

중학교 수학과 수행평가의 개발과 적용 효과에 관한 분석*

권 오 남** · 황 속 균*** · 권 기 순****

1. 서 론

최근 수학교육 평가는 기존의 표준화된 선다형 평가 방식이 단편적이고 낮은 수준의 지식, 이해력에 초점을 맞추어 평가한다는 비판을 받아왔다. 이에 따라 수학적 사고 과정과 문제해결 과정을 통한 종합적이고 고차적인 사고력과 문제해결력을 정의적 영역과 함께 평가해야 한다는 목소리가 높아지고 있다. 새로운 평가에 대한 요구와 평가 관점의 변화는 학습관 및 학습자관의 변화에 많은 영향을 받기도 하지만, 전통적인 평가 즉, 선다형 지필식 검사에 대한 비판의 소리가 높아졌기 때문이기도 하다. 많은 학자들이 평가의 형태가 선다형 지필식으로 한정됨으로써 생기는 문제점을 지적하고 있다. 평가에 있어 객관식 유형만을 고집함으로써 학생들은 객관식 대비 요령만을 익히려는 경향이 생겨나고, 이는 학습 방법과 수업 방법에 나쁜 영향을 미칠 수 있다. 결국 객관성의 확보라는 최대의 장점을 갖고 있는 선다형 검사는 그 취지와는 달리 더 많은 단점을 양산시키게 되었고, 이의 보완과 개혁이 필요하게 되었다.

Lester & Lambdin(1997)은 최근 수학교육

학자들이 수학을 고정적이고 변화하지 않는 사실과 기술들의 집합체로만 보았던 관점에서 수학을 추측하고 의사소통하고 문제해결하며, 논리적으로 추론하는 것을 강조하는 관점으로 변화하고 있다고 말한다. 이러한 변화의 조류는 다음의 세 가지 측면에서 이루어지고 있다. 이를 요약하면 <표 1>과 같다.

<표 1> 수학의 본질, 수학 교수·학습에 대한 관점 변화

관점	전통적인 관점	새로운 관점
수학의 본질	수학은 공식의 암기나 조건과 법칙의 목록일 뿐이다.	수학은 변화와 패턴을 탐구하고 예측하는 과학이다.
수학 교수의 본질	수학 교수는 학생의 학습 활동을 확실하게 설명하고, 감독하는 것, 오류를 수정하는 것과 관련된다.	수학 교수는 학생이 복잡한 체계를 구성하고 탐험할 수 있도록 하는 활동이다.
수학 학습의 본질	수학 학습은 수학적 기능, 법칙, 사실 등을 발전시키거나 삭제하고 오류를 수정하는 점진적인 과정일 뿐이다.	인간은 모델, 이론과 학습의 설계자다. 인간은 실제 혹은 가상 세계의 복잡한 체계를 설명하며 창조하고, 수정하여 적용하고, 예측하고, 조절하기 위해서 자신의 지식을 새롭게 구성한다.

* 본 논문은 교육부 초등교육정책과 열린 교육을 위한 연구 과제인 '중학교 영어·수학 교과에서의 열린 교육을 위한 수행평가 적용 및 효과 분석 연구'의 일부임.

** 이화여자대학교

*** 용곡중학교

**** 동작중학교

수학에 대한 관점의 변화는 수학교육의 목표, 내용, 평가 등과 관련하여 총체적인 변화를 가져왔다. 따라서, 현재 시행 중인 제 6 차 수학과 교육과정에서뿐만 아니라, 제 7 차 수학과 교육과정에서도 수학 학습의 평가가 학생 개인의 전인적인 성장과 교사 자신의 수업 방법 개선을 위한 것이어야 한다는 것을 강조하며 이를 위해서 지필 검사 이외에 관찰법 등 다양한 평가 방법을 사용할 것을 권장하고 있다.

이를 위해 요즘 많은 관심의 대상이 되고 활발히 연구되고 있는 것이 바로 '수행평가'이다. 1999년부터는 중학교에서도 수행평가를 전담으로 실시하도록 하고 있으나 아직은 대부분이 초등학교를 중심으로 한 연구들이 많고, 중학교 현장에서 수행 평가가 어떻게 이루어질 수 있고 어떤 효과를 기대할 수 있는지에 대한 적용 연구는 많이 부족한 실정이다. 따라서 본 연구에서는 수학과 수행평가를 정의하고, 수행평가를 위한 평가도구를 개발하고, 개발한 수행과제의 내용타당도를 검증하며, 채점의 신뢰도를 검증하는 등 현장에서 수행평가 적용의 효과를 분석하고자 한다.

2. 이론적 배경

(1) 수학과 수행평가의 정의

수행평가(performance assessment)는 특정한 과제에 대한 학생들의 수행 과정이나 결과에 대하여 교육적으로 판단하는 것을 의미한다. 평가자나 교사는 학생들이 수행하는 과정을 관찰하여 학생들의 산출물의 특징을 판단한다. 이런 의미에서 수행평가는 관찰과 판단에 근거한 평가라고 할 수 있다(Stiggins, 1994). 성태제

(1998b)는 평가방법의 유형으로 수행평가 여부를 판단하는 획일적 기준은 없으나 무엇을 평가하느냐, 곧 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 바람직하다고 말한다. 수행평가가 전통적인 평가의 대안으로 등장하게 된 궁극적인 쟁점은 평가도구의 형식보다는 평가의 내용이 무엇인가, 곧 무엇을 평가하는가에 있었기 때문이다. 성태제와 권오남(1995, 1999)는 기존의 평가관은 학생들의 성취 결과에만 초점을 두었지만, 결과뿐 아니라 그 과정이나 절차에 대한 평가 방법이 필요하며, 고등정신 능력의 측정을 위한 평가로서 수행평가를 제안하고 있다. 곧 수행평가는 전통적인 평가가 잘 다루지 못했던 복합적인 사고과정, 기술이나 기능 등의 고등정신 능력을 평가하고자 하는 것으로, 평가 형식보다는 평가의 내용이나 목적에 그 판단의 근거를 두는 것이라 할 수 있다.

본 연구에서는 수학과 수행평가를 다음과 같이 정의한다.

수학과 수행평가란 수학적 사고능력이나 문제해결 능력 등을 평가하기 위하여 학습자가 지니고 있는 수학적 지식이나 수학과 관련된 기능을, 얼마나 알고 있으며 얼마나 잘 수행(performing)하는가와 어떻게 잘 수행 할(how to perform) 것인가를 총체적으로 평가하는 방법으로서, 수학적 문제를 해결하는 과정과 결과, 수학적 원리를 찾아내는 과정과 결과, 그리고 실제 생활에 수학적 지식이나 기능을 적용하는 능력을 측정한다.

(2) 수학과 수행평가의 특징

Lesh & Lamon(1992)은 참된 수학적 활동의

평가는 단편적 지식만을 측정하는 좁은 평가여서는 안되고, 계산능력, 암기력만을 측정하는 등 평가 영역 상의 균형을 잃고 편중되어서도 안되며, 과거의 고전적인 이론과 방법을 강조하는 낡은 것이어서도 안 된다고 하였다. 예컨대, 교수 학문적으로 수학의 본질이나 유용함과 관련되는 실제적인 문제해결의 본질을 다루는 것이어야 한다. 즉, 수행평가를 통해 학생들은 공식이나 절차를 외워서 정답을 재생·기억해 내는 것보다는, 당면한 문제의 해결 과정을 표현할 것을 요구받음으로써 자신의 수학적 능력, 문제해결력 등을 평가받게 된다. 학생들은 실제 생활과 관련된 문제를 해결하면서 수학적 가치를 인식하게 될 것이며, 완성된 수학적 지식의 습득이 아니라 스스로 만들어 가는 수학을 경험할 수 있게 될 것이다. 학생들은 자신의 수행을 돌아봄으로써 부족한 점을 이해하게 되고 이는 이후 학습에 많은 도움을 줄 수 있다. 또한 교사는 학생들의 상태를 좀 더 정확하게 파악하게 되어 학생들에게 적합한 교수를 제공할 수 있고, 자신의 수업을 반성할 수 있다. 이러한 관점에서 수행평가는 수업의 연속선상에 있는 것이다.

이와 같은 사항들을 종합하여, 본 연구에서는 수학과 수행평가의 특징을 다음과 같이 요약한다.

첫째, 수학적 지식과 방법을 사용하여 나타난 활동과 산출물로 평가하는 것이다.

둘째, 문제해결의 과정이 드러나도록 하여 평가하는 것이다.

셋째, 창의력이나 논리적 사고력 등과 같은 고차적 정신능력을 평가하는 것이다.

넷째, 과정 중심적이고 평가 목표 지향적인 채점 방법을 사용하는 것이다.

다섯째, 실제 생활과 밀접한 관련이 있는 문제 상황(수학의 내적·외적 연결성을 강조

함)에서 평가하는 것이다.

여섯째, 협동학습과 개별학습이 평가와 병행하여 이루어지는 것이다.

일곱째, 평가가 수업의 과정 속에 이루어지는 것이다.

여덟째, 한 두 시간에서 수일에 걸치는 프로젝트나 탐구과제를 수행하는 것이다.

아홉째, 자와 컴퍼스 등과 같은 교구, 컴퓨터와 그래픽 계산기 등과 같은 테크놀로지를 비롯한 실생활적인 도구(realistic tools), 구체적 조작물(manipulatives)을 사용할 수 있는 평가라고 할 수 있다.

(3) 수학과 수행과제의 특징

수행과제(performance task)는 수행평가를 실시하기 위해 사용되는 과제로서 학생들에게 해결하도록 제시하는 문제이면서 동시에 해결 과정을 묻는 질문지이기도 하다. 답이 여러 개 있는 문제, 해결 접근법이 여러 가지인 문제, 실생활과 관련되고 적용이 가능한 문제, 개량된 주관식형의 문제 등이 여기에 해당된다.

Sternmark(1991)에 의하면 수학과 수행 과제는 다음과 같은 기준을 만족시키는 것이라고 한다. 과제는 교육과정의 핵심적 내용을 다루어야 하고, 다른 질문이나 문제를 이끌어 낼 수 있어야 하고, 고차적 사고를 유발할 수 있게 지속성을 가져야 한다. 또한 과제는 학생이 자신의 사고 및 의사의 결정자가 되도록 하고, 다른 학생과 상호작용할 수 있게 하는 것이어야 하며, 그 수업 내에 끝낼 수 있는 것이며, 안전한 것이어야 한다. 과제는 학생들이 긍정적인 태도를 가지도록 하는 것이어야 되고, 하나 이상의 해답을 가지고 모든 학생들에게 받아들여질 수 있는 다양한 접근 방법이 있는 것이어야 한다.

수학과 수행평가의 정의에서 볼 수 있듯이 수학과 수행평가는 수학적 지식이나 기능에 의한 산출물이나 정답여부에만 관심이 있는 것이 아니라 이를 도출하는 과정에 대해 관심이 있는 것이 그 특징이다. 그러므로 수행평가를 실시하기 위한 수학과 수행과제는 실제 생활과 관련이 있거나 실생활에 일어날 수 있는 문제 혹은 다양한 해결과정과 반응을 유도하는 개방형 과제(open-ended task)의 성격을 지닌다.

(4) 수학과 수행평가의 유형

수학과 수행평가는 학습자가 수학적 지식과 기능을 적용하여 실제 생활에서 요구되는 문제를 해결하는 수행 능력을 평가하는 것이므로 실생활과 관련되는 수행을 요구하며, 수학적 문제해결 과정이 관련되는 모든 것을 포함한다고 할 수 있다. 그 중에 수학적 사고와 다른 교과와 내용을 연결시켜 활동하게 함으로써 수학적 접근 방법의 중요성을 인식시키는 열린 과제를 다루는 프로젝트법과 학생들의 모든 작업을 모아 수학적 수행 과정을 평가하는 포트폴리오, 수학 문제해결을 위하여 사고하고 있는 개인, 소집단, 또는 학급에 대하여 관찰하고 기록하여 이를 평가하는 관찰법과 면접법 등을 중심으로 수학과 수행평가의 유형을 알아보도록 한다.

가. 프로젝트법

프로젝트는 열린 반응을 요구하는 일종의 수행 과제(open-response tasks)를 말한다. 다시 말하면, 학생들은 이러한 과제들을 수행하기 위하여 어떤 수학적 지식을 사용해야 하는지를 결정해야 할뿐 아니라 때로는 어떻게 접근해 나가야 할 것인가에 관한 수학적 방법까지도 결정해야 한다(Krulik & Rudnick, 1995). 프로젝트

과제는 주제선정이나 접근방법에 있어서 구조화된 방법과 비구조화된 방법으로 나눌 수 있다. 그러나 이러한 분류는 이원적으로 프로젝트 과제를 분리하는 것이어서는 안된다. 즉, 과제는 구조화의 정도에 따라 하나의 연속선상에서 고려되어야 하는 것이다. 교사는 학생들의 연령이나 학습 능력에 따라 어느 정도로 과제를 구조화하여 제시할 것인지를 결정해야 할 것이다.

나. 포트폴리오법

Sternmark(1989, 1991)와 성태제(1998a)에 의하면, 포트폴리오는 학생들의 모든 작업을 모아 두는 작업철로서, 이 속에는 다양한 형태의 과제, 프로젝트, 보고서, 그리고 글(writing)이 모아질 수 있다. 따라서 교사는 학생들의 포트폴리오 속에서 수학에 대한 태도와 이해에 대한 발전을 폭넓은 방법으로 평가해 볼 수 있다.

수학 포트폴리오에는 해마다 새로워지는 수학적 일대기, 배운 것과 아직 배우지 못한 것에 대한 학생들이 자기 보고서, 학생들이 수학에 대해 어떻게 느끼는지에 대한 글 등이 포함된다. 포트폴리오의 구성 내용은 교사가 목적하는 바에 따라 다양하게 선정할 수 있다.

다. 관찰법 및 면담법

관찰법은 수학 문제해결을 위하여 사고하고 있는 개인, 소집단, 또는 학급에 대하여 관찰한 것을 기록하여 이를 평가에 사용하는 것을 말한다(서울특별시교육청, 1997). Sternmark(1991)에 의하면, 관찰의 목표는 수학에 관한 인지적 영역에 관한 것일 수도 있고, 정의적 영역에 관한 것일 수도 있다. 면담법은 지필 검사나 관찰법과 더불어 학생들과 직접 대화함으로써 문제해결 상황에서 실제로 나타내 보인 행동이나 서면의 '결과'를 도출해 내기까지의 사고

‘과정’에 대한 통찰을 가능하게 하는 기법이다.

관찰과 면담의 방법은 다른 평가 방법에 비해 수학적 태도나 성향 등의 정의적 영역을 평가하기 위한 중요한 수단이 될 수 있다. 따라서 필요에 따라 다른 수행평가 방법과 병행하여 학생에 대한 정보를 더욱 풍부히 하는 데 사용하는 것이 적당하다.

(5) 수학과 수행평가의 채점 방법

지금까지 언급해온 수학과 수행평가의 특성과 방법을 고려하면 수행평가는 학생의 문제해결 과정과 사고능력을 살피는 것에 초점이 있었다. 따라서, 수행평가에서의 채점 방법은 기존의 선택형 문항을 중심으로 하는 평가의 채점 방법과 차이가 있다. Lester(1987)는 수학과 수행평가의 채점 방법으로 분석적 채점 방법과 총괄적 채점 방법 두 가지를 제시한다. 분석적 채점 방법(analytic scoring)은 문제해결을 몇 가지의 특정 과정 측면으로 나누어 보고 점수를 부여하는 것이며, 총괄적 채점 방법(focused holistic scoring)은 특별한 사고 과정에 관련된 증거를 기반으로 하여 문제 전체의 답안에 대한 점수를 부여하는 것이다.

가. 분석적 채점 방법

Lester(1987)는 분석적 채점 방법을 문제해결 과정을 여러 측면으로 나눠서 각 측면에 점수를 배당하는 것이라 했다. 보통은 문제해결을 문제 이해, 계획의 수립, 해답 도출의 세 과정으로 구분하고 점수의 범위는 0에서 2점 정도로 주기도 하지만, 이는 채점의 강조점을 어디에 두고자 하는지의 목표에 따라 교사마다 점수 부여가 다를 수 있다.

분석적 채점 방법은 답만이 아닌 문제해결에 대한 여러 측면을 고려할 수 있고, 학생의

작업에 대한 수치적 값을 할당할 수 있는 수단을 제공할 수 있다. 학생의 강점이나 취약점을 정확히 알 수 있는 수단을 제공하며, 다양한 교수 활동의 효과에 대한 특별한 정보를 제공하고, 척도를 구성하는 범주에 따라 가중치를 다르게 줄 수 있다는 장점을 가지고 있다. 그러나, 분석적 채점 방법 경우는 때로는 학생이 글로 쓴 답안에 학생의 문제해결 과정의 정보가 충분하지 않기 때문에 과제 해결 과정의 전체 내용을 파악하지 못하여 확신을 가지고 점수를 줄 수가 없다. 따라서, 분석적 채점 방법은 교사가 문제해결에 관련된 핵심 범주에서 학생의 수행에 대해 피드백을 주기를 원할 때, 학생들의 강점과 취약점에 대한 직접적인 정보를 얻고자 할 때, 학생이 쓴 답을 자세히 분석할 만큼 충분한 시간이 있을 때 사용하는 것이 바람직하다.

나. 총괄적 채점 방법

총괄적 채점 방법은 학생의 답안 전체에 대해 한 가지 점수를 주는 것이다. 분석적 채점 방법과는 달리 총괄적 채점 방법은 사고 과정의 여러 측면에 대해 점수를 주는 것이 아니라, 문제의 전반적인 해결에 초점을 맞추기 때문에 총괄적이다(Lester, 1987). 총괄적 채점 방법은 전체적 느낌에 의하여 점수를 부여하기 때문에 학생이 쓴 답안을 비교적 빠르게 평가할 수 있고, 답만이 아닌 문제해결에 사용된 사고 과정에 초점을 둔다는데 의의가 있다. 또한, 학생의 수행을 설명하는데 한 점수를 제공한다는 장점이 있다. 그러나, 한 가지 점수를 부여하고, 신속하게 채점을 하려다 보니 학생의 특정 취약점이나 강점을 지적해주지 못한다는 것과 문제해결과 관련되는 특성의 사고 과정에 차별을 둔 가중치를 부여할 수 없다는 단점이 있다.

(6) 수학과 수행평가에 대한 선행 연구

최근 교육학계에서 열린 교육의 실천에 대해 높은 관심을 보이면서 열린 교육의 한 방안으로 수행평가에 대한 연구도 활발해지고 있다. 최근 1990년대에 접어들어 미국에서는 이미 대안적인 평가의 한 형태로 수행평가를 도입한 것을 비롯하여 1992년에는 Indiana College Placement and Assessment Center에서 수행평가에 대한 연구를 실행하였고, 1992년에는 Wisconsin Performance Assessment Development Project의 시작으로 수행평가에 대한 연구가 이어지고 있으며, 캘리포니아와 Vermont주 등 미국의 21개 여의 다른 주에서도 수행평가의 실행 가능성에 대한 연구가 이루어지고 있으며, 또한 실시되고 있다(Smith, 1993).

우리 나라에도 1996년 이후, 국립교육평가원의 연구를 필두로 수행평가의 수용 분위기가 서서히 고조되고 있다. 중학교 수학과 수행평가의 연구로는 장경윤, 권오남, 최명례(1996)가 수학과 수행평가의 문항을 제작하고 평가 결과의 채점자간 신뢰도에 대해 연구한 것이 있다. 초등학교 수학과 수행평가의 연구로는 류희찬(1996)이 초등학교 고학년용에서 사용될 수 있는 수행 평가 문항을 개발하고 수행평가를 실시하는 절차와 채점 기준을 개발하고 이를 현장에 직접 실시하여봄으로써 우리 나라 학생들의 수학적 힘을 측정해 보고 우리 나라 학생들이 수행 과제를 다루는 데 있어서 범하는 오류에는 어떤 것이 있는지를 분석한 연구가 있다.

3. 연구 방법 및 절차

(1) 대상 및 기간

본 연구는 1998년 3월부터 1999년 2월까지 12개월 동안 이루어 졌다. 대상은 서울시 소재 공립 Y중학교 1학년 심화반 남녀 학생들로 이들은 수준별 이동수업을 실시하고 있는 상태였다. 학생들은 매번 중간고사와 기말고사 성적으로 반편성이 다시 이루어 졌으므로 실험기간 중 반의 이동이 있었던 학생들은 실험에 참여는 하였으나 결과분석에서는 제외시켰다. 따라서 실험에 일관되게 참여한 학생들은 실험집단에 23명, 통제집단에 25명으로 구성되었다. 중학교 1학년 수학 교과 각 단원의 대단원 별로 개발한 수행평가 도구를 대단원의 진도가 끝나고 나서, 1내지 2차시에 걸쳐서 시행하였다.

(2) 수학과 수행평가 도구개발 및 채점 기준 설정

본 연구에서 개발한 수행평가 도구는 앞서 언급한 수행평가의 유형 중 구조화된 프로젝트 과제에 해당한다. 프로젝트는 원래 학생들이 주제와 방법까지도 스스로 결정하여 주도적으로 학습하는 것이나, 연구의 대상이 되는 학생들이 중학교 1학년이고, 아직 새로운 평가 방법에 익숙하지 않은 점을 고려하여 구체적이고 구조화된 형태로 구현하여 적용하였다. 과제들은 모두 학생들이 지필 형태로 그 과정과 결과를 기록하여 제출하도록 하였다. 연구의 초기에는 2명에서 4명이 협동학습을 하고, 한 조가 하나의 활동지를 완성하도록 하였으나, 한 조 내에서 열심히 참여하지 않고 공부 잘하는 친구에게만 의존하는 학생이 생겨나서 이들을 똑같이 평가한다는 것에 문제점이 있었다. 따라서 이후에는 개별적으로 수행평가 문항을 해결하도록 하거나, 2내지 3명이 서로 의견 교환은 하되 활동지는 개별적으로 완성하도록 하였다. 본 연구에서 실시한 수행평가 도구에 대한 채

점은 총괄적 채점과 분석적 채점의 방법을 모두 사용하여 과제마다 두 차례의 채점을 실시하였다. 본 연구에서의 분석적 채점의 분석 측면은 수학적 문제해결의 과정과 수학적 의사소통의 측면을 평가하는 것으로 하였다. 다음은 개발된 9개의 수행과제 중 두 가지 예이다.

가. 「과제 6: 자료의 정리」의 평가 도구 개발
 학생들은 자료를 단순히 정리하는 데 그칠 것이 아니라 자료의 이면을 읽을 수 있어야 한다. 따라서 ‘초보기자의 기사쓰기’인 「과제 6」에서는 학생들과 가장 친근한 소재인 필기구를 조사하여 관련 자료를 정리하고, 이를 토대로 자신의 생각을 펼칠 수 있도록 개발되었다(<그림 1> 참조).

※ 용곡 일보의 수습기자인 여러분에게 숨은 실력을 발휘할 기회가 왔습니다. ‘학생들이 가지고 있는 필기구’에 대한 기사를 써달라는 의뢰가 온 것이죠. 무작정 쓰기는 어려워워서 우선은 실태를 조사하기로 했습니다. 다음은 우리 분단 친구들이 가지고 있는 필기구들이 어느 나라의 제품인지를 조사해 본 결과입니다.

- ① 이제까지 우리가 ‘자료의 정리’ 단원에서 배운 내용을 토대로 어떤 형태로든지 위의 자료를 보기 쉽게 정리하십시오. 교과서에서 배운 방법으로 정리하지 않아도 됩니다. 새로운 그래프를 만들어내어도 좋고, 신문에서 본 모양이어도 좋습니다. (예를 들면, 도수분포표, 막대그래프, 원그래프, 꺾은선 그래프, 수염막대그래프 등등)
- ② 자신이 정리한 표나 그래프를 가지고 여러분들이 문제를 낸다면 어떤 문제를 내겠습니까? (예를 들면, 어느 나라 제품이 가장 많습니까? 등등)
- ③ 이제는 여러분 자신들이 정리한 자료를 보고 느낀 점이나 자료에 대한 분석 등 자신의 주장, 생각으로 신문기사를 써 보십시오.

<그림 1> 「과제 6: 자료의 정리」의 문항

우선 학생들은 여러 친구들에게서 각각이 소유하고 있는 필기구를 종류별, 생산된 국가별, 상표별로 조사하고 분류한다. 그 다음, 자신이 조사한 자료를 자신의 주장하고자 하는 바에 맞게, 이제까지 배운 통계적 지식을 기초로 하여 가장 보기 쉽게 정리한다. 마지막으로 자신이 정리한 자료로 자신의 주장을 담은 기사를 써 보는 과제이다. 따라서 이 과제는 일정한 답안이 없는 열린 과제로 학생 반응의 자유도가 크고, 타교과(국어)와도 접목될 수 있는 과제이다. 「과제 6」은 다음을 평가하도록 개발되었다.

- 직접 조사한 자료들을 바르게 정리할 수 있는가
- 자료들의 의미를 분석하고, 자기 나름의 주장을 펼칠 수 있는가
- 자신의 주장에 적합한 표나 그래프를 선택하여 표현하는가

나. 「과제 6: 자료의 정리」의 채점 기준 설정
 「과제 6」에 대한 채점 방법은 주어진 자료를 올바르게 정리하고 분석하며, 문제의 해결 전략을 세우고 이를 정확하게 수행하는지에 대한 측면을 채점하고, 자신의 주장에 알맞은 형태의 표나 그래프를 만들고 자신의 의사를 정확하게 전달하기 위해서 사용하는 수학적 의사소통의 능력을 채점하도록 한다. 이 과제에 대한 분석적 채점 기준 가운데 수학적 의사소통에 대한 채점 기준은 <표 2>에서 제시하는 바와 같다. 또한 총괄적 채점 기준은 모든 과제에 공통적으로 적용되었는데, 이는 <표 3>과 같다.

<표 2> 「과제 6: 자료의 정리」의 수학적 의사소통에 대한 채점 기준

표나 그래프의 표현이 분명하고 명확함	1점	표나 그래프로 자료를 정리하지 못한다.
	2점	표나 그래프로 자료를 정리하고 있기는 하지만, · 축의 값이 제시되어 있지 않다. · 축에 대해서 변량의 수치들이 정확하게 표기되어 있지 못하다. 표나 그래프가 아닌 다른 독특한 방법으로 자료를 정리하기는 했지만, 정확하지가 않고, 부분적으로 오류가 있다.
	3점	그래프나 표가 정확하고 완벽하게 제시되어 있고, 상세하여 보기가 쉽고 명쾌하다.

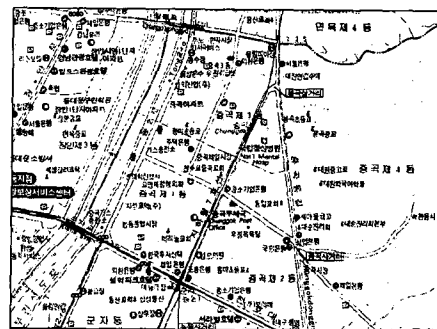
<표 3> 총괄적 채점 기준

0점	· 빈칸 · 문제의 정보를 순전히 베껴놓기만 했을 뿐 새로운 정보가 하나도 없다. · 오답 이외에 아무 것도 없다.
1점	· 자료를 베끼는 것 말고도 문제를 어느 정도 이해하고 풀기 시작했지만 본질적으로 평가하려고 했던 것 중 어떤 것의 준거도 만족시키지 못한다. · 부적절한 전략을 선택하여 실행했으나 다른 전략으로 바꾸지 못하고 중간에 포기하였다. · 하위 목표 달성을 위한 시도는 했으나 성공하지 못했다.
2점	· 적절한 전략을 선택하였지만 답을 구할 만큼 충분히 실행하지 않았다. · 적절한 전략을 선택하였지만 바르게 실행하지 못하여 오답에 이르렀다. · 정답을 썼지만, 과정을 쓰지 않았다. · 정답을 썼지만, 과정을 이해하지 못하게 썼다. · 해결 과정이 조건을 어느 정도는 만족시키지만, 문제에 내재되어 있는 수학적 추론이나 의사소통, 문제 해결 등의 측면에 있어서 만족스럽지 못한 수준이다.
3점	· 정답에 이르는 바른 전략을 사용했지만, 문제의 일부를 잘못 이해했거나 문제의 일부 조건을 무시해서 완벽한 답안 아니다. · 적절한 전략을 바르게 사용했지만 분명한 결과에 가서 틀리게 답하고 있다. · 적절한 전략을 바르게 사용했지만 답의 숫자부분은 맞지만, 단위가 없거나 잘못 쓴 경우가 있다. · 적절한 전략을 바르게 사용했지만 답을 쓰지 않았다. · 적절한 해결 전략의 수행과정에서 오류가 있지만 이 오류가 문제나 전략에 대한 이해 부족에서 비롯된 것이 아니라, 단순히 옮겨 쓰는 과정에서 발생한 오류이거나 계산상의 오류이다. · 정답을 썼고, 전략 사용의 측면도 적절하지만, 전략 수행의 과정에서 완전히 분명하지 않은 점이 있다.
4점	· 적절한 전략을 사용하여 이를 실행했고, 문제에 지시된 대로 답을 완벽하게 적었다.

다. 「과제 7: 도형의 기초」의 평가 도구 개발 「과제 7」은 학생들이 살고 있는 지역의 지도를 제시하고, 자신의 집을 찾아 번개 친 곳의 위치를 예측하도록 하여 학생들이 흥미롭게 과제를 해결할 수 있도록 하였다. 번개가 친 후 두 지역에서 동시에 천둥소리를 들었을 경우와 세 지역에서 동시에 천둥소리를 들었을 경우에 각각 번개가 쳤을 가능성이 있는 곳을 찾는 것이 주된 내용이다. 이는 2학년에서 배우게 될 의심과 의심원에까지 확장시켜 학생들이 수행평가를 하면서 자연스럽게 스스로 학습할 수 있도록 구성된 것이다(<그림 2> 참조).

2. 나()와 용곡이는 번개 친 날 밤, 집에서 서로 전화를 하고 있었습니다. 그런데 어느 순간 우리는 전화소리를 통해 우리 집과 용곡이네 집 두 곳에 동시에 천둥 소리가 들렸다는 사실을 알았습니다. (용곡이네 집은 용곡중학교 바로 옆에 있습니다.)

※ ()안에는 자신의 이름을 쓰고, 지도에 자신의 집을 표시하고(●모양으로 진하게 표시하세요) 문제를 푸세요. 자신의 집이 지도에 나타나 있지 않거나 용곡이네와 너무 가깝다고 생각될 경우에는 친구 이름을 쓰고, 친구 집을 대신 표시해서 문제를 푸세요.



(1) 번개 친 곳은 어디일까요? 지도 위에 표시해 주세요. 어떻게 찾았지요?

(2) 앞의 문제(1)에서 번개 친 곳이 여러 지점일 수 있을까요?

그렇다면, 번개 친 곳이라고 생각되는 모든 곳을 표시하고, 왜 그렇지 설명해 주세요. 그렇지 않다면, 왜 그렇지 않은지 설명해 주세요.

<그림 2> 「과제7: 도형의 기초」의 문항

「과제 7」은 주로 다음을 평가하도록 개발되었다.

- 번개와 천둥소리와 시간 간격과 천둥소리를 듣는 위치에서 번개 친 곳까지의 거리와의 관계를 이해할 수 있는가

- 이상의 사실을 적절히 이용할 수 있는가
- 수직이등분선의 성질을 이해하고, 적용하고, 이를 정확히 작도할(그려낼) 수 있는가

라. 「과제 7: 도형의 기초」의 채점 기준 설정
본 과제에 대한 채점 방법은 주어진 문제를 올바르게 이해하고 분석하여, 합당한 문제 해결의 전략을 세우고 이를 정확하게 수행하는지에 대한 측면을 채점한다. 또, 기하적인 내용을 수행함에 있어, 기호나 용어를 정확하게 사용하고 전달하는지에 관해 수학적 의사소통의 능력을 채점하도록 한다. 이 과제의 문제해결 과정에 대한 분석적 채점 기준은 <표 4>와 같으며, 「과제 7」에 대한 분석적 채점기준 전체는 [부록 1]에 실려 있다.

<표 4> 「과제7: 도형의 기초」의 문제해결과정에 대한 분석적 채점 기준

과제에 대한 이해	0점	무응답
	1점	문제를 완전하게 이해하지 못한 경우 · 문제에서 제시되지 않은, 관련 없는 생각들을 비논리적으로 나열해놓은 경우
	2점	문제를 부분적으로만 이해한 경우 · 얻어진 자료의 수치들을 그대로 베껴놓은 경우
	3점	이해된 경우
해결의 접근 방안, 전략, 절차의 선택	1점	전반적으로 상황에 알맞지 않은 전략을 사용하거나 논리적으로 부적절한 절차를 사용한 경우
	2점	부분적으로 적절한 접근이나 절차를 사용한 경우 · 위치를 발견하는 과정이 조건에 적절하지 못하고 비효율적인 경우
	3점	해결 전략을 알맞게 정하고, 절차가 모두 맞는 경우
전략의 실행	1점	자신이 세운 전략을 실행하지 못한 경우 · 자나 컴퍼스를 사용하여 위치를 나타내려 하나 무의미한 빗금만 친 경우 · 작도를 하는 과정에서 실수나 착오가 발생한 경우
	2점	작도를 정확하게 수행했으나, 문제 상황에 맞게 해석하지 못한 경우 · 작도는 정확치 않게 수행했으나, 문제 상황에 맞게 해석한 경우
	3점	작도를 정확하게 하고, 문제에서 요구하는 위치를 정확하게 도시한 경우

4. 결과 분석

(1) 수행과제의 내용타당도 분석

평가과제의 내용타당도는 16명의 교과전문가에게 단원별 수행과제의 내용타당도에 대한 점검표를 작성하게 하여 그 결과를 분석하였다. 「과제7」과 「과제8」은 경기도에서 1급 정교사 자격 연수를 받는 교사 75명이 점검표 작성에 참여하였다. 점검표는 Herman 등(1992)이 고안한 것을 수정하여 제작한 것으로, 점검표의 13문항 중 ‘과제와 평가목표간의 일치성’은 1, 2, 6, 12번 문항을 묶어서 분석하였고, ‘수학적 고등사고 능력의 평가’는 3번과 7번 문항, ‘과제의 실생활 반영도’는 4번과 8번 문항으로 묶어서 보았으며, ‘통합 교과적 접근’은 5번 문항, ‘평가과제의 공정성’은 9번 문항, ‘교과내용과의 관련성’은 10번 문항으로, ‘학생들의 참여 유도 정도’는 11번 문항으로, ‘과제의 실행 가능성’은 13번 문항으로 분석하였다. 내용타당도에 대한 점검표는 [부록 2]에 실려 있다. 다음은 9개 과제에 대한 내용타당도 분석 결과이다.

<표 5> 내용타당도 분석 결과 (단위: 명)

영역	과제와 평가목표 간의 일치성	고등사고 능력의 평가	과제의 실생활 반영도	통합 교과적 접근	평가 과제의 공정성	교과내용과의 관련성	학생들의 참여유도 정도	실행 가능성
과제 5	긍정 (87.5%)	26 (81.3%)	25 (78.1%)	12 (75%)	11 (68.8%)	15 (93.8%)	15 (93.8%)	5 (31.3%)
	부정 (6.3%)	4 (12.5%)	4 (12.5%)	3 (18.8%)	4 (25%)	0 (0%)	0 (0%)	9 (56.3%)
	무응답 (6.3%)	4 (6.3%)	2 (3.1%)	3 (18.8%)	1 (6.3%)	1 (6.3%)	1 (6.3%)	2 (12.5%)
과제 7	긍정 (87.8%)	142 (84.5%)	139 (82.7%)	67 (79.8%)	67 (79.8%)	64 (76.2%)	76 (90.5%)	74 (88.1%)
	부정 (8.9%)	30 (14.9%)	22 (13.7%)	15 (17.6%)	14 (16.7%)	17 (20.2%)	5 (6.0%)	6 (7.1%)
	무응답 (3.3%)	11 (2.4%)	4 (3.6%)	6 (3.6%)	3 (3.6%)	3 (3.6%)	3 (3.6%)	4 (4.8%)

내용타당도 결과에 따르면, 9개의 과제 모두는 몇몇을 제외한 모든 항목에서 70%이상의

긍정적인 반응을 얻었다. 특히 학생들의 참여 유도의 정도 항목에서는 9개의 과제 모두 87% 이상의 매우 높은 긍정적인 반응을 얻기도 했다. 그러나 「과제5」의 실행 가능성 항목에 대해서는 31.3%만이 긍정적인 반응을 하여 비율로는 가장 낮았다. 이는 점검표에 응답한 교과 전문가 중에는 CBR¹⁾에 대한 이해가 부족하거나 CBR을 직접 사용해본 적이 없기 때문인 것으로 보여진다. 그러나, 실제로 실험에 참여한 학생들은 교실에서 이 과제를 충분히 수행했던 것으로 보아 실행가능성에 대해서는 큰 문제가 되지 않을 것으로 보인다. 따라서 이상의 과제들은 긍정의 비율이 대부분 60%이상으로 내용 타당도가 대체로 높다고 할 수 있다.

(2) 수행과제의 채점자간 신뢰도 분석

과제의 평가 결과에 대한 채점자간의 신뢰도는 연구에 직접 참여하고 있는 교사 A와 연구자인 B, 그리고, 채점자 훈련을 거친 중학교 교사 C가 함께 채점하고, 그 채점 결과의 상관계수로 분석하였다. 채점자간 신뢰도는 「과제4」에서부터 「과제8」까지의 5개 과제에 대해서만 분석하였다. 다음은 각 과제별 채점자간 신뢰도 분석 결과이다.

<표 6> 채점자간 신뢰도 분석 결과(1)

과제	범 주	분석적 채점 방법			총괄적 채점 방법			
		A*B	A*C	B*C	A*B	A*C	B*C	
과제 4	수학적 문제해결	.925	.748	.771	.950	.841	.848	
	수학적 의사소통	.962	.717	.737				
과제 6	수학적 문제해결	.733	.447	.447	.765	.391	.579	
	수학적 의사소통	.885	.370	.264				
과제 7	2번	수학적 문제해결	.846	.737	.736	.724	.586	.410
		수학적 의사소통	.768	.757	.706			
	3번	수학적 문제해결	.722	.732	.540			
		수학적 의사소통	.783	.701	.802			
과제 8	1번	수학적 문제해결	.901	.540	.673	.504	.411	.715
		수학적 의사소통	.771	.700	.890			
	2번	수학적 문제해결	.726	.482	.770			
		수학적 의사소통	.906	.853	.853			
	3번	수학적 문제해결	.753	.443	.549			
		수학적 의사소통	.849	.800	.954			

1) CBR(Calculator-Based Ranger)은 계산기에 연결하여 시간에 따른 물체의 위치에 관련된 자료를 수집하는 센서이다.

<표 7> 채점자간 신뢰도 분석 결과(2)

과제	범 주	총괄적 채점 방법		
		A*B	A*C	B*C
과제 5	과제의 2번	.961	.668	.669
	과제의 3번	.755	.924	.742
	과제의 4번	1.000	.970	.969
	과제의 5번	.922	.770	.675
	과제의 6번	.865	.714	.739
	과제의 7번	.776	.783	.727
	과제	.808	.903	.869

채점자간 신뢰도 결과에 따르면, 각 과제에 대하여 채점자 A와 B사이의 신뢰도는 대체로 .7이상으로 높은 반면, 채점자 A와 C, B와 C간의 신뢰도는 .2에서 .9 정도로 과제마다 차이가 많이 났다. 이는 채점자 A와 B는 직접 연구에 참여하면서 과제를 제작하고, 채점기준을 설정 하였으므로, 채점에만 참여했던 채점자 C보다는 과제와 채점기준에 대한 이해가 높았기 때문으로 보여진다. 「과제7」의 채점자간 신뢰도 결과를 보면, 몇몇을 제외하면 분석적 채점방법에서 채점자간 신뢰도가 대체적으로 .6이상으로 높다고 볼 수 있으며, 총괄적 채점의 결과도 채점자간 신뢰도가 비교적 높았다고 볼 수 있다.

(3) 수행평가가 학업성취도에 미치는 영향 분석

수행평가를 실시함으로써 실험집단 학생들의 학교 수학 성적이 통제집단에 비해서 어떠한 변화를 보였는지 알아보았다. 학생들의 사전 학업성취도는 5월초에 실시한 중간고사로 하였고, 사후 학업성취도는 12월초에 실시한 기말고사로 하였다. 실험이 3월부터 실시되어

실험집단의 학생들은 이미 두 번의 수행평가를 경험한 상태였다. 그러나 이는 짧은 기간이었고, 수행평가의 실시 횟수도 2번밖에 되지 않으므로 이의 영향은 무시하였다. 수행평가 실험을 시행하기 전의 중간고사 성적과 사후의 기말고사 성적에 대한 실험집단과 통제집단에 대한 비교를 공분산분석(ANCOVA)을 통해서 알아보았다. 공분산분석 결과에 의하면, F통계 값은 .018로서 수행평가 실시 후의 통제집단과 실험집단의 학업 성취도에는 차이가 없었다(<표 8> 참조). 즉, 두 집단은 수행평가 실시 후에서도 기존의 수학 시험에서 비슷한 성적을 보인 것이라고 해석할 수 있다. 이는 수행평가가 기존의 지필식·선다형의 지식 암기 위주의 평가와는 다른 평가목표를 가지고 있음에도 불구하고 수행평가를 실시한 실험집단의 학생들이 기존의 방법으로 학습한 학생들과 비교하여 학업성취도의 면에서도 뒤지지 않았음을 의미한다.

<표 8> 수학과 실험·통제집단의 학업 성취도 검사에 대한 서술통계

	통제집단 (사례수: 25명)		실험집단 (사례수: 23명)	
	평균	표준편차	평균	표준편차
사전성취도(매개변수)	88.80	8.79	88.87	9.87
사후성취도(종속변수)	74.48	15.42	75.04	13.02
교정된 성취도	74.50	2.72	75.02	2.84

(4) 수학적 태도에 대한 사전·사후 검사의 결과 비교 분석

수행평가를 실시하기 전과 후에 학생들의 수학적 태도에 대한 변화를 분석하기 위해서 수학적 태도에 대한 설문지 조사를 실험반과 통제반에 각각 실시하였다. 설문 문항의 변인 설정은 제 6 차 수학과 교육과정에서 제시하는

정의적 영역의 평가 항목으로 수학적 자신감(5 문항), 수학적 호기심(5문항), 의지(5문항), 수학적 가치(5문항)의 20문항으로 구성하였다. 수학적 태도에 대한 통제집단과 실험집단의 차이 검증 결과는 <표 9>과 같다.

<표 9> 수학적 태도에 대한 통제집단과 실험집단의 차이 검증

	통제집단 (사례수: 23명)			실험집단 (사례수: 23명)			t
	사전점수	사후점수	차이	사전점수	사후점수	차이	
평균	70.48	63.87	-6.61	77.22	76.26	-.96	2.08*
표준편차	15.45	14.78	10.29	11.23	11.54	8.73	

* p < .05

두 집단의 수학적 태도의 변화에 차이가 있는지를 두 독립표본 t-검정으로 알아 본 결과, t 값은 2.08로 유의수준 .05에서 통계적으로 유의한 차이가 있었다. 즉, 수행평가를 실시한 실험집단의 수학적 태도의 변화에 비해서, 전통적 평가만을 실시한 통제집단의 수학적 태도는 통계적으로 유의한 정도로 부정적으로 변화했음을 알 수 있다. 각 변인에 대한 결과는 <표 10>에서 제시하는 것과 같다.

<표 10> 수학적 태도의 각 변인에 대한 통제집단과 실험집단의 차이 검증

영역	집단	통제집단 (사례수: 23명)			실험집단 (사례수: 23명)			t
		사전점수	사후점수	차이	사전점수	사후점수	차이	
자신감	평균	17.26	14.65	-2.61	19.13	18.04	-1.09	1.68
	표준편차	4.81	4.39	2.81	4.14	4.57	3.31	
의지	평균	17.74	15.78	-1.96	19.22	19.17	-.05	2.033*
	표준편차	4.15	4.25	3.47	3.53	3.41	2.88	
호기심	평균	17.13	16.04	-1.09	18.26	17.83	-.43	.733
	표준편차	4.28	4.06	2.94	3.05	3.59	3.06	
가치	평균	18.43	17.56	-.87	21.48	21.22	-.26	.599
	표준편차	4.74	4.18	4.31	3.53	2.19	2.28	

* p < .05

수행평가 실시 전·후의 통제집단과 실험 집단의 수학적 의지 영역의 변화는 두 독립표본 t-검증에서 유의수준 .05에서 통계적으로 유의한 차이가 있었고, 그 외의 영역에서는 통계적으로 유의한 차이는 없었다.

전통적 평가 방법을 실시한 통제집단과 수행평가를 실시한 실험집단에서의 수학적 태도의 변화 결과를 요약하면, 통제집단과 실험집단 모두는 수학적 태도에 있어 부정적인 태도로의 변화가 있었고, 그 변화 정도는 유의수준 .05에서 통계적으로 유의한 차이를 나타냈다. 특히, 각 변인 중에서는 수학적 의지 영역에서 유의한 차이를 나타냈다. 다시 말하면, 통제집단의 수학적 태도의 부정적 변화의 폭은 실험집단에 비해 통계적으로 유의한 정도로 크다는 것이다.

이러한 결과는 학년이 올라감에 따라 수학 학습에 대한 학생들의 태도가 부정적으로 변화하는 경향이 있다는 강승호(1997)의 연구 결과로 볼 때, 수행평가의 실시가 학습 내용이 어려워짐에 따라 나타날 수 있는 수학에 대한 부정적 태도를 방지하는 의미 있는 일이라고 할 수 있다.

(5) 수행평가에 대한 설문 조사 분석

9개의 수행과제를 모두 마치고 1999년 2월에 실험반 학생들을 대상으로 수행평가에 대한 설문을 실시하였다. 설문지는 Hackett(1994)이 사용한 설문 및 면담 문항을 참고로 하여 본 연구자가 수정·제작하였다. 설문지는 수행평가에 대한 선호도와 느낌, 수행과제에 대한 선호도, 수학학습에의 공헌 정도, 수학에 대한 관점의 변화 등을 묻는 12문항으로 구성되어 있

다. 설문조사의 결과는 다음과 같다(<표 11> 참고).

<표 11> 수행평가에 설문 조사 결과 (단위: 명)

문항	반응	긍정	부정	긍정과 부정	무응답	계 ¹⁾
수행평가에 대한 느낌		25(75.8%)	2(6.1%)	4(12.1%)	2(6.1%)	33
수행평가를 계속 하고자 하는가		26(78.8%)	6(18.2%)	0 (0%)	1 (3.0%)	33
기존의 평가를 수행평가로 대체하고자 하는가		19(57.6%)	14(42.4%)	0 (0%)	0 (0%)	33

다음은 학생들이 수행평가에 대해 가지고 있는 긍정적인 느낌들이다.

- 교과서만 가지고 공부를 하다보니 응용력이라든지, 수학공부의 재미를 별로 느끼지 못했는데, 이 수행평가를 매 단원 끝날 때마다 하니가 수학이 얼마나 신비스럽고 재미있는지 새삼 느끼게 되었다. 선생님! 2학년 때에도 이렇게 하나요?
- 뭔가 여태까지 했던 수업과는 달리 열린 교육을 한 것 같다. 또 그냥 막 푸는 문제가 아닌 배운 내용과 관련된 내용이어서 내가 배운 것을 응용하는 좋은 기회가 되었다. 또 내 실력을 마음껏 발휘할 수 있는 재미도 있는 수업이었다.

학생들은 수행평가에 대해서 대체로 재미있다, 수업에 더 열심히 참여하게 되었다 등의 긍정적인 반응을 보였다. 그러나 기존의 평가 방식을 수행평가로 바꾸는 문제에 대해서는 많은 학생들이 부정적인 생각을 가지고 있었다. 이런 학생들은 수행평가에 대해 바르게 인식하지 못한 경우이거나, 체점의 객관성을 그 이유로 들고 있다. 즉, 학생들은 수행평가에 대한 올바른 인식을 가지고 있지 못한 채 수행평가

1) 수준별 이동 수업으로 인하여 3월부터 실험에 계속 참여한 학생들은 23명이지만 설문조사 당시 심화반에 속한 학생들은 9가지의 수행과제 중 2가지 이상을 경험한 학생들이므로 33명 모두의 응답을 분석하였다.

가 또 다른 형태의 시험이라고 생각하고 부담을 느끼게 된 데 기입하는 것 같다. 다음은 수행평가에 대해 긍정적인 느낌을 가지고 있으면서도 기존의 평가를 대신하는 문제에 대해서는 반대하는 학생들의 답변이다.

- 반대이다. 만약 그렇게 되면 시험을 많이 친다는 생각이 들어서이다. 그리고 시험이 다가와도 공부를 많이 하는데 시험대신 수행평가를 한다면 공부를 더욱더 많이 해야 할 것 같아서
- 싫다. 기준을 정하기도 그렇고 중간고사나 기말고사를 대신해서는 하고싶지 않다.
- 반대. 시험을 보는데 있어서 수행평가를 한다면, 시간도 시간이지만, 제일먼저 시험은 자신이 배운 내용을 정리해 복습하는 것과 같이 테스트하는 것인데 시험에 수행평가를 실시한다면 시험 같은 것은 있으나마나 인 것 같다.

학생들은 수준에 따라 적당한 장애가 있으면서 학생들이 도전해 볼 만한 정도의 난이도가 있는 과제를 가장 선호했다. 또한 학생들의 직접적인 활동이 많이 포함된 과제일수록 더 선호하는 경향을 볼 수 있었다. 또한 학생들은 대부분 사고력의 신장, 수학적 의사소통 능력의 신장에 수행평가가 도움이 되었다고 판단하고 있었다. 또한 수행평가를 하면서 협동학습을 했던 것도 수학학습에 도움이 되었다고 판단하고 있었다. 전체적으로 학생들은 수행평가가 수학학습에 도움이 된다고 받아들이고 있었다. 수행평가를 경험하고 나서 학생들은 수학을 계산만 반복하고 실생활에서 필요 없는 과목이 아니라, 재미있고 실생활에서 많이 사용되고 논리를 키울 수 있는 과목이라는 인식을 하고 있었다. 다음은 수행평가를 경험한 후 수학에 대한 관점이 어떻게 변화하였나 하는 질문에 대한 학생들의 반응이다.

- 예전에는 기계처럼 계산만 하는 것이 수학을 알았는데 지난 1년간은 생각하는 수학을 하였다.
- '수학이란 생각하는 과목이다'라는 생각이 이전에는 그냥 생각하는 과목이란 말밖에 못했지만 지금은 예를 들어 친구에게 설명해 줄 수 있다.
- 전에는 수학문제를 무조건 문제 풀기만 하고 그런 줄만 알았는데 ... 지금은 수학이 무조건 문제 풀기는 아니고 수학도 실험(?)을 해 가면서 하는 게 재미있다.

5. 결론 및 제언

본 연구는 중학교 1학년 수학과 수행과제의 개발과 그 효과를 검증하기 위한 실험 연구로 이루어졌다. 수행평가는 중학교 1학년 심화반 학생을 대상으로 이루어졌고, 다음과 같은 결론을 얻었다.

첫째, 개발된 과제의 내용타당도는 과제와 평가 목표간의 일치성, 수학적 고등 사고 능력의 평가, 과제의 실생활 반영도, 통합 교과적 접근, 평가 과제의 공정성, 교과 내용과의 관련성, 학생들의 참여 유도 정도, 과제의 실행 가능성의 8개 항목으로 나누어 교과전문가에게 의뢰하여 분석하였다. 본 연구에서 사용한 9개의 과제 모두는 대부분 60%이상의 긍정적인 반응을 얻어 내용타당도가 높다고 볼 수 있다. 결국, 본 연구에서 제작·사용된 수행 과제들은 수행 과제가 갖추어야 할 기본 조건들을 만족한다고 볼 수 있다.

둘째, 수행평가의 채점자간 신뢰도는 연구에 참여한 연구자 사이에서는 높았으나, 연구자와 그렇지 않은 일반 교사 사이에서는 신뢰도가 .2에서 .9까지 분포하였다. 따라서, 일선 학교에서 수행평가를 실제로 적용하게 될 때에는, 가능한 대로 최대한 교사가 수행 과제를

직접 제작하고 채점하는 과정이 필요할 것이다.

셋째, 수행평가를 실시하기 전의 중간고사 성적과 사후의 기말고사 성적에 대한 실험집단과 통제집단에 대한 비교를 공분산분석을 통해서 알아보았다. 수행평가 실시 후의 통제집단과 실험집단의 학업 성취도는 통계적으로 유의한 차이가 없었다.

넷째, 수행평가를 실시한 실험집단과 전통적 평가 방법을 실시한 통제집단의 수학적 태도 변화를 검증하였다. 그 결과, 실험집단과 통제집단 모두 수학적 태도에 있어 사전·사후점사의 차이 검증에서 부정적인 변화를 나타냈고, 실험집단과 통제집단의 수학적 태도의 변화는 유의수준 .05에서 통계적으로 유의한 정도의 차이를 나타냈다. 특히, 수학적 태도의 영역 중에는 수학적 의지에 있어서 유의수준 .05에서 통계적으로 유의한 정도의 차이를 나타냈다. 수학 교과목의 경우, 학년이 올라감에 따라 수학적 태도가 부정적으로 변화한다는 경향에 비추어 볼 때, 수행평가의 실시가 학습 내용이 어려워짐에 따라 나타날 수 있는 수학에 대한 부정적 태도를 방지하는 의미 있는 일이라고 할 수 있다.

다섯째, 수행평가에 대한 설문 결과, 학생들은 수행평가에 대해서 대체로 긍정적인 반응을 보였고, 전체적으로 수행평가가 수학학습에 도움이 된다고 받아들이고 있었다. 학생들은 적당한 장애가 있으면서 학생들이 도전해 볼만한 정도의 난이도가 있는 과제를 가장 선호했고, 직접적인 활동이 많이 포함된 과제일수록 더 선호하는 경향을 볼 수 있었다. 또한 수행평가를 경험하고 난 후의 학생들은 수학을 단순한 계산이 아니라 유용성을 지니고 있고, 사고력과 창의력을 키워주며 그리고 흥미로운 과목으로 이해하는 경향이 많아졌다.

본 연구를 토대로 다음과 같은 제언을 하고자 한다.

첫째, 학생들의 수준에 맞는 수행과제가 많이 개발되어야 할 것이다. 본 연구에서는 1학년 심화반 학생을 대상으로 한 과제만이 개발되었지만, 현장에서 수행평가의 실시를 위해서는 다른 학년의 학생들 또는 다른 수준의 학생들을 대상으로 한 수행과제의 개발이 이루어져야 할 것이다. 과제의 개발이 어려운 일이지만, 동료 교사들이 공동으로 과제를 제작한다면 보다 손쉽게 과제를 제작할 수 있을 것이다.

둘째, 평가의 객관성 확보를 위한 채점기준과 채점 방법에 관련된 후속 연구가 뒤따라야 할 것이다.

셋째, 협동학습의 경우 공정한 평가 방법에 대한 후속 연구가 있어야 한다. 수행평가는 대부분 개별학습보다는 협동학습으로 이루어지는 경우가 많다. 본 연구의 초반에서 나타났듯이 한 조에서 하나의 결과물을 제출할 때 열심히 참여하지 않고 공부 잘하는 친구에게만 의존한다거나 협동학습에 열심히 참여하지 않는 학생의 공정한 평가에 대한 문제가 해결되어야 할 것이다.

넷째, 수행평가가 현장에서 적용되기 위해서는 교사와 학생, 학부모 모두의 인식전환이 필요하다. 설문조사 결과에서 나타났듯이 학생들이 수행평가에 대해 긍정적인 생각을 가지고 있음에도 불구하고, 실제 시행에 대해 부정적인 반응을 보이는 것은 수행평가에 대한 바른 인식이 자리잡기 못했기 때문이다. 상급학교로의 진학에 내신성적을 내기 위한 양적 평가보다는 교사와 학생들의 교수·학습 활동을 돕고, 학생들의 수학적 사고력과 수학적 힘을 짐작할 수 있는 질적 평가로 자리잡는 것이 바람직할 것이다.

참고문헌

- 강승호 (1997). 중·고교생의 수학적 성향 및 태도와 학업성취도와의 관계 분석. 운암 김연식교수 정년퇴임기념논총 (pp. 61-98).
- 권오남, 김경자 (공역) (1997). 초등 수학 수행평가 과제 제작 및 분석. 서울: 양서원. (Mathematical Sciences Education Board of the National Research Council. (1993). *Measuring up: Prototypes for mathematics assessment.*)
- 류희찬 (1996). 초등학교 고학년용 수학 수행능력 평가 문항 개발 연구. 교과 교육 공동 연구 결과 보고서.
- 서울특별시교육청 (1997). 창의력 신장을 돕는 중학교 수학과 학습 평가 방법. 서울: 서울특별시교육청.
- 성태제 (1995). 고등정신능력 신장을 위한 교육평가 방안 탐색. 교육의 국제 경쟁력 재고를 위한 평가방안 탐색. 전국 교육평가 심포지움 보고서 제 12편. 서울: 국립교육평가원.
- 성태제 (1998a). 교육평가 방법의 변화와 결과 타당도. 한국교육평가학회 창립 15주년 학술 세미나 발표논문집 (한국교육평가학회).
- 성태제 (1998b). 언어 수행평가의 과제와 전망. 1998 한국 응용언어학회 겨울 학술 대회 발표 논문집 (pp. 5-18). 서울: 한국 응용언어학회.
- 성태제, 권오남 (1999). 수학과 학업 성취도 평가를 위한 수행평가의 과제와 전망. 학교 수학 제1권 1호. pp.217-234.
- 장경윤, 권오남, 최명례. (1996). 중학교 수학 수행평가 문항의 개발 및 그 활용 가능성의 탐색. 교과 교육 공동 연구 결과 보고서.
- Hackett, R. K. (1994). *Doing mathematics portfolios in a middle school classroom: A case study exploring an emerging alternative assessment practice.* Ph.D., Stanford University.
- Herman, J. L., Aschbacher, P. R., & Winters, L. (1992). *A practical guide to alternative assessment.* Alexnadria, VA: Association for Supervision and Curriculum Development.
- Krulik, S., & Rudnick, J. (1995). Project in the middle school mathematics curriculum. *Connecting mathematics across the curriculum, 1995 Yearbook* (pp. 159-162). Reston, VA: NCTM.
- Lesh, R. & Lamon, S. (Eds.). (1992). *Assessment of authentic performance in school mathematics.* Washington, D. C.: AAAS Pub. Co.
- Lester, F. Randall, C., & O'Daffer, P. (1987). *How to evaluate progress in problem solving.* Reston, VA: NCTM.
- Lester, F. K., Lambidin, D. V., & Preston, R. V. (1997). A new vision of nature and purposes of assessment in the mathematics classroom. In G. D. Phye (Eds.), *Handbook of classroom* (pp. 287-319). New York: Academic Press.
- Krulik, S. & Rudnick, J. (1995). Project in the middle school mathematics curriculum. *Connectiing mathematics across the curriculum, 1995 Yearbook* (pp. 159-162). Reston, VA: NCTM.
- Smith, L. M., Kuhs, T. M. & Ryan, J. M. (1993). *Assessment of student learning in mathematics.* South Carolina University, Columbia.

- Stenmark, J. K. (1989). Assessment alternatives in mathematics. EQUALS staff & Assessment Committee of the California Mathematics Council Campaign for Mathematics.
- Stenmark, J. K. (1991). *Mathematics assessment*. Reston, VA: NCTM.
- Stiggins, R. (1994). *Student-centered classroom assessment*. New York:

Development and Implementation of Performance Assessment for Middle School Mathematics*

Kwon, Oh-Nam · Hwang, Sook-Gyun · Kwon, Kee-Soon

The purpose of this study is to define performance assessment of mathematics, to make a model of performance task of mathematics for the first grade middle school students in higher group, to examine validity and reliability of performance task, and to investigate the effects on the students' achievement and attitude. It first defines performance assessment and examines its main features and scoring methods. Based on these, nine performance assessment tasks and scoring criteria were designed for the first grade middle school students in higher group. The validity of performance tasks were examined by experts. The learning achievement and mathematical attitude test between two groups were performed as pre and post test. The thinking of students about performance assessment was investigated by attitude survey. The results of this study are as follows: First, the validity of the performance tasks is very high. Second, The control group. Forth, there is a difference in student's attitude about mathematics between scorer reliability was high due to the scorer training. Third, there is little difference in learning achievement between experimental and experimental and control group in 5% meaningful levels. That is, student's of both groups attitude about mathematics comes negatives, but the width of change of negative attitude in experimental group is less than control group. In the trend of negative attitude of mathematics as grade comes higher, this results showed that performance assessment of mathematics had positive influence on attitude of mathematics. The result of survey about mathematical performance assessment experience showed that students have positive attitude to performance assessment and recognized effectiveness of the performance assessment of mathematics.

* This research was supported by the Ministry of Education Research Fund '98.

** Ewha Womans University

*** Yong-Gok middle school

**** Dong-Jak middle school

[부록 1] 「과제 7」에 대한 분석적 채점 기준표

단원 : VI. 도형의 기초	학년 : 중학교 1 학년
<p>개관 : 본 과제에 대한 채점 방법은 주어진 문제를 올바르게 이해하고 분석하여, 합당한 문제해결의 전략을 세우고 이를 정확하게 수행하는지에 대한 측면을 채점하고, 기하적인 내용을 수행함에 있어, 기호나 용어를 정확하게 사용하고 전달하는지에 관해 수학적 의사소통의 능력을 채점하도록 한다.</p>	

가. 문제 해결 과정에 대한 채점 기준

문제 해결의 과정		
과제에 대한 이해	0점	무응답
	1점	문제를 완전하게 이해하지 못한 경우 · 문제에서 제시되지 않은, 관련되지 않은 생각들을 비논리적으로 나열해놓은 경우
	2점	문제를 부분적으로만 이해한 경우 · 얻어진 자료의 수치들을 그대로 베껴놓은 경우
	3점	이해된 경우
해결의 접근 방안, 전략, 절차 등의 선택	1점	전반적으로 상황에 알맞지 않은 전략을 사용하거나 논리적으로 부적절한 절차를 사용한 경우
	2점	부분적으로 적절한 접근이나 절차를 사용한 경우 · 위치를 발견하는 과정이 조건에 적절하지 못하고 비효율적인 경우
	3점	해결 전략을 알맞게 정하고, 절차가 모두 맞는 경우
전략의 실행	1점	자신이 세운 전략을 실행하지 못한 경우 · 자나 컴퍼스를 사용하여 위치를 나타내려하나 무의미한 빗금만 친 경우 · 작도를 하는 과정에서 실수나 착오가 발생한 경우
	2점	작도를 정확하게 수행했으나, 문제 상황에 맞게 해석하지 못한 경우 · 작도는 정확하지 않게 수행했으나, 문제 상황에 맞게 해석한 경우
	3점	작도를 정확하게 하고, 문제에서 요구하는 위치를 정확하게 도기한 경우

나. 수학적 의사소통에 대한 채점 기준

수학적 의사소통		
기하적인 기호나 용어의 표현이 분명하고 명확함	1점	합당한 이유가 없이 빈개 친 곳을 '여기' 혹은 '저기' 등의 지시어로 사용하여 답안을 제시함
	2점	빈개 친 곳을 점(기호화)으로 표시하기는 했지만, · 위치를 추리하기까지 기호를 사용함에 거리가 같은 지점이라는 것을 명확하게 표현하지 않아 답안이 완성되지 않은 경우 · 기호를 전혀 동원하지 않고 서술로만 의사소통을 하였으나 빈개 친 곳의 위치가 바르게 지적된 경우(서술적인 의사소통의 방법이 성공한 경우)
	3점	빈개 친 곳을 정확하게 점으로 표현하고 그 위치를 추리하게 되기까지 과정에서 기하적인 기호를 세련되고 정확하게 사용한 경우

**** 수학적 의사소통의 방법은 여러 가지가 있지만, 본 과제에서는 주로 기하 내용과 관련된 기호나 용어를 사용하는 측면의 의사소통을 평가하도록 한다. 따라서, 기호나 용어를 제시할 수 있는 능력에 국한해서만 채점하도록 한다.

[부록 2] 내용타당도 점검표

점 검 표

번호	질 문	네	아니오
1	과제가 구체적인 교수 의도(평가의 구체적인 목표)와 일치하는가?		
2	과제가 학생들에게 성취하길 기대하는 내용과 기술을 적절하게 표현하고 있는가?		
3	과제를 통해 학생들의 능력이나 능력의 향상을 입증할 수 있는가?		
4	과제는 학생들의 실제 생활을 반영하거나 근접한가?		
5	과제가 통합교과적인 접근을 하고 있는가?		
6	과제를 통하여 학생들이 다양한 목표와 측면을 성취했는지 알 수 있는가?		
7	평가의 목표들은 분석력, 종합력과 같은 복잡한 사고 능력을 반영하는가?		
8	과제는 학생들이 학교와 미래의 생활에서 만날 수 있는 문제의 유형들과 관련이 있는가?		
9	과제는 모든 학생들에게 공정하고, 편파적이지 않은가? (예를 들어, 성별, 지역별, 문화적, 경제적(ex, 어떤 수업자료는 돈이 많이 드는 경우)인 측면에서 특정 학생들에게 유리하거나 불리하지 않은가?)		
10	과제는 학생, 교사, 학부모 등에게 신뢰로운가? (과제가 교과내용과 관련이 깊은가?, 의미 있고 도전적인 과제인가?)		
11	과제는 학생들이 그들의 능력을 보여줄 수 있도록 의미 있는 것이며, 학생의 참여를 유도하는 것인가?		
12	과제는 가르칠만한 가치를 담고 있는가? 과제의 내용은 학생들이 획득할 수 있으며, 교수자료에 있는 기술과 지식을 대표하는가?		
13	과제들은 공간, 장비, 시간, 비용 등의 요소들을 고려할 때, 학교 혹은 학급에서 실행할 수 있는 것인가?		