 이런 점에서 전문성이 양호하다고 ...” (영어)</p> <p>“평가하고자 하는 기준에 맞는 문항 제작하고 시행하고 있으니까 그렇다고 생각하는 것 같아요...” (수학)</p> <p>“평가가 매우 중요하다고 생각하고, 최근 평가문항 제작에 대한 연수가 많으며 담당자들을 통해 시험을 칠 때마다 주의사항과 전달 연수를 받아 문항을 제작하므로 나름의 노하우를 갖고 있으나 전문적이라고 생각하지는 않는다.” (국어)</p>	<p>“서술형 평가문항 제작을 평가의 목적·목표와 연계하는데 미숙, 유형설정과 제작 미숙” (영어)</p> <p>“우리학교 교사들은 서술형 평가 및 수행평가문항의 제작에 대해 전문적 지식과 기능을 가지고 있고 이를 능숙하게 활용한다.” (수학)</p> <p>“양호하고... 다양한 독서지도의 경험이 있고 오랜 세월 지도의 노하우(20년차 이상)” (국어)</p> <p>“대부분의 교사들이 서술형 평가의 중요성(주 목적이 고등정신능력을 평가하는 것임)을 깊이 이해하고 있기 때문에” (사회)</p> <p>“서술형 평가문항에 대한 학생들의 낮은 관심과 평가에 대한 낮은 응답으로 인해 양질의 평가문항이 개발되더라도 실제 평가를 실시하면 긍정적인 답이 안 나오는 경우가 많음.” (과학)</p>

	“학생의 사고력, 문제해결력, 창의력 등 고등 사고 기능을 평가할 수 있는 서술형·논술형 평가 문항의 중요성을 충분히 인지하고 있어 평가문항 제작의 전문성을 가지고 있다고 할 수 있을 것 같은데 ... ” (사회)	
--	-------------------------------------------------------------------------------------------------------------------------	--

(3) 서술형문항 제작절차

“선생님이 있는 학교 또는 교과 교사들이 서술형 ‘평가문항 제작의 절차’를 숙지하고 그에 따라 평가문항을 제작하는 정도는 어떠합니까?” 또는 “선생님학교 교사들은 ‘제작 절차’를 숙지하고 그에 따라 서술형 평가문항을 제작하십니까?”라는 질문에 대하여 중학교 교사들은 ‘절차를 잘 따르고 있다’, ‘절차를 따라 제작하고 있다.’고 인식하고 있는 반면에 고등학교 교사들은 ‘절차를 따라 제작하지 않는다.’ ‘절차를 따라 제작하는 것이 어렵다.’고 인식하고 있다. 굳이 수치로 표현하자면 6점 기준으로 중학교 교사들은 3.07점, 고등학교교사들은 2.80점 정도이다. 이 질문에 대한 교사들의 면접 내용은 다음과 같다.

학교급 항목	중학교	고등학교
서술형문항 제작절차	<p>“서술형 평가 문항을 제작함에 있어 평가의 내용이나 방법에 집중하지, 절차에 대해서는 커다란 관심을 기울이지 않아요” (영어)</p> <p>“성취 기준을 확인하고 성취 기준에 알맞은 문항 제작하는데 다양한 유형의 문항으로 작성하려 노력하고 ... 그리고 공동 출제로 내용을 확인하고 점검하는 과정을 거치고 있어요...” (수학)</p> <p>“교육목표 확인, 평가문항 제작의 대상 교과서 단원의 내용 확인, 수업시간의 학생활동 참고, 학생활동 중에 서술형 문항으로 적합한 내용 확인한 후 문항을 제작” (국어)</p> <p>“서술형 평가의 중요성을 충분히 인지하고 있으며 교과협의회를 통한 평가계획 수립하고 성취기준 도달 정도를 측정하는 평가도구 제작한 후 공동 검토 및 수정하는 과정을 거치고 있지요.” (사회)</p>	<p>“이원목적분류표를 먼저 작성하는 것의 불편함.” (영어)</p> <p>“절차 자체에 대해 정확히 숙지하지 못하고 일반 객관식 평가도구 출제하듯이 하고 있고 절차를 따를 시간적 여유가 없고 절차교육이 없었으며 노력이 지나치게 크다...” (국어)</p> <p>“계획서작성, 목표설정, 문항 초안 작성, 채점 기준표 및 모범답안 작성, 문항검토 및 수정, 최종문항 완성 이런 일련의 절차를 따르고 ... 고등정신능력을 평가하는 서술형 평가의 중요성을 깊이 이해하고 있기 때문에 ... ” (사회)</p> <p>“절차를 따르려고 노력은 하지만 잘 안 되는 것 같습니다. 특히 교과협의회를 통한 협의보다는 개인적으로 문항을 제작하고 교과협의회는 형식적으로 실시하는 경향이 있습니다. 다른 교사들도 그런 편입니다. 학업성적관리 규정에는 관련 내용이 있지만 엄격히 지키고 있는 것 같지는 않습니다.” (과학)</p>

(4) 서술형에 성취기준 및 수준 적용하기

“선생님이 있는 학교 또는 교과 교사들의 서술형 ‘평가문항제작’에 있어서 ‘성취기준과 성취수준’을 적용하는 정도에 점수를 매긴다면 6점 만점에 몇 점 정도라고 생각하십니까?” 또는 “선생님학교 교사들은 서술형 평가문항을 제작할 때 ‘성취기준과 성취수준’을 우선적으로 적용합니까?”라는 질문에 대하여 중학교 교사들은 ‘성취기준을 잘 적용하고 있다’고 응답한 반면에 고등학교 교사들은 ‘성취기준보다는 교과서와 수능기출문제에 따른다.’고 응답하고 있다. 굳이 수치로 표현하자면 6점 기준으로 중학교 교사들은 4.67점, 고등학교 교사들은 3.20점 정도이다. 이 질문에 대한 교사들의 면접 내용은 다음과 같다.

학교급	중학교	고등학교
항목		
서술형에 성취기준 및 수준 적용하기	<p>“성취기준을 전체적으로 숙지하고 그를 바탕으로 평가문항을 제작하기 보다는 주로는 평가문항을 제작하고 그에 맞는 성취기준을 찾는 것이 통례이다. 평가문항제작은 주로 교육청에서 제시하는 기준이나 예시문항을 참고로 하여 교과서에 맞게 수정 보완하는 형태로” (영어)</p> <p>“평가문항 제작에 있어 성취기준과 수준보다 단원의 내용을 기준으로 제작한다.” (국어)</p> <p>“학년별, 교과별 성취기준 도달 여부를 평가할 수 있는 출제 계획 수립 후, 서술형 평가문항을 제작하고 있으므로 ‘성취기준과 성취수준’을 잘 적용하고 있다고 보여지고, 평가문항을 제작하는 과정에서 교육부 및 교육청 개발 자료를 참고하고 그것을 기준으로...” (사회)</p>	<p>“수업과 평가의 괴리, 문항 선정과 문항 제작의 어려움으로 성취기준을 적용을 소홀히 하고, 문항제작의 기준은 교과와 각 단원에 제시된 학습목표, 제작의 편의성, 채점의 용이성, 답안의 명확성을 중심으로 ... ” (영어)</p> <p>“수능 기출문제 유형을 시험 범위에 적용, 형성평가 요소 중 출제” (수학)</p> <p>“성취기준과 현실 학습 내용이 직접 연결이 안 된다. 제작기준은 텍스트 자체의 중심 내용, 중요 포인트, 기출된 내용, 모의, 수능 출제 가능성 위주로 ... ” (국어)</p> <p>“교사들이 성취기준과 성취수준을 모두 알고 있다고 여기고 ... 교육부 및 교육청 개발 자료를 참고하여 제작” (사회)</p> <p>“성취기준과 성취수준은 학기가 시작하기 전에 작성을 하므로 평가문항 제작할 때는 바로 사용할 수 있어요 ... ” (과학)</p>

(5) 수행평가 전문성

“선생님이 있는 학교 또는 교과 교사들의 ‘수행평가 전문성’은 어느 정도라고 생각하십니까?” 또는 “선생님학교 교사들은 수행평가에 대해 전문적 지식과 기능을 가지고 있고 이를 능숙하게 활용합니까?”라는 질문에 대해 중학교 교사들은 ‘능숙하게 활용한다.’ ‘수행평가 전문성 높다.’고 응답한 반면에 고등학교 교사들은 ‘수행평가 전문성이 양호하는 의견과 그렇지 않다.’로 응답이 나뉜다. 이를 수치로 표현하자면 6점 기준으로 중학교 교사들은 5.00점, 고등학교교사들은 3.80점 정도이다. 이 질문에 대한 교사들의 면접 내용은 다음과 같다.

학교급 항목	중학교	고등학교
수행평가 전문성	<p>“교과의 특성에 맞게 수행평가 영역을 구성하고 실시하며 ... 학생들의 배움을 점검하는 과정으로서의 수행평가를 실시하므로 전문성이 매우 높다고 ...” (영어)</p> <p>“평가윤리, 평가목적과 목표의 설정, 평가유형과 방식의 이해 및 선정에 대하여 잘 알고 있으므로” (국어)</p> <p>“수행평가의 공정성, 객관성이 유지되도록 연초에 수행평가 계획을 수립하고 연구·실험보고서, 프로젝트, 역할극, 협력학습, 포트폴리오 등 과정 중심의 다양한 수행평가 도구 활용함으로 전문성이 높은 것으로 생각...” (사회)</p>	<p>“일반적으로 영어듣기, 태도, 노트검사, 보고서 작성, 독후감 제출 등을 편의적 기준에 따라 사용 ... ” (영어)</p> <p>“오랜 시간 평가한 경험과 평가 방법 다양화 시도하고 있어 수행평가 전문성은 양호 ... ” (국어)</p> <p>“수행평가 출제 및 채점, 시행에 대해 시간과 노력이 과도하게 소요 ... ” (사회)</p> <p>“실습중심의 교과들은 학생들의 수행평가에 대한 관심과 응답률이 높고, 교사들도 학생들의 전공교과에 대한 전문성을 높이기 위한 다양한 방법을 실시하고 있어서 ” (과학)</p>

(6) 수행평가 실행 전문성

“선생님이 있는 학교 또는 교과 교사들의 '수행평가 실행의 전문성'은 어느 정도라고 생각하십니까?” 또는 “선생님학교 교사들은 수행평가를 실행함에 있어서 전문적 지식과 기능을 가지고 있고 이를 능숙하게 활용하니까?”라는 질문에 대하여 중학교 교사들은 ‘능숙하게 활용한다.’ ‘수행평가 실행의 전문성 높다.’고 응답한 반면에 고등학교 교사들은 ‘수행평가 실행의 전문성이 양호하는 의견과 그렇지 않다.’으로 응답이 나뉜다. 굳이 수치로 표현하자면 6점 기준으로 중학교 교사들은 3.84점, 고등학교교사들은 4.00점 정도이다. 이 질문에 대한 교사들의 면접 내용은 다음과 같다.

학교급 항목	중학교	고등학교
2-1-6) 수행평가 실행 전문성	<p>“동과 교사끼리 연초에 수행평가를 계획하고 계획에 맞게 수행평가를 실행해 가고 계획대로만 진행되는 것이 아니라 수행평가를 시행하면서도 중간 중간 교사 간, 교사-학생 간 피드백을 통해 수행평가를 유동적으로 조정한다. 이러한 면에서 수행평가를 실행함에 있어 전문성을 가진다고 말할 수 있다...” (영어)</p> <p>“평가가 매우 중요하다고 생각하고, 최근 평가문항 제작에 대한 연수가 많으며 담당자들을 통해 시험을 칠 때마다 주의사항과 전달 연수를 받아 문항을 제작하므로 나름의 노하우를 갖고 있거나 전문적이라고 생각하지는 않는다.”</p>	<p>“수행평가의 목적·목표에 적합한 평가 방식 또는 평가유형을 설정하는데, 그에 따른 채점기준이나 루브릭 작성에 어려움” (영어)</p> <p>“교직 경력 동안 확보한 경험과 프로젝트 발표 수행의 도입으로 인한 다양성 확보로 수행평가 실행의 전문성은 양호하다고 ... ” (국어)</p> <p>“대부분의 교사들이 수행 평가의 중요성, 특히 지식과 기능을 활용할 수 있는 능력이 어느 정도인지를 평가할 수 있는 방법을 깊이 이해하고 있기 때문에 실행도 ... ” (사회)</p> <p>“전공교과의 수행평가는 정해진 절차와</p>

	(국어) “수행평가 실행은 성취기준에 적합한 학습과제 수행 과정 및 결과를 직접 관찰하고 그 결과를 판단하므로 그 중요성을 충분히 인지하고 있어 전문성이 있다고 ... ” (사회)	방법을 그대로 사용해야 하기 때문에 나름 내실 있게 진행되고 있고 ” (과학)
--	---------------------------------------------------------------------------------------------------------	--------------------------------------------------

(7) 수행형 문항 제작절차

“선생님이 있는 학교 또는 교과 교사들이 수행형 ‘평가문항 제작의 절차’를 숙지하고 그에 따라 평가문항을 제작하는 정도는 어떠하다고 생각하십니까?” 또는 “선생님학교 교사들은 수행평가의 ‘제작 절차’를 숙지하고 그에 따라 수행형 평가문항을 제작하니까?”라는 질문에 대하여 중학교 교사들은 ‘절차에 따라서 수행형 문항을 제작하는 경우’와 ‘알고는 있지만 절차를 전혀 따르지 않는 경우’로 나누어지고 고등학교 교사들은 ‘제작절차에 따라 수행형 문항을 만들지 않는다.’ 라고 응답하고 있다. 굳이 수치로 표현하자면 6점 기준으로 중학교 교사들은 3.15점, 고등학교교사들은 3.00점 정도이다. 이 질문에 대한 교사들의 면접 내용은 다음과 같다.

학교급 항목	중 학교	고 등 학 교
수행형 문항 제작절차	<p>“수행평가 또한 문항을 제작함에 있어 평가의 영역, 내용이나 방법에 집중하지, 절차에 대해서는 커다란 관심을 기울이지 못하는 편이다.” (영어)</p> <p>“교육목표를 확인, 평가문항 제작의 대상인 교과서 단원의 내용 확인, 수업시간의 학생활동 설계, 평가 계획 작성, 문항 제작후 평가의 순으로 한다고 생각하지만 그 절차를 따르지는 않는다.” (국어)</p> <p>“교과협의회를 통한 평가계획을 수립하고 성취기준 도달 정도를 측정하는 평가도구 제작하여 공동 검토 및 수정하고 평가시행 및 채점하는 과정은 서술형 평가와 별 차이가 없다.” (사회)</p>	<p>“특별한 절차를 따르지 않음.” (영어)</p> <p>“절차 자체에 대해 정확히 숙지하지 못하고 있음: 일반 객관식 평가도구 출제 하듯이 하고 있고 ... 서술형 평가 자체에 대해 큰 부담을 느끼고 있어서 ... ” (국어)</p> <p>“절차를 따르려고 노력은 하지만 잘 안 되는 것 같습니다. 특히 교과협의회를 통한 협의보다는 개인적으로 문항을 제작하고 교과협의회는 형식적으로 실시하는 경향이 있습니다. 다른 교사들도 그런 편이고 ... 학업성적관리 규정에는 관련 내용이 있지만 엄격히 지키고 있는 것 같지는 않고 ... ” (과학)</p>

(8) 수행평가에 성취기준 및 수준 적용하기

“선생님이 있는 학교 또는 교과 교사들의 ‘수행평가 실행’에 있어서 ‘성취기준과 성취수준’을 적용하는 정도는 어떠하다고 생각하십니까?” 또는 “선생님학교 교사들은 수행평가를 실행함에 있어서 ‘성취기준과 성취수준’을 우선적으로 적용하십니까?”라는 질문에 대하여 중학교 교사들은 수행평가에 ‘성취기준 및 수준을 잘 적용한다.’라고 응답한 반면에 고등학교 교사들은 수행평가에 ‘성취기준과 성취수준을 잘 적용하지 않는다.’라고 응답하고 있다. 굳이 수치로 표현하자면 6점 기준으로 중학교 교사들은 4.25점, 고등학교교사들은 3.00점 정도이다. 이 질문에 대한 교사들의 면접 내용은 다음과 같다.

학교급 항목	중학교	고등학교
수행 평가에 성취기준 및 수준 적용하기	<p>“성취기준에 맞게 학생들의 배움을 점검하고 피드백하기 위해서는 수행평가가 다양한 방식으로 자주 일어나야 하나 실제적으로는 많은 학생들의 수행평가를 자주 실시하기에는 교사들의 시간이 절대적으로 부족하다. 문항을 제작하는 것 또한 힘들며, 실시함에 있어 수업시간이 줄으로써 평가시간 확보도 어렵고, 채점하는 데도 시간이 많이 들며 결정적으로 수행평가에 대해 개별 교사에게 주어지는 결정권도 많이 확보되지 않는 편이다. ... 성취기준을 토대로 교과서에서 주어지는 4영역에 대한 다양한 활동들을 참고한다. 이를 수정 보완하여 영역에 맞는 수행평가를 실시하려 노력한다.” (영어)</p> <p>“학년별, 교과별 성취기준 도달 여부를 평가할 수 있는 출제 계획 수립 후, 수행평가를 실행하고 있으므로 ‘성취기준과 성취수준’을 잘 적용하고 있고 ... 교육부 및 교육청 개발 자료를 참고하여 평가문항을 제작...”(사회)</p>	<p>“수행평가에서 성취기준·수준의 의미를 찾지 못함, 수업과 수행평가 목적과의 괴리... 수행평가실행의 기준은 태도, 교과서 문장 외우기, 발표하기 등을 내용, 논리, 전달력 등에 따라서 다양하나 일관된 성취기준 없이 실행 ...” (영어)</p> <p>“수행평가 시행 시 주먹구구식으로 시행함 ... 수행평가실행의 기준은 텍스트 자체의 중심 내용, 중요 포인트, 기출 내용, 모의, 수능 출제 가능성 위주로 ...” (국어)</p> <p>“교사들이 성취기준과 성취수준을 모두 알고 있다고 생각하고 ... 수행평가실행의 기준은 수행 평가의 중요성 인식 제고, 수행 평가문항 제작 절차 숙지, 성취기준과 성취수준과의 연계 수행 평가문항 제작 ...” (사회)</p> <p>“성취기준과 성취수준이 적용되지 않는 수행평가는 실시하지 못해요 ...” (과학)</p>

(9) 수행형 및 서술형 문항제작에 이원목적분류표 적용하기

“선생님이 있는 학교 또는 교과 교사들이 서술형 평가문항 및 수행형 평가문항을 ‘이원목적분류표’에 따라 제작·활용하는 것은 어느 정도라고 생각하십니까?” 또는 “선생님학교 교사들은 ‘이원목적분류표’에 따라 서술형 평가문항 및 수행형 평가문항을 제작·활용하십니까?”라는 질문에 대하여 중학교 교사와 고등학교 교사 모두 평가문항을 제작하는데 ‘이원목적분류표를 잘 적용하지 않는다.’라고 응답했다. 이원목적분류표를 활용했다고 응답한 경우에도 제작 후 적용한 것으로 해석된다. 굳이 수치로 표현하자면 6점 기준으로 중학교 교사들은 3.33점, 고등학교교사들은 2.50점 정도이다. 이 질문에 대한 교사들의 면접 내용은 다음과 같다.

학교급 항목	중학교	고등학교
수행형 및 서술형 문항제작에 이원목적분류표 적용하기	<p>“지필평가와 달리 서술형과 수행 평가는 채점 기준을 가지고 용인되는 범위 내에서 다양한 학생들의 답을 인정해주는 평가이므로 이원목적분류표로 점수를 매기기에는 한계가 있는 것 같아요.</p>	<p>“이원목적분류표에 따른 제작 경험부족, 수업과 평가의 괴리 및 분리, 업무과다로 평가에 몰입할 시간부족과 피로도 가중... 평가문항 제작 후 이원목적분류표 틀에 맞추어 넣기를 ...” (영어)</p>

	<p>기본적으로 이원목적분류표에서 모범답안을 제시하나 채점기준도 같이 제공하여 기준에서 용인되는 범위에서 다양한 학생답을 인정해주고 있어요.” (영어)</p> <p>“지식 적용 이해에 따라 성취 수준에 맞추어 활용하고 ... ” (수학)</p> <p>“교사들이 서술형 평가문항 및 수행형 평가문항의 제작에 이원목적(내용 및 행동)을 모두 이해하고 적용한다고 여기고 있으므로 이원목적분류표 작성 후 평가 문항 제작은 현실적으로 어렵고 ... 제작·활용하는 과정에서 ‘이원목적분류표’의 활용도가 낮아요.” (사회)</p>	<p>“출제 후 이원목적분류표를 작성하는 경우가 있다.” (수학)</p> <p>“이원목적분류표의 여러 기준들, 평가 기준, 난이도 등을 최대한 준수하여 작성 ... 되도록 평가기준을 먼저 설정한 후 모범 답안을 구성 ... ” (국어)</p> <p>“교사들이 서술형 평가문항 및 수행형 평가문항의 제작에 이원목적(내용 및 행동)을 모두 이해하고 적용한다고 여기고 ... 거의 활용하지 않고 ... ” (사회)</p> <p>“먼저 이원목적분류표를 제작하고 그 내용에 따라서 서술형 평가문항 및 수행형 평가문항을 제작하게 하고 있으나 실제 운영은 먼저 문항을 제작하고 그 문항에 맞추어 이원목적분류표를 후에 작성하고 있는 대부분의 실정이고 ... 이론적으로는 알고 있으나 실제로 평가 문항을 제작할 때는 교과 내용을 보고 문항을 먼저 출제한다.” (과학)</p>
--	---------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------	------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------

나) 채점기준 및 루브릭 설정

(1) 채점기준 및 루브릭 설정에 이원목적분류표 적용하기

“선생님께서 몸담고 있는 학교 또는 교과 교사들의 서술형 평가 및 수행평가문항의 ‘채점기준 또는 루브릭’을 이원목적분류표에 따라 작성하는 정도는 어떠합니까?” 또는 “선생님학교 교사들은 서술형 평가 및 수행평가문항의 ‘채점기준 또는 루브릭’을 이원목적분류표에 따라 작성합니까?”라는 질문에 대하여 중학교 교사들은 채점기준 설정과 루브릭 작성에 ‘이원목적분류표를 잘 적용한다.’고 응답한 반면 고등학교 교사들은 ‘이원목적분류표를 잘 적용하지 않는다.’라고 응답했다. 이를 수치로 표현하자면 6점 기준으로 중학교 교사들은 5.00점, 고등학교교사들은 3.20점 정도이다. 이 질문에 대한 교사들의 면접 내용은 다음과 같다.

학교급	중학교	고등학교
항목		
채점기준 및 루브릭 설정에 이원목적분류표 적용하기	<p>“지필평가와 달리 서술형과 수행평가는 채점기준을 가지고 용인되는 범위 내에서 다양한 학생들의 답을 인정해주는 평가이므로 채점기준을 세우고 평가하는 것이 적절하다고 생각하지만 최대한 실질적인 내용에 맞게 서술한 학생들의 답에 점수를 주기 위한 객관적 기준을 마련하는 것이 현실적으로 어려운 것 같아요.” (영어)</p>	<p>“경험에 기초하여 문항의 난이도를 기준으로 설정 ... 이원목적분류표에 따른 실제 경험부족, 수업과 평가의 괴리 및 분리, 업무과다로 평가에 몰입할 시간부족과 피로도 가중 ... ” (영어)</p> <p>“문제가 요구하는 답안과 허용되는 동일 의미의 유사답안을 만점 기준으로 설정하고 기준에 일부 미달할 때와 출제 시 문제에 조건 제시한 것에 따라</p>

	“성취기준과 성취수준을 중심으로 그 달성 정도에 따라 채점기준표(루브릭) 작성한다고 여기는데 ...” (사회)	감점 ... ” (국어) “성취기준과 성취수준을 중심으로 그 달성 정도에 따라 채점 기준표(루브릭) 작성 ...” (사회) “채점기준은 나름 철저하게 준비한다.” (과학)
--	---------------------------------------------------------------	-------------------------------------------------------------------------------------------------------

(2) 채점기준 및 루브릭의 수정·보완

“선생님께서 몸담고 있는 학교 또는 교과 교사들이 서술형 평가 및 수행평가의 ‘채점기준 또는 루브릭’을 학생들의 반응에 따라 수정·보완하는 정도는 어떠하다고 생각하십니까?”라는 질문에 대하여 중학교 교사들과 고등학교 교사들 공히 채점기준과 루브릭을 ‘학생들의 반응에 따라 잘 수정·보완한다.’고 응답하였으나 고등학교 교사들이 수정·보완하는데 더 보수적이었다. 이를 수치로 표현하자면 6점 기준으로 중학교 교사들은 5.00점, 고등학교교사들은 4.02점 정도이다. 이 질문에 대한 교사들의 면접 내용은 다음과 같다.

학교급 항목	중학교	고등학교
채점기준 및 루브릭의 수정·보완	<p>“교사가 의도했던 답과는 다르지만 학생들이 문제의 요구에 알맞게 다양한 답을 제시하였다면 미처 세우지 못한 채점기준을 보완하고 ...” (영어)</p> <p>“사전에 서술형 평가 및 수행평가 채점 기준표를 작성하나 문항에 대한 예상치 못한 학생들의 다양한 반응에 대응하기 위해서 이미 제시한 채점기준표에 근거하여 다소 수정·보완하고 ... ” (사회)</p>	<p>“유사답안의 발생, 채점의 정확성, 평가 신뢰도 향상을 위해 수정·보완 ... ” (영어)</p> <p>“국어과의 특성상 다양한 어휘 개념이 유사하게 정답처리 될 수 있고 ... 채점 기준안 공개 이후 이의 신청 받고 학생들의 반응을 모아 기준안 수정 후에 채점, 그리고 채점 내용 공개 ... ” (국어)</p> <p>“서술형 평가 및 수행평가에 대한 학생들의 반응은 다양하고 예상하지 못한 응답을 하기 때문에 그에 따라 채점기준(루브릭)을 수정·보완 ...” (사회)</p> <p>“바람직하지 않는 방법이라고 생각함. 가급적 계획대로 진행하는 것이 원칙임. 한 번 원칙을 지키지 않으면 추후 많은 문제가 ... ” (과학)</p>

(3) 이원목적분류표에 따라서 분할점수 산출하기

“선생님이 있는 학교 또는 교과 교사들이 ‘성취기준과 성취수준, 그리고 이원목적분류표’를 기준으로 서술형 평가와 수행평가의 분할점수를 산출할 수 있는 정도는 어떠합니까?” 또는 “우리학교 교사들이 ‘성취기준과 성취수준, 이원목적분류표’를 기준으로 서술형 평가와 수행평가의 분할점수를 산출할 수 있다.”는 질문에 대하여 중학교 교사들과 고등학교 교사들 공히 이원목적분류표에 따라서 ‘분할점수를 산출하는데 어려움이 있다.’라고 응답하였다. 이는 중학교는 고정분할점수를 사용하기 때문에 당연한 결과이나 고등학교 교사들이 중학교 교사들 보다 분할점수 산출을 더

못한다는 것은 의외이다. 이를 수치로 표현하자면 6점 기준으로 중학교 교사들은 2.67점, 고등학교 교사들은 2.60점 정도이다. 이 질문에 대한 교사들의 면접 내용은 다음과 같다.

학교급 항목	중학교	고등학교
이원목적분류표에 따라서 분할점수 산출하기	<p>“문제가 요구하는 바에 있어서 수행 비율에 따라 분할 점수를 설정한다. 실질적으로 분할 점수를 부여하지 않을 경우 실질적인 내용은 통하나 사소한 형식(예, 스펠링이나 대문자) 때문에 점수를 아예 받지 못하는 경우도 생길 수 있기 때문이다.” (영어)</p> <p>“고정 분할 점수 사용하기 때문에 분할 점수를 산정할 필요가 없어요.” (수학)</p> <p>“제시한 문항별 배점과 채점기준표에 근거하여 분할점수를 산출 ... ” (사회)</p>	<p>“분할 점수는 편의적으로 산출 ... 고정분할점수를 사용하는 경우 평가문항을 쉽게 출제하기도 ... ” (영어)</p> <p>“고정식 분할점수와 단위학교 산출식 분할점수를 시행하고 있다.” (수학)</p> <p>“산출 프로그램을 이용하나 분할 점수는 거의 감으로 설정하는 수준 ... 각 문항의 난이도는 개별교사의 직관적 판단으로 설정하고 ... ” (국어)</p> <p>“총점을 기준으로 분할점수를 산출하기 때문 ... ” (사회):</p> <p>“학교 자체 분할 점수 산출은 실시하고 있지 않고 고정 분할 점수 활용 ... ” (과학)</p>

3) 종합결과

가) 평가도구개발 영역에 대한 결론

평가도구개발의 핵심인 이원목적분류표의 내용영역 및 행동영역에 따라서 평가문항을 개발하는 것에 대하여 정량적 분석에서 중학교 교사들의 평균이 4.30으로 고등학교 교사의 평균인 3.64보다 .36만큼 높고 유의미한 차이($t=8.033$, $p<.05$)가 있는 것으로 나타난다. 수행형 또는 서술형 문항제작 전문성에 관련된 정성적 분석에서 나타난 ‘배운 내용을 토대로 성취기준을 고려하여 ...’, ‘평가하고자 하는 기준에 맞는 문항을 제작하고 ...’, ‘오랜 세월 지도(평가)의 노하우’ 등의 표현에 알 수 있듯이 중등학교 교사들은 평가문항제작에서 나름대로의 전문성을 가지고 있다고 자부한다. 그러나 개별 교사들의 응답에서 알 수 있듯이 성취기준, 평가기준, 수업기준, 경험적 기준 등 통일되지 않은 다양하고 제각각의 기준에 의해 평가 문항 개발·제작되고 있다.

그리고 ‘서술형 및 수행형 평가문항의 제작에 이원목적분류표를 거의 활용하지 않고 ...’ ‘실제로는 먼저 문항을 제작하고 그 문항에 맞추어 이원목적분류표를 후에 작성하고 ...’ ‘이원목적분류표를 먼저 작성하는 것이 불편하고...’ ‘평가문항을 제작함에 있어서 절차에는 관심을 기울이지 않고 ...’ 등의 응답은 중등교사들이 이원목적분류표를 평가문항개발의 핵심도구로 사용하지 않고 절차를 무시하고 있으며 학생의 인지발달을 이끄는 행동영역에 대한 관심이 부족하다는 간접적인 증거이고 평가의 목적과 목표, 평가유형과 문항 간의 정합성과 통일된 일관성의 부재를 드러내는 증거이기도 하다. 따라서 평가문항 제작의 절차와 이원목적분류표의 원리에 의하여 평가도구를 개발·제작하는 과정은 교사들의 평가전문성 제고를 위한 첫 번째 핵심영역일 것이다.

평가도구개발의 또 다른 기준인 성취기준 및 성취수준을 문항개발 과정에서 학생의 수준을 고려하여 재설정할 수 있는지에 대하여 중학교 교사의 평균은 3.70이고 고등학교 교사의 평균은 3.57로 이들 간의 차이는 없다. 5.00을 최고점으로 하는 척도에서 이 수치들은 상당히 높다고 할 수 있다. 즉 중등학교 교사들은 평가도구개발에서 성취기준과 성취수준을 상당한 수준으로 활용하는 있다고 할 수 있다. 그러나 일반적으로 교사집단의 자기평가 또는 동료평가의 평균이 4.00을 상회한다는 점을 고려하면 또 다른 해석이 가능하다. 이를 수행형 또는 서술형 문항제작 전문성에 관련된 정성적 분석에서 살펴보면 ‘주로 평가문항을 제작하고 그에 맞는 성취기준을 찾는 것이 통례이다.’ ‘성취기준과 현실 학습내용이 직접 연결이 안 된다.’ ‘성취기준에 맞게 해야 하나 실제로 실시하기에는 시간이 절대적으로 부족하고 ...’ 등의 응답에서 알 수 있듯이 중등교사들이 상당한 정도로 성취기준 및 성취수준을 서술형 문항개발이나 수행형 문항 개발·실행에 활용하지 않고 있다는 것이다. 따라서 성취기준과 성취수준을 활용한 수행형 및 서술형 문항개발제작은 교사들이 평가전문성을 신장하기 위한 두 번째 핵심영역일 것이다.

성취수준별 분할점수를 산출하는 것은 교사들의 평가전문성을 가늠해 볼 수 있는 중요한 척도이다. 분할점수 산출을 위해서는 성취기준에 따른 성취수준을 설정할 수 있어야 하고 개별 평가문항의 난이도와 학생들의 능력이라는 요소를 비교 판단할 수 있어야 하기 때문이다. a 문항개발과정에서 성취수준별 분할점수를 산출하는 것과 관련하여 중학교 교사의 평균은 3.64이고 고등학교 교사의 평균은 3.30로 이 문항은 다른 문항에 비하여는 가장 낮은 값을 보이는 문항 중 하나이다. 즉 중등학교 교사들이 가장 관심이 없거나 가장 이해하지 못하는 전문성 영역이다. 분할점수 산출과 관련된 정성적 분석을 살펴보면, ‘수행비율에 따라 분할점수를 설정한다.’ ‘분할점수는 거의 감으로 설정...’ ‘총점을 기준으로 분할점수를 산출...’ ‘고정분할점수를 사용...’ 등의 응답은 학교현장에서 성취수준별 분할점수를 산출하는 과정이 형식으로 무의미하게 시행되고 있다는 것을 보여준다. 따라서 성취수준별 분할점수의 산출원리에 대한 심층적 이해와 실습은 교사들의 평가전문성 제고를 위한 세 번째 필수 요소일 것이다.

나) 채점기준 설정 영역에 대한 결론

정량적 분석에서 성취기준과 성취수준 그리고 이원목적분류표의 내용영역과 행동영역에 따라서 채점기준을 설정하는 것에 대하여 중학교의 평균은 각각 4.25와 4.18이고 고등학교의 평균은 3.75와 3.61이다. 중학교의 채점기준설정은 4.00을 상회하여 우수한 것을 인식되며 고등학교는 중학교에 비하여 저조하다. 성취기준 또는 이원목적분류표에 따른 채점기준설정과 관련한 정성적 분석에서도 중학교 교사들은 ‘성취기준에 따라서 평가기준(채점기준)을 중심으로 채점...’ ‘성취기준과 성취수준을 잘 적용하고 있고.. 중심으로 채점기준표(루브릭)을 작성하고...’ ‘수행평가와 서술형 평가에 대하여 객관적인 채점기준을 마련하기가 어려울 것 같다.’라고 응답하여 긍정적 진솔이 강하고, 고등학교 교사들은 ‘이원목적분류표에 따른 채점기준 설정에 대한 경험부족’ ‘문제가 요구하는 답안을 만점으로 하고 기준미달시 감점...’ ‘채점기준은 나름대로 ...’라고 부정적인 응답이 강하다. 성취기준에 따른 채점기준을 설정하는데 있어서 중학교와 고등학교의 이러한 차이는 성적 부담이 적고 자유학기제를 시행하는 중학교와 성적에 대한 부담감과 대학입시 중심의 평가가 일

반적인 고등학교의 환경에 기인한 바가 많을 것이다.

학생의 반응에 따라 채점기준을 수정·보완하는 경우를 살펴보면 정량적 분석에서 중학교는 평균이 4.08이고 고등학교는 평균이 3.69으로 차이가 난다. 정성적 분석에서도 ‘의도했던 답과 다르지만 학생들이 문제의 요구에 알맞게 답했다면 채점기준을 보완하고..’ ‘사전에 작성한 채점기준표를 예상치 못했던 학생들의 반응에 대응하여 수정보완하고 ...’와 같은 중학교 교사들의 응답과는 달리 ‘다양한 유사정답... 학생의 반응에 따라 채점기준안 수정’ ‘한 번 원칙을 지키지 않으면 추후 많은 문제가 ...’와 같은 고등학교 교사들의 응답은 고등학교 교사들이 미세하지만 다소 완고함을 보여준다. 아니면 아래와 같이 중학교의 학생들이 더 어리기 때문에 중학교 교사들이 성적의 수정보완에 대해 더 수용적일 수도 있을 것이다. ‘사실 중학교의 많은 선생님들이, 예를 들면 나는 채점기준을 이렇게 만들었는데, 아이(학생이 아닌)가 다른 방법으로 풀었어요, 틀리진 않았어요. 그런데 문제에서 인수분해를 사용해서 풀라고 했지만 아이(학생이 아닌)가 그렇게 풀지 않고 새로운 것을 뭔가 해서 풀었다. ... 채점기준표를 사실은 수정해요. 이게 나가기 전에 ...’ 따라서 성적부담이 큰 고등학교의 경우에는 중학교보다 엄격한 채점기준이 적용되고 있다. 중학교와 고등학교의 이러한 차이에도 불구하고 일관되고 정합성이 있는 채점기준설정은 앞선 평가도구개발에서와 같이 성취기준과 이원목적분류표에 대한 정확한 이해와 이를 채점기준설정에 활용하는 능력을 제고하는 것일 필수적이다.

4) 우수사례 적용 및 분석

가) 기존 우수사례와 그 요소

아래 서술형 평가문항은 2011년 충청남도 교육청의 영어과 서술형 평가 문항자료집(고등학교용)에 수록된 것으로 ‘평가문항 유형중 대표적인 문항을 선정하여 잘못된 서술형 평가문항과 바람직한 서술형 평가문항을 함께 제시하여 출제 과정에서 범할 수 있는 가능한 오류를 쉽게 파악할 수 있도록 구성한’ 것이다. 이 문항은 크게 서술형 평가문항카드, 평가문항, 그리고 채점기준표 및 채점유의사항, 세 부분으로 이루어져있다. 첫째, 문항카드에 성취기준, 평가요소, 난이도 그리고 배점을 제시하고 있다. 둘째, 평가문항에서는 문제, 조건을 제시하고 있다. 셋째, 채점기준표에는 예시답안, 채점영역 및 채점기준 및 배점 그리고 채점유의사항을 제시하고 있다.

이 최종문항이 우수문항이라고 제시하는 이유는 첫째, 아래 [그림 III-4] 서술형 평가 문항카드를 보면 같은 자료집 개발문항에서의 성취기준, ‘주어진 그림을 보고 자기생각을 글로 쓸 수 있다.’를 최종문항의 성취기준, ‘말하고자 하는 바를 상황에 알맞게 여러 가지 방법으로 표현한다.’로 수정하였다. 평가요소도 ‘자신의 생각쓰기’에서 ‘자신의 의견 말하기’로, 평가영역은 ‘쓰기(말하기)’에서 ‘말하기’로 난이도는 ‘보통’에서 ‘어려움’으로, 배점은 ‘6점’에서 ‘9점’으로 수정하였다.

서술형 평가 문항 카드

과 목	영 어	학 년	고등학교 1학년
성취기준	말하고자 하는 바를 상황에 알맞게 여러 가지 방법으로 표현한다.		
평가요소	자신의 의견 말하기		
평가영역	난이도	배 점	출 처
말하기	어려움	9점	천재교육(이재영 외)

[그림 III-4] 서술형 평가 문항카드

둘째, 평가문항 측면에서 보면 문항질문인 문두가 ‘다음 그림을 보고 어느 그림이 fairness를 나타내는지 자신의 의견을 쓰시오.’에서 아래 [그림 III-5] 평가문항에서처럼 ‘다음은 학생들이 책상을 나르고 있는 두 가지 유형의 그림이다. 어느 그림이 fairness를 나타내는지 자신의 의견을 쓰시오.’로 수정하였다. 그리고 조건이 ‘주어진 그림을 바탕으로 문장을 작성해야 함.’에서 ‘1)주어진 그림을 바탕으로 문장을 작성할 것 2) 의견에 대한 이유를 반드시 제시할 것’으로 수정되었다. 수정한 이유는 다음 세 가지 이다. 1)자신의 의견에 대하여 그 이유를 쓰도록 함으로써 답안 작성 방법을 구체적으로 제시하였다. 2)각 그림에 번호를 부여하여 그림 선택을 명확하게 하도록 하였다. 3)문항에 그림에 대한 상황을 설명하여 fairness에 대한 판단을 할 수 있도록 하였다.

1 평가문항

- 다음은 학생들이 책상을 나르고 있는 두 가지 유형의 그림이다. 어느 그림이 fairness를 나타내는지 자신의 의견을 쓰시오.

조 건

- 주어진 그림을 바탕으로 문장을 작성할 것
- 의견에 대한 이유를 반드시 제시할 것

1)

2)



정답

I think _____.

[그림 Ⅲ-5] 평가문항

셋째, 채점 기준표 및 채점시 유의사항(에서는 예시답안을 ‘I think the first picture represents fairness more than the second one. or I think every man has to do the same job.’에서 ‘(I think) the first picture represents fairness more than the second one because every man has to do the same job.’으로 수정했다. 평가영역과 평가요소라는 항목 이름을 채점 영역과 채점기준으로 수정했으며 과제완성의 채점기준, ‘그림의 내용을 참고하여 fairness에 대한 자신의 생각을 완전한 문장으로 서술 함’을 ‘그림의 내용을 참고하여 fairness에 대한 자신의 생각을 완전한 문장으로 작성하였다.’로 수정했다. 내용의 채점기준, ‘관련 내용이 완벽하고 구체적이며 잘 전개됨’을 ‘관련 내용이 완벽하고 구체적이며 잘 전개되었다.’로 수정하였다. 언어사용의 채점기준, ‘영어만을 사용하고 어휘 사용이 적절함. 문법, 철자, 구두점의 오류가 없다.’를 ‘영어만을 사용하고 어휘 사용이 적절하다. 문법, 철자, 구두점의 오류가 없다.’로 바꾸었다. 채점영역에 대한 채점기준이 최종문항에서 거의 달라진 것이 없다. 예시답안과 배점을 수정한 이유로는 각각 다음과 같다. 1) 예시답안에 생각과 이유를 모두 제시한 완벽한 문장으로 수정하여 제시하였다. 1) 타당한 이유를 포함하여 자신의 의견을 제시하도록 함으로서 난이도를 어려움으로 조정하여 9점 배점으로 하였다([그림 Ⅲ-6] 채점 기준표 및 유의사항).

2

채점 기준표

예시 답안

(I think) the first picture represents fairness more than the second one because every man has to do the same job.

채점 영역	채점 기준
과제완성	· 그림의 내용을 참고하여 fairness에 대한 자신의 생각을 완전한 문장으로 작성하였다.
내 용	· 관련 내용이 완벽하고 구체적이며 잘 전개되었다.
언어사용	· 영어만을 사용하고 어휘 사용이 적절하다. · 문법, 철자, 구두점의 오류가 없다.

영역 배점	과제완성	내용	언어사용
9	fairness에 대한 자신의 생각, 이유를 완전한 문장으로 서술함	관련 내용이 완벽하고 구체적이며 잘 전개됨	문법, 철자, 구두점의 오류가 전혀 없음
6	fairness에 대한 자신의 생각, 이유 중 일부가 없거나, 불완전한 문장으로 서술함	관련성이 있으나 완벽하지 못하며 잘 전개되지 못함	문법, 철자, 구두점의 오류가 일부 발생하나 의미 전달에 문제가 없음
3	fairness와 어울리지 않는 생각을 썼으며 완전한 문장으로 서술하지 못함	관련성이 없고 잘 전개되지 않았으며, 완성되지 않음	문법, 철자, 구두점의 오류가 발생하여 의미 전달에 다소 문제가 있음
0	과제를 수행하지 못하거나 부적절한 내용(낙서, 욕 등), 무응답, 전혀 상관없는 다른 내용을 씀	관련성이 전혀 없고 잘 전달되지 않음	언어적 오류가 아주 심하여 의미전달에 부정적 영향을 줌

3

채점시 유의사항

- 그림의 내용을 참고하여 관련 내용을 자신의 생각으로 쓸 수 있는 능력을 평가하는 문항이므로 대소문자 구분은 채점 기준에 포함하지 않는다.
- 문법, 철자 오류 1개 당 0.2점씩 감점한다.

[그림 Ⅲ-6] 채점 기준표 및 유의사항

이상에서 우수문항을 문항카드, 평가문항, 그리고 채점기준 및 유의사항 측면에서 고찰해 보았다. 자세하고 성실한 문항제작과 채점기준의 설정은 서술형 및 수행평가의 모범이 되기에 부족함

이 없을 것이다. 다만 의문이 드는 것은 여기에서 제시하고 있는 우수하고 모범적인 평가문항 제작과 채점기준안은 물론이고 현재까지 다양한 서술형 평가 및 수행평가 문항제작과 채점기준 설정 방안들에도 불구하고 학교 현장과 교사들의 평가문항 및 채점기준의 어려움은 해소되지 않고 그리고 교사들은 평가문항제작 및 채점기준 설정의 전문성, 즉 평가과정 전문성(EPP)은 향상되고 있지 못하는가하는 점이다.

5) 평가 전문성 제고 방안

가) Bloom의 이원목적분류표 및 성취기준에 따른 문항 제작

Bloom의 평가유형(<표 II-22> 참조)은 내용영역인 지식과 행동영역인 인지과정으로 구성되어 있다. Bloom의 평가유형의 지식차원과 인지단계를 평가요소로 하는 평가문항을 제작한다. 지식차원에는 다양한 교과와 다양한 내용들이 사용될 수 있다. 다양한 교과와 다양한 내용의 구체적이고 명세화된 지식이 성취기준의 한 축을 구성하고 발달단계별 인지과정이 성취기준의 다른 한 축을 구성한다. 따라서 평가문항의 제작은 성취기준의 확인에서 시작하고 이것은 곧 평가유형으로 이어진다. 이것은 무엇을 평가할 것인가라는 평가목적이나 목표와 직접적으로 연결되어 교사들이 평가 목적 및 목표를 설정하느라 시간과 노력을 낭비할 필요가 없다. 평가문항은 평가목적 및 목표와 자연스럽게 정렬되고 정합성을 갖게 될 것이다. 또 개별 성취기준이나 평가문항이 이원목적분류표의 어떤 단계에 속하는지를 알 수 있을 때 여러 평가문항을 포함하는 전체 평가도구가 평가의 목적과도 정합성을 갖게 될 것이다. 이원목적분류표와 성취기준을 전체 평가도구제작에 다음과 같이 활용할 수 있다.

“성취기준이라는 것 자체가 제시가 됐잖아요. 옛날엔 제가 이 문제가 참 괜찮네. 라고해서 문제를 내다보면 알고 보면 이 같은 영역을 묻는 문제가 많아지거든요. 지금 같은 경우는 저희가 이제 성취기준을 가지고 하기 때문에 처음부터 저희가 이원목적분류표에 성취기준을 먼저 뿌려놔요. 이거 한 두개 묻고 이거 한 세 개 묻고 뿌려놓으면 이거에 해당되는 문제를 나중에 선택하게 되니까 골고루 이제 가는 거죠. 그러면 뭐 이 영역에서 꽤 괜찮고 재밌는 문제들이 많지만 버려야 되는 거죠. 그게 오히려 그 쪽에는 더 맞는 것 같아요. 그니까 옛날에는 수학문제가 굉장히 어려운 문제가 많이 나오거나 한 영역에 치우쳐진 게 많이 나올 수 있다면 지금은 그게 이제 좀 없어진 거는 맞는 거 같아요.” (교사 Y)

성취기준에서 구체적인 평가요소를 추출하는 것이 가능해야 한다. 앞의 최종문항에서 성취기준은 ‘말하고자 하는 바를 상황에 알맞게 여러 가지 방법으로 표현한다.’이고 여기서는 성취기준에서 측정가능한 평가요소를 추출할 수가 없다. 평가요소인 ‘자신의 의견 말하기’도 무엇에 대해서, 또 어떤 내용이나 주제에 대해 자신의 의견을 말한다는 것인지 알 수 없다. 즉 ‘자신의 의견말하기’는 제2유형인지 제5유형인지 제8유형인지 알 수가 없다, 즉 세 유형 모두와 관련이 된다(<표 5-2>). 이는 인지적 평가요소만이 있고 지식/내용적 평가요소가 없기 때문에 평가문항 제작이 거의 불가능하게 된다. 따라서 성취기준은 지식/내용영역과 인지/행동영역이 모두 관찰가능하고 측정·가능한 수준으로 기술되어야 한다. 이에 반하여 동일 예시에서 평가문항의 문두, ‘다음은 학

생들이 책상을 나르고 있는 두 가지 유형의 그림이다. 어느 그림이 fairness를 나타내는지 자신의 의견을 쓰시오.’는 내용인 ‘fairness’와 ‘의견쓰기’라는 행동영역 모두 관찰가능하고 구체적으로 측정이 가능하다. 즉 성취기준은 평가적 측면에서의 학습목표이다. 학습목표가 구체적이지 않다면 그 수업이 구체성을 가질 수 없듯이 성취기준이 구체적으로 관찰가능하고 측정가능하지 않다면 평가는 불가능할 것이다. 결론적으로 앞의 예시는 성취기준, 평가요소, 그리고 평가문항 간의 정렬이 어긋나 있고 정합성이 낮다.

<표 III-15> Bloom의 이원목적분류표에 따른 평가문항 제작

지식(내용) 차원	인지발달 단계		
	이해(·기억)하기	적용하기	분석(·평가·창안)하기
사실 지식	제1 유형	제2 유형 자신의 의견말하기	제3 유형
개념 지식	제4 유형	제5 유형 자신의 의견말하기	제6 유형
절차 지식(·메타지식)	제7 유형	제8 유형 자신의 의견말하기	제9 유형

나) Bloom의 이원목적분류표 및 성취기준에 따른 채점의 영역과 기준 설정

평가문항에 대한 채점의 영역은 평가영역에서 나온다. 이원목적분류표의 지식과 인지단계가 채점의 영역이다. 이는 성취기준의 영역이고, 채점기준은 성취기준에 따른 성취수준 3단계를 적용한다. 예를 들면 성취기준, ‘두 그림 중 어느 그림이 fairness를 나타내는지에 대하여 자신의 의견을 쓰시오.’를 평가문항으로 설정하고 이 기준에 따른 성취수준 상중하를 상정할 수 있다. 이것은 다시 이원목적분류표의 난이도와 연계가 된다. 물론 ‘Bloom의 평가유형’에 따른 성취기준 자체에 난이도와 성취수준, 즉 채점기준이 중층적으로 적용되어야 할 것이다(<표 III-16>).

<표 III-16> Bloom의 이원목적분류표에 따른 채점기준 설정

지식(내용) 차원	인지발달 단계		
	이해(·기억)하기	적용하기	분석(·평가·창안)하기
사실 지식	제1 유형	제2 유형	제3 유형
개념 지식	제4 유형	제5 유형 ‘두 그림 중 어느 그림이 fairness를 수준 상	제6 유형

		나타내는지 에 대하여 자신의 의견을 쓰시오.’	수준 하	
절차 지식(·메타지식)	제7 유형	제8 유형	제9 유형	

앞의 ‘[그림 III-6] 채점 기준표 및 유의사항’ 예시에서 채점 영역은 과제완성, 내용, 언어사용으로 구성되어있다. 형식적으로 볼 때에도 과제완성은 내용과 언어사용을 모두 포괄하는 개념이다. 또 내용의 채점기준을 보면 ‘관련 내용이 완벽하고 구체적이며 잘 전개되었다.’라고 기술하고 있다. 이는 이원목적분류표에서의 내용영역은 아니고 오히려 채점기준의 수준영역에서의 정의에 가깝다. 과제완성의 채점기준인 ‘그림의 내용을 참고하여 fairness에 대한 자신의 생각을 완전한 문장으로 작성하였다.’는 평가문항의 문두와 같고 이는 다시 말하면 성취기준이고 이는 곧 상수준의 성취수준이다. 언어사용의 채점기준, ‘1) 영어만을 사용하고 어휘 사용이 적절하다. 2) 문법, 철자, 구두점의 오류가 없다.’는 원 평가문항의 성취기준과 별도의 성취기준이나 평가문항이 되어야 할 요소들로 여겨진다. 따라서 채점기준의 영역과 기준이 평가문항과 성취기준과 일관성 있게 정렬되어 있지 않고 정합성이 떨어진다. 채점의 효율성이 떨어지는 것은 말할 것도 없을 것이다.

라. 평가결과 분석

1) 정량적 분석

가) 학교별 분석

평가결과에 대한 두 가지 평가전문성 핵심요소는 평가결과의 분석과 그에 따른 해석인데, 아래 표 <표 4-1>에서 보듯이 평균이 3.5를 넘어 교사들이 평가결과를 학생의 학업성취와 학업성장과 관련하여 분석하고 해석하는 전문적인 능력, 즉 ‘평가결과 전문성(ERP)’이 우수하다고 해석될 여지가 있다. 그러나 이 수치는 일반적으로 교사집단이 자기평가와 동료평가에서 보여주는 평균적 수치($M > 4.0$)에 비하여 낮게 나타난다는 점에 유의할 필요가 있다. 이는 전반적으로 중등교사들이 평가결과에 대한 전문성이 떨어진다는 것을 인식하고 있다는 것을 의미한다고 해석할 수 있다. 중학교 교사와 고등학교 교사의 평가결과 전문성(ERP)을 비교해보면 <표 4-1>에서와 같이 중학교 교사가 고등학교 교사보다 평균이 더 높고 그 차이($d = .4037$, $p < .05$)가 유의미하다. 그리고 전체평가 전문성(TEP)과 평가결과 전문성(ERP)을 비교해보면 전체에 비하여 평가결과 전문성이 더 낮다. 중학교와 고등학교 두 집단 모두에 같은 현상이 나타난다. 이는 중등학교 교사들이 평가전문성의 다른 영역보다 평가결과 전문성 영역, 즉 평가결과의 통계적 분석이나 통계결과를 성취기준 및 수준 또는 이원목적분류표에 따라서 해석하는 것을 소홀히 하거나 그 능력에 대한 향상을 필요로 한다는 것을 의미한다.

<표 III-17> 평가결과 전문성(ERP)에 대한 학교 급별 차이

		N	평균	표준편차	F
전체	중학교	489	3.8915	.47496	70.821*
	고등학교	531	3.4283	.53497	
	합계	1,020	3.6504	.55687	
평가 결과	중학교	489	3.7221	.65452	33.637*
	고등학교	531	3.3164	.63489	
	합계	1,020	3.5109	.67468	

나) 담당 과목별 분석

아래 <표 III-18>은 평가결과에 대한 분석과 평가결과를 성취기준 및 수준 그리고 이원목적분류표에 따라서 해석하는 평가결과 전문성(ERP)에 대하여 중등학교 교사들의 과목별로 유의미한 차이(Wilks' $\lambda = .928$, $p < .05$)가 있다는 것을 보여준다. 수학과 다른 교과(국어, 영어, 사회, 과학) 간의 차이를 보여주고, 특히 수학과 영어 간의 교과간의 차이($d = .483$, $p < .05$)가 가장 크다는 것을 알 수 있다. 그러나 아래 <표 III-19>에서 알 수 있듯이 국어와 영어, 국어와 과학, 국어, 국어와 사회 등의 다른 교과 간에는 유의한 차이가 없다. 그리고 평가결과 전문성(ERP)이 3.5109로 앞서 분석한 평가과정 전문성(EPP)의 4.0178보다 현저히 낮다. 이는 교사들이 평가결과를 평가의 목

적 및 목표와 연계하는 능력을 제고해야 할 필요가 있다는 것을 명백히 한다.

<표 III-18> 평가결과 전문성(ERP)에 대한 과목별 차이 사후검증

과목(I)	N/ 평균	과목(J)	N	평균	평균차(I-J)	Wilks' λ	F
수학	165/ 3.8308	국어	186	3.4481	.383*	.928*	2.604
		영어	234	3.3479	.483*		
		과학	150	3.3795	.451*		
		사회	192	3.4767	.354*		

*p<.05

<표 III-19> 평가결과 전문성(ERP)에 대한 과목별 차이 사후검증

과목(I)	N/ 평균	과목(J)	N	평균	평균차(I-J)	Wilks' λ	F
국어	186/ 3.4481	영어	234	3.3479	.1003	.928*	2.604
		수학	165	3.8308	-.3826*		
		과학	150	3.3795	.0686		
		사회	192	3.4767	-.0285		

*p<.05

다) 문항별 분석

아래 <표 III-20>은 평가결과의 분석 및 해석에 대한 문항별 분석이다. 평가결과 분석과 관련하여 22번 문항, ‘평가결과에 대하여 양적·질적으로 통계분석을 한다.’라는 설문에 대하여 중학교와 고등학교 간의 차이($d=.35$)가 유의미($t=4.074$, $p<.05$)하다. 결과해석과 관련된 23번 문항, ‘평가의 목적에 따라서 평가결과를 해석한다.’라는 설문에서는 중학교 교사들이 고등학교 교사들보다 전문성($d=.44$)이 높고 유의미한 차이($t=4.990$, $p<.05$)가 있는 것으로 나타난다. ‘평가결과에 따라 학생의 강점과 약점을 해석한다.’라는 24번 문항에서도 중학교와 고등학교는 유의미한 차이($d=.37$, $t=4.311$, $p<.05$)를 보인다. 평가결과 전문성 영역 중 ‘성취기준 및 성취수준에 근거하여 평가결과를 해석한다.’는 25번 문항에서 중학교가 고등학교 간의 평균차이가 가장 이 높고($d=.52$) 유의미한 차이($t=4.941$, $p<.05$)를 보인다.

<표 III-20> ‘평가결과 분석 및 해석’에 대한 문항별 분석

문항번호	학교 급	N	평균	표준편차	평균차(d)	t
22번	중학교	489	3.58	.800	.35	4.074*
	고등학교	531	3.23	.789		

23번	중학교	489	3.77	.828	.44	4.990*
	고등학교	531	3.33	.794		
24번	중학교	489	3.88	.746	.37	4.311*
	고등학교	531	3.51	.820		
25번	중학교	489	3.93	.836	.52	5.691*
	고등학교	531	3.41	.837		

*p<.05

2) 정성적 분석

가) 평가결과 분석 및 해석

(1) 수행 및 서술형 평가의 통계적 분석 및 해석

“선생님이 있는 학교 또는 교과 교사들이 서술형 평가 및 수행평가 결과에 대하여 ‘통계적 분석과 해석’을 할 수 있는 정도는 어떠하다고 생각하십니까?” 또는 “선생님학교 교사들은 서술형 평가 및 수행평가 결과에 대하여 ‘통계적 분석과 해석’을 할 수 있습니까?”라는 질문에 대하여 중학교 교사들은 ‘나이스 상에서 제공하는 통계분석을 이용한다.’라고 응답하였고 고등학교 교사들도 대체로 ‘나이스 통계를 활용하고 그 외는 필요를 느끼지 못한다.’라고 하고 있다. 이를 수치로 표현하자면 6점 기준으로 중학교 교사들은 3.75점, 고등학교교사들은 2.00점 정도이다. 이 질문에 대한 교사들의 면접 내용은 다음과 같다.

학교급	중학교	고등학교
항목		
수행 및 서술형 평가의 통계적 분석 및 해석	<p>“사실상 지필평가에 대해서는 시험 후 정답률비교분석이나 도수분포표를 보고 분석을 하기도 하나, 서술형이나 수행에 대해서는 수치적인 분석이 이루어지고 있지는 않다. 그러나 통계적 분석이 가능하다면 실시해보는 것도 학생들에게 주는 피드백을 생각해 볼 때 유의미한 것이라 생각한다.” (영어)</p> <p>“나이스에서 제공하는 성적 분석 항목을 이용하고 ... ” (수학)</p> <p>“학급별 표준편차, 분산, 최상위 10% 이내 학생 등 활용하고 ... ” (국어)</p> <p>“나이스 상의 평균점수, 표준편차, 그리고 문항별 정답률 정도를 활용하고 있고 일반적으로 통계적 분석과 해석 및</p>	<p>“나이스 상의 평균과 표준편차, 과목별 등수, 원점수를 이용하고 모의고사 등 외부 학생평가에 의존하고 있어서 통계적 분석 능력을 계발할 필요를 느끼지 못하고 ... ” (영어)</p> <p>“평가 결과를 가지고 통계적 분석과 해석을 시행하지 않는 경향이 있다.” HKJ(수학)</p> <p>“나이스 상에서 통계를 확인하나 통계자료 해석에 대한 전문성 부족하고 ... ” (국어)</p> <p>“통과율, 변별도 반응분포를 활용하고 있으나 일반적으로 통계적 분석과 해석 및 활용의 중요성에 대한 인식이 미흡 ... ” (사회)</p>

	활용의 중요성에 대한 이해가 부족하고 또한 이에 대한 구체적인 연수기회가 없었기 때문에 평가결과 분석에 대한 연수가 필요할 것 같아요 ...” (사회)	“통계적 분석은 반별 평균 비교하여 10점 이상 차이가 나는 사유를 제출하는 것 말고는 특별한 통계적 분석은 실시하고 있지 않고 굳이 통계적 분석이 필요하다고 여기지 않고 ... ” (과학)
--	--------------------------------------------------------------------------------------	------------------------------------------------------------------------------------------------------------

(2) 수행 및 서술형 평가의 성취기준 및 성취수준에 따른 해석

“선생님께서 몸담고 있는 학교 또는 교과 교사들이 서술형 평가 및 수행평가 결과를 ‘성취기준과 성취수준’에 따라서 해석할 수 있는 정도에 점수를 매긴다면 6점 만점에 몇 점 정도라고 생각하십니까?” 또는 “선생님학교 교사들은 서술형 평가 및 수행평가 결과를 ‘성취기준과 성취수준’에 따라서 해석할 수 있습니까?”라는 질문에 대하여 중학교 교사들은 ‘성취기준 및 수준에 따라서 해석하는 것이 좋은 것 같다.’라는 정도로 응답하였고 고등학교 교사들도 대체로 ‘평가결과를 성취기준에 따라 해석하려는 인식이 거의 없다.’라고 응답하고 있다. 이를 수치로 표현하자면 6점 기준으로 중학교 교사들은 3.67점, 고등학교교사들은 2.04점 정도이다. 이 질문에 대한 교사들의 면접 내용은 다음과 같다.

학교급 항목	중학교	고등학교
수행 및 서술형 평가의 성취기준 및 성취수준에 따른 해석	<p>“서술형 및 수행 평가 또한 학생들의 배움의 정도를 알아보는 것이라면, 학생들의 배움을 염두에 두었을 때 애초에 생각했었던 성취기준과 성취수준을 가지고 해석하는 것이 적절하다고 생각한다.” (영어)</p> <p>“점수 결과에 따라 점수가 높은 학생은 성취수준이 높다고 판단하고 낮은 학생은 성취수준이 낮다고 보고 ... ” (국어)</p> <p>“성취기준 도달 정도를 파악하기 위해 수준별 학생 분포 비율 파악으로 평가 시행 전후를 확인하는 정도이다.” (사회)</p>	<p>“평가 결과를 해석할 이유가 별로 없고 점수는 학생의 변별과 등수로 이해하고 학생의 학업성취 및 발달과 관련하여 해석하지 않고 개별 문항의 정답률을 문항 난이도와 관련하여 사고할 뿐 학생의 성취기준·수준과 연계하지 않는다.” (영어)</p> <p>“평가 결과를 성취기준과 성취수준에 따라 해석을 시도하지 않는 경향이 있다.” HKJ(수학)</p> <p>“성취기준 자체에 대한 인식이 거의 없음” (국어)</p> <p>“평가 점수가 곧 성취수준을 의미한다고 생각하고 ... ” (사회)</p> <p>“평가 전에 평가계획서의 완성도 여하에 따라서 다르다고 생각하고 ... 성취기준과 성취수준에 따라서 명확한 채점기준을 가지고 있다면 평가 결과 해석에 무리가 없을 같다.” (과학)</p>

(3) 수행 및 서술형 평가의 이원목적분류표에 따른 해석

“선생님이 있는 학교 또는 교과 교사들이 서술형 평가 및 수행평가 결과를 ‘이원목적분류표’에 따라서 해석할 수 있는 정도는 어떠하다고 생각하십니까?” 또는 “선생님학교 교사들은 서술형 평가 및 수행평가 결과를 ‘이원목적분류표’를 기준으로 능숙하게 해석할 수 있습니까?”라는 질문에

대하여 중학교 교사들은 ‘이원목적분류표에 따라 해석해야 하는 이유를 알 수 없다.’라고 응답하였고 고등학교 교사들도 대체로 ‘평가결과를 이원목적분류표에 따라 해석하지 않거나 의미가 없다.’라고 응답하고 있다. 이를 수치로 표현하자면 6점 기준으로 중학교 교사들은 2.75점, 고등학교 교사들은 2.06점 정도이다. 이 질문에 대한 교사들의 면접 내용은 다음과 같다.

학교급 항목	중학교	고등학교
수행 및 서술형 평가의 이원목적분류표에 따른 해석	<p>“굳이 이원목적분류표로 해석한다면 가능하겠지만, 서술형과 수행평가는 성취기준에 도달했는지의 여부를 다양한 가능성을 두고 치는 평가이므로 채점기준을 가지고 해석하는 것보다 그 의미는 떨어진다고 본다.” (영어)</p> <p>“이원목적분류표에 따라 해석해야 하는 이유를 알 수 없고 한 번도 해 본 적 없다.” (국어)</p> <p>“양적 평가가 이원목적분류수준에의 도달을 의미한다고 생각함” (사회)</p>	<p>“개별 문항의 정답률을 문항 난이도와 관련하여 사고할 뿐 이원목적분류표와 연계하지 않는다.” (영어)</p> <p>“평가 결과를 이원목적분류표에 따라 해석한다는 것의 의미를 이해하기 어렵다.” HKJ(수학)</p> <p>“서술형 평가는 학생 정답률 자체가 지나치게 낮아 해석의 여지가 제한되어 있다.” (국어)</p> <p>“교과협의회를 이원목적분류표를 작성하고 여기에 맞게 문항을 출제하고 검토를 한다면 서술형 평가 및 수행평가 결과해석에 큰 문제가 없겠지만 대부분의 교사들은 문항 출제를 먼저 하기 때문에 이원목적분류표의 해석은 의미가 없다.” (과학)</p>

3) 종합결과

가) 평가결과 분석 영역에 대한 결론

평가결과에 대한 통계적 분석을 사용하는 것에 대하여 정량적 분석에서 중학교 교사들의 평균이 3.58로 고등학교 교사의 평균인 3.23보다 .35만큼 높고 유의미한 차이($t=4.074$, $p<.05$)가 있는 것으로 나타난다. 교사들의 통계적 분석 정도를 묻는 정성적 분석에서 ‘서술형이나 수행에 대해서는 수치적인 분석이 이루어지고 있지는 않다.’, ‘일반적으로 통계적 분석과 해석 및 활용의 중요성에 대한 이해가 부족하고’ 등의 표현에 알 수 있듯이 중학교 교사들은 평가결과에 대한 통계적 분석을 어려워하고 있다. 고등학교 교사들도 ‘평가 결과를 가지고 통계적 분석과 해석을 시행하지 않는 경향이 있다.’ ‘나이스 상에서 통계를 확인하나 통계 자료 해석에 대한 전문성 부족’을 드러내고 있다.

나) 평가결과 해석 영역에 대한 결론

평가결과를 어떻게 해석하느냐는 평가도구의 타당도를 높일 뿐만 아니라 학생의 학업능력의 현재와 미래를 판단하는 중요한 요소이다. 첫째 평가목표에 따라서 평가결과를 해석하는 정량적 분석에서 중학교 교사들의 평균이 3.77로 고등학교 교사의 평균인 3.33보다 .44만큼 높고 유의미

한 차이($t=4.990$, $p<.05$)가 있는 것으로 나타난다. 평가결과를 통해서 학생의 강점과 약점을 해석하는 정량적 분석에서 중학교 교사들의 평균은 3.88로 고등학교 교사들의 평균인 3.51보다 .37만큼 높고 유의미한 차이($t=4.311$, $p<.05$)가 있다. 이와 관련하여 평가결과를 이원목적분류표에 따라서 해석하는지를 묻는 정성적 분석에서 중학교 교사들은 ‘굳이 이원목적분류표로 해석한다면 가능하겠지만, ... 채점기준을 가지고 해석하는 것보다 그 의미는 떨어진다...’ ‘이원목적분류표에 따라 해석해야 하는 이유를 알 수 없고’ 등의 표현에 알 수 있듯이 중학교 교사들은 평가의 목적이거나 이원목적분류표에 따라서 평가결과를 해석하지 않는다. 고등학교 교사들도 ‘평가 결과를 이원목적분류표에 따라 해석한다는 것의 의미를 이해하기 어렵다.’ ‘대부분의 교사들은 문항 출제를 먼저 하기 때문에 이원목적분류표의 해석은 의미가 없다.’ 등의 응답에서 알 수 있듯이 중학교 교사들과 유사한 결론에 도달한다.

평가결과를 해석하는 또 하나의 기준인 성취기준 및 성취수준을 활용하는지를 묻는 질문에 대하여 중학교 교사의 평균은 3.93이고 고등학교 교사의 평균은 3.41로 이들 간의 차이는 .52로 유의미하다($t=5.691$, $p<.05$). 없다. 성취기준 및 수준을 통해서 평가결과를 해석하는 정성적 분석에서 중학교 교사들의 ‘... 학생들의 배움을 염두에 두었을 때 애초에 생각했었던 성취기준과 성취수준을 가지고 해석하는 것이 적절하다고 ...’ ‘점수 결과에 따라 점수가 높은 학생은 성취수준이 높다고 판단하고 ...’라는 응답에서 알 수 있듯이 성취기준 및 수준을 평가결과를 이해하고 해석하는 의미 있는 도구로 사용하고 있지 않다. 그리고 ‘평가 결과를 ... 학생의 성취기준·수준과 연계하지 않는다.’ ‘평가 결과를 성취기준과 성취수준에 따라 해석을 시도하지 않는 경향이 있다.’ 등 고등학교 교사들의 응답도 같은 결론에 이른다.

4) 우수사례 적용 및 분석

평가 결과 영역에서의 우수사례에 대한 적용 및 분석은 평가 과정 영역에서 충분히 설명하였기에 여기서는 생략하도록 한다.

5) 평가 전문성 제고 방안

학교현장에서 교사들이 평가결과에 대한 분석은 주로 지필평가와 관련하여 이루어진다. 평균, 표준편차, 정답률, 난이도, 등급, 과목별 등수를 나이스 시스템을 통해서 알 수 있으나 실제 학생상담에서 이용하는 것은 등급과 등수, 그리고 정답률이다. 학생들이 일반적으로 많이 틀리는 문항에 대한 분석이 이루어지고 있으나 이것도 주로 시험문제 중 몇 번 문항이 어려워져 정답률이 어느 정도인가를 서술할 뿐 학생들의 학업성취의 관점이나 평가의 목표와 관련된 서술은 없다.

가) 성취기준 및 수준에 따른 평가결과 분석 및 해석

본 연구에서는 평가결과의 분석과 해석을 준거를 성취기준 및 수준 그리고 이원목적분류표로 제시한다. 성취기준에 의한 평가결과의 분석과 해석은 학생들이 도달해야 할 성취기준과 그에 따

른 성취수준을 상세화하고 그에 따라 학생들의 학업성취를 분석하고 해석하여 학생의 강점과 약점을 제시함으로써 학생의 학업능력을 향상시키는데 기여할 수 있을 것이다. 예를 들어 아래 [그림 4-1]에서 성취기준 ‘일반적 주제에 관한 글을 읽고 글의 연결 관계를 파악한다.’라는 평가문항을 학생이 틀렸다면 학생의 학업능력에 관한 분석을 ‘주제’, ‘일반적 주제에 관한 글에 대한 이해’, 그리고 ‘글의 연결 관계 파악’에 대한 분석이 필요하고 ‘일반적 주제에 관한 글을 읽고 글의 연결 관계를 파악한다.’는 성취기준에 비추어서 학생의 학업능력이 어느 정도인지를 해석해야 할 것이다.

나) 이원목적분류표에 평가결과 분석 및 해석

이원목적분류표에 따라서 평가결과를 분석하고 해석할 수 있을 것이다. 이것은 본 연구에서는 성취기준과 연결되며 동시에 학생의 학업능력에 대한 더 상세한 분석과 해석을 가능하게 할 것이다. Bloom의 평가유형에 따라 평가문항이 설계되었다면, 예를 들어서 ‘두 문장을 읽고 연결 관계 파악하여 문장 완성하시오.’가 평가문항이라면 학생은 ‘연결사’에 대한 지식을 가지고 있어야 하고 ‘두 문장의 관계’를 구성하는 인지능력이 있어야 할 것이다. 이는 ‘Bloom의 평가유형’의 개념적 지식차원과 분석적 인식단계로 이루어지는 제6 유형에 해당한다(<표 II-22>). 이를 준거로 하여 학생의 지식 정도와 인지발달의 정도를 해석할 수 있다. 따라서 성취기준 및 수준과 이원목적분류표에 대한 역량강화가 교사들의 평가결과 전문성(ERP)을 제고하는 데 필수적이다.

서술형 평가문항 카드			
과 목	영 어	학 년	중학교 3학년
성취기준	일반적 주제에 관한 글을 읽고 글의 연결 관계를 파악한다.		
평가요소	두 문장을 읽고 연결 관계 파악하여 완성하기		
평가영역	난이도	배 점	출 처
쓰기	쉬움	4점	천재교육(김덕기 외)

[그림 IV-1] 성취기준과 평가결과의 분석 및 해석

<표 III-21> Bloom의 이원목적분류표에 따른 평가문항 해석

지식(내용) 차원	인지발달 단계		
	이해(·기억)하기	적용하기	분석(·평가·창안)하기
사실 지식	제1 유형	제2 유형	제3 유형
개념 지식	제4 유형	제5 유형	제6 유형 평가문항: ‘두 문장을 읽고 연결 관계 파악하여 문장 완성하시오.’
절차 지식(·메타지식)	제7 유형	제8 유형	제9 유형

마. 평가 환류 분석

1) 정량적 평가

평가환류에 해당하는 6개 문항 전체와 각 문항에 대하여 학교 급, 교과목, 교직경력에 따른 차이를 분석하였다.

가) 학교 급

평가환류 전체 문항에 대하여 학교 급에 따른 기술통계 및 차이 검정을 실시하였다. 그 결과 학교 급에 따라 교사의 평가환류에 대한 평가 전문성 지수에 통계적인 차이가 나타났으며, 중학교 교사가 고등학교 교사에 비하여 전문성이 높은 것으로 나타났다.

〈표 III-22〉 평가환류 전체 문항에 대한 기술통계 및 학교 급에 따른 차이 검정

학교 급		평균	표준편차	N	F
평가환류	중학교	3.6613	.76684	489	23.517***
	고등학교	3.2845	.66538	531	
	전체	3.4652	.73920	1,020	

***p<.001

평가환류에 해당하는 6개 문항 각각에 대하여 학교 급에 따른 차이를 분석하였다. 그 결과, 모든 문항에서 중학교 교사와 고등학교 교사의 평가환류에 대한 전문성은 통계적인 차이를 나타내고 있다. 또한 모든 문항에서 중학교 교사가 고등학교 교사에 비하여 높은 전문성을 가지는 것으로 인식하고 있었다.

〈표 III-23〉 평가환류 개별 문항에 대한 기술통계 및 학교 급에 따른 차이 검정

학교 급		N	평균	표준편차	t
26	중학교	489	3.75	.810	4.815***
	고등학교	531	3.33	.802	
27	중학교	489	3.85	.766	5.026***
	고등학교	531	3.40	.868	
28	중학교	489	3.54	.911	3.860***
	고등학교	531	3.18	.786	
29	중학교	489	3.69	.848	4.154***
	고등학교	531	3.32	.792	
30	중학교	489	3.80	.808	5.180***
	고등학교	531	3.36	.785	
31	중학교	489	3.57	.934	4.570***
	고등학교	531	3.14	.807	

***p<.001

나) 담당 교과목

평가환류 전체 문항에 대하여 담당 교과목에 따른 기술통계 및 차이 검정을 실시하였다. 그 결과, 담당 교과목에 따라 교사의 평가환류에 대한 평가 전문성 지수에 통계적인 차이가 나타났으며, 담당 교과목에 따른 구체적인 차이를 확인하기 위하여 Schéffe 사후검정을 실시하였다. 그 결과, 수학 교과를 담당하는 교사의 평가환류에 대한 평가 전문성 지수가 국어, 영어(외국어), 과학 교과를 담당하는 교사보다 전문성이 높은 것으로 나타났다.

〈표 III-24〉 평가환류 전체 문항에 대한 기술통계 및 담당 교과목에 따른 차이 검정

	교과목	평균	표준편차	N	F	사후검정
평가환류	국어	3.3897	.69921	186	3.379**	③>①,②,④
	영어(외국어)	3.3629	.72434	234		
	수학	3.7975	.70190	165		
	과학과	3.3632	.64576	150		
	사회과	3.3830	.84751	192		
	기타	3.6181	.67838	93		
	합계	3.4652	.73920	1,020		

** p<.01

평가환류에 해당하는 6개 문항 각각에 대하여 담당 교과목에 따른 차이를 분석하였다. 그 결과, 31번 문항을 제외한 모든 문항에서 담당 교과목에 따른 교사의 평가환류에 대한 전문성은 통계적인 차이를 나타내고 있다. 26번 문항 “서술형 평가 결과에 기초하여 수업계획을 수정·개선한다.”의 경우는 수학 교과를 담당하는 교사가 국어, 영어(외국어) 교과를 담당하는 교사에 비해 높은 전문성을 가지는 것으로 인식하고 있었다. 27번 문항 “서술형 평가 결과에 기초하여 학생들의 약점에 대한 개선안을 마련한다.”, 28번 문항 “서술형 평가 결과를 성취기준 및 성취수준을 정교화 하는데 활용한다.” 30번 문항 “서술형 평가 결과를 학급 및 학년별 교수 방법 개선 등을 위하여 학급 및 학년수준에서 활용한다.”의 경우, 수학 교과를 담당하는 교사가 영어 교과를 담당하는 교사에 비해 높은 전문성을 가지는 것으로 인식하고 있었다. 29번 문항 “서술형 평가 결과를 학생 개인의 성장 정도 확인 등을 위하여 학생수준에서 활용한다.”의 경우, 수학 교과를 담당하는 교사가 국어, 영어(외국어), 과학 교과를 담당하는 교사에 비해 높은 전문성을 가지는 것으로 인식하고 있었다.

〈표 III-25〉 평가환류 개별 문항에 대한 기술통계 및 담당 교과목에 따른 차이 검정

	교과목	N	평균	표준편차	F	사후검정
26	국어	186	3.40	.819	2.475*	③>①,②
	영어(외국어)	234	3.40	.827		
	수학	165	3.85	.803		

	과학과	150	3.54	.734		
	사회과	192	3.52	.926		
	기타	93	3.61	.761		
	합계	1,020	3.54	.832		
	국어	186	3.53	.863		
	영어(외국어)	234	3.50	.864		
	수학	165	3.95	.803		
27	과학과	150	3.52	.789	2.341*	③>②
	사회과	192	3.58	.887		
	기타	93	3.71	.783		
	합계	1,020	3.61	.849		
	국어	186	3.29	.912		
	영어(외국어)	234	3.23	.852		
	수학	165	3.69	.820		
28	과학과	150	3.24	.771	3.180**	③>②
	사회과	192	3.23	.921		
	기타	93	3.65	.755		
	합계	1,020	3.35	.866		
	국어	186	3.39	.817		
	영어(외국어)	234	3.41	.859		
	수학	165	3.89	.786		
29	과학과	150	3.29	.771	3.752**	③>①,②,④
	사회과	192	3.44	.852		
	기타	93	3.63	.809		
	합계	1,020	3.49	.838		
	국어	186	3.47	.824		
	영어(외국어)	234	3.38	.825		
	수학	165	3.87	.818		
30	과학과	150	3.50	.814	2.902*	③>②
	사회과	192	3.61	.828		
	기타	93	3.74	.729		
	합계	1,020	3.57	.826		
	국어	186	3.19	.902		
	영어(외국어)	234	3.22	.883		
	수학	165	3.56	.977		
31	과학과	150	3.32	.935	1.823	
	사회과	192	3.34	.781		
	기타	93	3.60	.855		
	합계	1,020	3.34	.895		

*p<.05, **p<.01

다) 교직 경력

평가환류 전체 문항에 대하여 교직 경력에 따른 기술통계 및 차이 검정을 실시하였다. 그 결과, 교직 경력에 따라 교사의 평가환류에 대한 평가 전문성 지수는 통계적인 차이가 나타나지 않았다.

〈표 III-26〉 평가환류 전체 문항에 대한 기술통계 및 교직 경력에 따른 차이 검정

	교육경력	평균	표준편차	N	F
평가환류	1년 미만	3.5714	.59331	21	1.398
	1~6년	3.4749	.84990	141	

6~11년	3.3138	.85969	126
11~16년	3.6698	.70538	171
16~21년	3.4996	.77179	75
21년 이상	3.4197	.67629	486
합계	3.4652	.73920	1,020

평가환류에 해당하는 6개 문항 각각에 대하여 교직 경력에 따른 차이를 분석하였다. 그 결과, 모든 문항에서 교육 경력에 상관없이 평가환류에 대한 전문성은 통계적인 차이를 나타내지 않았다.

〈표 III-27〉 평가환류 전체 문항에 대한 기술통계 및 교직 경력에 따른 차이 검정

	교육경력	N	평균	표준편차	F
26	1년 미만	21	4.00	.816	1.972
	1~6년	141	3.68	.935	
	6~11년	126	3.31	.749	
	11~16년	171	3.68	.783	
	16~21년	75	3.60	1.000	
	21년 이상	486	3.47	.797	
	합계	1,020	3.54	.832	
27	1년 미만	21	3.43	.535	1.363
	1~6년	141	3.70	.907	
	6~11년	126	3.52	.890	
	11~16년	171	3.84	.774	
	16~21년	75	3.64	.952	
	21년 이상	486	3.54	.835	
	합계	1,020	3.61	.849	
28	1년 미만	21	3.29	.756	1.750
	1~6년	141	3.41	.956	
	6~11년	126	3.12	.889	
	11~16년	171	3.60	.842	
	16~21년	75	3.20	.816	
	21년 이상	486	3.34	.843	
	합계	1,020	3.35	.866	
29	1년 미만	21	3.71	.756	1.307
	1~6년	141	3.51	.953	
	6~11년	126	3.36	.850	
	11~16년	171	3.68	.805	
	16~21년	75	3.64	.810	
	21년 이상	486	3.42	.814	
	합계	1,020	3.49	.838	
30	1년 미만	21	3.57	.787	.644
	1~6년	141	3.47	.952	
	6~11년	126	3.48	.833	
	11~16년	171	3.72	.750	
	16~21년	75	3.52	.918	
	21년 이상	486	3.58	.802	
	합계	1,020	3.57	.826	
31	1년 미만	21	3.43	1.134	.747
	1~6년	141	3.34	1.006	
	6~11년	126	3.29	.929	

11~16년	171	3.54	.914
16~21년	75	3.40	.913
21년 이상	486	3.28	.836
합계	1,020	3.34	.895

2) 정성적 평가

가) 채점기준 공지

학교급 항목	중학교	고등학교
채점기준을 사전에 공지	<p>“서술형 또는 수행평가를 시행하기 전에 채점기준을 공개하고 있습니다. 공개하는 이유는 학생들의 심리적 부담을 줄이고 배움의 요소에 대한 점검을 하는데 있어 조금이라도 도움을 주고 싶은 의도입니다.” (영어)</p> <p>“평가 문항에 대한 학생 신뢰도를 확보하기 위하여 채점 기준을 평가 시행 이전에 공지하고 있습니다만, 아주 간략하게 제시하는 수준입니다. 채점 기준을 모두 공개하는 것은 시험문제를 그대로 노출시키는 것과 같다고 생각합니다.” (수학)</p> <p>“서술형평가 및 수행평가 계획을 학년 초에 학교교육계획서에 포함하고 학교 홈페이지를 통해 사전에 공지하고 채점 기준은 평가 전에 학급에 게시하고 있습니다.” (사회)</p>	<p>“채점기준을 공개하지 않는다. 채점기준 자체가 문제에 대한 힌트가 될 수 있다. 또한 기준안이 학생들의 답안 때문에 수정되는 경우도 많기 때문이다.” (국어)</p> <p>“거의 공지하지 않는다. 채점기준을 공지하는 것이 시험문제를 가르쳐주는 것과 유사하다고 느껴집니다.” (영어)</p> <p>“서술형평가 및 수행평가의 채점기준은 사전에 공지하여 학생들이 준비할 수 있게 합니다.” (사회)</p>

나) 서술형평가 및 수행평가 결과 분석 및 공지

학교급 항목	중학교	고등학교
서술형평가 및 수행평가 결과 분석 및 공지	<p>“서술형평가를 제외한 지필평가는 객관적 통계 자료를 통해 교과협의를 거쳐 분석을 하나, 실제로 서술형이나 수행평가는 그러한 분석이 잘 이루어지지 않는다.” (영어)</p> <p>“교과 별로 확인하고 있지만 학생들에게 따로 공지하지는 않습니다.” (수학)</p> <p>“서술형평가 및 수행평가 결과를 분석하</p>	<p>“공지하고 있지만 분석은 거의 하지 않습니다. 분석할 의무가 없으며, 분석 자체가 서술형평가나 수행평가의 문제를 해결해 준다고 생각하지는 않습니다.” (국어)</p> <p>“평가 결과는 단순 평균 비교나 석차를 내는 용도 외에는 사용하지 않습니다.” (과학)</p>

	고 그 결과를 학생들에게 공지하는 것은 교사의 당연한 의무라고 여기고, 평가의 공정성 및 객관성 확보를 위해 서술형평가 및 수행평가의 채점 결과를 개인별로 확인하게 하고 있습니다.” (사회)	“서술형평가 및 수행평가 결과를 분석하고 그 결과를 공지하는 것은 교사의 당연한 의무라고 여기고, 평가의 공정성 및 객관성 확보를 위해 결과를 공지하고 있습니다.” (사회)
--	------------------------------------------------------------------------------------------------------------	--------------------------------------------------------------------------------------------------

다) 서술형평가 및 수행평가 결과를 학생의 학업 능력 향상에 활용하는 정도

학교급 항목	중학교	고등학교
서술형평가 및 수행평가 결과를 학생의 학업 능력 향상에 활용하는 정도	<p>“현실적으로 시험은 학생들의 학습을 유도하는 기능을 가지고 있다. 이런 면에서 본다면 서술형이나 수행 평가는 지필평가에서 가지지 못하는 알파를 가졌다고 본다. 일회성의 지필평가가 아니라 꾸준한 수업참여나 노력, 배움의 과정을 지켜보고 배움을 다양한 방식으로 수행해보게 하는 기회 제공이 서술형이나 수행 평가가 아닐까 생각한다. 수업시간에 중점적으로 다루었던 활동이나 프로젝트에 대해서는 서술형이나 수행에서 한번쯤은 다루어지는 것이 좋으며, 교과서 본문을 텍스트 삼아 아니면 교과서 외의 텍스트를 가지고 성취기준의 배움을 도출해내게 하는 것도 좋은 방법이라고 생각한다.” (영어)</p> <p>“기본 성취 수준 미달인 학생을 선별하여 방과 후 수업이나 보충 수업 등을 통해 해당 학생들의 학업 능력을 향상시키는데 활용하고 있습니다. 그러나 요즘은 방과 후에 학생을 남겨서 학습결손을 보완하는 것이 쉽지는 않습니다. 학원과 같은 사교육 시간과 겹치는데 학교에 남겨서 보충하는 것에 대하여 학부모의 항의를 받은 적도 있습니다.” (수학)</p> <p>“학생의 학업 능력 향상을 위해 학습수행과정 및 결과를 서술형평가 및 수행평가에 반영하는 것이므로 학업 능력 향상에 도움이 된다고 생각합니다.” (사회)</p>	<p>“서술형평가 및 수행평가 결과에 따라 훌륭한 결과는 학습특기 사항에 기재하고, 미흡한 결과는 개별학생을 불러 개별 지도하고 있습니다.” (사회)</p> <p>“결과 분석이 미비하여 학업능력 향상 활용도가 낮습니다.” (영어)</p> <p>“교과에서 알아야 할 중요개념을 묻는 문제를 출제하고 있으며, 개념을 명확히 숙지한 학생들이 득점할 수 있기에 학생 간 수준차이를 명확히 판별할 수 있게 합니다. 이를 근거로 평가 결과에 따라 개별 학생에게 피드백하고 있습니다.” (국어)</p>

라) 서술형평가 및 수행평가 결과를 교사의 수업 개선에 활용하는 정도

학교급 항목	중 학교	고 등 학 교
서술형평가 및 수행평가 결과를 교사의 수업 개선에 활용하는 정도	<p>“평가의 결과를 통해 학생들의 배움이 어떤 방향으로 일어나고 있는가를 살피는 것은 중요한 포인트가 됩니다. 서술형이나 수행 평가의 결과를 보고 학생들이 성취기준에 맞는 배움의 과정을 유도하도록 수업 설계를 계속 수정하고 있습니다.” (영어)</p> <p>“ 서술형평가 혹은 수행평가의 각 문항의 결과를 분석하여 통과 비율이 낮은 경우, 추후 수업에 이를 반영하여 학생들이 기본 성취 수준을 도달할 수 있도록 수업방법 및 내용을 수정하는 편입니다.” (수학)</p> <p>“서술형평가 및 수행평가 결과로 학생들이 어려워하는 부분, 성취기준에 도달하지 못한 부분을 파악하여 성취기준 도달 정도를 보완할 수 있는 방법을 고민하고 있습니다. 그리고 학생들이 작성한 오답의 특징을 분석하여 이해부족이나 오개념 혹은 문항의 보완점을 모색하고 있습니다.” (사회)</p>	<p>“서술형 평가를 출제하면서 수업시간에 주의·환기의 도구로 사용합니다. 즉 서술형 평가 출제를 예고하고 수업을 함으로써 방어적 수업전략을 사용하고 있습니다. 또한 수행평가에 태도점수를 넣어서 수업분위기 개선을 시도합니다.” (국어)</p> <p>“ 서술형평가 및 수행평가 결과에 따라 학생들이 우수한 영역은 빨리 지나가고, 미흡한 영역에 대하여 집중적으로 수업을 진행하고 있습니다.” (사회)</p> <p>“서술형 평가의 경우, 답안을 작성하지 못하는 학생의 수가 많습니다. 그로 인하여 일부 학생의 결과를 제외하고는 수업개선에 활용하기가 현실적으로 어렵습니다. 그러나 수행평가의 경우는 동적인 활동을 통해 학생들의 학습능력 향상에 도움이 된다고 느낍니다. 따라서 교사가 다양한 교수학습 방법을 도입해야 할 필요를 느끼고 실제로 많은 교사들이 다양한 교수학습 방법을 활용하고자 노력하고 있습니다.” (과학)</p>

마) 서술형평가 및 수행평가 결과에 따라 성취기준 및 성취수준을 수정·조절하는 정도

학교급 항목	중 학교	고 등 학 교
서술형평가 및 수행평가 결과에 따라 성취기준 및 성취수준을 수정·조절하는 정도	<p>“평가 결과를 보고 성취기준을 다소 조정하는 것은 현실적으로 우리 학교 학생의 수준과 환경에 맞추어 배움이 일어나도록 하겠다는 의지로 보입니다. 학력이 다소 낮은 학교의 경우에는 학생들의 평가 후에 결과를 보고 성취기준을 재조정하고 그에 맞는 수업설계를 새로 해나가는 노력은 필요하다고 생각합니다.” (영어)</p> <p>“ 문항의 난이도를 조절하거나, 성취기준 및 성취수준을 조절하지는 않습니다. 성취기준 및 성취수준을 조정한다는 것은 자칫 성적 부풀리기와 같은 부</p>	<p>“학교 현장에서는 수업과 평가가 따로 분리된 것처럼 느끼는 경우가 많습니다. 다시 말해 해당 수업의 목표를 성취기준과 성취수준에 맞추지 않는 경우가 많아 보입니다.” (영어)</p> <p>“학기 초 계획 수립 때, 계획한 성취기준 및 성취수준을 수정 조절하지 않는 경향이 있다.” (수학)</p> <p>“현재 성취기준 및 성취수준을 수정하여 반영하지 않습니다. 학기 초에 계획된 내용으로 진행하고 다음 해에는 다시 새로운 기준에 맞춰 조절합니다. 매</p>

	<p>정행위와 직결될 수도 있다고 생각합니다.” (수학)</p> <p>“현장에서 서술형평가나 수행평가를 실시해보면 답안을 작성하지 못하는 경우와 과제의 수행수준이 현저히 낮을 때가 많습니다. 이런 경우는 학생을 제대로 평가하기 어렵습니다. 따라서 이전 평가의 결과를 반영하여 성취수준을 다소 낮게 수정·조절하는 경우도 있습니다.” (사회)</p>	<p>년 다른 학년을 지도하는 경우가 많다 보니 연속성이 부족하여 실제적으로 성취기준 및 성취수준을 수정·조절을 실시하기 어렵습니다.” (과학)</p>
--	---------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------	--------------------------------------------------------------------------------------

바) 서술형평가 및 수행평가 결과를 학생수준에서 활용하는 정도

학교급 항목	중학교	고등학교
서술형평가 및 수행평가 결과를 학생수준에서 활용하는 정도	<p>“학력이 다소 낮은 우리 학교의 경우에는, 학생들의 평가 후에 결과를 보고 성취기준을 재조정하고 학생들의 수준에 맞는 수업설계를 새로 해나가는 노력을 하고 있습니다.” (영어)</p> <p>“ 성취 수준이 미달인 학생에 대하여 보충 수업 실시하고 있습니다.” (수학)</p> <p>“서술형평가 및 수행평가 결과를 통해 학생이 어려워하는 부분, 성취기준에 도달하지 못한 부분을 파악하여 피드백하고 있습니다. 또한 평가결과를 교수·학습방법이 학생들의 이해수준에 적합한가에 대한 점검의 준거로 활용하고 있습니다.” (사회)</p>	<p>“서술형평가나 수행평가의 결과는 개인별로 통지하는 것을 제외하고는 별다른 조치가 없습니다.” (과학)</p>

사) 서술형평가 및 수행평가 결과를 학급·학년·학교수준에서 활용하는 정도

학교급 항목	중학교	고등학교
서술형평가 및 수행평가 결과를 학급·학년·학교수준에서 활용하는 정도	<p>“서술형평가 및 수행평가의 결과를 학년수준에서 활용하기 위해 학년협의회를 하는 학년도 있습니다. 학습공동체로서 특정 교과나 학습 요소에 대해 부족한 반이 있다면 통합교과로 접근할 수도 있으며 담당 선생님이 아닌 다른 선생님들의 공동 지원이 필요할 수도 있기 때문에 학년 수준의 활용도 가능하다고 생각합니다.” (영어)</p>	<p>“평가 결과에 따라서 미흡한 교과는 방과 후 수업 등을 개설하여 학생들의 학업 능력을 보완하고자 하고 있습니다.” (사회)</p>

	<p>“ 성취 수준이 미달인 학생에 대하여 학교 차원의 보충 수업 실시하고 있습니다.” (수학)</p> <p>“서술형평가 및 수행평가 결과를 학급·학년·학교수준에서 활용하는 정도는 낮습니다. 지필평가의 비중이 높으므로 주로 지필평가 결과를 기준으로 활용합니다. 서술형평가 및 수행평가 결과는 학기말에 주로 입력되므로 활용도가 낮습니다.” (사회)</p>	
--	-------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------	--

3) 종합결과

서술형 및 수행평가에 관한 평가 환류는 교사측면, 학생측면, 교육이해 당사자 측면 등에서 문제점을 고찰할 수 있다.

첫째, 일선 학교 교사는 학생 개인별 서술형 평가와 수행평가 결과를 종합·분석하고 개선방안을 도출하고 있지 않다. 즉, 교사는 서술형 평가와 수행평가 결과를 토대로, 학생들의 재능과 학업성취 정도를 파악하고 학생의 꿈과 끼 실현을 위한 학습계획을 구체적으로 세우고 있지 않으며, 서술형 평가와 수행평가 시행 취지에 부합한 학생들의 강점과 약점을 파악하고 맞춤형 교수학습 방법의 개선을 위한 노력이 전혀 이루어지고 있지 않다.

둘째, 일선 학교 교사는 서술형 및 수행 평가를 통하여 학생 스스로 성취기준에 제시된 평가내용과 능력(이해, 적용, 분석)을 파악하여 강점과 약점을 고찰하고 있지 못하고, 심지어 교사들은 학생들조차 서술형 및 수행 평가에 대한 학업흥미와 동기를 전혀 유발하지 못하며 오히려 학업부담만 가중시킨다고 인지하고 있다고 밝히고 있다. 나아가 이를 통한 학습 능력 신장에 필요한 정보와 자기 주도 학습이 가능하게 할 수 있는 정보를 담지 못함을 발견할 수 있었다.

셋째, 일선 학교에서는 서술형 및 수행 평가 자료가 누적적으로 구비되어 있지 못한 실정이다. 즉, 이를 토대로 학생들의 학력 향상, 공교육 정상화와 사교육 경감 등의 교육정책방향에 구체적으로 활용되고 있지 못한 실정이다. 서술형 및 수행 평가 자료를 토대로, 학생들의 학력신장에 얼마나 기여했는지에 대한 정보를 누적적으로 관리되어질 필요성이 제기된다.

4) 개선방안

앞서 제시한 평가 환류 영역에서의 문제점을 중심으로 그 개선 방안을 제시하면 다음과 같다. 첫째, 학생들의 강점과 약점, 재능과 학업성취 정도를 파악하여 교수학습방법의 개선과 학생의 꿈과 끼 실현을 위한 학습계획을 세워야 한다.. 또한 서술형 평가와 수행평가 결과를 활용하여, 학생의 창의성 신장을 위한 효과적인 교육활동을 구상하고, 학생 발달과정을 이해하고 지속적이고 안정적인 성장을 도모

해야한다. 즉, 서술형 평가와 수행평가 결과가 단순한 학생의 성취 정도만 파악하는 단순한 평가에서 그칠 것이 아니라 학생들의 학습동기와 학습 환경의 개선을 위한 자료로 활용되어야 한다.

학생 스스로 무엇을 알고, 무엇을 할 수 있는지에 대한 구체적인 분석을 통하여 학생 스스로 문제점을 파악하여 자기 주도적인 학습이 이루어질 수 있도록 가능하게 할 수 있다. 이는 곧 학생들의 학력 향상, 공교육 정상화와 사교육 경감 등에 기여할 수 있다.

IV. 결론 및 제언

1. 결론

가. 평가 환경

현 일선학교의 서술형 평가와 수행평가에 관한 평가 환경은 교육여건(교육과정), 교수학습·방법, 교육평가 세부여건 등으로 고찰해 볼 수 있다.

첫째, 서술형 및 수행 평가에 관한 대부분의 일선 학교교사는 이러한 유형의 평가방식은 학생들의 창의적인 학습과 개념적 지식을 이해하는데 활용 가치는 있지만, 현 교육여건 상 학생들의 학습 부담과 교사의 업무 부담으로 작용함을 피력하고 있다. 이런 점에서 서술형 및 수행 평가에 관한 교사들 간의 협업은 사실상 불가능하고, 명확한 서술형 및 수행평가 기준을 정립할 수 없는 한계점을 지적하고 있다. 가장 시급한 것은 서술형 및 수행 평가에 적합한 즉, 학생을 위한 수업이 이루어질 수 없는 학교 분위기를 언급하고 있다. 무엇보다 현 입시제도에서 서술형 및 수행 평가 본래 취지에 맞는 시험을 준비는 학생들의 공감을 얻지 못하고 있다. 이런 문제점을 해결하기 위해서 정부는 교사가 서술형 평가와 수행평가에 몰입할 수 있는 적합한 평가 환경을 조성해 주어야 한다.

둘째, 일선학교 교사들은 학생들의 수준을 고려하여 평가 기준을 수정하다 보면, 객관적인 기준 자체가 모호해 지는 경향이 있음을 지적하고 있다. 서술형 및 수행 평가는 평가기준이 불분명한 경우가 많고, 이의 비율조정 등 평가에 관한 교사 자율성 여건 형성이 유연하지 못하다. 이런 점에서 서술형 및 수행 평가를 채점하는데 많은 시간과 노력이 필요함을 토로하고 있다. 즉, 협의의 과정(부분점수 기준설정 등)을 거치면서 시간 부족과 업무 부담만을 가중시킨다고 인식하고 있는 것이 일선학교의 현실임을 발견할 수 있었다. 정부는 이런 점들을 절실히 인지할 필요가 있으며, 교사의 업무부담 경감과 서술형 및 수행 평가 기준(안)을 참고할 수 있는 자료제공이 시급한 실정이고, 단위별로 구체적으로 제공 되어질 필요성이 제공된다. 나아가 서술형 및 수행 평가만을 위한 교수학습·방법 개선을 위한 정책적 방안 등을 고찰해 보아야 한다.

나. 평가 투입

서술형평가 및 수행평가의 전문성을 향상시키기 위한 시발점은 평가투입이라고 할 수 있다. 이에 대한 논의 및 결론은 다음과 같다.

첫째, 대다수의 교사들이 평가의 목표설정에 있어서 스스로는 나름의 준거를 가지고 평가의 목

표를 설정하는 것으로 인식하고 있으나 그 준거가 명확하지 않다. 이러한 문제로 인하여 문항 개발과 채점기준 마련에 어려움을 느끼는 것으로 판단된다. 다시 말하면, 학교 현장에서 많은 교사들이 경험을 바탕으로 문항개발과 채점기준을 마련한다는 것이다. 또한 교사들은 성취기준을 중시하지 않는다는지, 교육과정에 제시된 성취기준을 무작정 따르려는 경향을 보이고 있다. 교육과정에 제시된 성취기준은 포괄적이고 추상적이기 때문에 학생의 수준에 걸맞게 성취기준을 수정·보완 기회가 마련되어야 할 것이다. 따라서 서술형평가 및 수행평가가 제 기능이 잘 발휘되게 하기 위해서는 교육과정에 제시된 성취기준을 평가 목표에 맞도록 수정·보완하는 것이 필수 선결과제라 하겠다.

둘째, 교사들이 서술형 및 수행평가의 문항 유형을 선정하는 준거가 명확하지 않다. 대다수의 현장 교사들은 ‘자료 제시형’, ‘응답 제한형’ 등과 같은 문항의 형식이나 학생의 응답 범위를 준거로 문항의 유형을 선정하는 것으로 보인다. 그러나 명확한 평가 요소를 정확히 측정할 수 있는 즉, 평가 목표에 부합하는 문항을 선정하는 것이 문항의 유형 선정에 있어서 가장 중요한 준거라 할 수 있다. 이에 본 연구에서는 ‘Bloom의 평가 유형’을 제안하고 있다.

다. 평가 과정

서술형 및 수행평가에 있어서 현재 학교현장에서 교사들이 가장 많은 시간을 투입하고 있는 부분은 평가문항을 제작하고 그 채점기준을 설정하는 일일 것이다. 이 연구에서 서술형 및 수행평가 전문성의 영역에서 평가과정은 평가문항의 제작과 분할점수산출 그리고 채점과정이고 이에 대한 논의와 결론은 다음과 같다.

첫째, ‘교사들이 성취기준 및 이원목적분류표의 내용영역인 지식차원과 행동영역인 인지단계를 핵심평가요소로 하는 평가문항을 제작하는가?’라는 평가문항 제작의 전문성에 대한 논의를 살펴보면, 성취기준에 대해서는 고등학교 교사들보다 중학교 교사들이 더 잘 이해하고 더 많이 활용하고 있지만 중등학교 교사 전체적으로 평가문항을 제작하는데 일정 수준이상으로 활용하고 있다고 보인다. 그러나 중학교나 고등학교 모두 이원목적분류표를 문항개발의 핵심도구로 사용하지 않고 있음은 명백하다. 이는 이원목적분류표의 내용영역/지식차원과 행동영역/인지단계를 평가문항의 평가요소로 구체화되지 못하고 있음을 알 수 있다.

둘째, ‘교사들이 평가문항의 채점기준을 설정하는데 채점요소를 구체적이고 명확하게 설정하는가?’라는 채점 기준 설정에 대한 논의를 살펴보면, 기존의 평가요소 분석에서 보듯이 채점기준은 내용영역이 중복되는 형식을 취하고 있어서 기준의 타당성이 떨어진다. 채점기준을 설정하는데 있어서 중학교가 고등학교보다 더 잘하고 있다. 성적 부담이 적고 자유학기제를 시행하는 중학교는 여유가 있고 대학입시로 인하여 내신성적에 부담이 큰 고등학교의 경우에는 중학교보다 엄격한 채점기준이 적용되고 있다는 의미일 것이다. 중학교와 고등학교의 이러한 차이에도 불구하고 중고등학교 모두 평가문항의 평가요소와 일관되고 정합적인 채점기준설정이 되지 못하고 있다.

셋째, 이상적인 평가문항 제작절차를 현실에 적용하는 것은 노력에 비해 효율이 떨어지는 것으로 나타난다. 현실적으로는 이원목적분류표의 이해와 작성, 서술형 및 수행평가 문항의 제작 그리

고 채점기준표의 설정을 중심으로 하는 핵심적 제작절차나 채점절차를 적용하는 것이 비용대비 효과가 크다. 따라서 교사들에게 핵심적 서술형 및 수행평가 문항 제작절차에 관한 연수와 매뉴얼 제공이 기초적으로 제공되어야 할 것이다. 문항제작 및 채점기준 설정의 핵심적 절차과정 중에 성취기준 및 수준과 이원목적분류표에 따른 별 분할점수를 산출하는 것은 개별 평가문항의 난이도와 학생들의 능력이라는 요소를 비교 판단하여 학생의 학업능력을 신장하는 중요한 절차이므로 이에 대한 심도 있는 교사연수가 지속되어야 함은 물론 학생평가(서술형 및 수행평가) 전문성 척도 가운데 이를 포함하여야 할 것이다.

라. 평가 결과

평가결과는 평가목적과 성취기준 및 수준 그리고 이원목적분류표와의 정합성, 성취기준과 평가문항 간의 정합성, 그리고 평가문항과 채점기준과의 일관성 및 정합성에 따른 학생평가결과 의 분석과 해석에 대한 고찰이다. 평가결과의 분석은 결과의 양적 분석 및 통계적 분석으로 문항의 정답률, 난이도, 평균을 문항자체를 중심으로 하고 평가결과의 해석은 성취기준 및 수준의 달성정도를 중심으로 개별 학생의 발달정도를 중심으로 한다.

첫째, 평가결과에 대한 교사의 통계적 분석은 나이스에서 제공하는 지필평가에 관한 것이 대부분이었고 서술형 및 수행평가에 관한 통계적 분석은 거의 없었다. 그리고 평가에 대한 전문성(TEP) 전체에 대한 교사의 인식과 평가결과에 대해 분석하는 능력정도를 비교해보았을 때 평가결과의 분석에 대한 능력은 낮게 나타난다. 그리고 평가결과에 대한 분석능력이 평가과정, 즉 평가문항제작과 채점기준설정 능력보다 현저히 낮다. 중고등학교 교사들이 평가전문성의 다른 영역보다 서술형 및 수행평가의 결과전문성 영역, 즉 평가결과의 통계적 분석 능력을 우선적으로 신장하고, 평가결과를 평가의 목적 및 목표와 연계하는 능력을 제고하도록 교육통계 분석과정을 제공해야 한다. 또한 나이스(NEIC)시스템 상에서 서술형 및 수행평가에 대한 독립적 통계분석 지원하는 것도 동시에 진행되어야 할 것이다.

둘째, 서술형평가 및 수행평가 결과를 이원목적분류표를 기준으로 해석할 수 있는지, 또는 성취기준 및 수준을 기준으로 해석할 수 있는지에 대한 논의에 대하여 입시 중심의 고등학교는 물론 자유학기제를 시행하고 있는 중학교에서조차도 그 해석능력이 현저히 떨어졌다. 이원목적분류표와 성취기준 및 수준에 의한 해석은 학생의 지식 정도와 인지발달의 정도를 파악하고, 학생의 강점과 약점을 파악하여 개별화된 학업향상 처방을 제시할 수 있는 근거가 될 것이다. 따라서 교사들의 평가결과 전문성(ERP)을 제고하기 위해서 성취기준 및 수준과 이원목적분류표에 대한 역량강화를 필수적 과정으로 제공해야 할 것이다. 또한 서술형 및 수행평가를 포함하는 교육(학생)평가전문성 인정제도나 교육(학생)평가 분석사 국가자격증 제도를 도입한다면 이러한 목표를 달성하는데 한층 강력한 동기를 제공할 수 있을 것이다.

마. 평가 환류

현 일선학교의 서술형 평가와 수행평가에 관한 평가 환류는 교사측면, 학생측면, 교육이해 당

사자 측면 등에서 문제점을 고찰할 수 있다.

첫째, 일선 학교 교사는 학생 개인별 서술형 평가와 수행평가 결과를 종합·분석하고 개선방안을 도출하고 있지 않다. 즉, 교사는 서술형 평가와 수행평가 결과를 토대로, 학생들의 재능과 학업성취 정도를 파악하고 학생의 꿈과 끼 실현을 위한 학습계획을 구체적으로 세우고 있지 않으며, 서술형 평가와 수행평가 시행 취지에 부합한 학생들의 강점과 약점을 파악하고 맞춤형 교수학습 방법의 개선을 위한 노력이 전혀 이루어지고 있지 않다. 더욱 심각한 것은 서술형 평가 문제를 출제 할 때, 학생 창의성 신장을 도모하기 위하여 출제를 하고 있지만 이에 적합한 효과적인 교육활동을 구상하고 있지 못한 점과 이의 결과를 활용하여, 학생 발달과정을 지속적이고 안정적으로 확인하고 개선방안을 모색하고 있지 않는 점이다. 정부는 이러한 점을 근간으로 서술형 평가와 수행평가의 본래 취지를 위한 평가 환류 측면의 정책적 개선노력이 절실히 필요하다.

둘째, 일선 학교 교사는 서술형 및 수행 평가를 통하여 학생 스스로 성취기준에 제시된 평가내용과 능력(이해, 적용, 분석)을 파악하여 강점과 약점을 고찰하고 있지 못하고, 심지어 교사들은 학생들조차 서술형 및 수행 평가에 대한 학업흥미와 동기를 전혀 유발하지 못하며 오히려 학업부담만 가중시킨다고 인지하고 있다고 밝히고 있다. 나아가 이를 통한 학습 능력 신장에 필요한 정보와 자기 주도 학습이 가능하게 할 수 있는 정보를 담지 못함을 발견할 수 있었다. 정부는 서술형 평가와 수행평가의 본래 취지를 위한 지속적이고 안정적인 학생 성장과정을 도모하기 위한 정책적 대안마련이 시급한 실정이다.

셋째, 일선 학교에서는 서술형 및 수행 평가 자료를 누적적으로 구비되어 있지 못한 실정이다. 즉, 이를 토대로 학생들의 학력 향상, 공교육 정상화와 사교육 경감 등의 교육정책방향에 구체적으로 활용되고 있지 못한 실정이다. 서술형 및 수행 평가 자료를 토대로, 학생들의 학력신장에 얼마나 기여했는지에 대한 정보를 누적적으로 관리되어질 필요성이 제기된다. 정부는 학력 평가관리 시스템을 적용하고 활용할 수 있는 여건마련과 이를 토대로 교사들의 평가 전문성 향상을 위한 프로그램 개발 등이 시급한 실정이다. 즉, 정부는 학교현장의 안정적이고 지속적인 평가환류체계 구축을 위한 인적, 물적 여건 조성이 필요하다. 특히, 평가 환류에 적합한 맞춤형 교육과정 편성 및 운영, 평가 전문성 향상을 위한 전문가 협의체 구성 및 연수의 필요성 인지, 안정적이고 지속적인 학력 평가관리 및 평가실행 시스템 구축이 필요하다.

2. 제언

본 연구에서 수행한 서술형 평가 및 수행 평가의 평가전문성 제고 방안을 통해 향후 정책 분석 및 개발, 연구에 활용될 수 있는 다양한 정책 제언이 도출되었다. 정책 제언은 본 연구의 전문성 신장 모형을 토대로, 크게 평가 환경, 투입, 과정, 결과, 환류에 근거하여 평가 전문성과 관련한 교육정책 개선 제언 중심으로 구성되었다.

평가 환경은 교사들의 학생 평가 전문성 신장을 위하여 수행되어야 하는 일련의 활동들에

영향을 미치는 전반적인 것을 의미한다. 즉, 교사가 서술형 평가 및 수행 평가의 본질적인 목표달성을 위해 수행해야 하는 제반적인 활동에 대하여 몰입할 수 있도록 주어져야 하는 여건을 의미하는 것이다. 일선 학교에서 현 평가 환경의 문제점은 서술형 평가와 수행평가에 적합한 교육여건(교육과정), 교수학습·방법, 교육평가 세부여건 등에서 문제점을 고찰할 수 있는데 이를 개선시키기 위한 제언을 하면 다음과 같다.

첫째, 교육여건을 개선하기 위하여 정부는 우선적으로 교사의 업무 부담을 경감시킬 수 있는 정책적 방안 모색이 절대적으로 필요하다.

둘째, 정부는 서술형 및 수행 평가 취지에 맞는 교육과정이 개편시킬 필요가 있다. 나아가 장기적으로는 이러한 평가를 대입제도 등에 적용시킬 방안 모색도 필요하다.

셋째, 일선 학교 현장에서 교사는 명확한 서술형 및 수행 평가 기준 정립에 어려움을 토로하고 있다. 이를 개선시키기 위해서는 이와 관련한 실질적 실효성을 얻을 수 있는 연수 등의 방안을 적극적으로 모색해야 할 필요성이 제기된다. 나아가 서술형 평가 기준(안)을 참고할 수 있는 다양하고 구체적인 자료제시 방안 모색이 필요하다.

넷째, 일선 학교에서는 서술형 및 수행 평가를 위한 교사 협업이 적극적으로 이루어질 수 있는 학교분위기를 모색할 필요성이 제기된다.

다섯째, 정부는 서술형 및 수행 평가를 위한 교수학습·방법개선을 위한 정책적 방안 모색도 필요하다. 현 정부의 교육정책 즉 “꿈과 끼를 끌어내는 행복교육”에 부합하는 서술형 및 수행 평가를 위한 다양한 교수학습·방법을 창의·인성 프로그램과 연계하여 지속적이고 안정적으로 정착시킬 수 있는 방안 모색이 필요하다. 이를 위하여 서술형 및 수행 평가를 위한 교수학습·방법과 관련한 지속적인 우수사례를 발굴할 수 있는 정책 방안 모색이 절대적으로 필요하다.

여섯째, 서술형 및 수행 평가를 위한 다양한 교수학습·방법을 활용할 수 있는 수업시간확보와 평가시기 결정 등의 합의점 방안 모색이 필요하다.

일곱째, 교육평가 세부여건 부분에서는 서술형 및 수행 평가 기준을 명확히 정립할 수 있는 안내 책자 제공 및 지속적인 홍보를 통하여 서술형 및 수행 평가 채점 시 교사들 간의 견해 차이를 조율하기 위한 방안 모색이 필요하다. 그리고 서술형 및 수행 평가만을 위한 수업시간 및 시험시간이 확보 방안 마련이 필요하다. 또한 서술형 및 수행 평가의 유연한 평가비율 조정, 시기와 범위문제, 출제와 채점시간에 부담감을 해소시킬 수 있는 방안 모색도 필요할 것이다. 마지막으로 서술형 및 수행 평가에 관한 교사 안내책자뿐만 아니라 학생 답안 작성 방법 등을 담은 학생 교육용 안내 책자 제공도 필요할 것으로 사료된다.

평가 투입은 교사들이 학생 평가를 위한 시작점이라 할 수 있다. 평가 투입 단계에서는 평가 목표의 설정, 평가 혹은 평가 문항의 유형 선정이 이루어진다. 평가 과정은 투입 단계에서 설정된 평가 목표에 적합한 문항이 개발되어야 하며, 평가 결과는 평가 목표를 고려하여 학생의 성취 수준을 정확히 파악할 수 있어야 함을 의미한다. 평가 투입-과정-결과에서의 문제점은 경험적인 평가 목표 설정과 평가 목표를 고려하지 않은 평가 혹은 평가 문항의 유형을 선정함으로 인하여 평가 목표에 따라 학생의 성취수준을 정확히 측정하지 못하는 것이라 하겠다. 이를 개선하기 위한

제언을 하면 다음과 같다.

첫째, 성취기준과 평가문항 간의 정합성과 일관성에 대한 교사들의 인식과 관심을 강화하기 위해서 이원목적분류표와 평가문항과 관계와 제작원리를 안내하는 것이 교사들의 평가전문성 제고를 위하여 반드시 필요하다. 그리고 교사들이 평가문항의 제작을 원활하게 할 수 있도록 Bloom의 평가유형에 따른 평가문항의 모듈화를 제안한다.

둘째, 성취기준 및 성취기준 그리고 이원목적분류표에 대한 정확한 이해와 이를 채점기준설정 에 활용하는 능력을 신장시키는 위하여 성취기준·수준 및 이원목적분류표에 따른 평가문항 제작 과 채점기준의 일관성과 정합성을 신장하는 연수와 매뉴얼이 필요하다. 이를 근본적으로 가능하게 하기 위하여 수업과 평가를 통합하는 교수학습과정이 설계되어야 할 것이다.

셋째, 성취기준 및 수준과 이원목적분류표에 따른 분할점수를 산출에 대한 심도 있는 교사연수가 지속되어야 함은 물론 학생평가(서술형 및 수행평가) 전문성 척도 가운데 이를 포함하여야 할 것이다.

넷째, 교사의 통계적 분석 능력을 높일 수 있는 방안 모색이 필요하다.

다섯째, 교사의 역량 강화를 위하여 교육(학생)평가전문성 인정제도나 교육(학생)평가 분석사 국가자격증과 같은 제도적 장치를 마련하는 방안에 대한 모색이 필요하다.

평가 환류는 교사들의 학생 평가를 통하여 수행되어졌던 일련의 활동들을 단계별로 점검 및 관리하는 것을 의미한다. 무엇보다 교사가 서술형 평가 및 수행 평가의 본질적인 목표달성을 위해 수행되어졌던 제반적인 활동에 대하여 시스템적으로 구조화 시키는 것을 의미한다. 일선 학교에서 현 평가 환류의 문제점은 교사측면, 학생측면, 교육이해 당사자 측면 등에서 문제점을 고찰할 수 있는데 이를 개선시키기 위한 제언을 하면 다음과 같다.

첫째, 정부는 교사들이 서술형 평가와 수행평가 결과를 종합·분석하여, 학생들의 강점과 약점을 파악하고 교수학습방법의 개선점을 도모할 수 있는 시스템 구비가 되어야 한다. 이를 통하여, 학생들의 재능과 학업성취 정도를 파악하고 학생의 꿈과 끼 실현을 위한 학습계획을 세울 수 있는 제도적 마련이 시급하다. 이런 제도적 마련은 교사에게 학생의 창의성 신장을 위한 효과적인 교육활동을 구상할 수 있게 하며, 누적적인 자료는 학생 발달과정을 이해할 수 있으며 학생의 지속적이고 안정적인 성장을 도모할 수 있을 것이다.

둘째, 정부는 교사가 학생에게 성취기준에 제시된 평가내용과 능력(이해, 적용, 분석)을 파악하여 강점과 약점을 고찰할 수 있고, 꿈과 끼 실현을 위한 필요영역을 제공해 줄 수 있는 시스템 마련이 시급하다. 이를 위하여 이원목적분류표를 더욱 구체적으로 살펴볼 수 있도록 하면서, 학생 스스로 무엇을 알고, 무엇을 할 수 있는지에 대한 구체적인 분석을 제공할 수 있도록 해야 한다. 이를 활용한다면, 무엇보다 학생의 학습 능력 신장에 필요한 정보를 누적적으로 살펴볼 수 있을 것이다. 나아가 학생입장에서는 스스로 문제점을 파악할 수 있으며, 자기 주도적인 학습이 가능하게 할 수 있을 것이다.

셋째, 정부가 만약 이런 일련의 시스템이 안정적으로 구비시킨다면, 학생들의 학력 향상, 공교육 정상화와 사교육 경감 등에 기여할 수 있을 것이다. 정부가 학력 평가관리 시스템을 적용하여

서술형·수행 평가를 실시하여, 학생들의 학력을 파악하고 분석하는 일련의 절차적 지식을 활용할 수 있으면, 서술형 평가와 수행평가 결과에 대한 신뢰도와 타당도가 높일 수 있고, 무엇보다 교육 이해 당사자들의 평가 전문성 신장에 도움이 될 수 있을 것이다. 무엇보다 정부는 안정적이고 지속적인 평가환류체계 구축을 위한 인적, 물적 여건 조성이 절대적으로 필요한 시점에 있다.

앞서 제시한 정책제언을 단기, 중기, 장기로 구분하여 제시한다면 다음 <표 IV-1>과 같다.

<표 IV-1> 시기별 정책제언

단기	<ul style="list-style-type: none"> · 평가문항의 제작을 원활하게 할 수 있도록 'Bloom의 평가유형'에 따른 평가문항의 모델화가 필요하다. · 이원목적분류표와 평가문항과의 관계 및 제작원리를 안내하고 성취기준·수준 및 이원목적분류표에 따른 평가문항 제작과 채점기준의 일관성과 정합성을 신장할 수 있는 매뉴얼과 교원연수가 필요하다. · 성취기준 및 수준과 이원목적분류표에 따른 분할점수 산출에 대한 체험적 실습연수가 필요하다. · 선다형 지필평가의 비중을 40%이하로 서술형 평가 및 수행평가를 각각 60%이상으로, 하고 지필평가를 학기당 1회로 한정하는 것이 필요하다. · 평가결과의 환류 일지 및 매뉴얼을 개발·제공하는 것이 필요하다. · 평가 실시 후, 평가결과에 대한 환류기간을 제공하는 것이 필요하다.
중기	<ul style="list-style-type: none"> · 서술형 및 수행 평가 취지에 맞는 교육과정이 개편시킬 필요가 있다. 나아가 장기적으로는 이러한 평가를 대입제도 등에 적용시킬 방안 모색도 필요하다. · 지속적인 교사의 평가전문성 제고를 위하여 교육(학생)평가전문성 인정제도(가칭)와 같은 제도적 장치가 마련되어야 한다. · 교사들이 학생별 장단점을 쉽게 파악할 수 있도록 서술형 평가와 수행평가 결과에 대하여 데이터베이스화한 평가결과 누적 시스템을 구축해야한다. · 교사들이 평가결과의 환류에 대한 내용을 기록할 수 있는 평가 환류 일지(가칭)를 나이스(NEIS) 상에서 활용할 수 있도록 시스템을 구축해야한다. · 교원 양성에서부터 교육평가의 중요성을 강조하고 교원 임용 시 '한국사 능력'과 같이 교육평가 능력에 대한 검증이 필요하다.
장기	<ul style="list-style-type: none"> · 학교의 교무행정 업무를 보조할 행정보조원 혹은 평가 및 교수학습을 보조할 수 있는 교수학습보조사의 채용 등과 같은 교사의 업무 부담을 경감시킬 수 있는 정책적 방안 마련이 절대적으로 필요하다. · 평가대상의 감축, 즉 학급당 학생 수의 감축이 필요하다. 구체적으로는 학급당 15명 내로 한정해야 평가 및 교수학습의 질을 보장할 수 있다. · 교원의 평가전문성 신장 및 평가환류를 위하여 주당수업시수를 감축할 수 있는 방안 에 대한 연구가 필요하다. · 교사의 평가권을 강화하고 자율성을 행정적으로나 법적으로 보장해야 한다. · 서술형 및 수행 평가 취지에 맞는 교육과정이 개편과 더불어 장기적으로는 이러한 교육과정 및 평가체제를 대입제도 등에 적용시킬 방안 모색도 필요하다. · 교육(학생)평가 분석사(가칭) 국가자격증과 같은 국가적 제도적 장치가 필요하다.

참고문헌

- 강신천(2014). 성취평가제에서의 이원목적분류표 작성 실태분석과 개선 방안 연구. 교육과정평가 연구, 17(2), 21~48.
- 강진구(2013). 고등학생들의 수학과 서술형 평가에 대한 인식과 학습 특성. 석사학위논문. 순천대학교 교육대학원.
- 강호수, 차정주, 하봉운(2014). 교사 전문성 개발 평가를 위한 Desimone 모형의 시사점 및 한계, 교육연구, 30(3), 134-154.
- 경기도교육청(2010). 2010 서술형 평가 문항의 실제.
- 고정화(2010). 평가 문항의 질 향상을 위한 문항 수정 유형 분석, 학교수학, 12(2), 113-136.
- 곽영순(2015). 미래 학교교육 변화 및 교육과정 재구성에 필요한 교사 전문성 탐색, 교과교육학연구, 19(1), 93-111.
- 권순달(2010). 비교과 교사의 평가 전문성 분석, 교육방법연구, 22(4), 219-236.
- 김경희 외(2006). 교사의 학생평가 전문성 기준 개발. 교육평가연구, 19(2), 89~112.
- 김경희, 김신영, 김성숙, 지은림, 반재천, 김수동(2006). 교사의 학생평가 전문성 기준 개발, 교육평가연구, 19(2), 89-112.
- 김명숙(2003). 국가수준 학업성취도 평가 연구의 주요 쟁점과 발전 방향. 교육평가연구, 16(1), 241-264.
- 김석우(2015). 교육평가의 이해. 서울: 학지사.
- 김선영(2014). 국어과 서술형 평가 문항과 학생 반응 분석 연구. 석사학위논문. 인하대학교 교육대학원.
- 김소라(2015). 고등학교 수학 I 서술형 평가에 관한 사례연구. 석사학위논문. 공주대학교 교육대학원.
- 김수동 외(2005). 교사의 학생평가 전문성 신장 연구(Ⅱ). 서울: 한국교육과정평가원.
- 김수동 외(2005). 영어 수업에서 학생평가를 잘 하려면. 서울: 한국교육과정평가원.
- 김수동, 이의갑, 김경희, 김선희, 박은아, 신명선, 김수진, 박가나, 서수현, 전영석 (2005). 교사의 학생평가 전문성 신장 연구(Ⅱ). 한국교육과정평가원 연구보고 RRE 2005-3.
- 김순남(2013). 창의인재 육성을 위한 학생평가 정책연구: 국제 사례를 중심으로. 한국교육개발원.
- 김순남(2013). 학생평가 방법 개선 연구. 현안보고 OR 2013-04. 서울: 한국교육개발원.
- 김신영(2002). 현장교사의 평가전문성 연구. 교육평가연구, 15(1), 67-85.
- 김신영(2007). 교사의 학생평가전문성과 중등교사 양성과정, 교육평가연구, 20(1), 1-16.
- 김신영(2014). 학생평가에 대한 교사의 인식과 실천, 교육평가연구, 27(1), 141-161.
- 김윤자(2015). 경기도 '상시평가'에 대한 초등교사의 인식 및 실태 분석. 석사학위논문. 경인교육대학교 교육전문대학원.
- 김재춘, 박소영, 손민호(2004). 학교생활기록부 전산처리 및 관리지침 개선방안 연구. 교육인적자

- 원부 교육정책연구 2004-일-03.
- 김정덕(2009). 루브릭에 대한 교사의 인식과 수행평가 실행과의 관계. 교육방법연구, 21(1), 159-173.
- 김정호, 박선미 (1998). 국가 교육과정에 근거한 평가기준 및 도구 개발 연구 : 고등학교 공통사회. 한국교육과정평가원.
- 김종수 외(2011). 서술형 평가 ROAD VIEW. 서울: 도서출판 박이정.
- 김지선, 김혁동, 이원재, 이동배(2014). 교사전문성을 위한 자기평가 준거 개발 연구, 한국교육교육연구, 31(4), 167-196.
- 김진아(2013). 서술형 평가 문항의 메타적 연구. 석사학위논문. 고려대학교 교육대학원.
- 남명호(1995). 수행평가의 타당성 연구. 박사학위논문. 고려대학교.
- 남명호(2006). 교사의 학생평가 전문성 신장 연구(III) (RRE 2006-5). 서울: 한국교육과정평가원.
- 박도순 외 24명(2007). 교육평가: 이해와 적용. 교육과학사.
- 박도순(2004). 대학입학전형제도의 바람직한 개선방향. 2004년도 교육혁신위원회·한국교육평가학회 공동개최 학술세미나 자료집. 공교육 정상화를 위한 대학입학선발제도와 평가체제의 패러다임 변화, 3-10.
- 박도순(2013). 학교 교육평가란?-평가의 철학과 개념을 중심으로. 충남교육, 제181호(봄).
- 박삼철(2013). 호주 IBP프로젝트가 한국교육에 주는 시사점. (현안보고 CP 2013-01-4)
- 박원택(2014). 대구광역시 고등학교 1학년 성취평가제 적용을 위한 서술형 평가 실태 조사. 석사학위논문. 경북대학교 교육대학원.
- 박인기 (1995). 언어기술 영역의 수행평가 방법. 초등교육연구 6, 49-68.
- 박정(2013). 초등교사의 학생평가전문성에 대한 자기평가 분석, 한국교육교육연구, 30(3), 357-376.
- 박혜인(2015). 과학과 서술형 평가에 대한 중·고등학교 교사와 학생의 인식 조사 및 실태 분석. 석사학위논문. 강원대학교 교육대학원.
- 배준형(2011). 핀란드의 교육체계를 통해 살펴본 우리나라 교육의 개선방안. 영남대학교 교육대학원.
- 배호순(2001). 단위학교 차원에서의 수행평가의 질 관리 방안 구축을 위한 기초연구. 교육평가연구, 14(2), 47-70.
- 백순근 (2006). 학생평가에 대한 교사들의 학생부 신뢰도 제고 방안. 교육혁신위원회, 한국교육평가학회 공동 개최 학술세미나 자료집. 교사의 학생평가 전문성과 학생부 신뢰도 제고 방안 탐색. 61-80.
- 백순근 편 (1998). 중학교 각 교과별 수행평가의 이론과 실제. 서울 : 원미사.
- 백순근(1995b). 교수-학습 평가를 위한 새로운 대안 : 수행평가를 중심으로. 초등교육연구 6, 3-20. 청주교육대학 초등교육연구소.
- 백순근(1998, 편집). 수행평가의 이론과 실제. 서울: 원미사.
- 백순근(2000). 교육평가의 개념에 대한 고찰: '교육적 가치'를 중심으로. 교육평가연구, 13(1), 1-10.

- 서울시교육청(2010). 창의력과 표현력을 키워주는 서술형 평가 장학 자료집.
- 성명진(2015). 국어과 서술형·논술형 평가에 대한 교사의 인식 및 평가 실태 조사. 석사학위논문. 아주대학교
- 성태제 외(2007). 최신 교육학개론. 서울: 학지사.
- 성태제(1999). “교육평가 방법의 변화와 결과타당도”. 교육학연구, 37(1), 197~218.
- 성태제(2010). 현대교육평가. 서울: 학지사.
- 성태제(2013). 현대교육평가. 서울: 학지사.
- 손원숙·신이나·배주현(2014). 초·중학교 교실평가 실제와 학생의 정의 및 인지적 성취 간의 관계: TIMSS 2011 한국, 싱가포르 및 핀란드. 교육평가연구 27(5).
- 손준중(2010). 교사 전문성 담론의 성격 분석 : 이명박 정부 교원정책을 중심으로, 교육행정학연구, 17(4), 91-119.
- 송인발, 박주호(2014). 교사의 전문성 개발 참여에 미친 학교 수준 변인의 영향, 한국교원교육연구, 31(4), 93-117.
- 송인수(2005). 교사별 평가의 당위성. (사) 참교육을 위한 전국학부모회·교육문화연구회 주최 심포지엄, 9-21.
- 양정훈(2014). 고등학교 수학과 수준별 서술형 평가에 대한 교사와 학생의 인식 조사. 석사학위논문. 한국교원대학교 교육대학원.
- 원효현(2002). 교수활동의 질 개선을 위한 교사 자기평가, 한국평가교육, 19(3), 147-168.
- 윤정일, 신효정(2006). 교사 전문성에 관한 교사, 학생, 학부모의 인식 연구, 한국교원교육연구, 23(2), 79-100.
- 윤지희(2005). 교사의 평가권에 대한 평가. (사) 참교육을 위한 전국학부모회·교육문화연구회 주최 심포지엄, 22-37.
- 이수희(2014). 중·고등학교 수학과 서술형 평가에 대한 교사와 학생의 인식 및 실태. 석사학위논문. 국민대학교 교육대학원
- 이양락, 이선경, 홍미영, 홍재식, 이미경 (1998). 국가 교육과정에 근거한 평가기준 및 도구 개발 연구: 고등학교 공통과학. 한국교육과정평가원.
- 이용관(2005). 교사별 평가의 조건과 여건. (사) 참교육을 위한 전국학부모회·교육문화연구회 주최 심포지엄, 38-44.
- 이윤미(2010). 북유럽에서의 교육평가와 선발: 한국교육에의 시사점.
- 이윤승(2015). 영어과 서술형 평가 확대가 고교 수업에 미치는 영향. 석사학위논문. 인하대학교 교육대학원.
- 이인제, 이범홍, 박정, 진재관, 김옥남, 서수현, 김신영(2004). 교사의 학생평가 전문성 신장 모형과 기준. 한국교육과정평가원. 연구보고 RRE 2004-5-2.
- 이지현(2013). 교육과정 부합(Alignment)에 의거한 서술형 평가 문항 분석. 석사학위논문. 서강대학교 교육대학원.
- 장선자(2014). 고등학생들의 쓰기 능력 향상을 위한 서술형 평가의 활용 방안 연구. 석사학위논문.

- 문. 창원대학교.
- 전보람(2013). 서술형 평가와 채점방법에 대한 중·고등학교 수학 교사들의 인식. 석사학위논문. 이화여자대학교 교육대학원.
- 정구향, 김성숙, 양길석, 전은화 (1998). 대학입학 면접전형자료 개발과 활용방안. 한국교육과정평가원.
- 정연주(2013). 수준별 서술형 평가에 대한 학생들의 인식분석과 효율적인 적용방안에 대한 연구. 석사학위논문. 고려대학교 교육대학원
- 정일용 · 김든(2012). 핀란드 교육의 특징과 교육혁신의 성공요인. 세계 교육정책 인포메이션 제 11호.
- 정홍명(2014). 중고등학교 수행평가 실태분석을 통한 총괄평가 작성법 개선방안. 석사학위논문. 단국대학교.
- 지스카와 마유 · 지스타와 코모토(2009), 송태욱 역. 핀란드 공부법. 서울: 문학동네.
- 한국교육과정평가원 (1999). 고등학교 공통 필수 10개 과목에 대한 평가 방법 개선 연구 (총 10 권). 서울시교육청 수탁과제 보고서.
- 한국교육과정평가원(2005). 교사의 학생평가 전문성 신장연구 (Ⅱ) (연구보고 RRE 2005-3). 서울: 한국교육과정평가원.
- 한국교육과정평가원(2010). 수행평가 현장 적용 개선 방안. 연구자료 ORM 2010-4-2. 서울: 한국교육과정평가원.
- 한국교육과정평가원(2010). 외국의 역량기반 교육과정 현장 적용 사례 연구:호주와 뉴질랜드, 캐나다, 영국의 사례를 중심으로. 연구보고. RRC 2010-2. 서울: 한국교육과정평가원.
- 한국교육과정평가원(2013). 2013 교육과정·교육평가 국제동향 연구-뉴질랜드·호주-(연구자료 ORM 2013-97-2). 서울: 한국교육과정평가원.
- 한국교육연구네트워크 총서기획팀(2010). 핀란드 교육혁명. 서울: 도서출판 살림터.
- 한은주(2015). 중·고등학교 중국어과 서술형 평가 실태 및 문항 분석. 석사학위논문. 한국외국어대학교 교육대학원
- 후쿠타 세이지(2009), 박재원 역. 핀란드 교실혁명. 서울: 비아북.
- Abedi, J. (2007). English Language Proficiency Assessment & Accountability under NCLB Title III: An Overview. In J. Abedi (Ed.), English Language
- American Federation of Teacher, National Council on Measurement in Education, & National Education Association(AFT, NCME, & NEA) (1990). Standards for teacher competence in educational assessment of students. Educational Measurement: Issues and practice. 9(4), 30-32.
- Atkin, J.M., Black, P., & Coffey, J.(2001). Classroom assessment and the National Science Education Standards. Washington, DC: National Academy Press.
- Attali, Y., & Burstein(2006), J., Automated essay scoring with e-rater v. 2.0, Journal of Technology Learning and Assessment 4(3), 1-30.

Basic Education Act 628/1998.

California Department of Education, California standards tests technical report Spring 2010 administration, 2011. (<http://www.cde.ca.gov/>)

California Department of Education, Teacher guide for the 2005 California writing standards test in grade seven, 2006. (<http://www.cde.ca.gov/>)

Council of Europe Press(1996). Secondary Education in Finland.

Finnish National Board of Education(2004). NATIONAL CORE CURRICULUM FOR BASIC EDUCATION 2004.

Finnish National Board of Education(2004). NATIONAL CORE CURRICULUM FOR GENERAL UPPER SECONDARY

Fisher, D. & Frey, N.(2007). Checking for understanding: Formative assessment techniques for your classroom. Alexandria, VA: ASCD.

Haberman, S., Use of e-rater in scoring of the TOEFL iBT writing test. Research Report, Princeton, NJ: ETS, 2011.

Jari Lavonen(2007). National science education standards and assessment in Finland, Making it Comparable Standards in Science Education, Chapter 6. Germany Waxmann Publishing Co.

Johanson, G. & Motlomelo, S.(1998). An Item Format Continuum for Classroom Assessment. ERIC Document ED 420-692.

McMillan (2004). Classroom assessment principle and practice for effective instruction. Boston : Allyn & Bacon.

McMillan, J.H.(2007). Classroom assessment: Principles and practice for effective standards-based instruction(4th ed.). Boston, MA: Pearson Education, Inc.

Meyer(1992). What's the difference between authentic and performance assessment? Educational Leadership.

Miller, M. D., Linn, R. L., Gronlund, N. E.(2009). Measurement and Assessment in Teaching (Tenthed). Upper Saddle River, NJ: Pearson.

No Child Left Behind Act of 2001(NCLB), Pub. L. No. 107-110, § 115 Stat. 1425, U.S. Department of Education, 2002 (<http://www.ed.gov>)

Popham, W. J(1995). Classroom Assessment: What teachers need to know. Boston: Allyn and Bacon.

Popham, W.J.(2008). Transformative assessment. Alexandria, VA: ASCD.

Proficiency Assessment in the Nation: Current Status and Future Practice (pp. 3 -10). Davis: University of California.

Stiggins, R.(1994). Student-Centered Classroom Assessment. New York: Merrill.

The Finnish Education Evaluation Council(2005). New Directions in Educational Evaluation.

부록

[부록1] 서술형 및 수행평가에 대한 중·고등학교 교사의 인식 설문지

서술형 및 수행평가에 대한 중·고등학교 교사의 인식 설문지

안녕하십니까?

이 설문지는 한국교육개발원에서 시행하는 ‘교육현장지원연구’의 일환으로 서술형 평가에 대한 선생님들의 고견을 듣고자 작성되었습니다. 학생 교육으로 바쁘신 중에도 본 조사 연구를 위하여 귀중한 시간을 할애하여 주신 선생님들께 깊은 감사드립니다. 선생님들께서 응답하신 결과는 ‘중등학교 교사의 평가 전문성 제고 방안’ 연구를 수행하기 위한 매우 유익한 자료가 될 수 있을 것이라고 생각하며, 아울러 본 연구 이외의 목적으로는 사용하지 않을 것임을 약속드립니다.

선생님들의 솔직하고 성의 있는 답변을 부탁드립니다. 감사합니다.

부산대학교 사범대학 교육학과
김 석 우 올림

※ 다음은 서술형 및 수행평가와 관련하여 선생님의 개인적 의견에 대한 질문입니다.

1. 서술형 및 수행평가의 실시 후 채점 결과에 대해 학생 또는 학부모가 이의를 제기한 적이 있습니까? 있다면 그 이유는 무엇입니까?

- ① 문제오류 ② 채점의 오류 ③ 유사답안 인정여부 ④ 부분점수 인정여부
⑤ 이의 제기 받은 적 없음

2. 서술형 및 수행평가에 관련한 연수 또는 교육을 받으신다면, 우선순위는 무엇입니까?

(순위를 고려하여 2 가지만)

(1순위: , 2순위:)

- ① 서술형 및 수행평가의 문항 제작
② 서술형 및 수행평가의 결과 분석 및 활용
③ 서술형 및 수행평가의 효율적인 채점 방법
④ 서술형 및 수행평가에 적합한 수업모형 개발
⑤ 서술형 및 수행평가의 신뢰성 확보를 위한 채점기준안 작성 방법

⑥ 서술형 및 수행평가를 위한 수업모형 및 교육과정 개발

4	문항 개발부터 평가에 이르기까지 서술형 평가를 위해 할애할 시간이 충분하다.					
5	현행 교육과정의 성취기준에 따라 수업을 진행한다.					
6	교육과정에서 제시하는 성취기준 코드를 이해하고 활용한다.					
7	성취기준에 따라 서술형 및 수행평가 문항의 유형을 적절하게 선택한다.					
8	서술형 및 수행평가를 시행함에 있어서 학생의 성별, 가정환경 등의 요인에 따른 편파성을 최소화하려고 노력한다.					
9	서술형 및 수행평가 문항 개발 과정에서 학생의 수준을 고려하여 성취기준과 성취수준을 재설정한다.					
10	서술형 및 수행평가 문항 개발 과정에서 교수·학습 개선을 위한 정보습득에 대하여 고려한다.					
11	서술형 및 수행평가 문항 개발 과정에서 학생의 학습능력 성장에 대한 정보습득을 고려한다.					
12	서술형 및 수행평가의 문항 개발의 절차를 이해하고 그에 따라 문항을 개발한다.					
13	이원목적분류표에 근거하여 서술형 및 수행평가 문항을 개발한다.					
14	문항 개발 과정에서 학생의 능력 집단별 난이도를 예측한다.					
15	문항 개발 과정에서 서술형 평가 문항의 변별도에 대하여 예측한다.					
16	문항 개발 과정에서 성취수준별 기준점(분할 점수)을 산출한다.					
17	문항 개발 과정에서 서술형 평가도구의 타당도 및 신뢰도를 검토한다.					
18	성취기준 및 성취수준에 따른 채점기준을 설정한다.					
19	이원목적분류표의 내용영역 및 행동영역에 따른 채점기준을 설정한다.					
20	학생들에게 서술형 평가의 채점기준을 정확하게 안내한다.					
21	학생의 응답 반응을 참조하여 서술형 평가의 채점기준을 수정·보완한다.					

22	서술형 및 수행평가 결과에 대하여 양적·질적으로 통계 분석을 한다.					
23	서술형 및 수행평가의 목적에 따른 평가결과를 해석한다.					
24	서술형 및 수행평가 결과에 따라 학생의 강점과 약점을 해석한다.					
25	성취기준 및 성취수준에 근거하여 평가결과를 해석한다.					
26	서술형 및 수행평가 결과에 기초하여 수업계획을 수정·개선한다.					
27	서술형 및 수행평가 결과에 기초하여 학생들의 약점에 대한 개선안을 마련한다.					
28	서술형 및 수행평가 결과를 성취기준 및 성취수준을 정교화하는데 활용한다.					
29	서술형 및 수행평가 결과를 학생 개인의 성장 정도 확인 등을 위하여 학생수준에서 활용한다.					
30	서술형 및 수행평가 결과를 학급 및 학년별 교수 방법 개선 등을 위하여 학급 및 학년수준에서 활용한다.					
31	서술형 및 수행평가 결과를 학교교육 계획 수립의 기초 자료로 활용한다.					

※ 다음은 선생님의 개인적 사항에 대한 질문입니다.

1. 선생님께서 어느 학교 급에서 근무하십니까?

① 중학교 ② 고등학교

2. 선생님께서 담당하시는 과목은 무엇입니까?

① 국어 ② 영어(외국어) ③ 수학 ④ 과학과 ⑤ 사회과 ⑥ 기타

3. 선생님의 교육 경력은?

① 1년 미만 ② 1~6년 ③ 6~11년 ④ 11~16년 ⑤ 16~21년 ⑥ 21년 이상

[부록2] 서술형 및 수행평가에 대한 중·고등학교 교사의 인터뷰 질문지

서술형평가 및 수행평가에 대한 중·고등학교 교사의
심층면접

안녕하십니까?

이 질문지는 한국교육개발원에서 시행하는 ‘교육현장지원연구’의 일환으로 서술형평가 및 수행평가에 대한 선생님들의 고견을 듣고자 작성되었습니다. 학생 교육으로 바쁘신 중에도 본 조사 연구를 위하여 귀중한 시간을 할애하여 주신 선생님들께 깊은 감사드립니다. 선생님들께서 응답하신 결과는 ‘중등학교 교사의 평가 전문성 제고 방안’ 연구를 수행하기 위한 매우 유익한 자료가 될 수 있을 것이라고 생각하며, 아울러 본 연구 이외의 목적으로는 사용하지 않을 것임을 약속드립니다.

선생님들의 솔직하고 성의 있는 답변을 부탁드립니다. 감사합니다.

부산대학교 사범대학 교육학과
김 석 우 올림

* 다음 각 문항의 해당란에 V표를 하세요.

성별	남(), 여()
나이	()세
담당학년	1학년 (), 2학년 (), 3학년 ()
근무 학교급	중학교 (), 고등학교 ()
담당교과	국어() 영어(외국어)() 수학() 사회() 과학() 기타()

1) 선생님께서 몸담고 있는 학교 또는 교과 교사들의 ‘평가 전문성’에 점수를 매긴다면 6점 만점에 몇 점 정도라고 생각하십니까?

번호	1점	2점	3점	4점	5점	6점
· 우리학교 교사들은 평가에 대해 전문적 지식과 기능을 가지고 능숙하게 활용한다.						

⇒ 교사들의 평가전문성 요인에는 교육과정의 이해, 평가환경, 평가윤리, 평가목적과 목표의 설정, 평가유형과 방식의 이해 및 선정, 평가도구 및 채점기준의 제작과 설정, 평가 결과의 분석과 해석, 평가환류 등이 있습니다.

⇒ 교사들의 평가전문성이 양호하다면, 그 원인으로 생각하시는 것을 3가지 정도 제시하신다면 어떤 것들입니까?

⇒ 교사들의 평가전문성이 양호하지 않다면, 그 원인은 어떤 것들입니까? 이를 개선하기 위한 방안을 3가지 정도 제시해 주신다면 어떠한 것들이 있을까요(자기평가와 연수 이외)?

2) 선생님께서 몸담고 있는 학교 또는 교과 교사들이 서술형평가 및 수행평가에 얼마나 우호적인지에 대하여 점수를 매긴다면 6점 만점에 몇 점 정도라고 생각하십니까?

번호	1점	2점	3점	4점	5점	6점
· 우리학교 교사들은 서술형평가 및 수행평가에 우호적이다.						

⇒ 교사들이 서술형평가 및 수행평가에 우호적이라면, 그 이유는 무엇이라고 생각하십니까? 혹은 비우호적이라면 그 원인은 무엇이며, 분위기를 개선할 수 있는 방안에 대하여 3가지 정도 제시해 주시겠습니까?

3) 선생님께서 몸담고 있는 학교의 학생들이 서술형평가 및 수행평가를 얼마나 신뢰하는지에 대하여 점수를 매긴다면 6점 만점에 몇 점 정도라고 생각하십니까?

번호	1점	2점	3점	4점	5점	6점
· 우리학교 학생들은 서술형평가 및 수행평가를 신뢰한다.						

⇒ 학생들이 서술형평가 및 수행평가를 신뢰한다면, 그 이유는 무엇이라고 생각하십니까? 혹은 신뢰하지 않는다면 그 이유는 무엇이며, 이를 개선할 수 있는 방안에 대하여 3가지 정도 제시해 주시겠습니까?

- 4) 선생님께서 몸담고 있는 학교의 학부모들이 서술형평가 및 수행평가를 얼마나 신뢰하는지에 대하여 점수를 매긴다면 6점 만점에 몇 점 정도라고 생각하십니까?

번호	1점	2점	3점	4점	5점	6점
· 우리학교 학부모들은 서술형평가 및 수행평가를 신뢰한다.						

- ⇒ 학부모들이 서술형평가 및 수행평가를 신뢰한다면, 그 이유는 무엇이라고 생각하십니까? 혹은 신뢰하지 않는다면 그 이유는 무엇이며, 이를 개선할 수 있는 방안에 대하여 3가지 정도 제시해 주시겠습니까?

- 5) 선생님께서 몸담고 있는 학교 또는 교과 교사들이 2009 개정교육과정의 기본 취지에 따라서 서술형평가 및 수행평가를 실시하고 있는지에 대하여 점수를 매긴다면 6점 만점에 몇 점 정도라고 생각하십니까?

번호	1점	2점	3점	4점	5점	6점
· 우리학교 교사들은 2009 개정교육과정의 기본 취지에 따라서 서술형평가 및 수행평가를 실시한다.						

- ⇒ 2009 개정교육과정의 기본 취지에 따라 서술형평가 및 수행평가를 잘 실시되고 있다면, 그 이유는 무엇이고 어떤 방식으로 실시되고 있나요? 혹은 그렇지 않다면, 그 이유는 무엇이며, 이를 개선할 수 있는 방안에 대하여 3가지 정도 제시해 주시겠습니까?

- 5) 선생님께서 몸담고 있는 학교 또는 교과 교사들이 성취평가제에 따라서 서술형평가 및 수행평가를 실시하고 있는지에 대하여 점수를 매긴다면 6점 만점에 몇 점 정도라고 생각하십니까?

번호	1점	2점	3점	4점	5점	6점
· 우리학교 교사들은 성취평가제에 대하여 이해하고 있다.						

- ⇒ 성취평가제에 따라 서술형평가 및 수행평가를 잘 실시되고 있다면, 그 이유는 무엇이고 어떤 방식으로 실시되고 있나요? 혹은 그렇지 않다면, 그 이유는 무엇이며, 이를 개선할 수 있는 방안에 대하여 3가지 정도 제시해 주시겠습니까?

6-1) 선생님께서 몸담고 있는 학교 또는 교과 교사들의 수업방식은 서술형평가에 적합한가요?

번호	1점	2점	3점	4점	5점	6점
· 우리학교 교사들의 교사들의 수업방식은 서술형평가에 적합하다.						

⇒ 서술형평가에 적합한 수업을 진행하고 있다면 어떤 수업방식인가요? 혹은 그렇지 않다면, 그 이유는 무엇이며, 이를 개선할 수 있는 방안에 대하여 3가지 정도 제시해 주시겠습니까?

6-2) 선생님께서 몸담고 있는 학교 또는 교과 교사들의 수업방식은 수행평가에 적합한가요?

번호	1점	2점	3점	4점	5점	6점
· 우리학교 교사들의 교사들의 수업방식은 수행평가에 적합하다.						

⇒ 수행평가에 적합한 수업을 진행하고 있다면 어떤 수업방식인가요? 혹은 그렇지 않다면, 그 이유는 무엇이며, 이를 개선할 수 있는 방안에 대하여 3가지 정도 제시해 주시겠습니까?

7) 선생님께서 몸담고 있는 학교 또는 교과 교사들이 교과협의회를 통하여 서술형 평가 문항 및 수행평가의 실시에 대한 충분한 논의가 이루어지는 정도에 대하여 점수를 매긴다면 6점 만점에 몇 점 정도라고 생각하십니까?

번호	1점	2점	3점	4점	5점	6점
· 우리학교 교사들은 교과협의회를 통하여 서술형 평가 문항 및 수행평가의 실시에 대한 충분히 논의한다.						

⇒ 교과협의회를 통하여 서술형 평가 문항 및 수행평가의 실시에 대한 충분한 논의가 이루어진다면 그 이유는 무엇이고 어떤 방식으로 이루어지나요? 혹은 그렇지 않다면, 그 이유는 무엇이며, 이를 개선할 수 있는 방안에 대하여 3가지 정도 제시해 주시겠습니까?

8) 선생님께서 몸담고 있는 학교에서 서술형평가 및 수행평가 우수교사에게 다양한 인센티브를 제공하는 정도에 대하여 점수를 매긴다면 6점 만점에 몇 점 정도라고 생각하십니까?

번호	1점	2점	3점	4점	5점	6점
· 우리학교에서는 서술형평가 및 수행평가 우수교사에게 다양한 인센티브를 제공한다.						

⇒ 인센티브를 제공한다면 이유는 무엇이고 어떤 방식으로 제공되나요? 혹은 그렇지 않다면, 그 이유는 무엇이며, 이를 개선할 수 있는 방안에 대하여 3가지 정도 제시해 주시겠습니까?

9) 선생님께서 몸담고 있는 학교에서 서술형평가를 지필평가와 따로 실시하는 정도에 대하여 점수를 매긴다면 6점 만점에 몇 점 정도라고 생각하십니까?

번호	1점	2점	3점	4점	5점	6점
· 우리학교에서는 서술형평가를 지필평가와 따로 실시하고 있다.						

⇒ 따로 실시한다면 그 이유는 무엇이고 어떤 방식으로 이루어지나요? 혹은 그렇지 않다면, 그 이유는 무엇이며, 이를 개선할 수 있는 방안에 대하여 3가지 정도 제시해 주시겠습니까?

10) 선생님께서 몸담고 있는 학교에서는 교사들이 서술형평가 및 수행평가 업무(출제 및 채점, 시행 등)에 전념할 수 있는 충분한 시간이 주어집니까?

번호	1점	2점	3점	4점	5점	6점
· 우리학교에서는 교사들이 서술형평가 및 수행평가 업무(출제 및 채점, 시행 등)에 전념할 수 있는 충분한 시간이 주어진다.						

⇒ 시간이 충분히 주어진다면, 그 이유는 무엇이고, 어떤 방식으로 시간이 주어지나요? 혹은 그렇지 않다면 그 이유는 무엇이고, 이를 개선할 수 있는 방안에 대하여 3가지 정도 제시해 주시겠습니까?

11-1) 선생님께서 몸담고 있는 학교 또는 교과 교사들의 '서술형평가 전문성'에 점수를 매긴다면 6점 만점에 몇 점 정도라고 생각하십니까?

번호	1점	2점	3점	4점	5점	6점
· 우리학교 교사들은 서술형평가에 대해 전문적 지식과 기능을 가지고 있고 이를 능숙하게 활용한다.						

- ⇒ 교사들의 '서술형평가 전문성'이 양호하다면, 그 원인으로 생각하시는 것을 3가지 정도 제시하신다면 어떤 것들입니까?
 ⇒ 교사들의 '서술형평가 전문성'이 양호하지 않다면, 그 원인은 무엇이라고 생각되십니까? 이를 개선하기 위한 방안을 3가지 정도 제시해 주신다면 어떠한 것들이 있을까요(자기평가와 연수 이외)?

11-2) 선생님께서 몸담고 있는 학교 또는 교과 교사들의 '수행평가 전문성'에 점수를 매긴다면 6점 만점에 몇 점 정도라고 생각하십니까?

번호	1점	2점	3점	4점	5점	6점
· 우리학교 교사들은 수행평가에 대해 전문적 지식과 기능을 가지고 있고 이를 능숙하게 활용한다.						

- ⇒ 교사들의 '수행평가 전문성'이 양호하다면, 그 원인으로 생각하시는 것을 3가지 정도 제시하신다면 어떤 것들입니까?
 ⇒ 교사들의 '수행평가 전문성'이 양호하지 않다면, 그 원인은 무엇이라고 생각되십니까? 이를 개선하기 위한 방안을 3가지 정도 제시해 주신다면 어떠한 것들이 있을까요(자기평가와 연수 이외)?

12) 선생님께서 몸담고 있는 학교 또는 교과 교사들이 학생의 수준에 따라 성취기준을 수정·보완하는 정도에 대하여 점수를 매긴다면 6점 만점에 몇 점 정도라고 생각하십니까?

번호	1점	2점	3점	4점	5점	6점
· 우리학교 교사들은 학생의 수준에 따라 성취기준을 수정·보완한다.						

- ⇒ 학생의 수준에 따라 성취기준을 수정·보완한다면 그 이유는 무엇이고 어떤 방식으로 수정·보완하나요? 혹은 그렇지 않다면 그 이유는 무엇이고, 이를 개선할 수 있는 방안에 대하여 3가지 정도 제시해 주시겠습니까?

13) 선생님께서 몸담고 있는 학교 또는 교과 교사들이 서술형평가 및 수행평가를 위하여 성취기준을 재구성하는 정도에 대하여 점수를 매긴다면 6점 만점에 몇 점 정도라고 생각하십니까?

번호	1점	2점	3점	4점	5점	6점
· 우리학교 교사들은 서술형평가 및 수행평가를 위하여 성취기준을 재구성한다.						

⇒ 교사들이 서술형평가 및 수행평가를 위하여 성취기준을 재구성할 수 있다면, 그 이유는 무엇이고, 어떤 방식으로 재구성하나요? 혹은 그렇지 않다면 그 이유는 무엇이고, 이를 개선할 수 있는 방안에 대하여 3가지 정도 제시해 주시겠습니까?

14-1) 선생님께서 몸담고 있는 학교 또는 교과 교사들이 성취기준에 부합하는 평가 유형(지필평가, 서술형평가, 수행평가)을 설정하는 정도에 대하여 점수를 매긴다면 6점 만점에 몇 점 정도라고 생각하십니까?

번호	1점	2점	3점	4점	5점	6점
· 우리학교 교사들은 성취기준에 부합하는 평가유형을 설정한다.						

⇒ 교사들이 성취기준에 부합하는 평가유형을 설정한다면, 어떤 방식으로 설정하십니까? 혹은 그렇지 않다면, 그 이유는 무엇이며, 이를 개선할 수 있는 방안에 대하여 3가지 정도 제시해 주시겠습니까?

14-2) 선생님께서 몸담고 있는 학교 또는 교과 교사들이 서술형평가 및 수행평가의 비율을 성취기준에 따라 설정하는 정도에 대하여 점수를 매긴다면 6점 만점에 몇 점 정도라고 생각하십니까?

번호	1점	2점	3점	4점	5점	6점
· 우리학교 교사들은 서술형평가 및 수행평가의 비율을 성취기준에 따라 설정한다.						

⇒ 교사들이 서술형평가 및 수행평가의 비율을 성취기준에 따라 설정한다면, 어떤 방식으로 비율을 설정하십니까? 혹은 그렇지 않다면, 그 이유는 무엇이며, 이를 개선할 수 있는 방안에 대하여 3가지 정도 제시해 주시겠습니까?

15) 선생님께서 몸담고 있는 학교 또는 교과 교사들이 성취기준 코드체계를 이해하고 활용하는 정도에 대하여 점수를 매긴다면 6점 만점에 몇 점 정도라고 생각하십니까?

번호	1점	2점	3점	4점	5점	6점
· 우리학교 교사들은 성취기준 코드체계를 이해하고 활용한다.						

⇒ 교사들이 성취기준 코드체계를 이해하고 활용하지 않는다면 그 이유는 무엇이며, 이를 개선할 수 있는 방안에 대하여 3가지 정도 제시해 주시겠습니까?

16-1) 선생님께서 몸담고 있는 학교 또는 교과 교사들이 성취기준에 적합한 서술형평가(문항) 유형을 선택하고 활용하는 정도에 대하여 점수를 매긴다면 6점 만점에 몇 점 정도라고 생각하십니까?

번호	1점	2점	3점	4점	5점	6점
· 우리학교 교사들은 성취기준에 적합한 서술형평가(문항) 유형을 선택하고 활용한다.						

⇒ 성취기준에 적합한 서술형평가(문항) 유형을 선택하고 활용한다면, 그 이유는 무엇이고, 선생님이 알고 계시는 서술형평가(문제) 유형에 대해 설명해 주시겠습니까? 혹은 그렇지 않다면, 그 이유는 무엇이며, 이를 개선할 수 있는 방안에 대하여 3가지 정도 제시해 주시겠습니까?

16-2) 선생님께서 몸담고 있는 학교 또는 교과 교사들이 성취기준에 적합한 수행평가 유형을 선택하고 활용하는 정도에 대하여 점수를 매긴다면 6점 만점에 몇 점 정도라고 생각하십니까?

번호	1점	2점	3점	4점	5점	6점
· 우리학교 교사들은 성취기준에 적합한 수행평가 유형을 선택하고 활용한다.						

⇒ 성취기준에 적합한 수행평가 유형을 선택하고 활용한다면, 그 이유는 무엇이고, 선생님이 알고 계시는 수행평가 유형에 대해 설명해 주시겠습니까? 혹은 그렇지 않다면, 그 이유는 무엇이며, 이를 개선할 수 있는 방안에 대하여 3가지 정도 제시해 주시겠습니까?

17) 선생님께서 몸담고 있는 학교 또는 교과 교사들이 서술형평가 및 수행평가를 시행함에 있어서 학생의 성별, 경제적, 문화적 요인에 따른 편파성을 최소화하려는 노력에 대하여 점수를 매긴다면 6점 만점에 몇 점 정도라고 생각하십니까?

번호	1점	2점	3점	4점	5점	6점
· 우리학교 교사들은 서술형평가 및 수행평가를 시행함에 있어서 학생의 성별, 경제적, 문화적 요인에 따른 편파성을 최소화하려고 노력한다.						

⇒ 서술형평가 및 수행평가를 시행함에 있어서 학생의 성별, 경제적, 문화적 요인에 따른 편파성을 최소화하려는 노력한다면, 그 이유는 무엇이고, 어떤 방식으로 노력하십니까? 혹은 그렇지 않다면, 그 이유는 무엇이며, 이를 개선할 수 있는 방안에 대하여 3가지 정도 제시해 주시겠습니까?

18-1) 선생님께서 몸담고 있는 학교 또는 교과 교사들의 서술형 '평가문항 제작의 전문성'에 점수를 매긴다면 6점 만점에 몇 점 정도라고 생각하십니까?

번호	1점	2점	3점	4점	5점	6점
· 우리학교 교사들은 서술형평가문항의 제작에 대해 전문적 지식과 기능을 가지고 있고 이를 능숙하게 활용한다.						

⇒ 교사들의 서술형 '평가문항 제작의 전문성'이 양호하다면, 그 원인은 무엇이라고 생각되십니까?

⇒ 교사들의 서술형 '평가문항 제작의 전문성'이 양호하지 않다면, 그 원인은 무엇이라고 생각되십니까? 이를 개선하기 위한 방안을 3가지 정도 제시해 주신다면 어떠한 것들이 있을까요?

18-2) 선생님께서 몸담고 있는 학교 또는 교과 교사들의 '수행평가 실행의 전문성'에 점수를 매긴다면 6점 만점에 몇 점 정도라고 생각하십니까?

번호	1점	2점	3점	4점	5점	6점
· 우리학교 교사들은 수행평가를 실행함에 있어서 전문적 지식과 기능을 가지고 있고 이를 능숙하게 활용한다.						

⇒ 교사들의 '수행평가문항 실행의 전문성'이 양호하다면, 그 원인은 무엇이라고 생각되십니까?

⇒ 교사들의 '수행평가문항 실행의 전문성'이 양호하지 않다면, 그 원인은 무엇이라고 생각되십니까? 이를 개선하기 위한 방안을 3가지 정도 제시해 주신다면 어떠한 것들이 있을까요?

19-1) 선생님께서 몸담고 있는 학교 또는 교과 교사들이 서술형 '평가문항 제작의 절차'를 숙지하고 그에 따라 평가문항을 제작하는 정도에 점수를 매긴다면 6점 만점에 몇 점 정도라고 생각하십니까?

번호	1점	2점	3점	4점	5점	6점
· 우리학교 교사들은 '제작 절차'를 숙지하고 그에 따라 평가문항을 제작한다.						

⇒ 선생님은 평가문항을 제작하는 과정에서 어떤 '제작의 절차'를 따르고 있습니까? 그 절차를 간략하게 기술하신다면 어떠합니까?

⇒ 교사들이 서술형 '평가문항 제작의 절차'를 숙지하고 그에 따라 평가문항을 제작하는 정도가 양호하다면, 그 이유는 무엇이라고 생각되십니까?

⇒ 교사들이 서술형 '평가문항 제작의 절차'를 숙지하고 그에 따라 평가문항을 제작하는 정도가 양호하지 않다면, 그 이유는 무엇이라고 생각되십니까? 이를 개선하기 위한 방안을 3가지 정도 제시해 주신다면 어떠한 것들이 있을까요?

19-2) 선생님께서 몸담고 있는 학교 또는 교과 교사들이 수행형 '평가문항 제작의 절차'를 숙지하고 그에 따라 평가문항을 제작하는 정도에 점수를 매긴다면 6점 만점에 몇 점 정도라고 생각하십니까?

번호	1점	2점	3점	4점	5점	6점
· 우리학교 교사들은 '제작 절차'를 숙지하고 그에 따라 평가문항을 제작한다.						

⇒ 선생님은 평가문항을 제작하는 과정에서 어떤 '제작의 절차'를 따르고 있습니까? 그 절차를 간략하게 기술하신다면 어떠합니까?

⇒ 교사들이 서술형 '평가문항 제작의 절차'를 숙지하고 그에 따라 평가문항을 제작하는 정도가 양호하다면, 그 이유는 무엇이라고 생각되십니까?

⇒ 교사들이 서술형 '평가문항 제작의 절차'를 숙지하고 그에 따라 평가문항을 제작하는 정도가 양호하지 않다면, 그 이유는 무엇이라고 생각되십니까? 이를 개선하기 위한 방안을 3가지 정도 제시해 주신다면 어떠한 것들이 있을까요?

20) 선생님께서 몸담고 있는 학교 또는 교과 교사들의 서술형 '평가문항제작'에 있어서 '성취기준과 성취수준'을 적용하는 정도에 점수를 매긴다면 6점 만점에 몇 점 정도라고 생각하십니까?

번호	1점	2점	3점	4점	5점	6점
· 우리학교 교사들은 서술형 평가문항을 제작함에 있어서 '성취기준과 성취수준'을 우선적으로 적용한다.						

⇒ 교사들의 서술형 평가문항제작에 있어서 '성취기준과 성취수준'을 우선적으로 적용하기 어렵다면, 그 원인은 무엇이라고 생각되십니까? 이를 개선하기 위한 방안을 3가지 정도 제시해 주신다면 어떠한 것들이 있을까요(자기평가와 연수 이외)?

⇒ 선생님은 평가문항을 제작하는 과정에서 어떤 기준을 또는 무엇을 참고로 제작하고 있습니까? 그 기준 또는 참고에 대해 간략하게 기술하신다면 어떠합니까?

21) 선생님께서 몸담고 있는 학교 또는 교과 교사들의 '수행평가 실행'에 있어서 '성취기준과 성취수준'을 적용하는 정도에 점수를 매긴다면 6점 만점에 몇 점 정도라고 생각하십니까?

번호	1점	2점	3점	4점	5점	6점
· 우리학교 교사들은 수행평가를 실행함에 있어서 '성취수준과 성취기준'을 우선적으로 적용한다.						

⇒ 교사들의 '수행평가 실행'에 있어서 '성취기준과 성취수준'을 우선적으로 적용하기 어렵다면, 그 원인은 무엇이라고 생각되십니까? 이를 개선하기 위한 방안을 3가지 정도 제시해 주신다면 어떠한 것들이 있을까요?

⇒ 선생님은 평가문항을 제작하는 과정에서 어떤 기준을 또는 무엇을 참고로 제작하고 있습니까? 그 기준 또는 참고에 대해 간략하게 기술하신다면 어떠합니까?

22) 선생님께서 몸담고 있는 학교 또는 교과 교사들이 서술형 평가문항 및 수행형 평가문항을 ‘이원목적분류표’에 따라 제작·활용하는 정도에 점수를 매긴다면 6점 만점에 몇 점 정도라고 생각하십니까?

번호	1점	2점	3점	4점	5점	6점
· 우리학교 교사들은 ‘이원목적분류표’에 따라 서술형 평가문항 및 수행형 평가문항을 제작·활용한다.						

⇒ 교사들이 ‘이원목적분류표’에 따라서 서술형 평가문항 및 수행형 평가문항을 제작·활용하기 어렵다면, 그 원인은 무엇이라고 생각되십니까? 이를 개선하기 위한 방안을 3가지 정도 제시해 주신다면 어떠한 것들이 있을까요?

⇒ 선생님은 서술형 평가문항 및 수행형 평가문항을 제작·활용하는 과정에서 ‘이원목적분류표’를 어떻게 활용하고 있습니까? 그 활용 방법을 간략하게 기술하신다면 어떠합니까?

23) 선생님께서 몸담고 있는 학교 또는 교과 교사들의 서술형평가 및 수행평가문항의 ‘채점기준 또는 루브릭’을 이원목적 분류표에 따라 작성하는 정도에 점수를 매긴다면 6점 만점에 몇 점 정도라고 생각하십니까?

번호	1점	2점	3점	4점	5점	6점
· 우리학교 교사들은 서술형평가 및 수행평가문항의 ‘채점기준 또는 루브릭’을 이원목적분류표에 따라 작성한다.						

⇒ 선생님은 서술형평가 및 수행평가문항의 ‘채점기준 또는 루브릭’을 어떻게 설정합니까?

⇒ 교사들이 ‘이원목적분류표’에 따라서 서술형평가 및 수행평가문항의 ‘채점기준 또는 루브릭’을 작성하기 어렵다면, 그 원인은 무엇이라고 생각되십니까? 이를 개선하기 위한 방안을 3가지 정도 제시해 주신다면 어떠한 것들이 있을까요?

24) 선생님께서 몸담고 있는 학교 또는 교과 교사들이 서술형평가 및 수행평가의 '채점기준 또는 루브릭'을 학생들의 반응에 따라 수정·보완하는 정도에 점수를 매긴다면 6점 만점에 몇 점 정도라고 생각하십니까?

번호	1점	2점	3점	4점	5점	6점
· 우리학교 교사들은 학생들의 반응에 따라 서술형평가 및 수행평가의 '채점기준 또는 루브릭'을 수정·보완한다.						

⇒ 교사들이 서술형평가 및 수행평가의 '채점기준'을 학생들의 반응에 따라 수정·보완한다면, 그 이유는 무엇이고, 어떤 방식으로 수정·보완하십니까? 혹은 그렇지 않다면, 그 이유는 무엇이고 이를 개선하기 위한 방안을 3가지 정도 제시해 주신다면 어떠한 것들이 있을까요?

25) 선생님께서 몸담고 있는 학교 또는 교과 교사들이 '성취기준과 성취수준, 그리고 이원목적분류표'를 기준으로 서술형평가와 수행평가의 분할점수를 산출할 수 있는 정도에 점수를 매긴다면 6점 만점에 몇 점 정도라고 생각하십니까?

번호	1점	2점	3점	4점	5점	6점
· 우리학교 교사들이 '성취기준과 성취수준, 이원목적분류표'를 기준으로 서술형평가와 수행평가의 분할점수를 산출할 수 있다.						

⇒ '성취기준과 성취수준, 그리고 이원목적분류표'를 기준으로 서술형평가와 수행평가의 분할점수를 산출할 수 있다면, 어떤 방식으로 산출하십니까? 혹은 그렇지 않다면, 그 이유는 무엇이고 이를 개선하기 위한 방안을 3가지 정도 제시해 주신다면 어떠한 것들이 있을까요?

26) 선생님께서 몸담고 있는 학교 또는 교과 교사들이 서술형평가 및 수행평가 결과에 대하여 '통계적 분석과 해석'을 할 수 있는 정도에 점수를 매긴다면 6점 만점에 몇 점 정도라고 생각하십니까?

번호	1점	2점	3점	4점	5점	6점
· 우리학교 교사들은 서술형평가 및 수행평가 결과에 대하여 '통계적 분석과 해석'을 할 수 있다.						

⇒ 선생님이 현재 평가와 관련하여 사용할 수 있는 통계적 분석은 무엇이고 그것을 어떻게 활용하는지 간략히 기술한다면, 어떠합니까?

⇒ 교사들이 선생님이 현재 평가와 관련하여 사용할 수 있는 통계적 분석을 이해하고 활용하는 것이 어렵다면, 그 원인은 무엇이라고 생각되십니까? 이를 개선하기 위한 방안을 3가지 정도 제시해 주신다면 어떠한 것들이 있을까요?

27) 선생님께서 몸담고 있는 학교 또는 교과 교사들이 서술형평가 및 수행평가 결과를 ‘성취기준과 성취수준’에 따라서 해석할 수 있는 정도에 점수를 매긴다면 6점 만점에 몇 점 정도라고 생각하십니까?

번호	1점	2점	3점	4점	5점	6점
· 우리학교 교사들은 서술형평가 및 수행평가 결과를 ‘성취기준과 성취수준’에 따라서 해석할 수 있다.						

⇒ 교사들이 서술형평가 및 수행평가 결과를 ‘성취기준과 성취수준’에 따라서 해석한다면, 그 이유는 무엇이고, 어떤 방식으로 해석하십니까? 혹은 그렇지 않다면, 그 이유는 무엇이고 이를 개선하기 위한 방안을 3가지 정도 제시해 주신다면 어떠한 것들이 있을까요?

28) 선생님께서 몸담고 있는 학교 또는 교과 교사들이 서술형평가 및 수행평가 결과를 ‘이원목적분류표’에 따라서 해석할 수 있는 정도에 점수를 매긴다면 6점 만점에 몇 점 정도라고 생각하십니까?

번호	1점	2점	3점	4점	5점	6점
· 우리학교 교사들은 서술형평가 및 수행평가 결과를 ‘성취기준과 성취수준’을 기준으로 능숙하게 해석할 수 있다.						

⇒ 서술형평가 및 수행평가 결과를 ‘이원목적분류표’에 따라서 해석한다면, 그 이유는 무엇이고, 어떤 방식으로 해석하십니까? 혹은 그렇지 않다면, 그 이유는 무엇이고 이를 개선하기 위한 방안을 3가지 정도 제시해 주신다면 어떠한 것들이 있을까요?

29) 선생님께서 몸담고 있는 학교 또는 교과 교사들이 학생들에게 서술형평가 및 수행평가의 채점기준을 사전에 공지하는 정도에 점수를 매긴다면 6점 만점에 몇 점 정도라고 생각하십니까?

번호	1점	2점	3점	4점	5점	6점
· 우리학교 교사들은 학생들에게 서술형평가 및 수행평가의 채점기준을 사전에 공지한다.						

⇒ 서술형평가 및 수행평가의 채점기준을 사전에 공지한다면, 그 이유는 무엇이고, 어떤 방식으로 공지하십니까? 혹은 그렇지 않다면, 그 이유는 무엇이고 이를 개선하기 위한 방안을 3가지 정도 제시해 주신다면 어떠한 것들이 있을까요?

30) 선생님께서 몸담고 있는 학교 또는 교과 교사들이 서술형평가 및 수행평가 결과를 분석하고 그 결과를 공지하는 정도에 점수를 매긴다면 6점 만점에 몇 점 정도라고 생각하십니까?

번호	1점	2점	3점	4점	5점	6점
· 우리학교 교사들은 서술형평가 및 수행평가 결과를 분석하고 그 결과를 공지한다.						

⇒ 교사들이 서술형평가 및 수행평가 결과를 분석하고 그 결과를 공지한다면, 그 이유는 무엇이고, 어떤 방식으로 공지하십니까? 혹은 그렇지 않다면, 그 이유는 무엇이고 이를 개선하기 위한 방안을 3가지 정도 제시해 주신다면 어떠한 것들이 있을까요?

31) 선생님께서 몸담고 있는 학교 또는 교과 교사들이 서술형평가 및 수행평가 결과를 학생의 학업 능력 향상에 활용하는 정도에 점수를 매긴다면 6점 만점에 몇 점 정도라고 생각하십니까?

번호	1점	2점	3점	4점	5점	6점
· 우리학교 교사들은 서술형평가 및 수행평가 결과를 학생의 약점 개선에 활용한다.						

⇒ 교사들이 서술형평가 및 수행평가 결과를 학생의 약점 개선에 활용한다면, 그 이유는 무엇이고, 어떤 방식으로 활용하십니까? 혹은 그렇지 않다면, 그 이유는 무엇이고 이를 개선하기 위한 방안을 3가지 정도 제시해 주신다면 어떠한 것들이 있을까요?

32) 선생님께서 몸담고 있는 학교 또는 교과 교사들이 서술형평가 및 수행평가 결과를 교사의 수업 개선에 활용하는 정도에 점수를 매긴다면 6점 만점에 몇 점 정도라고 생각하십니까?

번호	1점	2점	3점	4점	5점	6점
· 우리학교 교사들은 서술형평가 및 수행평가 결과를 교사의 수업 개선에 활용한다.						

⇒ 교사들이 서술형평가 및 수행평가 결과를 교사의 수업 개선에 활용한다면, 그 이유는 무엇이고, 어떤 방식으로 활용하십니까? 혹은 그렇지 않다면, 그 이유는 무엇이고 이를 개선하기 위한 방안을 3가지 정도 제시해 주신다면 어떠한 것들이 있을까요?

33) 선생님께서 몸담고 있는 학교 또는 교과 교사들이 서술형평가 및 수행평가 결과에 따라 성취기준 및 성취수준을 수정·조정하는 정도에 점수를 매긴다면 6점 만점에 몇 점 정도라고 생각하십니까?

번호	1점	2점	3점	4점	5점	6점
· 우리학교 교사들은 서술형평가 및 수행평가 결과에 따라 성취기준 및 성취수준을 수정·조정한다.						

⇒ 교사들이 서술형평가 및 수행평가 결과에 따라 성취기준 및 성취수준을 수정·조정한다면, 그 이유는 무엇이고, 어떤 방식으로 수정·조정하십니까? 혹은 그렇지 않다면, 그 이유는 무엇이고 이를 개선하기 위한 방안을 3가지 정도 제시해 주신다면 어떠한 것들이 있을까요?

34) 선생님께서 몸담고 있는 학교 또는 교과 교사들이 서술형평가 및 수행평가 결과를 학생수준에서 활용하는 정도에 점수를 매긴다면 6점 만점에 몇 점 정도라고 생각하십니까?

번호	1점	2점	3점	4점	5점	6점
· 우리학교 교사들은 서술형평가 및 수행평가 결과를 학생수준에서 활용한다.						

⇒ 교사들이 서술형평가 및 수행평가 결과를 학생수준에서 활용한다면, 그 이유는 무엇이고, 어떤 방식으로 활용하십니까? 혹은 그렇지 않다면, 그 이유는 무엇이고 이를 개선하기 위한 방안을 3가지 정도 제시해 주신다면 어떠한 것들이 있을까요?

35) 선생님께서 몸담고 있는 학교 또는 교과 교사들이 서술형평가 및 수행평가 결과를 학급·학년·학교수준에서 활용하는 정도에 점수를 매긴다면 6점 만점에 몇 점 정도라고 생각하십니까?

번호	1점	2점	3점	4점	5점	6점
· 우리학교 교사들은 서술형평가 및 수행평가 결과를 학급·학년·학교수준에서 활용한다.						

⇒ 교사들이 서술형평가 및 수행평가의 결과를 학급·학년·학교수준에서 활용한다면, 그 이유는 무엇이고, 각 수준에 대하여 어떤 방식으로 활용하십니까? 혹은 그렇지 않다면, 그 이유는 무엇이고 이를 개선하기 위한 방안을 3가지 정도 제시해 주신다면 어떠한 것들이 있을까요?

[부록3] 중등교사 평가 전문성 체크 리스트



중등교사 평가 전문성 체크 리스트

A. 평가 환경		아주 좋지 않음	좋지 않음	보통	좋음	매우 좋음	가중 계수	점수 (평정치× 가중계수)
		1	2	3	4	5		
A.1 교육여건	A.1.1 교육여건(교육과정 등)							
	A.1.1.1 서술형평가 및 수행평가에 우호적인 학교 분위기를 갖추고 있다.							
	A.1.1.2 서술형 평가문항을 검토하기 위한 교사 협업을 할 수 있는 교육여건을 갖추고 있다.							
	A.1.1.3 서술형 및 수행평가 기준을 명확히 정립할 수 있는 연수가 충분히 시행되고 있다.							
	A.1.1.4 서술형 평가 기준(안)을 참고할 수 있는 다양하고 구체적인 자료제시가 충분하다.							
	A.1.1.5 서술형 및 수행 평가 취지에 맞는 교교육과정이 적합하다.							
	A.1.1.6 서술형 및 수행 평가를 위한 교사의 업무 부담이 적정하다.							
A.2 교수학습 방법	A.1.2 교수학습 방법							
	A.1.2.1 서술형 및 수행 평가 취지에 맞는 성취기준을 명확히 정립할 수 있다.							
	A.1.2.2 서술형 및 수행 평가 취지에 맞는 성취기준에 따라 다양한 교육방법을 적용할 수 있다.							
	A.1.2.3 서술형 및 수행 평가 취지에 맞는 성취기준에 따라 다양한 수업설계를 적용할 수 있다.							
	A.1.2.4 서술형 및 수행 평가 취지에 맞는 교수학습·방법 개선을 위한 노력이 충분하다.							
	A.1.2.5 서술형 및 수행 평가 취지에 맞는 수업시간확보와 평가시기를 마련하고 있다.							
A.3 교육평가 세부여건	A.1.3 교육평가 세부여건							
	A.1.3.1 서술형 및 수행 평가 기준을 명확히 정립할 수 있는 교사 교육용 안내 책자 제공이 충분하다.							
	A.1.3.2 서술형 및 수행 평가를 위한 학생 답안 작성 방법 등을 담은 학생 교육용 안내 책자 제공이 충분하다.							
	A.1.3.3 서술형 및 수행 평가의 유연한 평가 비율 조정이 가능하다.							
			▶ 소 계					

B. 평가 투입		아주 좋지 않음	좋지 않음	보통	좋음	매우 좋음	가중 계수	점수 (평정치× 가중계수)
		1	2	3	4	5		
B.1 목표 설정	B.1.1 목표설정							
	B.1.1.1 성취기준 코드를 이해·활용할 수 있다.							
	B.1.1.2 학생 수준을 고려하여 성취기준과 성취수준을 재설정할 수 있다.							
	B.1.1.3 학생 수준을 고려하여 성취기준과 성취수준을 수정·보완할 수 있다.							
B.2 평가유형 선정	B.3.2 평가유형							
	B.1.2.1 성취기준에 따라 선다형, 서술형, 수행형 평가 유형을 선택할 수 있다.							
	B.1.2.1 성취기준에 따라 서술형평가 유형을 선택할 수 있다.							
	B.1.2.3 성취기준에 따라 수행평가 유형을 선택할 수 있다.							
B.3 평가 윤리	B.3.3 평가 윤리							
	B.1.3.1 서술형 및 수행평가를 시행함에 있어 학생의 성별, 경제적, 문화적 요인에 따른 편파성을 최소화 하려고 노력한다.							
		▶ 소 계						

C. 평가 과정		아주 좋지 않음	좋지 않음	보통	좋음	매우 좋음	가중 계수	점수 (평정치× 가중계수)
		1	2	3	4	5		
C.1 평가도구 개발	C.1.1 서술형 및 수행평가 도구 개발							
	C.1.1.1 성취기준 및 수준에 따라서 서술형 및 수행평가 문항을 개발할 수 있다.							
	C.1.1.2 이원목적분류표에 따라 서술형 및 수행평가 문항을 개발할 수 있다.							
	C.1.1.3 문항개발 과정에서 서술형 및 수행평가도구의 신뢰도와 타당도를 검토할 수 있다.							
C.2 채점기준 정립	C.2.1 서술형 및 수행평가 채점기준 정립							
	C.1.2.1 성취기준 및 수준에 따라 서술형 및 수행평가의 채점기준을 설정할 수 있다.							
	C.1.2.2 이원목적분류표에 따라서 서술형 및 수행평가의 채점기준을 설정할 수 있다.							
	C.1.2.3 학생의 학업능력 또는 반응을 예측하여 채점기준을 설정할 수 있다.							
C.3 분할점수 산출	C.3.3 서술형 및 수행평가 분할점수 산출							
	C.1.3.1 이원목적분류표에 서술형 및 수행평가의 분할점수를 산출할 수 있다.							
	C.1.3.2 성취기준 및 수준에 따라 서술형 및 수행평가의 분할점수를 산출할 수 있다.							
	C.1.3.3 학생의 학업능력 또는 반응을 예측하여 분할점수를 설정할 수 있다.							
		▶ 소 계						

D. 평가 결과		아주 좋지 않음	좋지 않음	보통	좋음	매우 좋음	가중 계수	점수 (평정치× 가중계수)
		1	2	3	4	5		
D.1 평가결과 분석	D.1.1 평가결과 분석							
	D.1.1.1 서술형 및 수행평가 결과를 양적으로 분석할 수 있다.							
	D.1.1.2 서술형 및 수행평가 결과를 질적으로 양적으로 분석할 수 있다.							
	D.1.1.3 평가요소에 따라서 서술형 및 수행평가 결과를 분석할 수 있다.							
D.2 평가결과 해석	D.3.2 평가결과 해석							
	D.3.2.1 성취기준 및 수준에 따라서 서술형 및 수행평가 결과를 해석할 수 있다.							
	D.3.2.2 이원목적분류표에 따라서 서술형 및 수행평가 결과를 해석할 수 있다.							
	D.3.2.3 평가결과에 따라서 학생의 학업능력을 해석할 수 있다.							
		▶ 소 계						

E. 평가 환류		아주 좋지 않음	좋지 않음	보통	좋음	매우 좋음	가중 계수	점수 (평정치× 가중계수)
		1	2	3	4	5		
E.1 교사측면	E.1.1 교사측면							
	E.1.1.1	서술형평가와 수행평가 결과를 종합·분석하여, 학생들의 강점과 약점을 파악하고 교수학습방법의 개선점을 도모할 수 있다.						
	E.1.1.2	서술형평가와 수행평가 결과를 활용하여, 학생들의 재능과 학업성취 정도를 파악하고 학생의 꿈과 끼 실현을 위한 학습계획을 세울 수 있다.						
	E.1.1.3	서술형평가와 수행평가 결과를 활용하여, 학생의 창의성 신장을 위한 효과적인 교육활동을 구상할 수 있다.						
	E.1.1.4	서술형평가와 수행평가 결과를 활용하여, 학생 발달과정을 이해하고 지속적이고 안정적인 성장을 도모할 수 있다.						
	E.1.1.5	서술형평가와 수행평가 결과를 활용하여, 학생 수준을 파악하고 수업계획과 목표를 수립할 수 있다.						
	E.1.1.6	서술형평가와 수행평가 결과를 활용하여, 학생들의 학습동기와 학습환경 개선을 위한 자료로 활용할 수 있다.						
E.2 학생측면	E.1.2 학생측면							
	E.1.2.1	학생 스스로 성취기준에 제시된 평가내용과 능력(이해, 적용, 분석)을 파악하여 강점과 약점을 고찰할 수 있고, 꿈과 끼 실현을 위한 필요영역을 제공할 수 있다.						
	E.1.2.2	학생 스스로 창의력과 문제해결 능력을 고취시킬 수 있고, 흥미와 학업 동기를 유발시킬 수 있다.						
	E.1.2.3	학생 스스로 무엇을 알고, 무엇을 할 수 있는지에 대한 구체적인 분석을 제공 할 수 있으며, 무엇보다 학습 능력 신장에 필요한 정보를 제공할 수 있다.						
	E.1.2.4	학생 스스로 문제점을 파악하여 자기 주도적인 학습이 가능하게 할 수 있다.						
E.3 교육이해 당사자측면	E.1.3 교육이해 당사자 측면							
	E.1.3.1	학생들의 학력 향상, 공교육 정상화와 사교육 경감 등에 기여할 수 있다.						
	E.1.3.2	학교교육수립에 반영하여 교육과정 개선에 기여할 수 있다.						
	E.1.3.3	서술형평가와 수행평가 결과에 대한 신뢰도와 타당도가 높일 수 있고, 무엇보다 교육이해 당사자들의 평가 전문성 신장에 도움이 될 수 있다.						

	E.1.3.4 서술형평가와 수행평가 결과를 토대로 학생들의 학력신장에 대한 정보를 누적적으로 관리할 수 있다.						
	E.1.3.5 학력 평가관리 시스템을 적용하여 서술형·수행 평가를 실시하여 학력을 파악하고 분석하는 일련의 절차적 지식을 활용할 수 있다.						
	E.1.3.6 학력 평가관리 시스템 적용과 활용은 교사들의 평가 전문성 향상을 위한 연수 프로그램 등에 활용할 수 있다.						
	E.1.3.7 교육이해 당사자에게 안정적이고 지속적인 평가환류체계 구축을 위한 인적, 물적 여건 조성의 필요성을 인지시킬 수 있다.						
						▶ 소 계	

※ 점수 합계

항목	점수
A. 평가환경 합계(○○○점 만점)	
B. 평가투입 합계(○○○점 만점)	
C. 평가과정 합계(○○○점 만점)	
D. 평가결과 합계(○○○점 만점)	
E. 평가환류 합계(○○○점 만점)	
총 합계(○○○점 만점)	