

한국과 미국의 초등수학 서술형 평가의 제도 및 문항 비교¹⁾

김민경²⁾ · 조미경³⁾ · 김래영⁴⁾ · 김구연⁵⁾ · 노선숙⁶⁾

본 연구는 한국과 미국의 국가, 교육청, 학교 수준에서의 서술형 평가와 관련한 제도 및 문항에 관한 국제 비교를 함으로써 두 나라에서 서술형 평가가 실행되는 수준별 현황을 살펴보고, 서술형 평가의 효율적인 적용을 위한 시사점을 찾아보고자 한다. 비교 결과, 한국과 미국의 국가 수준의 교육과정에서는 공통적으로 문제해결 과정에 대한 평가와 교수·학습 과정과 평가 간의 높은 연계성을 갖는 평가 방식의 중요성이 강조되고 있다. 그러나 교육청 수준에서 한국은 문서상에서 서술형 평가 실시 비율을 제시하는 데에 치중하는 경향이 짙은 반면, 미국의 각 주(State)에서는 교실 현장에서 실제로 활용할 수 있는 예시문항과 채점기준을 제시하여 활용을 도울 수 있도록 한 차이를 나타냈다.

주요용어 : 서술형 평가, 초등수학, 국제비교

I. 시작하며

지식기반사회에 살아가는 우리들은 지식이 셀 수 없이 많고 새로운 지식이 생성되는 속도가 매우 빠른 사회에 살아가고 있기 때문에, 과거와 달리 모든 지식을 학교교육을 통해서 배울 수 없다. 따라서 과거의 학교 교육에서 지식 전달을 목표로 했던 것과는 달리 앞으로는 자기주도적인 문제해결능력을 갖추고 지식정보의 활용능력에 기초하여 사고력, 비판력 등을 갖춘 창의 인재 육성에 학교 교육은 힘써야 한다(신현석, 박균열, 엄준용, 2008). 이런 변화가 수학 교육 분야에서는 전 세계적으로 1990년대 이후로 학교 수학 교육에서 수학적 문제해결력, 추론 능력, 의사소통 능력과 같은 수학적 능력 신장의 강조로 나타나고 있다(교육과학기술부, 2008).

수학 교육의 변화는 국제 학업성취도 비교 연구에도 반영되었는데, 국제 학업성취도 비교 연구에서 우리나라 학생들의 수학과 성취도 결과를 비교하면 다음과 같다. PISA 2009년도

-
- 1) 이 논문은 2011년도 정부(교육과학기술부)의 재원으로 한국연구재단의 지원을 받아 연구되었음 (NRF-2011-32A-B00216)
 - 2) 이화여자대학교(mkkim@ewha.ac.kr)
 - 3) 이화여자대학교 대학원(cmk0530@hanmail.net)
 - 4) 이화여자대학교(kimrae@ewha.ac.kr) 교신저자
 - 5) 서강대학교(gokim@sogang.ac.kr)
 - 6) 이화여자대학교(noh@ewha.ac.kr)

결과 분석에 따르면(김경희 외, 2010), 우리나라는 수학적 과정의 전 영역에서 전반적으로 상승했으나, 하위 영역의 정답률을 보면 재생 75.9%, 연결 56.6%, 반성 49.7%로, 보다 문항이 복잡해지고 독창적이고 고차원적인 사고 능력을 요하는 반성 영역이 가장 낮게 나타나고 문항별 정답률 차이도 가장 큰 것으로 나타났다. 또한 상황과 맥락의 측면에서 수학적 소양이 향상된 것으로 나타났지만, 개인적 상황에서 2006년과 비교했을 때 유의한 하락폭을 나타냈고 이는 7차 교육과정 이후 일상생활에서의 문제해결능력을 지속적으로 강조하고는 있지만 학생들은 여전히 일상생활에서 수학을 접하는 상황에 익숙하지 않음을 나타내는 것으로 분석되었다. 그리고 TIMSS 2007 결과에서는 수학 인지영역별 정답률 분석 결과, 회상, 분류 또는 정렬하기 등의 알기가 79.1%, 모델화하기 또는 정형적인 문제해결하기의 적용하기가 75.2%, 일반화하기, 정당화하기, 비정형적인 문제해결하기의 추론하기가 65.9%로 나타났다(김경희, 김수진 외, 2008). 학생들에게 답과 풀이과정을 함께 제시하도록 하여 사고의 논리적 전개 및 개념의 활용과 더불어 의사소통 능력을 평가하는 TIMSS의 추론 영역과 PISA의 반성 영역의 문항의 결과를 보면, 우리나라 학생들에게는 그러한 능력이 부족한 것으로 볼 수 있다(박경미, 김동원, 2011). 또한 우리나라에서 수학적 문제해결력 신장은 제4차 교육과정 이래로 강조해오고 있지만 실질적으로 문제해결력이 증진되었다고 보기 어려운 실정이다.

이러한 학교 수학 교육의 문제를 개선하고자 2007년 및 2009년 개정교육과정에서는 교육 내용, 교수·학습 방법, 평가에 일관되게 수학적 능력의 신장을 더욱 강조하고 있다. 특히, 제7차 교육과정에서는 현대 사회에서 강조되고 있는 수학적 의사소통 능력이 그다지 강조되지 않았기에, 개정교육과정에서는 수학적 의사소통 능력의 신장을 위해 수학적 표현을 이해하고 정확히 사용하게 하고, 수학적 아이디어를 말과 글로 설명하고, 수학을 표현하고 토론하는 것을 통해 자신의 사고를 명확히 하고 반성해 보도록 하고 있다(교육과학기술부, 2008, 2009). 특히, 평가 영역에서는 다양한 평가 방법의 사용을 권장하고, 문항 성격의 중요성을 언급하며 학생들의 사고 과정과 결과를 확인할 수 있는 질 높은 문항들이 포함되어 의미 있는 평가가 수행되어야 한다고 하였다.

서술형 평가는 축적한 지식의 양과 습득 지식을 기억하여 인출하는 능력을 측정하는 과거의 평가 형태와는 달리, 알고 있는 정보나 지식 등을 활용하여 문제 상황을 분석, 설명, 해석하거나, 창의적으로 해결할 수 있는 능력을 측정하는 평가 유형이다(김주환, 2010). 박정과 박경미(2004)가 TIMSS-R 결과 분석에서 단답형·서술형 문항은 선다형 문항에 비해 학습자의 능력을 좀 더 정확하게 추정할 수 있고 상대적으로 적은 수의 문항으로도 학습자의 능력을 추정 가능하다고 하고, 노국향과 박정(2001)이 단답형·서술형 평가 문항이 선다형 문항에 비해 학습자의 능력에 관해 더 많은 정보를 제공하고 더 정확하게 추정하는 경향이 있다고 한 것을 토대로 볼 때, 서술형 평가는 학습자의 능력을 좀 더 정확하게 파악하고 학습자의 이해 정도를 파악하여 결국 평가와 교수·학습 과정이 연결되어 평가가 보다 발전된 교수적 기능을 수행하는 데에 큰 역할을 할 것이다.

그러나 우리나라에서는 1997년 수행평가가 도입된 이래 서울특별시교육청은 2005년부터 중·고등학교의 경우 서술형이나 논술형 평가를 확대함으로써 학교교육의 내실화를 추구하고자 하였으나 초등학교의 경우 별도의 기준이 없었고(서울특별시교육청, 2010a), 2011년도에는 주요업무계획의 ‘서울교육 방향’에서 기본교육의 내실화와 자기주도 학습역량 강화를 위한 실천과제로 과중 중심의 질적 평가 확대를 내세우며 문제해결력 신장을 위한 단답형을 제외한 서술형·논술형 평가 문항을 확대하도록 한 바와 같이(서울특별시교육청, 2011), 정

책 시행을 위한 구체적인 방침이나 노력은 뒤따르지 못한 것으로 보인다. 따라서 서술형 평가가 가지고 있는 장점들에 기초하여 서술형 평가 실행이 어떻게 이루어지고 있는지 알아보는 것이 필요하다.

이에 본 연구에서는 한국과 다른 나라의 서술형 평가에 관한 제도 및 문항을 비교해 봄으로써 현재 서술형 평가에 관한 제도 및 문항을 객관적인 관점에서 바라보고자 한다. 미국은 수행평가를 가장 먼저 교육 현장에 도입한 나라로 알려져 있고 완전히 새로운 개념의 평가가 아니라 학생 평가에 보고서, 쓰기 자료, 발표 등을 포함시키는 방식으로 이해하고 1980년대 후반부터는 교수·학습 개선의 목적으로 구조적으로 수행평가를 시행하고 있는 측면이 현재 우리나라에서 추구하고 있는 수행평가 및 서술형 평가의 확대 방향과 같은 맥락으로 이해되어 본 연구의 비교 대상으로 선정하였다. 또한, 수학 교과와 여러 분야에서 오랜 기간 동안 다양한 관점을 가지고 연구를 진행해왔고 전 세계의 수학 교육에 큰 영향을 미쳤다고 판단된 미국 NCTM의 자료를 중심으로 비교하고자 한다. 이를 통해 우리나라에서 시대적 필요와 요구에 의해 확대 실시되고 있는 서술형 평가의 효율적인 적용을 위한 시사점을 찾을 수 있을 것으로 기대한다.

II. 한국과 미국의 국가 수준의 평가 비교

전통적인 평가방법은 단편적인 지식과 정보를 기억하는 양을 파악하여 상대적인 서열을 매기는 행정적 기능으로 평가를 활용하고, 수학적 지식을 잘게 쪼갠 다음 각각을 하나의 평가 목표로 간주하여 문항을 출제하고 선택형 검사를 통해 양적인 평가를 하였다. 따라서 선다형, 참과 거짓 고르기, 단답형 등의 전통적인 평가는 학생들이 배운 것을 제대로 평가하지 못하고, 단지 시험 치는 기술만을 측정한다는 비판을 받아왔다(Idris, 2007).

그런데 수학교육의 현대화 운동 이후, 세계 각국의 초·중등학교 수학 교육과정에는 많은 변화가 있었고, 이러한 교육과정 상의 변화는 수학 교과와 평가관 또는 종래의 평가 방법에서 변화를 도모하게 되었다. 즉, 평가의 관점이 ‘학습의 평가’에서 ‘학습을 위한 평가’, 더 나아가 ‘학습으로서의 평가’로 변하고 있다(Black 외, 2003; Brookhart, 2005; Leahy 외, 2005). 오늘날 평가는 학생과 교사 간에 피드백이 가능하게 하는 것으로 학생의 학습동기를 유발하는 교수적 기능으로 활용되고, 추론 능력, 의사소통 능력, 다른 수학적 지식 및 타교과와 일반적 지식까지도 연결할 수 있는 능력과 창조성을 포함하는 종합적이며 고차적인 사고능력 및 문제해결과 문제 만들기 능력을 평가하는 질적인 평가를 하고자 한다(김수환 외, 2009; 백순근, 2000). 이러한 질적 평가의 일환으로 세계 각국에서는 다양한 형태의 수행평가를 시행하고 있는데, 수학 교과에서는 문항출제가 비교적 용이하고 학생들의 문제해결 과정을 분석할 수 있다는 이유로 서술형 평가의 활용도가 높은 편이다(성태제, 2000; 허경철 외, 1999; 황혜정, 2003). 초등 수학교육과정 개발자 Bell과 Isaacs(2007)은 현장에서 널리 활용되고 있는 객관식 평가, 단답형 평가 보다는 학생들의 문제 이해, 문제해결을 위한 사고 및 의사소통 과정을 통합적으로 표현할 것이 요구되는 서술형 평가가 과정평가에 적합한 평가 유형이라고 언급하였다. 수학적 개념이나 원리를 논리적으로 타당한 이유나 근거를 들어 설명하기, 수식의 의미를 그림이나 문장으로 설명하기, 수학적 계열성 및 관계성 설명하기, 문제 만들기 및 해결 방법에 대해 설명하기, 수학적 용어를 실생활 장면에서 활용하기, 자료

해석하기 등에 대해 자신이 생각하는 지식이나 의견을 그림이나 문장 등을 활용하여 직접 서술하기 등이 서술형 평가의 문제 유형으로 포함된다는 점을 고려해보면(박금란, 방정숙, 2008), 서술형 평가가 과정평가에 적합한 이유를 쉽게 이해할 수 있다.

이러한 평가 동향이 한국의 2007년 및 2009년 개정교육과정과 미국 NCTM의 2000년도 평가 기준을 중심으로 어떻게 반영되고 있는지 비교하고, 두 국가의 평가 방향과 서술형 평가의 적합성에 대해 분석해 보고자 한다.

먼저, 한국의 수학 교과 교육과정에서 평가가 어떻게 변천되었는지를 간단하게 살펴보면, 제3차 교육과정의 시기까지는 평가에 대한 내용이 제시되어 있지 않고, 제4차부터 ‘지도 및 평가상의 유의점’ 항목으로 명칭이 바뀌면서 평가시 고려할 점들이 제시되고 있다. 제4차 교육과정에서는 조작 활동에 대한 관찰 기록도 평가로 포함할 수 있고 형성 평가를 통해 보충과 심화를 요구할 수 있음을, 제5차 교육과정에서는 학습 목표를 다양한 방법을 통해 바르게 평가하고, 평가 결과를 학생, 교사, 학부모에게 유용한 정보로 제공될 수 있도록, 제6차와 제7차 교육과정에서는 평가 도구의 다양화와 문제해결력에 대한 평가를 강조하며 문제의 이해 능력과 문제해결 과정을 파악할 수 있는 주관식 평가를 위주로 하도록 제시하였다(교육과학기술부, 2008).

이와 같이 문제해결 과정을 중시하는 평가의 방향이 2007년 및 2009년 개정교육과정에서는 인지적 영역의 평가에서 학생들의 수학적 사고력 신장을 위하여 결과뿐만 아니라 과정도 중시하여 평가하여야 한다고 하며, 보다 구체적으로 ① 수학의 기본적인 개념, 원리, 법칙을 이해하고 적용하는 능력, ② 수학적 표현의 의미를 이해하고 정확하게 사용하는 능력, ③ 수학적 지식과 기능을 활용하여 타당하게 추론하는 능력, ④ 다양한 상황에서 발생하는 여러 가지 문제를 수학적으로 사고하여 해결하는 능력, ⑤ 생활 주변 현상, 사회 현상, 자연 현상 등의 여러 가지 현상을 수학적으로 사고하여 해결하는 능력, ⑥ 수학적 사고 과정과 결과를 합리적으로 의사소통하는 능력으로 평가하여야 할 세부 능력을 밝혔다(교육과학기술부, 2008, 2009).

한편, 미국의 경우, 구체적인 지식과 분리된 기술만을 측정했던 과거와 달리 수학적 힘을 총체적으로 평가해야 한다고 하며, NCTM(1995)은 정형적인 문제와 비정형적인 문제를 해결할 때 필요한 수학적 개념과 기술을 활용하듯이 개념과 기술을 이해하였는지, 수학적 탐구 과정상에서 활용한 전략과 추론 과정을 효과적으로 나타낼 수 있는지, 실생활에서 접하는 상황을 수학적으로 해석하고 이해하는 데에 자신감이 있는지 등을 평가해야 한다고 하였다. 이후, 2000년도 ‘학교 수학을 위한 원리와 기준’에서 평가는 학생들의 절차적 기능뿐만 아니라 이해에 초점을 두고, 시험과 퀴즈 같은 형식적인 평가뿐만 아니라 수업하는 동안의 질문, 개별학생과의 면담, 학생들의 쓰기 과정 등을 통해 학생들의 발전에 관한 정보를 모을 수 있어야 한다고 하였다(NCTM, 2000). NCTM에서도 계속적으로 학생들의 수학적 사고 과정이나 문제해결 과정을 강조하고 있음을 알 수 있다. 또한 최근에는 48개 주의 교육과정의 잘 된 내용들과 NCTM에서 강조하고 있는 내용들을 선택하여 새롭게 구성한 The Common Core State Standards(CCSS)의 수학 교육과정에서도 문제를 해결하는 과정을 중시하며, 구성형 문항(constructed response item)을 포함하고 표준화된 전통적인 평가 방식과는 다른 방식의 컴퓨터에 기반을 둔 평가 방식을 제안하고 있다(Wikipedia, n.d.).

한국의 2007년 및 2009년 개정교육과정과 미국 NCTM의 평가 기준에서 공통적으로 평가해야 할 대상으로 추구하는 바는 단편적인 수학적 지식이 아니라 수학적 지식을 활용하는 방법과 문제해결 과정에 중점을 두고 있음을 알 수 있다. 이는 과거에 많이 활용하였고 정

답 여부에만 많은 관심을 두었던 단답형이나 선택형 문항으로는 평가할 수 없고(성태제, 1999), 학생의 생각이나 의견을 직접 서술하도록 하는 서술형 평가를 통해 학생의 창의성, 문제해결력, 비판력, 판단력, 통합력, 정보 수집력 및 분석력 등 고등사고능력을 쉽게 평가할 수 있고 학생들이 개념이나 과정을 어떻게 공식화하고 조직하고 내면화하여 설명하는가를 평가할 수 있다(백순근, 2000; 이대현, 박배훈, 1999). 또한 서술형 평가를 통해서 학생이 서술한 문제해결과정을 분석함으로써 학생들이 문제를 이해하는지, 해결과정에서 무엇을 잘못 알고 있는지를 파악하여 오류를 바로잡아 줄 수도 있다(정동권 외, 2002; NCTM, 2001). 이와 같이 문제해결 과정을 평가할 수 있다는 특징을 지닌 서술형 평가의 가치는 여러 연구에서 나타났다. 초등학교 6학년 학생들을 대상으로 서술형 평가를 시행하여 수학 교과 기본 지식에 대한 이해의 실태를 조사한 박금란과 방정숙(2008)은 학생들이 전형적인 알고리즘에 의한 계산은 매우 익숙하게 수행하였으나 개념이나 원리를 설명하는 데에는 매우 부족한 것으로 드러났고, 이러한 부족함은 서술형 평가를 통해 교수·학습 과정에 지속적인 피드백을 제공함으로써 채워질 수 있고 더 나아가서는 평가의 본질을 회복하는 데에 큰 역할을 할 수 있을 것으로 기대하고 있다. 좌동지(2001)는 초등학생들로 하여금 문제해결과정을 서술하게 하는 것이 수학적 문제해결력을 신장시키는데 긍정적인 역할을 하고, 정현도, 강신포, 김성준(2010)은 4학년 학생들을 대상으로 서술형 평가를 시행하고 학업성취도에 따라 나타나는 오류 유형을 분석함으로써 평가를 통한 피드백이 효과적인 수학학습지도로 연결되기를 기대한다고 하였다. 따라서 서술형 평가는 한국의 2007년 및 2009년 개정교육과정과 미국 NCTM에서 수학적 지식을 활용하는 방법과 과정에 중점을 두는 평가를 지향하는 방향에 매우 유용한 방법으로 볼 수 있겠다.

또한, 한국의 2007년 및 2009년 개정교육과정에서는 평가가 교사와 학생 모두에게 유용한 정보를 제공함으로써 교수·학습 개선에 기여해야 하고 의미 있는 수학 학습을 뒷받침해야 한다고 하며, 평가를 통해 학생이 이미 학습한 수학 내용 중에서 강점을 갖는 부분이 어디이며 취약한 부분이 어디인가를 확인할 수 있고 평가 결과에서 나타난 학생들의 오개념 분석을 통해 수업의 방향을 결정할 수 있다고 하였다(교육인적자원부, 2008).

미국의 경우, 평가 규준에서 과거에 교육과정이나 교수·학습 과정과 평가를 독립된 것으로 여겼던 관점과 달리, NCTM(1995)은 교육과정이나 교수·학습 과정과 일관된 관점을 가지고 평가를 바라봐야 하고 학생들의 수학적 힘을 신장시키기 위해서는 학생 개인의 수학 학습을 계속적으로 지원해줘야 한다고 하였다. 또한 깊이 있고 수준 높은 학습을 위해 평가와 교수·학습은 통합되어야 하고, 이러한 평가는 교사가 적절한 수업 상의 결정을 내리는 데 필요한 정보를 제공한다고 하였다(NCTM, 2000). 이후, NCTM은 2001년에는 대안적인 방법으로 행할 수 있는 새로운 평가의 접근 방법들을 소개하고, 새로 전학온 학생들을 학급 생활에 어떻게 적응시킬까를 그 학생의 평가 결과를 통해 결정할 수 있었던 예시나 개별 면담을 통해 학생들의 수학적 이해에서 발전을 돕기 위해 평가 루브릭을 활용한 예시 등을 제시하며 평가 결과를 교수·학습 과정에 어떻게 활용할 수 있는지, 활용하는 과정에서 의문점 등을 엮거나(NCTM, 2001), 2005년에는 내용 영역별로 평가 과제 예시들을 제시하고, 학생들의 수학적 의사소통 능력을 평가하고 수학적 사고 능력을 키워줄 수 있도록 기존의 선다형 문항과 단답형 문항을 어떻게 수정, 보완할 수 있는지에 대한 가이드라인을 제시하고, 학생 응답 예시와 채점을 통해 어떻게 교수·학습 과정을 발전시킬 수 있는지를 제시함으로써(NCTM, 2005) 평가 전문성 신장을 위해 지속적으로 노력하고 있다.

여기서 공통적으로 나타나는 특징은 평가와 교수·학습 과정 간의 연계성을 높게 보고 있

다는 점이다. 이는 과거에는 각 학생의 성취 정도를 다른 학생과 비교하여 상대적인 위치를 중요하게 여기는 상대 평가를 지향하고 서열화 시키는데 중점을 두었지만, 지금은 학생이 교육목표에 도달하였는지, 내용을 이해하였는지를 확인하는 것을 목표로 하는 절대 평가를 지향하면서 중간고사, 기말고사로 보던 평가 방식이 교사가 학생들의 진행 정도나 향상 상태를 확인하고 수업을 올바른 방향으로 진행하는 데 도움을 주는 형성평가(교육인적자원부, 2008)와 필요에 따라 수시로 평가하는 수시평가의 형태로 바뀌고 있는 것과 같은 맥락이며, 이를 통해 학생의 학습에 대한 정보를 필요에 따라 그것에 기초하여 교수·학습의 방향을 결정하는 것이다. 이는 박경미와 김동원(2011)이 서술형 평가를 활용함으로써 평가가 교수·학습의 결과를 점검하게 되고 나아가 수학 수업을 개선할 수 있다고 밝힌 바와 같은 맥락이다.

마지막으로 한국의 2007년 및 2009년 개정교육과정에서는 획일적인 방법을 지양하고 지필 평가, 관찰, 면담, 자기 평가 등의 다양한 평가 방법을 활용해야 한다고 하며, 지필 평가 방법 자체가 미흡한 평가 방법은 아니며 지필 평가에서 구현하고 있는 문항이 수학적 사고 과정과 결과를 확인할 수 있는 질 높은 것이면 의미 있는 평가가 가능하다고 하였다(교육인적자원부, 2008, 2009). NCTM(1995, 2000)은 과거에 수학적 지식을 증명해 보이는 차원에서 단일한 방법을 활용했던 것과는 달리, 수학적 힘을 충분하게 보일 수 있는 다양한 기회를 제공해야 한다고 하였고, 개방형 질문, 구성형 과제, 선택형 문항, 수행 과제, 관찰, 대화, 일기, 포트폴리오 등 많은 평가 기법을 이용할 수 있고, 구성형 과제나 수행과제는 복잡하거나 새로운 상황에서 수학을 적용하는 학생들의 능력을 더 잘 드러낼 수 있다고 하였다.

여기에서 공통적으로 고려해야 하는 부분은 평가에 활용되는 문항이나 과제의 성격으로 볼 수 있는데, Lappan, Phillips 그리고 Fey(2007)는 학습자를 평가하기 위해서는 ‘좋은 문항’을 개발하여야 한다고 하며, 중요한 수학적 아이디어가 내재되어 있고 고등 수학적 사고와 다양한 해결 전략을 활용하여 문제를 해결하는 과정을 필요로 하는 문항을 좋은 수학과 평가문항이라고 하였다. 이러한 문항의 성격은 우리나라에서 최근 활용이 강조되고 있는 서술형 평가 문항에 반영시켜야 할 부분이다.

III. 한국의 지역 교육청과 미국의 각 주(States) 서술형 평가 시행 방향 비교

국가 교육과정 수준에서 세운 평가 원리를 바탕으로 두 나라의 교육청 수준에서는 학교급별로 서술형 평가의 어떠한 시행 방침을 안내하고 있는지 한국은 서울특별시교육청과 경기도 교육청을 중심으로, 미국은 NCTM의 분석적 채점 방법과 California의 총체적 채점 방법의 활용을 중심으로 비교하고자 한다.

1. 한국의 지역 교육청의 경우

한국에서 수행평가는 1999년부터 전국적으로 실시, 확대되었다고 볼 수 있는데, 수행평가의 전국적 실시를 앞두고 수행평가의 개념이 명료하게 제시되지 않았고 다만 주관식 서술형 평가와 객관식 평가 비율을 적정 수준으로 유지하라는 방침을 제시하였다. 교육부는 학부모 홍보자료, 교사 연수자료 등을 통해 수행평가 홍보에 나섰고, 여기에서 수행평가는 서술형이

나 논술형, 그리고 완성형이나 단답형 등 낮은 수준의 방법들까지도 포함시키고 있다고 하였다. 이렇게 시작된 수행평가는 단위 학교의 자율성을 최대한 살린다는 측면에서 실질적인 결정을 단위 학교에서 모두 결정하도록 하였는데, 이는 수행평가에 관한 공정성이나 객관성 문제가 꾸준히 제기됨으로써 그 한계를 드러냈다(서지영 외, 2008; 허경철 외, 1999)

수행평가 중에서도 서술형 평가를 특히 강조하게 되었는데, 서울특별시교육청에서는 2007년도부터 서술형·논술형 평가 확대 정책과 관련하여 비중을 전체의 50% 이상을 반영하도록 하고 교과별 특성을 고려하여 방법 및 비율을 학교 자율로 결정하도록 하였다. 그런데 교사들은 학교마다 서술형·논술형 평가를 수행평가의 일종으로 보기도 하고 그렇지 않기도 하여 평가 영역 및 배점 안배에 혼란이 있다고 지적하였다(서지영 외, 2008).

이러한 혼란은 최근 각 교육청의 노력으로 점차 자리를 잡아가고 있는 것으로 보인다. 서울특별시교육청(2010a)은 2010년 2월 ‘창의성 계발을 위한 평가 개선 기본 계획’에서 미래사회가 요구하는 창의적인 인재 육성을 위해 창의성, 표현력, 문제해결력 신장을 위해 평가방법을 전환함으로써 수업방법의 변화를 도모하고자 하였다. 이를 위해 서술형 문항 반영비율을 초등학교 3학년부터 2012년까지 단계적으로 높여 50%까지 확대해가도록 하였고, 교과별 특색 있는 수행평가를 활성화하는데 수학과에서는 결과뿐만 아니라 과정을 중시할 수 있는 방법을 지향하도록 하였다. 또한 2011년도 주요업무계획의 ‘서울교육 방향’에서는 기본교육의 내실화와 자기주도 학습역량 강화를 위한 실천과제로 과중 중심의 질적 평가 확대를 내세웠다(서울특별시교육청, 2011). 여기서는 문제해결력 신장을 위한 서술형·논술형 평가 문항을 출제하되 단답형은 제외하도록 하였고, 총 배점의 30% 이상을 포함하도록 하였다.

경기도 교육청(2010)도 2010년 10월 ‘학생 창의력 신장을 위한 평가방법 혁신 기본 계획’을 통해 서술형, 논술형 문항비율 확대 및 다양한 평가방법의 적용으로 문제해결력을 갖춘 창의적인 인재 육성을 위하여 학교급별, 교과별 특성을 고려하여 서술형 문항비율을 2010년도 20%에서 2013년도 35~40%까지 점진적으로 확대하고자 하였다. 이후에는 2011학년도 경기도 초등학교 창의·서술형 평가 실시 요강(경기도 교육청, 2011)에서 단순 지식 암기 위주의 교육을 탈피하고 자기주도적 학습력·문제해결력 등 미래핵심역량 강화를 위해 서술형 평가 컨설팅단을 조직, 운영하고, 교과 전문성과 평가 전문성 함양을 위해 워크숍 및 2009 개정교육과정 적용에 따른 교과별 평가 방법을 개발하여 희망교, 희망 학년을 대상으로 해당 학년의 교육과정 범위 내 기본 원리에 대한 이해, 실생활 적용, 창의적 결론 도출 등의 능력을 고루 측정할 수 있는 문항을 출제하여 2011년도 11월에 평가를 실시하기도 하였고, 이는 2012년도 11월에도 계획되어 있다(경기도 교육청, 2012b). 최근 2012년도 경기교육 기본 계획(경기도 교육청, 2012a)에서는 서술형·논술형 평가를 확대하는 평가방법 혁신을 통해 창의지성교육으로 미래형 학력 신장을 6대 중점정책 중 하나로 꼽으며, 서술형 평가의 중요성을 강조하였다.

이외에도 여러 지역 교육청들은 서술형 평가의 확대를 위해, 문항을 개발하여 보급하거나 교사 연수를 실시하고 있다. 예를 들면, 지난 2007년 서울특별시교육청은 초등 학업성취도 수학과 평가 문항 개발 지침에서 결과보다는 과정을 중시하고 논리적으로 추론하여 문제를 해결할 수 있는 문항을 개발하도록 하며, 각 과목별로 예시문항과 채점기준을 제시하기도 하였고, 2010년에는 서울특별시교육연구정보원을 통해 ‘창의력과 표현력을 키워주는 초등 서술형 평가 장학자료집(3학년용, 4학년용)’을 배포하고, 단위별로 서술형 평가 예시문항과 채점기준, 유의사항을 제시하여 서술형 평가의 실시를 돕고자 하였다. 여기서 제시된 예를 보면, <표 III-1>처럼 총체적 채점 기준의 방법으로 각 문항에 대한 구체적인 채점기준을 제

시하며 하단에는 유의사항을 제시하여 교사에게 채점의 융통성을 부여하고 있다. 또한 대구 교육과학연구원(대구광역시교육청, 2011)에서는 2011학년도에 초등학교 5, 6학년 학업성취도 평가 문항을 개발하여 시내 전체 초등학교에 배부하고 홈페이지에 탑재하였는데, 전체를 서술형 평가 문항으로 개발하고 예시 답안 및 해설과 채점 기준을 제시하였다.

<표 III-1> 서울특별시교육청의 총체적 채점 기준 예시(서울특별시교육청, 2010b, p.96)

번호	채점기준	
	배점	특징
4	4	규칙을 찾아 설명하는 과정이 논리적이고 답이 맞는 경우
	3	규칙을 찾아 설명하는 과정이 논리적이거나 답은 틀린 경우
	2	답은 맞았으나 규칙을 찾아 설명하는 과정이 논리적이지 못함
	1	규칙을 찾아 설명하였으나 논리적이지 못하고 답도 틀린 경우
	0	무응답
유의사항	답을 구하는 것보다 규칙을 찾는 것을 중점으로 하는 문항이므로, 규칙을 찾는 과정이 논리적으로 기술되어 있는지 확인하는 것이 중요함	

강원도(강원도교육청, 2011)는 2011년도에 서술형 평가문항 컨설팅 연구회를 조직하여 각급 학교의 서술형 평가 문항을 분석하고 체계적인 평가 문항 관리 지원 시스템을 구축하여 교사의 서술형 평가문항 제작 전문성을 신장하고 우수한 평가모델을 개발 보급하고자 하는 노력을 펼치기도 하고, 서울특별시교육청은 2010년에는 2년간 서울특별시교육청 지정 연구학교를 통해 서술형 평가문항 개발 연구를 시행하기도 하였다.

2. 미국의 주(States)의 경우

미국은 국가 정책적 차원에서 교과 내용과 학생 수행 수준을 표준화하는 교육 개혁이 논의되면서 여러 주에서 수행평가 도입을 공식적으로 추진했는데, 1980년대 중반부터 학생의 쓰기 평가를 위하여 선택형 문항 대신에 학생이 직접 쓴 결과를 이용하였고, 1980년대 후반 1990년대 초에는 몇 개의 주들이 학생 수행에 근거한 평가의 개발과 도입을 추진하였다(서지영 외, 2008).

1992년 Wisconsin Performance Assessment Development Project[WPADP]를 시작으로 수행평가에 대한 연구가 본격적으로 시작되었다고 볼 수 있는데, 이 프로젝트에서는 수행평가를 ‘학습자들의 지식이나 기술을 증명하는 산물이나 해답을 학습자들이 창조해내도록 요구하는 평가방법’이라고 정의한다. 이 관점은 문제해결과정을 중시하여 평가한다(좌동지, 2001). 그리고 Alaska 주에서는 개념이해, 문제해결전략, 의사소통, 논리와 추론의 영역을 포함하여 수행 채점표를 만든 후, 학생들의 이해를 기준으로 그들의 용어로 재진술하여 학생들에게 제시하였다. 이를 통해 학생들은 자신의 모든 작업을 이해하였고, 비록 정확한 답을 하지 못하였더라도 점수를 얻을 수 있고 수학은 답이 전부가 아니라는 사실을 이해하였다(최승현, 1999).

이외에도 각 주별로 <표 III-2>에 제시한 바와 같이 문제해결과정을 평가하는 루브릭에 각 하위영역을 포함시킨 분석적 채점 기준표를 제시하거나, Maryland, Kentucky,

California, North Carolina, Maine 주(State)는 총체적 채점 기준표를 제시하였다(Chicago Public Schools, n.d.).

<표 III-2> 각 주별 문제해결 평가 기준 예시

분석적 채점	주 (State)	NCTM	Vermont	QUASAR	Illinois	Oregon
	하위 영역	문제이해 해결전략 찾기 답구하기	문제이해 문제해결과정 의사결정 수학적 추론	수학적 지식 전략적 지식 의사소통	수학적 지식 전략적 지식	개념적 이해 절차적 지식 문제해결전략 의사소통
영역별 척도		0-2	1-4	0-4	0-4	1,3,5
총체적 채점	주 (State)	Maryland	Kentucky	California	North Carolina	Maine
	영역별 척도	0-4	1-5	1-6	0-3	0-4

<표 III-2>에 제시된 예 중에서 NCTM의 분석적 채점 기준(analytical scoring)을 자세히 살펴보면 <표 III-3>과 같다(Charles, Lester & O'Daffer, 1987, p.44).

<표 III-3> NCTM의 분석적 채점 기준(Charles, Lester & O'Daffer, 1987, p.44)

평가기준	척도	특징
문제이해	2	문제를 완전히 이해함
	1	문제를 부분적으로 잘못 이해하거나 해석함
	0	문제를 완전히 잘못 이해함
해결전략 찾기	2	적절하게 실행하기만 하면 정답을 이끌어낼 수 있음
	1	문제를 올바르게 해석한 것에 기초하여 부분적으로 옳은 해결 전략을 세움
	0	문제해결 시도가 없거나, 완전히 부적절한 전략을 세움
답구하기	2	정답과 올바른 단위를 씀
	1	잘못 옮겨 쓰거나 계산의 오류, 답이 여러 개인데 답의 일부만 씀
	0	답을 쓰지 않았거나, 부적절한 전략에 따라 오답을 씀

<표 III-3>에 제시된 것과 같은 분석적 채점 기준은 평가기준의 하위 영역별로 척도를 정해 각 영역별로 점수를 부여하는 것으로, 각 영역별로 학생의 강점이나 취약점을 파악하기에 용이하다는 장점이 있다(Charles, Lester & O'Daffer, 1987). 이러한 장점은 서술형 평가의 결과를 즉각적으로 교수·학습 과정에 반영시켜 학습자의 학습의 이해를 높이는 데 도움을 줄 수 있을 것이다. 그러나 각 학습자의 응답에 대한 채점 결과 각 영역별 점수가 다르더라도 총점이 같을 수 있기 때문에 학습자 사이의 점수를 비교할 때에는 매우 신중해야 한다(Charles, Lester & O'Daffer, 1987).

<표 III-2>에 제시된 예 중에서 California의 총체적 채점 기준(holistic scoring)을 자세히 살펴보면 <표 III-4>와 같다(Chicago Public Schools, n.d.).

<표 III-4> California의 총체적 채점 기준(Chicago Public Schools, n.d.)

척도	특징
6	예시답안: 명확한 설명과 함께 완전한 정답을 제시함 문제 상황에 제시된 수학적 개념을 이해함 문제 상황에 제시된 중요한 정보들을 인식함 풀이과정을 타당하게 논증함
5	유능한 답안: 합리적인 설명과 함께 상당히 완전한 답을 제시함 문제 상황에 제시된 수학적 개념을 이해함 문제 상황에서 제시하는 중요한 정보들의 대부분을 인식함 풀이과정을 비교적 타당하게 논증함
4	사소한 오류가 포함된 만족스러운 답안: 문제를 해결하였지만 설명이 약간 부족함 풀이과정을 타당하게 논증하지 못함 수학적 개념을 이해하고 효과적으로 사용하기 위해 노력함
3	심각한 오류가 포함되어 만족스럽지 못한 답안: 문제해결 전략을 적절하게 실행하지 못해 문제를 해결하지 못함 문제해결에 요한 수학적 개념을 완전히 이해하지 못함 수학적 용어를 잘못 사용하거나 치명적인 계산 실수가 있음
2	문제해결을 시도하였으나 해결하지 못한 답안: 문제 해결 과정의 설명을 이해할 수 없음 문제 상황을 전혀 이해하지 못함
1	문제해결을 시도하지 못한 답안: 문제 상황을 전혀 이해하지 못함 문제 상황을 적었으나 문제 해결을 이어지지 못함

<표 III-4>에 제시된 것과 같은 총체적 채점 기준은 과제에 대한 응답을 전체적으로 판단하여 하나의 점수를 부여한다. 이는 수행과정의 부분적인 측면에 대한 정보보다는 전체를 중시하는 방법으로 문제해결과정과 답을 전체적으로 바라보는 관점을 취한다. 이는 결과를 중시했던 과거의 평가 방법과 달리 과정을 중시하는 요즘의 평가 방향에 같은 맥락으로 적용가능하며, 학습자로 하여금 문제해결 후 문제해결과정을 반성해보는 단계까지 확장시켜 지도하는데 도움을 줄 수 있을 것이다.

한국의 시·도 교육청과 미국의 각 주별로 서술형 평가를 시행해 온 발자취를 보면, 예시 문항과 함께 평가기준이나 채점기준들이 함께 제시되고 있다. 특히, 평가 기준 및 채점기준에 포함된 하위영역이나 척도별 특징을 살펴보면, 정답 자체보다는 정답을 구하는 과정에서 수학적 개념의 이해 정도나 타당하게 추론하는 능력, 수학적 사고 과정을 의사소통하는 능력 등의 문제해결과정이 중시되고 평가의 대상이 되고 있음을 알 수 있다. 또한 두 나라 모두 지역의 자율성에 기초하여 자율적으로 시행하도록 하고 있다는 공통점을 찾아볼 수 있는데, 이와 같이 지역의 자율성에 기초하는 것은 학교와 지역의 상황을 최대한 고려하여 평가를 형식적으로 행하기보다는 의미 있게 활용하려는 의도로 해석할 수 있겠다. 그러나 한국의 경우, 수학 교과목의 평가에서 서술형 문항의 포함 정도나 문항 형태가 문서상에서 강조되고 있고, 실제적인 문항이나 채점기준은 학년별, 단원별, 문항별로 구체적으로 제시되어 그

단원에 대한 활용이 아니면 활용하는 데에 제한이 있는 것에 반해, 미국의 각 주는 서술형 문항을 채점할 때 고려할 평가기준을 일반적인 관점에서 제시한 후 구체적인 예시들이 함께 제공되고 있어, 교사의 입장에서는 변형하여 활용하기가 더욱 쉬울 것으로 생각된다.

IV. 학교 수준에서 활용되고 있는 서술형 평가 문항 비교

앞서 문헌상에 밝혀진 바에 따르면 공통적으로 서술형 평가가 과정 중심의 평가 기법이라는 측면에서 필요와 요구가 증대되고 있음을 알 수 있다. 이렇게 문헌상에서 밝혀진 바가 실제로 학교에서 활용하고 있는 서술형 평가 문항이 학생들의 문제해결과정을 평가할 수 있도록 구성되어 있는지를 중심으로 살펴보기 위해, 두 나라의 학교 수준에서 활용하고 있는 서술형 평가의 예시문항과 채점기준을 예시로 제시하여 비교하고자 한다. 한국의 경우, 서울의 C초등학교에서 5학년 2학기 '6. 자료의 표현과 해석' 단원에서 실제로 활용했던 문항이고, 그 문항에 대한 학생들의 답안을 예시로 제시하여 어떻게 채점하였는지를 제시하였다. 미국의 경우, 실제 학교에서 활용하고 있는 문항을 구하기 어려워 Everyday Mathematics Grade 5 - Assessment Handbook(The University of Chicago School Mathematics Project, 2007)에 제시된 문항을 예시로 활용하였다.

1. 한국의 서술형 평가 문항 및 평가 기준

<표 IV-1>은 서울의 C초등학교 5학년 2학기 '6. 자료의 표현과 해석' 단원에서 실제로 활용했던 서술형 평가 예시문항과 평가 기준이다.

<표 IV-1> 서울 C초등학교의 서술형 평가 예시 문항과 평가기준

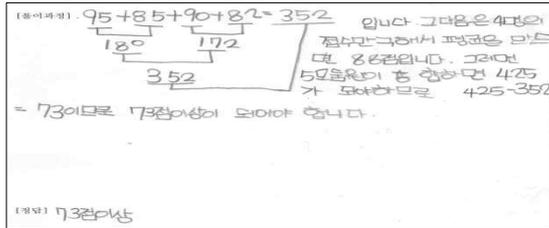
성취 기준	평균의 의미를 이해하고 활용하여 문제를 해결할 수 있다.																
문항 예시	창수네 모둠원들의 10월 수학점수는 다음과 같습니다.																
	<table border="1" style="margin-left: auto; margin-right: auto;"> <tr> <td>모 둠 원</td> <td>창 수</td> <td>아 영</td> <td>선 진</td> <td>영 훈</td> <td>현 욱</td> </tr> <tr> <td>점수</td> <td>95</td> <td>85</td> <td>90</td> <td>82</td> <td>60</td> </tr> </table>						모 둠 원	창 수	아 영	선 진	영 훈	현 욱	점수	95	85	90	82
모 둠 원	창 수	아 영	선 진	영 훈	현 욱												
점수	95	85	90	82	60												
	창수네 모듬은 다음달 11월 평가에서 평균을 85점으로 올리려고 합니다. 현욱이의 점수를 올려 평균을 높이려고 할 때, 현욱이의 성적은 몇 점 이상이 되어야 합니까? (다른 모듬원들의 10월 성적은 유지한다고 가정합니다) (풀이과정을 쓰고 답을 쓰시오.)																
채점 기준	배점	채 점 기 준															
	A 매 우	<ul style="list-style-type: none"> · 문제 상황 및 문제 속에 내포된 평균의 개념을 바르게 이해함 · 제시된 표에서 문제해결에 적절한 정보를 선택하여 활용함 (모듬원들의 수학점수를 모두 활용, 평균 구하는 개념을 활용) · 다음의 2가지 전략 중 한 가지를 바르게 기술함 															

<p>잘 합</p>	<p>1) 모든 모둠원들의 점수를 합하면(95+85+90+82+60=412) 총 412점이 나오는데, 평균이 85점이 되려면 총 425점이 필요하여, 필요한 점수가 13이상이 되므로 73점 이상이 된다는 것을 기술하는 전략(총점 이용)</p> <p>2) 모든 모둠원들의 점수를 합해서 5로 나누면 평균이 82.4이고, 85점 이상이 되려면 2.6점 이상을 올리면 되므로 2.6과 모둠원의 수인 5를 곱하여 13을 구하여 73점 이상으로 기술하는 전략(평균 이용)</p> <ul style="list-style-type: none"> · 답을 알아내는 모든 단계나 과정을 명확하고 논리적으로 충분히 설명함 · 문제해결과정에 사용된 수학적 용어 및 기호 등의 표현이 정확함
<p>B 잘 합</p>	<ul style="list-style-type: none"> · 문제상황 및 문제 속에 내포된 평균의 개념을 바르게 이해함 · 제시된 표에서 문제해결에 적절한 정보를 선택하여 활용함(모둠원들의 수학점수를 모두 활용, 평균 구하는 개념을 활용) · A단계에서 제시된 전략을 찾아 실행하고 있으나 답이 부정확함 <p>a) 사소한 오차나 옮겨 쓰기 과정에 오류가 있음</p> <p>b) 답을 쓰지 않음</p> <ul style="list-style-type: none"> · 답을 알아내는 전체 과정을 논리적으로 설명하고 있으나 약간의 비약을 포함하고 있음 · 문제해결과정에 사용된 수학적 용어 및 기호 등의 표현에서 약간의 오류를 나타냄
<p>C 보 통</p>	<ul style="list-style-type: none"> · 문제 상황 및 문제 속에 내포된 평균의 개념을 부분적으로 이해함 · 문제의 조건을 바르게 이해하지 못하여 문제해결에 적절한 정보 중 일부만을 바르게 선택하여 활용함 <p>a) 문제에서 10월 성적은 유지한다고 가정하였으나 이러한 조건을 바르게 이해하지 못함</p> <p>b) 현욱이의 성적을 구하지 않고 다른 학생의 성적 향상 점수를 구하였음</p> <ul style="list-style-type: none"> · A단계에서 제시한 전략을 찾아 시도하고 있으나 끝까지 실행하지 못하여 정답에 이르지 못함 <p>a) 모든 모둠원들의 점수를 합산하여야 하는데, 일부 모둠원의 점수만 합산하였음</p> <p>b) 총점을 구했으나, 몇 점을 올려야 하고, 현욱이의 점수가 몇 점 이상이 되어야 하는지 바르게 구하지 못함</p> <ul style="list-style-type: none"> · 답을 알아내는 과정 설명에 비약이 많거나, 문제해결과정 없이 정답(73점 이상)만 기술함 · 문제해결과정에 사용된 수학적 용어 및 기호 등의 사용에 오류 및 부정확한 표현이 많이 포함됨
<p>D 노 력 요 합</p>	<ul style="list-style-type: none"> · 문제 상황 및 문제 속에 내포된 평균의 개념을 이해하지 못함 · 문제의 조건을 이해하지 못하여 문제해결과 관련 없이 제시된 기록들을 활용함 · 문제해결과정에 대한 설명이 불완전하거나 명료하지 않아 이해하기 어려움 · 문제해결을 전혀 시도하지 못함 <p>a) 백지 또는 오답 이외에 아무것도 쓰지 않음</p> <p>b) 문제와 관련없는 것만을 기술함</p>

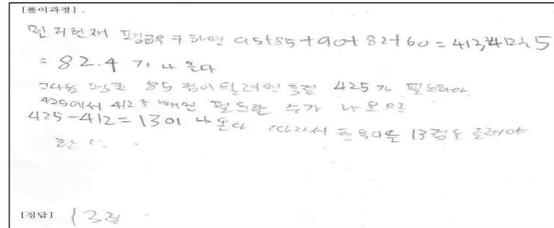
서울 C초등학교에서 활용하고 있는 이 문항은 문제 상황을 해결하기 위해 평균의 개념을 이해하고 적용해야 한다. 평균의 개념을 이해한다는 것은 평균 구하는 계산을 할 수 있어야 할 뿐만 아니라 평균이 적용되는 상황에 평균의 개념을 어떻게 적용할지를 알아야 하는 것이다. 그러나 위와 같은 성취기준(학습목표)을 평가하기 위해 과거에 주로 사용했던 평가 문항들은 ‘평균을 구하시오’와 같이 문항을 제시하고 평균을 구하는 계산에 초점을 두었던 것과는 달리, 위의 예시와 같은 서술형 평가 문항에서는 풀이과정을 서술하도록 문제에 명시하고 이해한 바를 적용하는 과정이 평가의 대상이 된다.

그리고 채점기준은 문제를 해결하는데 필요한 수학적 개념에 관한 부분, 문제를 해결하기 위해 문제에서 제시된 정보를 활용하는 정도, 문제 해결 전략을 실행하는 부분, 자신의 풀이과정을 서술하는 과정에서 논리성과 표현의 정확성에 관한 부분의 정도에 따라 1-4점으로

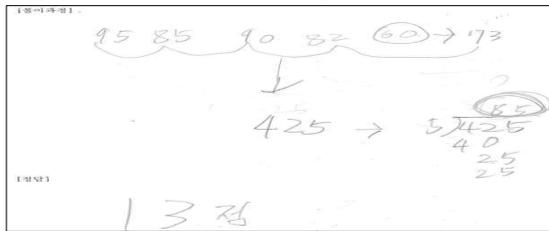
나누어 제시하는 총체적 채점 방법을 활용하고 있다. 평가 기준과 채점기준을 학생들의 문제해결과정 평가를 위해 어떻게 활용하였는지 척도별 학생 응답 예시를 제시하여 살펴보고자 한다.



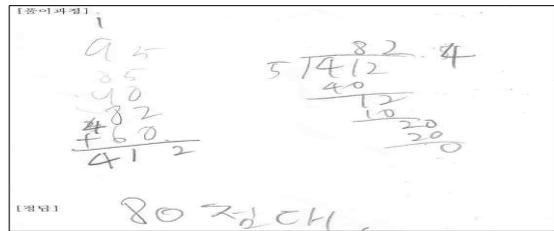
[그림 IV-1] A(매우잘함) 단계의 학생 응답 예시



[그림 IV-2] B(잘함) 단계의 학생 응답 예시



[그림 IV-3] C(보통) 단계의 학생 응답 예시



[그림 IV-4] D(노력요함) 단계의 학생 응답 예시

[그림 IV-1]은 <표 IV-1>에서 제시한 문항의 A(매우잘함) 단계에 해당하는 학생 응답 예시인데, 문제를 해결하기 위해 활용해야 하는 수학적 개념(평균)에 기초하고 문제 상황에서 제시하는 정보 중에서 필요한 것을 골라내어 선택한 전략에 따라 명확하게 풀이과정을 제시하였기에 A(매우잘함)으로 채점하였다. [그림 IV-2]는 [그림 IV-1]에 나타난 예시와 비교할 때 문제 해결에 필요한 수학적 개념(평균)을 이해하고 적절한 전략을 찾아 실행하고 있으나 답이 부정확하기에 B(잘함)으로 채점하였다. [그림 IV-3]은 문제 해결에 필요한 수학적 개념(평균)을 활용하여 적절한 전략을 선택하여 문제 해결을 시도한 것으로 보이나 답을 알아내는 과정에 대한 설명이 부족하여 문제 해결 과정과 정답과의 관계 이해하기에 어려우며, 정답도 부정확하기에 C(보통) 단계로 채점하였다. [그림 IV-4]는 문제 해결에 필요한 수학적 개념(평균)을 활용하려고 노력은 하였으나 문제에서 요구하는 사항이 무엇인지 제대로 파악하지 못한 것으로 보이며, 답도 부정확하기에 D(노력요함) 단계로 채점하였다.

2. 미국의 서술형 평가 문항 및 평가 기준

<표 IV-2>는 미국의 Everyday Mathematics Grade 5 - Assessment Handbook(The University of Chicago School Mathematics Project, 2007)에 제시된 문항의 예시이다. 이 문제는 풀이과정과 왜 그렇게 풀었는지 학생의 생각을 서술하도록 문제에 명시함으로써 학생들에게 풀이과정을 쓰도록 하였다. 그리고 위에는 제시하지 않았지만 채점기준 이외에 예시답안을 제시하였고, 총체적 채점 방법을 활용하고 있다. 이 문제는 우리나라 교과서에는 수식만을 제시한 것과 달리 문제가 발생된 배경까지 함께 포함되어 있다. 이는 김민경 외

(2011)에서 우리나라 수학 교과서에 제시된 문제가 실세계와 관련되기보다 단순 적용의 문제로 제시되는 한계점을 안고 있다고 지적한 부분에서도 드러난 바와 같이 우리나라 교과서 문제가 가지고 있는 한계로 볼 수 있다.

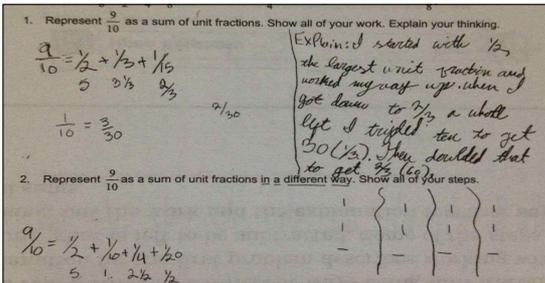
<표 IV-2> Everyday Mathematics Grade 5 - Assessment Handbook에 제시된 서술형 평가 예시 문항과 평가기준(UCSMP, 2007, pp. 112-114)

성취 기준	단위분수와 단위분수의 합을 이용하여 분수를 다른 방법으로 표현할 수 있다.								
문항 예시	<p>5학년 학생들이 분수에 관한 조사를 통해, 고대 이집트 사람들은 모든 분수를 반복되지 않는 단위분수의 합으로 나타냈다는 것을 발견했다. 예를 들면, $\frac{3}{4} = \frac{1}{4} + \frac{1}{4} + \frac{1}{4}$로 나타낼 수 없고, $\frac{1}{4} + \frac{1}{2}$로 나타낼 수 있다.</p> <p>이렇게 어떤 분수를 단위분수의 합으로 나타내는 알고리즘을 찾는 방법을 ‘Greedy Method’라고 한다.</p> <p>분수 X를 Greedy Method로 나타내기 위해서는 X보다 작은 분수 중에서 가장 큰 단위분수로 시작한 후, 나머지 값을 같은 방법을 반복적으로 적용하면 된다.</p> <p>예를 들어, $\frac{7}{8}$을 단위분수의 합으로 나타내기 위해서는</p> <p>첫째, $\frac{7}{8} = \frac{1}{2} + \frac{3}{8}$ [$\frac{1}{2}$은 $\frac{7}{8}$보다 작은 분수 중에 가장 큰 단위분수임]</p> <p>둘째, $\frac{7}{8} = \frac{1}{2} + \frac{1}{4} + \frac{1}{8}$ [$\frac{1}{4}$은 $\frac{3}{8}$보다 작은 분수 중에 가장 큰 단위분수임]</p> <p>(1) $\frac{9}{10}$를 단위분수의 합으로 나타내시오. 풀이과정을 쓰고, 왜 그렇게 풀었는지 설명하시오.</p> <p>(2) $\frac{9}{10}$를 다른 방법을 이용하여 단위분수의 합으로 나타내시오. 풀이과정을 함께 쓰시오.</p>								
채점 기준	<table border="1"> <thead> <tr> <th data-bbox="319 1265 383 1310">배점</th> <th data-bbox="383 1265 1345 1310">채점 기준</th> </tr> </thead> <tbody> <tr> <td data-bbox="319 1310 383 1444">4</td> <td data-bbox="383 1310 1345 1444">합이 $\frac{9}{10}$가 되도록 2가지 방법의 단위분수의 합으로 나타냄 단위분수의 합으로 나타낼 때 단위분수를 반복해서 사용하지 않음 어떤 단위분수의 조합이 $\frac{9}{10}$를 만드는지 찾아낸 방법을 완벽하게 설명함</td> </tr> <tr> <td data-bbox="319 1444 383 1601">3</td> <td data-bbox="383 1444 1345 1601">합이 $\frac{9}{10}$가 되도록 적어도 1가지 방법의 단위분수의 합으로 나타냄 단위분수의 합으로 나타낼 때 단위분수를 반복해서 사용하지 않음 어떤 단위분수의 조합이 $\frac{9}{10}$를 만드는지 찾아낸 방법을 설명하였으나, 약간 보충이 필요함 2번 문제는 문제해결을 시도하였으나 완벽하지 못하고 오류를 포함함</td> </tr> <tr> <td data-bbox="319 1601 383 1794">2</td> <td data-bbox="383 1601 1345 1794">합이 $\frac{9}{10}$가 되도록 1가지 방법의 단위분수의 합으로 나타냄 단위분수의 합으로 나타낼 때 단위분수를 반복해서 사용하지 않음 약간의 오류가 있음 어떤 단위분수의 조합이 $\frac{9}{10}$를 만드는지 찾아낸 방법을 설명하려고 시도하나, 설명이 부족함</td> </tr> </tbody> </table>	배점	채점 기준	4	합이 $\frac{9}{10}$ 가 되도록 2가지 방법의 단위분수의 합으로 나타냄 단위분수의 합으로 나타낼 때 단위분수를 반복해서 사용하지 않음 어떤 단위분수의 조합이 $\frac{9}{10}$ 를 만드는지 찾아낸 방법을 완벽하게 설명함	3	합이 $\frac{9}{10}$ 가 되도록 적어도 1가지 방법의 단위분수의 합으로 나타냄 단위분수의 합으로 나타낼 때 단위분수를 반복해서 사용하지 않음 어떤 단위분수의 조합이 $\frac{9}{10}$ 를 만드는지 찾아낸 방법을 설명하였으나, 약간 보충이 필요함 2번 문제는 문제해결을 시도하였으나 완벽하지 못하고 오류를 포함함	2	합이 $\frac{9}{10}$ 가 되도록 1가지 방법의 단위분수의 합으로 나타냄 단위분수의 합으로 나타낼 때 단위분수를 반복해서 사용하지 않음 약간의 오류가 있음 어떤 단위분수의 조합이 $\frac{9}{10}$ 를 만드는지 찾아낸 방법을 설명하려고 시도하나, 설명이 부족함
배점	채점 기준								
4	합이 $\frac{9}{10}$ 가 되도록 2가지 방법의 단위분수의 합으로 나타냄 단위분수의 합으로 나타낼 때 단위분수를 반복해서 사용하지 않음 어떤 단위분수의 조합이 $\frac{9}{10}$ 를 만드는지 찾아낸 방법을 완벽하게 설명함								
3	합이 $\frac{9}{10}$ 가 되도록 적어도 1가지 방법의 단위분수의 합으로 나타냄 단위분수의 합으로 나타낼 때 단위분수를 반복해서 사용하지 않음 어떤 단위분수의 조합이 $\frac{9}{10}$ 를 만드는지 찾아낸 방법을 설명하였으나, 약간 보충이 필요함 2번 문제는 문제해결을 시도하였으나 완벽하지 못하고 오류를 포함함								
2	합이 $\frac{9}{10}$ 가 되도록 1가지 방법의 단위분수의 합으로 나타냄 단위분수의 합으로 나타낼 때 단위분수를 반복해서 사용하지 않음 약간의 오류가 있음 어떤 단위분수의 조합이 $\frac{9}{10}$ 를 만드는지 찾아낸 방법을 설명하려고 시도하나, 설명이 부족함								

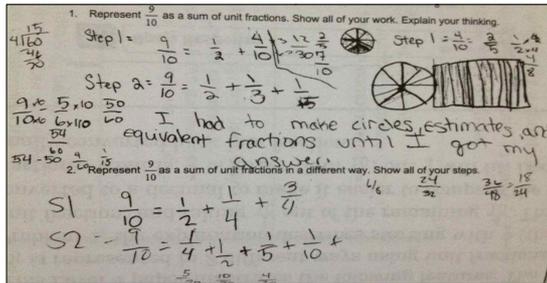
한국과 미국의 초등수학 서술형 평가의 제도 및 문항 비교

	고 오류를 포함함
1	합이 $\frac{9}{10}$ 가 되도록 1가지 방법의 단위분수의 합으로 나타냄 어떤 단위분수의 조합이 $\frac{9}{10}$ 를 만드는지 찾아낸 방법을 설명하려고 시도하나, 그 설명이 문제 상황과 연결되지 않음
0	문제해결을 시도하지 않음

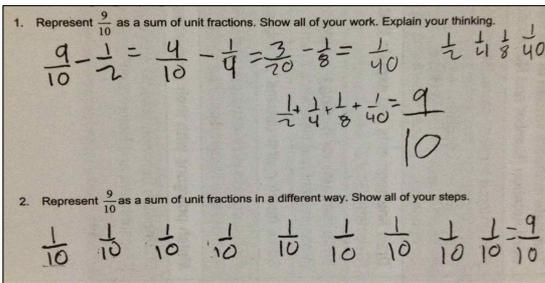
<표 IV-2>에 제시된 채점기준이 실제적인 채점에 활용된 예시를 살펴보면 다음과 같다. [그림 IV-5]는 2가지 방법의 단위분수의 합으로 문제를 해결하였고 문제 상황에서 제시한 Greedy Method를 문제 해결에 어떻게 적용하였는지 잘 설명하였기에 4점으로 채점하였다. [그림 IV-6]은 [그림 IV-5]처럼 2가지 방법의 단위분수의 합으로 문제를 해결하였으나 자신의 문제 해결 방법을 설명한 것에 보충이 필요하기 때문에 3점으로 채점하였다. [그림 IV-7]은 2가지 방법으로 단위분수의 합으로 문제를 해결하려고 시도하였으나 (2)번에서 제시한 답은 문제의 요구 사항에 맞지 않으며, 자신의 문제 해결 과정을 설명하지 못하였기에 2점으로 채점하였고, [그림 IV-8]은 단위분수는 분자가 1인 분수라는 개념은 이해했으나 분수의 개념을 문제 해결에 어떻게 활용하는지 전혀 설명이 이루어지지 않아 1점으로 채점하였다.



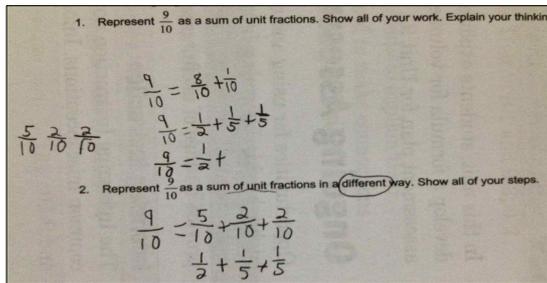
[그림 IV-5] 4점의 학생 응답 예시
(UCSMP, 2007, p.115)



[그림 IV-6] 3점의 학생 응답 예시
(UCSMP, 2007, p.116)



[그림 IV-7] 4점의 학생 응답 예시
(UCSMP, 2007, p.115)



[그림 IV-8] 1점의 학생 응답 예시
(UCSMP, 2007, p.117)

한국과 미국의 학교 수준에서 활용하고 있는 서술형 평가 문항과 평가기준의 예시를 분석한 결과, 복잡하거나 새로운 상황에 수학적 개념이 내재되어 있고 다양한 해결 전략을 활용할 수 있는 좋은 문항이 평가 문항이 되어야 한다고 제시된 것처럼 평가에서 활용하고 있는 문항의 질과 형태가 중요함을 알 수 있다.

또한 채점기준을 적용한 사례를 보면, 학생들의 문제해결과정을 평가의 대상으로 삼음으로써 각 척도별 특징에 따라 학생들에게 개별화 또는 차별화된 도움을 줄 수 있고 이는 학생들의 수학적 개념의 이해도를 높일 것으로 생각된다. 따라서 평가와 교수·학습의 연계성도 높아지게 될 것이다.

V. 마치며

본 연구는 학생들의 창의성 및 문제해결력 신장과 함께 기본교육 내실화를 위한 방안 중 하나로 실행되고 있는 서술형 평가가 국가, 교육청, 학교 수준에서 실행되고 있는 실재를 한국과 미국의 경우를 예로 들어 비교하고, 계속적인 필요와 요구에 의해 확대 실시되고 있는 서술형 평가의 효율적인 적용을 위한 시사점을 찾아보고자 하였다.

첫째, 한국과 미국 NCTM의 수학 교과 교육과정의 평가에서 지향하는 바를 분석한 결과, 모두 결과뿐만 아니라 과정 중심의 평가를 지향하고 있고 교수 학습의 개선을 위해 평가 결과를 활용할 것을 강조하고 있다. 이를 위해 시·도 교육청에서는 학교 수준에서의 서술형 평가 확대를 돕기 위해 교사 연수를 실시 또는 교사 자치 활동을 장려하거나, 문항을 개발하여 보급하고 있기는 하지만, 김민경 외(2012)에서 제안한 바처럼 과정 중심의 평가가 수학 교육의 궁극적 목표인 수학적 사고력이나 창의력, 문제해결력 등과 같은 수학적 힘의 신장을 위한 방법으로 실제적으로 행해지고 있는지, 정현도, 강신포, 김성준(2010)이 밝힌 바와 같이 평가의 결과가 교수 학습의 개선을 위한 구체적인 방향을 제시하고 있는지, 서술형 평가의 의도에 맞추어 교실 현장에서 서술형 평가를 시행하고 있는지, 서술형 평가를 통해 평가와 교수·학습 과정 간의 연계를 높이기 위한 노력을 교사들이 실제적으로 펼치고 있는지 등에 관한 심층적인 연구를 진행하고 연구 결과를 수학 교육 현장에 활용함으로써 평가 전문성을 키워야 할 것이다.

둘째, 한국과 미국의 서술형 평가 활용 실태를 교육청 수준에서 비교한 결과, 한국은 서술형 평가의 실시를 확대하기 위해 문서상에서 평가에서 포함해야 하는 서술형 평가의 비율 정도에 그치지만, 미국은 주(States)별로 문제 해결 과정을 채점하기 위한 구체적인 루브릭을 제시하여 각 학교들이 활용하기 쉽도록 돕고 있다. 물론, 최근 한국에서도 서울특별시교육청과 서울특별시교육연구정보원(<http://www.serii.re.kr>)이 홈페이지와 학교로 배부된 책자를 통해 창의적인 사고력과 문제해결력을 키워주는 교육 자료를 지원하며 서술형 평가를 확대코자 학년별, 과목별로 다양한 수준의 수행평가 및 서술형 평가 문항을 제작하여 보급하고는 있지만, 여전히 서술형 평가를 위한 문항 및 평가기준의 개발과 보급에 대한 현장의 요구가 높은 실정이다(김민경 외, 2012; 노선숙 외, 2008; 도주원 외, 2009). 서울특별시교육청의 연구 지정학교에서 서술형 평가 연구를 수행한 한 교사가 언급하였듯이(김민경 외, 2012), 전문 기관에서 지속적으로 문항을 개발하여 문항의 성취 기준이나 난이도에 대한 객관적인 정보와 교사들이 쉽게 접근할 수 있는 접근성, 문항지 만들기 등의 편리한 기능 등

을 함께 제시하여 교사들의 서술형 평가에 대한 접근 용이성과 활용 가능성을 높임으로써 서술형 평가의 확대를 도울 수 있어야 할 것이다. 따라서 우리나라에서도 서술형 평가의 비중과 같은 실시 방침만을 밝히는 행정적 수준에서 벗어나 각 학교에서 실제 실행하는데 직접적으로 도움이 되는 구체적인 방안을 마련하여 제시해줌으로써 학교별로 다양한 수준과 형태로 실행하고 있는 서술형 평가가 신뢰성 있는 평가로 자리 잡을 수 있도록 도와야 할 것이다.

참고문헌

- 강원도교육청(2011). 강원도 서술형 평가문항 컨설팅 연구회 연구보고서. Retrieved from 강원도교육청 홈페이지, <http://www.gwe.go.kr>
- 경기도 교육청(2010). 학생 창의력 신장과 교사 평가전문성 향상을 위한 평가방법 혁신 기본 계획. Retrieved from 경기도 교육청 홈페이지, <http://www.goe.go.kr>
- 경기도 교육청(2011). 2011학년도 경기도 초등학교 창의·서술형 평가 실시 요강. Retrieved from 경기도 교육청 홈페이지, <http://www.goe.go.kr>
- 경기도 교육청(2012a). 2012 경기교육 기본계획. Retrieved from 경기도 교육청 홈페이지, <http://www.goe.go.kr/>
- 경기도 교육청(2012b). 2012년 경기도 초등학교 서술형·논술형 평가 시행 계획. Retrieved from 경기도 교육청 홈페이지, <http://www.goe.go.kr>
- 교육과학기술부 (2008). 초등학교 교육과정 해설 IV-수학, 과학, 실과. 서울: 교육과학기술부.
- 교육과학기술부 (2009). 2009 개정교육과정-수학과 교육과정. 서울: 교육과학기술부.
- 김경희, 김수진, 김남희, 박선용, 김지영, 박효희, 정송 (2008). 수학·과학 성취도 추이변화 국제비교 연구-TIMSS 2007 결과보고서. 서울: 한국교육과정평가원.
- 김경희, 시기자, 김미영, 옥현진, 임해미, 김선희, 정송, 정지영, 박희재 (2010). OECD 학업성취도 국제비교 연구(PISA 2009) 결과 보고서-연구보고 RRE 2010-4-2. 서울: 한국교육과정평가원.
- 김민경, 조미경, 주유리 (2012). 서술형 평가에 대한 인식 및 실태에 관한 조사연구-서울시 소재 초등교사를 중심으로. 한국초등수학교육학회지, 16(1), 63-95.
- 김수환, 박성택, 신준식, 이대현, 이의원, 이종영, 임문규, 정은실 (2009). 초등학교 수학교육론. 서울: 동명사.
- 김주환 (2010). 서술형 평가 Road View-서술형 평가 문항 제작, 컨설팅 그리고 채점까지. 서울: 박이정.
- 노국향, 박정 (2001). 문항의 형태와 배점에 따른 검사 정보의 비교. 교육평가연구, 14(2), 173-191.
- 노선숙, 김민경, 조성민, 정연숙, 정윤아 (2008). 중등 수학과 서술형 평가의 현황 분석 연구. 한국학교수학회논문집, 11(3), 377-397.
- 대구광역시교육청(2011). 대구교육과학연구원, 서술형 평가 문항 제공으로 교사는 업무 경감 학생은 사고력 개발. Retrieved from 대구광역시교육청 홈페이지, <http://www.dge.go.kr>
- 도주원, 오지연, 공정인, 주미정, 김미영, 이대현, 박만구 (2009). 초등 교사들의 수학과 서술

- 형 평가에 대한 인식 및 실태. *초등수학교육*, 12(2), 63-80.
- 박경미, 김동원 (2011). 우리나라 수학교육의 문제점 진단을 위한 조사 연구. *수학교육*, 50(1), 89-102.
- 박금란, 방정숙 (2008). 서술형 평가를 통한 초등학교 6학년 학생들의 수학과 기본 지식 이해에 관한 실태 연구. *수학교육*, 47(2), 181-195.
- 박정, 박경미 (2004). 수행형 문항과 선다형 문항의 수학적 능력 추정 효율성 비교. *수학교육*, 43(2), 151-162.
- 백순근 (2000). 수행평가의 원리. 서울: 교육과학사.
- 서울특별시교육청 (2007). 학습의 질 관리를 위한 초등 학업성취도 평가 문항 개발-초등 교원 평가전문성 신장 연수 교재. 서울특별시교육청.
- 서울특별시교육청 (2010a). 창의성 계발을 위한 평가 개선 기본 계획. Retrieved from 서울특별시교육청 홈페이지, <http://www.sen.go.kr>
- 서울특별시교육청(2010b). 창의성과 표현력을 키워주는 초등 서술형 평가 장학자료집(4학년용). Retrieved from 서울특별시교육연구정보원 홈페이지, <http://www.serii.re.kr>
- 서울특별시교육청 (2011). 2011 주요업무계획. Retrieved from 서울특별시교육청 홈페이지, <http://www.sen.go.kr>
- 서지영, 김소영, 신명선, 홍수진 (2008). 학교교육 내실화를 위한 수행평가 개선 연구(I) (연구보고 RRE 2008-1). 서울: 한국교육과정평가원.
- 성태제 (1999). 교육평가 방법의 변화와 결과타당도에 의한 고려. *교육학연구*, 37(1), 197-218.
- 성태제 (2000). 초등교육 수행평가의 장애요인 분석과 개선안. *교육학 연구*, 38(1), 153-184.
- 신현석, 박균열, 엄준용 (2008). 창의적 인재양성을 위한 교육체제의 구축. *인력개발연구*, 10(2), 83-108.
- 이대현, 박배훈 (1999). 서술형 수행평가의 개념과 활용에 관한 연구. 추계 대한수학교육학회 연구발표대회 논문집, 179-192.
- 정동권, 송상현, 김홍구, 이용길, 김성만, 정주자, 안승학, 박정수 공저 (2002). 제7차 교육과정 수행평가 시리즈-수학과 수행중심 평가. 서울: 학문출판.
- 정현도, 강신포, 김성준 (2010). 초등수학 서술형 평가에서 나타나는 오류 분석. *한국초등수학교육학회지*, 14(3), 885-905.
- 좌동지 (2001). 서술형 평가가 수학적 성향 및 문제해결력에 미치는 효과-초등학교 수학과를 중심으로. 부산대학교 교육대학원 석사학위논문.
- 최승현 (1999). 외국에서의 수학과 수행평가: 수행평가 현장 정착을 위한 세미나 자료집-초·중등학교 교과별 수행 평가의 실제(5) 수학(pp.25-58). 서울: 한국교육과정평가원.
- 허경철, 백순근, 박경미, 최미숙, 양길석, 김광주 (1999). 수행평가 정책 시행 실태 분석과 개선 대책 연구 (연구보고 CRE 99-2). 서울: 한국교육과정평가원.
- 황혜정 (2003). 수학과 수행평가에 관한 이해의 혼돈. *수학교육*, 42(2), 159-176.
- Bell, M. & Isaacs, A. (2007). The case of Every Mathematics. In C. R. Hirsch (ed), *Perspectives on the design and development of school mathematics curricula*. Reston, VA: NCTM.
- Black, P., Hariison, C., Lee, C., Marashall, B., & Wiliam, D. (2003). *Assessment for learning: putting it into practice*. Berkshire: Open University Press.

- Brookhart, S. M. (2005). Research on formative classroom assessment. In Paper presented in the symposium, formative classroom assessment: research, theory, and practice, at the annual meeting of the American Educational Research Association Montreal, Canada.
- Charles, R., Lester, F., & O'Daffer, P. (1987). How to evaluate progress in problem solving. Reston: VA: National Council of Teachers of Mathematics. 강완, 김진호, 신혜진(역)(1997). 문제 해결 과정의 평가 기법. 서울: 동명사.
- Chicago Public Schools (n.d.). Retrieved from Office of Performance website, <http://research.cps.k12.il.us/cps/accountweb/Assessment/Resources/IdeasandRubrics/RubricBank/>
- Idris, N. (2007). Classroom assessment in mathematics education. Kuala Lumpur: McGraw-Hill Education.
- Lappan, G., Phillips, E. D., & Fey, J. T. (2007). The case of Connected Mathematics. In C. R. Hirsch(Ed.), Perspectives on the design and development of school mathematics curricula pp.67-80. Reston, VA: NCTM.
- Leahy, S., Lyon, C., Thompson, M., & William, D. (2005). Classroom assessment: Minute by minute, day by day. Educational Leadership, 63(3), 18-24.
- National Council of Teachers of Mathematics (1995). Assessment standards for school mathematics. Reston, VA: Author.
- National Council of Teachers of Mathematics (2000). Principles and standards for school mathematics. Reston, VA: Author.
- National Council of Teachers of Mathematics (2001). Mathematics assessment: Cases and discussion questions for grades K-5. Reston, VA: Author.
- National Council of Teachers of Mathematics (2005). Mathematics assessment sampler : grades 3-5. Reston, VA: Author.
- The University of Chicago School Mathematics Project (2007). Everyday mathematics assessment handbook-Grade 5. Chicago, IL: McGraw-Hill.
- Wikipedia (n.d.) Common Core State Standards Initiative. Retrieved from http://en.wikipedia.org/wiki/Common_Core_State_Standards_Initiative#Mathematics

A Comparative Analysis of System and Problems of Descriptive Assessment in Elementary School Mathematics Between Korea and U.S.A.⁷⁾

Kim, Min Kyeong⁸⁾ · Cho, Mi Kyung⁹⁾ · Kim, Rae Young¹⁰⁾
Kim, Gooyeon¹¹⁾ · Noh, Sunsook¹²⁾

Abstract

The present study examined the differences of system and problems of descriptive assessment at the national, district level, and school's levels between Korea and U.S.A. (focused on NCTM perspective). As results, both Korea and U.S.A. showed that their goals for descriptive assessment at the national level pointed out the importance of assessment of process rather than result. In addition, States of U.S.A. demonstrated concrete rubrics and examples to help many school teachers to use easily, while many Districts of Education in Korea presented implementation rate in school in order to give a official direction to teachers.

Key Words : descriptive problems, elementary school, international comparative study

7) This work was supported by the National Research Foundation of Korea Grant funded by the Korean Government (NRF-2011-32A-B00216)

8) Ewha Womans University (mkkim@ewha.ac.kr)

9) Ewha Womans University (cmk0530@hanmail.net)

10) Ewha Womans University (kimrae@ewha.ac.kr)

11) Sogang University (gokim@sogang.ac.kr)

12) Ewha Womans University (noh@ewha.ac.kr)