대한수학교육학회지 <학교수학> 제 6 권 제 1 호
Journal of Korea Society of Educational Studies in Mathematics
School Mathematics Vol.6, No.1, 21-35. Mar 2004

삼각함수 단원의 수행평가 도구 개발 및 적용[1)]

고 상 숙* · 백 정 환**

고등학교 수학내용에서 학생들이 가장 어려워하는 삼각함수단원을 중심으로 수행평가 도구를 개발하고 이를 현장에 적용하여 효과를 파악하고자 하였다. 평가도구는 Bloom의 인지적 영역을 중심으로 총 12개의 문항으로 구성되었으며 이에 대한 채점기준표가 구성되었다. 개발된 도구의 적용성을 알기위해 문항에 대한 양호도검사와 다양한 영역에서 성취도 검사가 고등학생, 208명을 대상으로 2003학년 11월 중에 조사되었다. 양호도 검사에서는 본 평가도구가 우수한 것으로 나타났으며, 성취도 검사에서는 남학생이 여학생보다, 과학고 학생이 일반고 학생보다 우수하게 나타났으며, 지역별의 차이는 나타나지 않았다.

Ⅰ. 서 론

1. 연구의 필요성 및 목적

세계화와 정보화로 특징 지워지는 현대 사회에서는 지식을 기억하고 재생하는 능력보다는 사고력과 창의력 문제해결능력이 중요시된다. 학교 수학교육도 학습자의 다양한 개성과 잠재능력을 길러주고, 사고력과 문제해결능력을 키우며, 인성 및 창의성을 신장시키는 방향으로 이루어져야 한다. 따라서 학교에서의 평가도 기억력만을 재는 선다형 평가가 아니라 창의력과 문제해결능력을 길러주고, 개성과 다양성을 조장하는 평가, 석차 매기기를 위한 평가가 아니라 학습 활동을 돕기 위한 평가체제로 변환되지 않으면 안 될 것이다(백순근, 1998). 현실적으로 평가는 접하고 있는 지식·정보화 사회에 적응해야하고 학교교육의 정상화를 도모해야하며 지식관·학습관의 변화에 따른 새로운 교육과정에 부응해야 할 필요성이 있다.

이런 최근의 경향을 반영하는 수학과의 평가는 학습의 결과보다는 과정에, 학습의 양보다는 질에, 특정 시점에 이루어지는 일회성의 총괄평가보다는 수시 평가에 초점을 맞추게 되고, 교사가 학생의 성취도 점수를 매기던 데서 학습자의 주체적인 자기 평가까지도 강조하게 된다(박경미, 1998). 평가는 처음에 이루고자 한 목표에 비추어 계획된 교육 과정이 제대로 이루어졌는지를 파악하고 이를 통해 교육의 질을 개선하는데 그 목적이 있다(정동권 외7인, 2002). 따라서 모든 과정에는 평가가 반드시 포

* 단국대(sangch@dankook.ac.kr)
** 단국대 대학원(soohak1@hanmail.net)
1) 이 연구는 2003년 단국대학교 대학연구비의 지원으로 연구되었음.

함된다. 교사는 평가를 통해 학습자의 석차를 매기는 데 그치지 않고 자신이 행하는 수업 방법의 효율성과 적절성, 학생의 학습태도, 문제에 대한 학생의 접근 방법, 학생들이 범하기 쉬운 오류의 유형 등을 제대로 파악할 수 있어야 한다. 또 이러한 평가자료를 바탕으로 하여 교사의 수업을 질적으로 개선해 나가야 한다.

1989년 "학교수학을 위한 새로운 교육과정과 평가에 관한 규준집" 발간을 시작으로 '평가'를 주제로 한 부록을 발간했던 미국수학교사협의회(NCTM)는 1995년의 평가 규준집을 통해 새로운 세기에 걸맞는 평가 방향의 변화를 제시하고 있다. 즉, 학생들에게 탐구하고 추측하고 논리적으로 추론하며, 비정형적인 문제를 해결하는데 다양한 수학적 방법을 효과적으로 사용하는 등의 수학적 힘을 길러주는 것을 수학교육의 목표로 삼고 있다.

이러한 수학교육의 새로운 목표는 교수·학습에서 강조하는 것과 마찬가지로 평가에서도 일관성 있게 강조하여 실행될 때, 즉 교수·학습과 평가가 동일하게 변화될 때 비로소 구현될 수 있는 것이다. 기존의 객관식 선다형 검사는 일반적인 지식이나 기능을 위주로 평가하거나, 학생이 이해하지 못한 채 정답에 표시를 할 가능성이 있고 특히, 문제해결력이나 추론능력, 수학적인 사고력을 측정할 수 없다는 문제점을 가지고 있다. 따라서 우리나라의 제 7차 수학과 교육과정에서는 제 6차 교육과정에서도 이미 제기한 수학적인 사고와 문제해결력의 향상을 지향하면서 전통적인 지필검사 이외에 여러 가지 평가 방법을 적용하고 그 평가결과를 다양하게 활용한다는 기본 입장에서 한 걸음 더 나아가, "객관식 선다형 위주의 평가를 지양하고 주관식 지필 검사, 관찰, 면담 등 다양한 평가방법을 활용하여 종합적인 수학학습 평가가 이루어 질 수 있게 한다"(교육부, 1997)와 같이 보다 명시적이고 구체적인 평가방법을 제안하고 있다. 앞에서 언급한 바와 같이, 평가를 제시된 목표의 달성을 위한 교수·학습 과정의 한 부분으로 보았을 때, 창의력이나 문제해결력을 개발하기 위한 구체적인 방법 및 교수 방법의 개선과 학생들의 수학적 능력의 발전 및 성취를 어떻게 측정하고 평가할 것인가는 교사들의 당면한 과제이자 당연한 관심사라 할 수 있다.

따라서 본 연구는 학생들이 일반적으로 많이 어려워하는 고등학교 삼각함수 단원에서 Bloom(1956)이 제시한 교육목표 분류학에 의거하여 수행평가의 문항을 개발하고 채점기준(rubric)을 제시하며 이 평가도구의 현장 적용 가능성을 파악하고자 다음과 같은 연구문제로 조사되었다.

2. 연구문제

연구문제 1. 삼각함수 단원에서 평가도구를 개발하여 제시한다.

연구문제 2. 본 연구에서 개발한 삼각함수 단원의 수행 평가 문항에 대한 신뢰도, 타당도, 난이도와 변별도를 알아본다.

연구문제 3. 본 연구에서 개발한 삼각함수 단원의 수행 평가 문항에 대한 수행수준을 성별에 따라 조사한다.

연구문제 4. 본 연구에서 개발한 삼각함수 단원의 수행 평가 문항에 대한 일반학생의 수행수준을 지역별에 따라 조사한다.

연구문제 5. 본 연구에서 개발한 삼각함수 단원의 수행 평가 문항에 대한 과학고와 일반고 학생의 수행 수준을 조사한다.

II. 이론적 배경

대부분의 추상적인 용어들이 그렇듯이 '수행평가'라는 용어의 경우도 그 의미를 고정된 하나, 또는 몇 개의 단어로 명료하게 정의할 수는 없다. 그러나 인간사회에서 의사소통은 이루어져야 하므로 동일한 의미의 공유를 위한 노력 중 한 방법은 우리가 사용하고자 하는 용어에 대한 일종의 사전적 정의들을 설정하고, 그 정의를 수용하는 일이다.

일반적으로 수행(performance)이란 구체적인 상황 하에서 실제로 행동을 하는 과정이나 그 결과를 의미한다. 수행평가라는 용어는 일상적 용어가 아니므로 사전에 그 의미가 정의되어 있지 않다. 수행평가라는 용어는 현재 그 의미를 규정해 나가려는 학자들에 의하여 그 의미가 형성되어 가고 있는 용어이다. 그러므로 수행평가에 대해서는 하나의 정의만 있는 것이 아니다. 학자들의 관점이나 강조점에 따라 그 정의는 달라진다. 그러나 물론 이러한 의미의 변용은 어느 범위 안에서 이루어지는 것이다. 그렇지 않을 경우 그 정의는 더 이상 의사소통의 기능을 상실하기 때문이다(허경철 외,2000).

Stiggins와 Bridgeford(1982)는 수행평가를 "새로운 문제나 특정의 과제를 해결하는 능력을 측정하기 위한 체계적 시도로서 실제 또는 모의 상황에서 피평가자들이 나타내 보이는 반응들을 전문가인 평가자가 직접 관찰하고 판단함으로써 이루어지는 평가방식"이라고 정의하고 있다. Haertel(1992)은 "피평가자들로 하여금 특정의 산출물을 만들어 내거나 구체적인 활동을 수행하게 함으로써 자신들의 능력을 직접적으로 나타내 보일 것을 강조하는 평가방식"이 수행평가라고 정의하고 있다. 한편, Mehrens(1992)는 수행평가란 "피평가자의 반응을 평가하는 과정에서 평가자의 관찰과 전문적인 판단을 특히 중시하는 평가방식"을 의미한다고 하였다.

국내 학자의 경우, 성태제(1999)와 최연희, 권오남, 성태제(1998)는 수행평가를 습득한 지식, 기능이나 기술을 실제 상황이나 인위적 평가상황에서 얼마나 잘 수행하는지(doing, performing), 또는 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 방법으로 정의하고 있다. 박도순(1995)은 수행평가를 "학생 개개인이 처한 '상황'이라는 요인을 고려하여 과제 수행 과정과 결과에 대한 포괄적인 정보수집 방식"이라고 정의하고 있으며, 백순근(1998)은 "학생 스스로가 자신의 지식이나 기능을 나타낼 수 있도록 산출물을 만들어 내거나, 답을 작성(구성)하도록 요구하는 평가방식"으로 정의하고 있다. 또한 양길석(1999)은 "수행에 근거하여 관찰과 판단을 통해 이루어지는 평가"로 수행평가를 정의하고 이 때 수행이란 의미를 "학습과제 및 문제해결 상황에서 학습자가 보여주는 구성적 반응으로서 관찰 가능한 학습자의 행동이나 학습의 산출물 및 그 기록물"이라고 부언 설명하고 있다.

이상 수행평가의 정의에서 공통적으로 중시되는 개념들은 '과제의 수행', '과정과 결과', '관찰', '전문적 판단' 등임을 감안하여 본 연구에서는 이 개념들에 충실한 문항 개발을 목적으로 한다. 이를 위해 단계에 따른 차별화를 고려한 채점기준이 고려되었다.

Bloom(1956)은 교육목표 분류학에서 교육목표를 크게 인지적 영역(cognitive domain), 정의적 영역(affective domain), 신체적 영역(psychomotor domain)으로 나누었다. 본 연구에서는 이 중 일차 영역으로 합리적·지적 사고 작용이

일어나는 인지적 영역을 대상으로 문항을 개발하였다. 수학 · 역사 · 화학 · 영어 등과 같은 전통적인 학문적 내용을 학습하는 것이 이에 해당된다. Bloom의 인지적 영역은 '단순한 행동'으로부터 '복잡한 행동'으로의 위계성을 지니며 다음과 같이 이루어진다.

(1) 지식(knowledge) : 인지나 재생에 의하여 과거에 학습한 아이디어나 자료 또는 현상을 기억해내는 행동을 말한다. 특수사상에 관한 지식, 특수사상을 다루는 방법과 수단에 관한 지식, 보편적, 추상적 사상에 관한 지식 등으로 다시 분류되며 인지적 영역 중 가장 낮은 수준의 행동으로 분류된다. 이 행동은 재생에 의하여 학습된 내용을 기계적으로 기억해 내는 행동을 말한다.

(2) 이해(comprehension) : 이는 자료나 기호, 언어의 의미를 파악하는 능력을 말한다. 문장 독해력에서처럼 읽기와 관련되는 경우가 많으나 이보다 더 넓은 뜻으로 사용된다. 의미파악과 의사소통의 자료는 구두나 서면으로 제시될 수도 있고, 그 밖의 언어적 형태나 기호적 형태로도 제시될 수 있다. 자료라는 용어를 좀 더 넓게 해석하면 종이 위에 제시되는 자료는 물론 물리실험, 현장방문에서 관찰한 현상, 어떤 건축 상의 특색을 보여주는 건물 등 구체적인 형태를 갖춘 자료를 총칭한다. 이해는 다시 번역, 해석, 추론 등으로 분류된다.

(3) 적용(application) : 과거에 학습한 자료나 내용을 새로운 구체적인 사태에 사용할 줄 아는 능력을 말한다. 과거에 학습된 개념, 방법, 법칙, 원리, 이론에 관한 지식을 새로운 사태에 적용하는 문제해결능력을 의미한다. 이는 우리가 흔히 '응용'이라고 말하는 능력과 같은 성질의 것이다.

(4) 분석(analysis) : 이 행동유형은 이해나 적응보다는 높은 수준의 행동으로서, 자료를 구성부분으로 분해하고 부분간의 관계와 그것이 조직되고 있는 방법을 발견하는 것을 중요한 내용으로 한다. 교육목표로서의 분석력은 세 수준으로 나눌 수 있는데 첫째, 자료를 그 구성 성분으로 나누고, 자료의 요소를 발견하고 분류하는 능력(요소의 분석력), 둘째, 요소간의 관계를 분명히 하는 상호관계를 발견하는 능력(관계의 분석력), 셋째, 자료를 전체로 하여 묶고 있는 조직원리, 즉 배열과 구조를 인지하는 능력(조직원리의 분석력)이다.

(5) 종합(synthesis) : 이 행동은 여러 가지 요소나 부분을 이전까지는 분명하지 않았던 어떤 하나의 구조나 형태로 결합하는 행동이다. 이것은 인지적 영역에 있어서 학생에게 창의적 행동을 가장 분명하게 제공하는 유형이다. 종합은 독특한 의사전달과 자료의 창조, 계획 및 절차의 창안, 그리고 추상적 관계의 도출 등으로 분류한다.

(6) 평가(evaluation) : 이 행동은 특정한 목적을 위해서 사용된 자료나 방법의 가치를 판단하는 행동을 말한다. 이 판단은 양적일 수도 있고 질적일 수도 있다. 또한 어떤 특수한 사상들이 얼마나 정확하고 효과적인가를 감정하기 위해 준거를 활용하는 행동도 여기에 포함된다. 이 준거는 학생들에 의해 결정될 수도 있고 교사에 의해 주어질 수도 있다. 평가는 내적 준거에 의한 판단과 외적 준거에 의한 판단으로 분류된다.

본 연구는 Bloom에 의해 학생을 평가하기 위한 문항의 영역을 구성하였다. 또한, 이들 문항의 수행에 필요한 채점기준의 마련도 문항개발의 중요한 부분이다. NCTM(1999)에서는 총체적 채점기준(holistic rubric)과 분석적 채점기준(analytic rubric)을 제시하였는데 '단순한 행동'으로부터 '복잡한 행동'으로의 위계성을 지니며 발달하는 블룸의 인지적 영역은 본 연구가 지향하는 수행평가의 개념을 포함함과 동시에 채점기준의 구성에 필요한 이론적 틀을 제공한다.

III. 연구방법 및 절차

본 연구는 크게 두 가지 국면으로 구성된다.

첫째 국면은 문헌고찰에 따른 선행연구의 분석을 통해 평가도구를 개발하는 것이고, 두 번째 국면은 이 개발된 도구를 현장에 사용하여 평가도구로서의 적용 가능성을 조사하는 함으로써 수정을 거친 보다 나은 도구 개발이 이루어지도록 설계되었다.

1. 연구대상

개발한 평가 도구의 현장 적용성을 조사하기 위해 참여한 연구 대상은 경기도, 대구광역시, 충남시의 00고등학교 8개 학급에서 208명으로 남자 72명, 여자 136명이며, 이중 과학고, 15명, 일반고는 나머지 193명으로 구성되었다. 여학생이 남학생보다 많은 이유는 남녀 혼성 고등학교 외에 여자 고등학교가 참여하였기 때문이며, 또한 과학고 학생은 작은 수이지만 전집에 대한 대표성을 고려하여 포함되었다.

2. 평가도구 개발

수행평가는 학생의 인지적인 영역(창의성이나 문제해결력 등 고등사고 기능을 포함)뿐 만 아니라, 학생 개개인의 행동발달 상황이나 흥미·태도, 신체적인 발달 등 종합적이고 전인적인 평가를 중시하는 평가방식이다. 연구자의 연구범위에는 전 영역을 포함하고 있으나 본 논문에서는 일차 영역으로 인지적 영역만을 대상으로 문항을 개발하였다. 요즘 학교 현장에서는 인지적인 영역에 대한 수행평가가 이루어지고는 있으나 평가에 대한 어려움 때문에 주로 단답형 문제만이 실시되는 경향이다. 본 연구도구로써는 학교 현장에서 수행평가 실시에 따른 이런 어려움을 해결할 수 있는 검사지가 개발되었다. 이를 위해 Bloom의 교육목표 분류학 중 인지적 영역, '지식·이해·적용·분석·종합·평가' 단계가 이론적 배경으로 사용되었다. 하지만 아직 과인수 학급의 운영체제를 벗어나지 못하고 있는 현 고등학교 과정에서 이 여섯 단계 모두를 적용하기는 무리가 따르므로 상위 수준에 해당하는 분석·종합·평가를 종합이라는 한 단계로 축소하여 지식·이해·적용·종합이라는 전체 네 단계로 구성하여 평가기준(<표Ⅲ-1>참고)을 마련하고 이를 바탕으로 삼각함수를 중심으로 문항의 초안이 작성되었다. 각 문항에 따라 평가기준과 배점을 바탕으로 채점 기준표를 구성하였다. 특히, 본 연구에서 사용하고 있는 Bloom의 인지적 영역의 교육목표는 공동 연구자가 현재 재직하고 있는 학교에서 정기고사 문제 출제 시 문항별로 해당란을 표기하게 되어있는, 지식, 이해, 적용, 표현으로 구성된 평가의 목표이원분류표 구성에 사용되고 있을 뿐만 아니라, 최근 들어 Wheeler & Eschberger (2003)의 연구의 이론적 틀을 구성하고 있는 등 교과교육연구에 여전히 활발하게 적용되고 있음이 고려되어 도입되었다.

작성된 초안을 바탕으로 본 연구의 전 과정을 통해 각 문항은 계속된 수정을 통해 다듬어졌다. 이 과정에는 문항에 대한 채점기준의 구성도 포함되었다. NCTM(1999)가 제시한 분석적 채점기준(analytic rubric)과 총체적 채점기준(holistic rubric)에 의하면, 본 연구의 각 문항은 이들 단계에 따라 1점씩 부여되는 분석적 채점기준이 적용되었다. 뿐만 아니라 상위 단계(n단계)의 채점기준 속에 하위 단계(n-1 단계)의 요소를 포함하여서 이들 배점간에 위계성(난이도)에 의한 합산이 가능하므로 문항별로는 총체적 채점기준을 적용하는 형식이 함께 택해졌다. 즉, 위계성에 따라 세 단계와 네 단계로 나

누어 해당 영역을 만족할 시에 각각 중복되지 않게 1점, 2점, 3점, 4점으로 채점기준을 부여하였다(<표Ⅲ-2>참고).

본 연구도구에서 각 단계에 따른 신뢰도는 <표Ⅲ-3>과 같이 조사되었다.

위의 기준에 따라 구성된 문항의 평가 기준에 의해 지식, 이해, 적용, 종합의 내용별 세부 평가 기준은 <표Ⅲ-2>와 같다. 이들간의 위계성으로 인해 평가기준이 상위수준, n단계는 하위수준, n-1 단계를 포함하는 것을 원칙으로 하기 때문에 채점기준표가 총체적 채점표가 되는 틀을 제공한다.

<표Ⅲ-1> 평가 기준표

평 가 기 준 표	
지식	내용을 기억하고 재생할 수 있다.
이해	문제를 이해하고 추론할 수 있다.
적용	응용하고 활용할 수 있다.
종합	내용의 요소들을 모아 새로운 전체로 만들어 낼 수 있다.

<표Ⅲ-2> 지식, 이해, 적용, 종합의 내용별 세부 평가 기준

수행 평가 문항 분류	영역	평가 기준표
지식	일반각과 호도법	문제에서 사용된 용어의 정의를 기억할 수 있다.
		관련된 공식을 나타낼 수 있다.
		올바른 답을 구할 수 있다.
이해	삼각함수의 성질	문제에서 사용된 용어의 정의를 기억할 수 있다.
		문제에 관련된 공식을 나타낼 수 있다.
		문제를 자신의 방법으로 나타낼 수 있다.
		올바른 설명과 문제풀이를 할 수 있다.
적용	Sin & Cos 법칙	관련 공식들을 나타낼 수 있다.
		문제의 내용을 자신의 방법으로 나타낼 수 있다.
		관련 공식들을 추론하고 유도할 수 있다.
		공식들을 응용하고 활용할 수 있다.
종합	삼각함수의 응용	관련된 공식들을 나타낼 수 있다.
		문제의 조건들을 자신의 방법으로 나타낼 수 있다.
		조건과 공식들을 활용하여 필요한 식을 세울 수 있다.
		문제의 조건에 맞게 정확하게 증명할 수 있다.

<표Ⅲ-3> 수행 평가 검사지 구성

수행 평가	문항수	관련문항번호	신뢰도 계수
지식	3	1~3	0.84
이해	3	4~6	0.71
적용	3	7~9	0.76
종합	3	10~12	0.72
계	12		0.85

3. 검사의 실시 및 자료 분석 방법

개발된 평가도구의 효과를 파악하기 위해 표집집단을 대상으로 2003년 11월 1일부터 11월 30일까지의 기간 중 2일 간에 걸쳐서 수학수업 시간에 평가가 실시되었다. 문항의 양호도 분석 중 문항 내적 일관성 신뢰도와 변별도를 구하기 위하여 SPSS 10.0K를 사용하여 Cronbach α를 구하였고, 내적 타당도와 난이도를 구하기 위하여 문항 반응 이론 중 Rasch의 1-모수 문항 반응 모형에 근거하여 BIGSTEPS(Livacre & Wright, 1994, 2003)를 사용하여 분석하였다.

또한, 고등학생들의 수행수준이 수행 평가 문항에 나타내는 성별과 지역별에 미치는 영향을 알아보기 위해 수행 평가 수준은 각 문항별과 지식 · 이해 · 적용 · 종합으로 분류하여 t-검증을 사용하여 조사하였다.

IV. 연구 결과

연구문제 1. 삼각함수 단원에서 평가도구를 개발하여 제시한다.

선행연구를 바탕으로 마련된 평가기준표, <표 Ⅲ-1>과 <표 Ⅲ-2>에 의해 12개 평가문항(부록 참고)을 개발하였고 이들 채점에 필요한 채점기준표를 다음과 같이 구성하였다.

<표Ⅳ-1> 채점기준표

문항번호	채점기준표	배점
1	미지수를 문자로 표기하고 조건에 맞는 식이나 문장을 나타낸 경우	1
	올바른 문제풀이를 하였으나 계산에 오류가 있는 경우	2
	올바른 문제풀이와 정답을 구한 경우.	3
2	문제의 조건에 맞는 식이나 문장을 나타낸 경우	1
	올바른 문제풀이를 하였으나 계산에 오류가 있는 경우	2
	올바른 문제풀이와 정답을 구한 경우	3
3	제2사분면의 각을 일반각으로 나타낸 경우	1
	각각의 경우를 부분적으로 설명하였거나 계산에 오류가 있는 경우	2
	올바른 설명과 문제풀이를 한 경우	3
4	문제에 맞게 그림을 그려놓았거나 식의 일부를 나타낸 경우	1
	문제의 조건에 맞는 식과 부채꼴의 넓이 공식을 모두 나타낸 경우	2
	올바른 풀이를 하였으나 계산에 오류가 있는 경우	3
	올바른 풀이와 정답을 모두 구한 경우	4
5	sin법칙이나 cos법칙 둘 중 하나를 나타낸 경우	1
	sin법칙과 cos법칙을 모두 기술한 경우	2
	올바른 풀이를 하였으나 계산에 오류가 있는 경우	3
	올바른 풀이와 정답을 구한 경우	4
6	sin법칙이나 cos법칙 둘 중 하나를 나타낸 경우	1
	문제의 조건에 맞는 그림과 공식들을 나타낸 경우	2
	올바르게 풀이를 하였으나 문자의 표기에 오류가 있는 경우	3
	문제의 조건에 맞게 문제를 해결한 경우	4
7	제2 cos법칙 또는 삼각형의 넓이 공식을 나타낸 경우	1
	제2 cos법칙을 이용하여 대각선의 길이를 구한 경우	2
	구한 조건에 맞게 올바르게 풀이를 하였으나 계산에 오류가 있는 경우	3
	문제의 조건에 맞게 올바르게 문제를 해결한 경우	4
8	sin법칙을 기술한 경우	1
	중심각이 원주각의 두 배임을 나타낸 경우	2
	올바르게 풀이를 하였으나 계산에 오류가 있는 경우	3
	문제의 조건에 맞게 올바르게 문제를 해결한 경우	4
9	sin법칙 또는 제2 cos법칙을 나타낸 경우	1
	문제의 조건과 sin법칙, cos법칙의 공식들을 연립한 경우	2
	올바르게 풀이를 하였으나 문자의 표기에 오류가 있는 경우	3
	올바른 풀이와 정답을 구한 경우	4
10	제2 cos법칙을 나타낸 경우	1
	각 사분면의 cos의 부호를 알고 있는 경우	2
	올바른 증명을 하였으나 모든 경우를 증명하지 못한 경우	3
	문제의 조건에 맞게 모든 경우를 올바르게 증명한 경우	4
11	삼각형의 넓이 공식을 나타낸 경우	1
	넓이 공식을 이용하여 조건에 맞게 식을 나타낸 경우	2
	올바른 증명을 하였으나 문자의 표기에 오류가 있는 경우	3
	문제의 조건에 맞게 올바르게 증명한 경우	4
12	문제의 조건에 맞게 그림을 그려 놓은 경우	1
	삼각형의 넓이 공식을 기술하고 식들을 연립한 경우	2
	올바른 증명을 하였으나 문자의 표기에 오류가 있는 경우	3
	문제의 조건에 맞게 올바르게 증명한 경우	4

연구문제 2. 본 연구에서 개발한 삼각함수 단원의 수행 평가 문항에 대한 신뢰도, 타당도, 난이도와 변별도를 알아본다.

가. 문항 내적 일관성 신뢰도

검사의 신뢰도를 위하여 문항 내적 일관성 신뢰도인 Cronbach α를 구하였다. 수행 평가 문항 12개에 대한 신뢰도 계수는 0.85(<표Ⅲ-3> 참고)이다. 이 검사는 수행 평가를 측정하는 검사인데, 이러한 계수가 나온 것은 비교적 양호한 것이라고 볼 수 있으며, 이는 채점 기준을 보다 객관적으로 만들었기 때문이라고 판단된다.

나. 문항 적합도 지수로 본 내적 타당도

검사 문항에 대한 내적 타당도는 문항 반응 이론 중 Rasch의 1-모수 문항 반응 모형에 근거하여 모수치를 측정하고 문항 분석을 하도록 하는 컴퓨터 프로그램인 BIGSTEPS를 사용하여 문항들의 적합도 지수를 산출하였다. 사용된 분석 모형은 부분점수(Partial Credit) 모형이다. 대개 문항의 적합도 지수가 1.2 보다 큰 경우에는 그 문항이 사용된 분석모형에 적합하지 않은 피험자 반응을 가지고 있음을 의미한다. 보다 관대한 기준을 세울 경우에는 1.5까지의 적합도 지수는 모델에 적합한 것으로 수용된다. 문항별 적합도 지수를 측정할 때에는 문항 1, 문항 7, 문항 8은 문항의 적합도 지수 1.2를 상회하는 것으로 나타났다. 이러한 문항은 지식을 측정하는 문항으로서 많은 학생들이 쉽게 해결 할 수 있는 문항으로 보여지고, 문항 7과 문항 8은 축척에 대한 이해가 부족하여 이러한 결과가 나왔으나 Infit과 Outfit 지수가 모두 1.5 보다 높은 문항은 하나도 없으므로 이 문항들도 관대한 기준하에서는 분석모형에 적합한 것으로 수용할 만한 문항이라고 볼 수 있다.

다. 난이도

문항 난이도는 문항의 어렵고 쉬운 정도를 나타내는 것으로서 본 연구에서는 Rasch의 1-모수 문항 반응 모형에 근거하여 계산하였다. 문항 난이도가 0.0인 것은 문항들 중에서 평균 정도라는 것을 의미하며 양수(+)값을 가질수록 어려운 문항이다. 본 수행평가 검사에서는 로짓트 점수로 본 난이도는 -1.43에서 2.72까지 분포하고 있으며 언어적 표현으로 나타내면 <표Ⅳ-4>와 같은 분포를 가지고 있다. 난이도 척도상에 각 문항의 난이도를 나열해 보았을 때, 문항간의 난이도의 차이는 로짓트 점수가 12번 문항을 제외한 나머지가 -2.0에서 2.0까지의 범위에서 골고루 분포되어 있고, 문항 신뢰도 지

<표Ⅳ-2> 수행평가 문항 적합도 지수

문항	1	2	3	4	5	6	7	8	9	10	11	12
Infit	1.28	.92	1.15	.98	1.01	.72	1.16	1.23	.96	.78	.70	.82
Outfit	1.75	1.07	1.14	.82	0.89	.67	1.34	.99	.69	.57	.35	.27

<표Ⅳ-3> 수행평가 문항 난이도

문항	1	2	3	4	5	6	7	8	9	10	11	12
난이도	-1.43	-1.19	-0.35	-.19	-.16	.68	.01	.07	-.36	.00	.20	2.72

수가 모두 0.72보다 높은 것(<표Ⅲ-3>참고)으로 나타나 사용된 문항들은 서로 잘 분리되어 학생의 수행 능력을 추정하고 변별하는데 무리가 없다고 볼 수 있다.

문항 난이도는 <표Ⅳ-3>에서와 같이 문항 12, 문항 6, 문항 11, 문항 8, 문항 7, 문항 10 순으로 나타나고 있다. 다만 2.72인 어려운 문항 12번은 헤론의 공식을 유도하는 문항으로 매우 어려운 문항으로 학생들은 생각하고 있으며 과학고 학생들을 제외하고는 이 문항을 풀이한 학생은 일반 학교에서는 없었다. 이러한 결과는 검사 제작자의 의도가 애초부터 수학 수행 평가 검사 도구가 아주 뛰어난 학생들도 고려하였기 때문이라고 생각한다.

라. 변별도

수행 평가 문항의 변별도는 점이연 상관(point-biserial correlation)에 의하여 분석하였다. 점이연상관은 해당 문항 점수와 총점과의 상관으로서 유수(-)값을 나타내는 문항은 능력이 높은 피험자와 낮은 피험자를 제대로 변별하지 못하는 문항이라 할 수 있다. 점이연상관이 음수로 산출된 문항들의 경우는 대부분 그동안의 지식을 바탕으로 쉽게 점수를 받을 수 있는 문항이기 때문에 문항을 변별해 주기에는 부적절함을 의미하나 본 연구에서는 음수로 산출된 문항이 없어 모든 문항이 학생들의 삼각함수 능력을 변별해 줄 수 있을 것으로 보인다.

수행 평가 문제에서 능력이 높은 학생들이 오히려 낮은 점수를 받게 되는 것은 아니지만 수행 평가 문제에서 능력 높은 학생과 낮은 학생들을 변별하는데 별로 도움이 되지 않는 문항 즉, 점이연 상관 계수가 0에 가까운 문항이 없는 것으로 나타났다. 따라서 모든 문항이 학생들의 수행 능력을 변별해 줄 수 있을 것으로 보인다.

연구문제 3. 본 연구에서 개발한 삼각함수 단원의 수행 평가 문항에 대한 수행수준을 성별에 따라 조사한다.

수행 평가 문항에서 성별의 차이를 검증하기 위하여 t 검증을 하여 보았다. 고등학교 2학년에서 남학생이 여학생보다 문항 7을 제외하고 모든 문항에서 평균과 총합에서 높게 나타나고 있고 문항3, 문항 4, 문항 5, 문항 6, 문항 10, 문항 11과 총합에서 $p<.05$ 수준에서 통계적으로 의미가 있게 나타나고 있다. 문항7은 여학

<표Ⅳ-4> 언어적 표현에 의한 수행평가 문항 난이도

언어적 표현	문항 난이도 지수	해당 문항
매우 쉽다	-2.0 이하	없음
쉽다	-2.0 ~ -.5	1, 2
중간이다	-.5 ~ +.5	3, 4, 5, 7, 8, 9, 10, 11
어렵다	+.5 ~ +2.0	6
매우 어렵다	+2.0 이상	12

<표Ⅳ-5> 각 문항의 점수와 총점 간의 점이연 상관 계수

문항	1	2	3	4	5	6	7	8	9	10	11	12
변별도	.38	.53	.59	.67	.66	.73	.55	.58	.70	.73	.75	.57

생이 남학생보다 평균 0.09 높게 나타나고 있으며 p<.05 수준에서 통계적으로 의미가 있게 나타나고 있다.

수행 평가 세부 문항에서 성별의 차이를 검증하기 위하여 t 검증을 하여 보았다. 고등학교 2학년에서 남학생이 여학생보다 지식, 이해, 적용과 종합에서 평균이 높게 나타나고 있고 p<.05 수준에서 통계적으로 의미가 있게 나타나고 있다.

이것은 남학생이 여학생보다 본 연구에서 개발한 수행 평가에서도 높게 나타나고 있다는 것을 보여 주고 있다.

연구문제 4. 본 연구에서 개발한 삼각함수 단원의 수행 평가 문항에 대한 일반학생의 수행수준을 지역별에 따라 조사한다.

지역별 평가에서 충남 지역 15명을 제외하면 G 지역과 D 지역이다. 수행 평가 문항에서 지역별의 차이를 검증하기 위하여 t 검증을 하여 보았다. 고등학교 2학년에서 G 지역이 D 지역보다 문항 1, 문항 5, 문항 6, 문항 8과 문항 9에서 평균이 높게 나타나고 있으나 문항 9를 제외하고는 p<.05 수준에서 통계적으로 의미가 없게 나타나고 있다. D 지역이 G 지역보다 문항 2, 문항 3, 문항 4, 문항 7, 문항 10, 문항 11에서 평균이 높게 나타나고 있으나 문항3과 문항 7을 제외하고는 p<.05 수준에서 통계적으로 의미가 없게 나타나고 있다.

연구문제 5. 본 연구에서 개발한 삼각함수 단원의 수행 평가 문항에 대한 과학고와 일반고 학생의 수행수준을 조사한다.

<표Ⅳ-6> 성별에 따른 수행 평가 수행 수준

	문항 1	문항 2	문항 3	문항 4	문항 5	문항 6	문항 7	문항 8	문항 9	문항 10	문항 11	문항 12	지식	이해	적용	종합	합
남성	2.04	2.04	1.50	1.37	1.15	.75	.64	.86	1.11	1.34	.97	.39	5.58	3.28	3.61	2.71	15.18
여성	1.94	1.66	.81	.61	.67	.21	.73	.54	.96	.31	.23	.00	4.41	1.49	2.22	.54	8.65
평균차	0.10	0.38	0.69	0.76	0.48	0.54	-0.09	0.32	0.15	1.03	0.74	0.39	1.17	1.79	1.39	2.17	6.53
t-검증	0.60	0.67	0.01*	0.01*	0.00*	0.01*	0.00*	0.13	0.54	0.00*	0.01*	0.07	.01*	.00*	.01*	.00*	0.00*

* p<.05

<표Ⅳ-7> 지역별에 따른 수행 평가 수행 수준

	문항 1	문항 2	문항 3	문항 4	문항 5	문항 6	문항 7	문항 8	문항 9	문항 10	문항 11	문항 12	지식	이해	적용	종합	합
G 지역	1.94	1.66	.81	.61	.67	.21	.73	.54	.96	.31	.23	.00	4.41	1.49	2.22	.53	8.65
D 지역	1.84	1.79	1.28	1.02	.64	.18	1.40	.47	.49	.72	.39	.00	4.91	1.84	2.37	1.10	10.23
평균차	0.10	-0.13	-0.47	-0.41	0.03	0.03	-0.67	0.07	0.47	-0.41	-0.16	0.00	-0.50	-0.35	-0.15	-0.57	-1.58
t-검증	.66	.56	.02*	.09	.92	.76	.01*	.73	.04*	.06	.36		.276	.409	.751	.215	.17

* p<.05

수행 평가 문항에서 집단별의 차이를 검증하기 위하여 t 검증을 하여 보았다. 고등학교 2학년에서 과학 고등학교 학생이 일반계 고등학교 학생보다 모든 문항에서 평균, 지식, 이해, 적용, 종합과 총합에서 높게 나타나고 있으며 $p<.01$ 수준에서 통계적으로 의미가 있게 나타나고 있다. 이는 과학고 학생들이 수행 평가에서 일반계 학생들 보다 더 잘 수행 한다고 볼 수가 있다. 또 본 연구는 Bloom의 교육목표 분류학 중 인지적 영역을 다루기 때문에 과학고 학생들이 우수할 것이라는 것은 자연스럽게 예측할 수 있는데, <표Ⅳ-8>을 통해 예측대로 과학고 학생이 일반고 학생들보다 높게 나타나고 있다고 할 수 있다.

V. 결 론

본 연구의 목적은 고등학교 삼각함수 영역에서 수행평가를 위한 평가 도구를 개발하여 이를 현장 수업에 적용해보아 문항에 대한 양호도 검사와 다양한 영역에서 학생의 성취도를 검사하여 개발된 평가도구의 효과를 파악해 보는 것이다. 이를 위해 연구자는 2003년 1학기 중에 Bloom의 인지적 영역의 분류에 따라 하위 요소인 지식, 이해, 적용, 종합에서 각각 3문제씩 모두 12문제를 계속된 수정 작업을 거쳐 고안하였다. 평가 문항으로써 이 개발된 자료의 효과를 검증하기 위해 우리나라 고등학교 2 학년 학생, 208명을 대상으로 2003년 2 학기에 검사를 실시하였다.

문항에 대한 양호도 검사를 위해 신뢰도, 타당도, 난이도, 변별도에 대한 조사가 문항 반응이론에 근거한 분석으로 이루어졌는데 이들 조사에서 본 도구, 검사지는 현장교육에서 활용할 수 있는 평가도구로서의 가치가 꽤 높게 나타났다.

수행 평가 문항의 성별 차이도에서는 문항 7을 제외한 거의 모든 문항(문항 3-6, 문항 10-11)에서 남학생이 여학생보다 높게 나타났는데 이는 수학에서 남학생이 여학생보다 우수하다는 일반적인 사회적 통념이 반영된 결과라 할 수 있다. 이런 성별차이는 사회적, 문화적, 환경적 요인이 복합적으로 학습 성취도에 영향을 끼치는바 여전히 수학에서 남녀의 성 차별을 극복할 수 있는 개선 안이 수학교육의 평가에서뿐만 아니라 교육과정과 그에 대한 교수방법 등의 측면에서도 함께 병행하여 꾸준히 연구되어야 함을 함축한다. 오히려 평가는 이 두 측면에 의한 산물이기 때문에 더욱 그렇다.

<표Ⅳ-8> 과학고 학생과 일반고 학생간의 수행 평가 수행 수준

	문항 1	문항 2	문항 3	문항 4	문항 5	문항 6	문항 7	문항 8	문항 9	문항 10	문항 11	문항 12	지식	이해	적용	종합	합
과학고	2.80	3.00	2.33	2.73	3.07	2.93	2.53	2.33	3.47	3.73	3.20	1.87	8.13	8.73	8.33	8.80	34.00
일반고	1.91	1.70	.95	.73	.66	.20	.93	.52	.82	.43	.28	.00	4.56	1.59	2.26	.71	9.12
평균차	0.89	1.30	1.38	2.00	2.41	2.73	1.60	1.81	2.65	3.30	2.92	1.87	6.57	7.14	6.07	8.09	24.88
t-검증	.00**	.00**	.00**	.00**	.00**	.00**	.00**	.00**	.00**	.00**	.00**	.00**	.00**	.00**	.00**	.00**	.00**

** $p<.01$

수행평가 문항의 지역별 차이에서는 문항 9에서 경기지역이, 문항 3과 7에서 대구 지역이 높게 나타나나($p< .05$) 전반적으로는 거의 지역별 차이를 드러내지 않는 것으로 나타났다. 또, 과학고 학생과 일반고 학생간의 차이는 모든 문항에서 차이를 보이고 있어($p< .01$) 인지적 영역에 대한 과학고 학생의 우수성을 나타내므로 이는 문항 개발이 적절함을 내포한다.

이상 문항개발에 필요한 여러 분석 측면을 조사해본 결과, 본 연구에서 개발된 수행 평가 검사지는 문항으로서의 적절성을 가지고 있는 것으로 사료되었다. 최근의 수학교육은 학생들의 창의력과 문제해결력 개발에 깊은 관심을 갖고 교수·학습 과정에 이를 실현 가능한 방법을 찾기 위해 특히 평가 영역에서 여러 가지 많은 연구가 이루어지고 있다. 따라서 수학교육의 기본이라 할 수 있는 학습자의 수학적 능력을 제대로 평가하기 위해서는 전통적인 평가 방법인 객관식 지필검사 방법에서 벗어나 제시된 문제해결에서 나타나는 학생의 사고 과정까지 확인, 평가할 수 있어야 한다. 또한, 앞으로 Bloom의 다른 영역인 학습자의 관심과 태도, 신체적 활동 등을 교사, 동료 학생 또는 학생 자신이 스스로 평가하는 과정도 병행할 수 있어야 할 것이다. 따라서, 인지적 영역뿐만 아니라 정의적, 신체적 영역에 대한 평가를 포함한 수행평가를 위한 도구 개발이 더욱 활발히 이루어져 교실 현장에서 평가의 목적을 실현해나가는 데 기여할 수 있길 기대해본다.

참고문헌

교육부(1997). **수학과 교육과정, 제7차 교육과정 교육부 고시 제1997-15호 [별책8]**. 서울: 대한 교과서 주식회사.

박경미(1998). 수학. 백순근(편). **중학교 각 교과별 수행평가의 이론과 실제**, 239-295. 원미사.

박도순(1995). 연구의 최근 동향과 수행평가의 문제. **교육진흥 27.** 중앙교육진흥연구소.

백순근(1998). 총론. 백순근(편), **중학교 각 교과별 수행평가의 이론과 실제**, 19-75. 원미사.

성태제(1999). **수행평가의 실제**. 경남교원연수원.

양길석(1999). **수행평가의 개념에 대한 고찰**. 한국교육과정평가원 토요 포럼 발표 논문.

정동권 외 7인(2002). **수학과 수행중심 평가**. 학문출판(주).

최연희·권오남·성태제(1998). **중학교 영어·수학 교과에서의 열린 교육을 위한 수행평가 적용 및 효과분석 연구**. 교육부 초등교육정책과 연구과제.

허경철 외 9인(2000). **수행평가와 교과교육**. 한국교원대학교 출판부.

Bloom, B. S.(Ed.) (1956). *Taxonomy of Educational Objectives: The Classification of Educational Goals: Handbook I, Cognitive Domain.* New York: David McKay Company Inc.

Haertel, E. (1992). Performance measurement. In M. Alkin (Ed.), *Encyclopedia of Educational Research* (6th ed.). London: MacMillan Publishing.

Livacre, J. M., & Wright, B. D. (1994, 2003). A user's guide to *BIGSTEPS rasch-model computer programs.* Winsteps.com.

Mehrens, W., (1992). Using performance assessment for accountability purpose. *Educational Measurement: Issues and Practice.* Spring.

National Council of Teachers of Mathematics. (1989). *Curriculum and evaluation standards for school mathematics*. Reston, VA.: the Author.

National Council of Teachers of Mathematics. (1995). *Assessment standards for school mathematics*. Reston, VA.: the Author.

National Council of Teachers of Mathematics. (1999). *Mathematics assessment a practical handbook for grades 9-12*. Reston, VA.: the Author.

Stiggins, R., & Bridgeford, N. (1982). The role of performance assessment in day to day classroom assessment and evaluation. *Papers presented at the NCME conference*. New York.

Wheeler, J., & Eschberger, D. (2003). Benchmark mathematics performance assessments. *Programs presented at NCTM's 81th Annual Meeting(#678)*. San Antonio, Texas.

The Development of a Tool and Its Application to High Schools for the Assessment in Trigonometry

Choi-Koh, Sang Sook (Dankook University)
Baek, Jung Whan (Dankook Unviersity, Graduate School)

This article was to develop a tool and apply it to the high school classrooms for the performance assessment in trigonometry. Bloom(1956)'s cognitive domain and holistic rubric and analytic rubric(NCTM, 1999) were used to guide the development of 12 problems. To find validity and credibility of this developed tool, Cronbach α and Rasch's BIGSTEPS were used with the samples of high students, 208, using SPSS 10.0K. The results from the investigation, indicated that the tool was very worth assessing students' achievement and there was no difference between the areas where students lived, but were differences between genders as well as between a specialized high school and preparatory high schools.

* key words: performance assessment(수행평가), tool(도구), Bloom's cognitive domain(블룸의 인지적 영역), rubric(채점기준표), trigonometry(삼각함수)

논문접수 : 2004. 1. 30
심사완료 : 2004. 3. 8

<부록> 수학 수행 평가 검사지

수학 수행 평가 검사지

(　　) 고등학교 (　)학년 (　)반 (　)번　이름 (　　) 성별(남, 여)

주의 사항 : 문항 1에서 문항 12까지 모두 풀이 과정을 써라.

1. 다섯 명의 학생이 패스트푸드점에서 각각 한 개 씩 햄버거를 주문했다. 종류는 불고기버거와 새우버거 두 가지가 있다. 불고기버거는 개당 3,000원이고 새우버거는 개당 2,500원이다. 25%의 세금이 붙어 지불한 총 금액이 17,500원이라 하면 학생들은 각각 몇 개의 햄버거를 구입했는지 구하여라.

2. 명철이와 영희는 페인트 작업 아르바이트를 하고 있다. 그러던 중 어느 날 사무실 페인트 작업 일거리가 들어왔다. 각자 혼자서 사무실 한 개를 페인트 작업할 때 명철이는 4시간이 소요되고 영희는 12시간이 소요된다. 시간을 절약하기 위해 함께 작업을 하기로 결정을 했다. 그렇다면 사무실 한 개의 페인트 작업을 모두 마치는데 소요되는 시간을 구하여라.

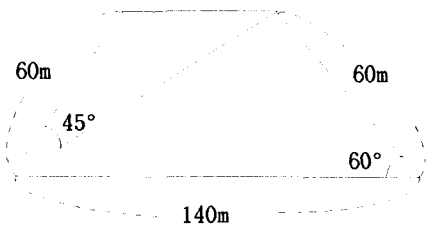
3. θ가 제 2사분면의 각일 때, $\frac{\theta}{3}$는 제 몇 사분면 각인지 구하여라.

4. 부채꼴에서 둘레의 길이가 일정할 때, 넓이가 최대가 되는 중심각의 크기를 구하여라(단, 단위는 라디안).

5. $\triangle ABC$에서 두 꼭지각의 크기 A, B에 대하여 $\sin A=\frac{\sqrt{3}}{2}$, $\sin B=\frac{\sqrt{3}}{3}$이고 변 AC의 길이가 4일 때, 이 삼각형의 넓이를 구하여라.

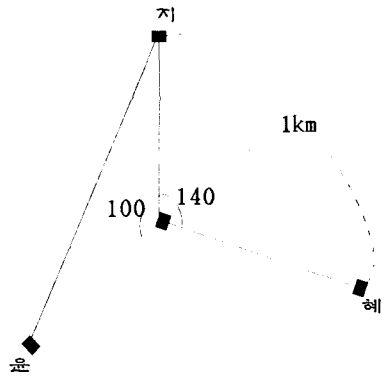
6. $\triangle ABC$의 넓이 S를 변의 길이 a와 두 각 B, C로 나타내어라.

7. 의정부시 k여고의 운동장은 아래 그림과 같다. 이 운동장의 넓이는 얼마인지 구하여라.



8. 지선, 혜인, 윤정의 집은 우리 학교를 중심으로 같은 거리에 있다. 또 아래와 같이 우리 학교를 중심으로 지선과 혜인이의 집 사이의 각도는 140°, 지선과 윤정의 집 사이의 각도는 100°를 이루고 있다. 지선과 혜선이의 집 사이의 직선거리가 1km일 때 지선과 윤정이의 집 사이의 직선거리는 얼마일까요?

(단, $\cos 40°=0.7660$　$\sin 70°=0.9397$)



9. $\triangle ABC$에서 다음 관계가 성립할 때, 이 삼각형은 어떤 삼각형인가?

$$\sin A=2\cos B\sin C$$

10. $\triangle ABC$에서, 코사인(cosine) 제 2 법칙을 이용하여 다음을 증명하여라.

A가 예각 $\Leftrightarrow$ $a^2<b^2+c^2$

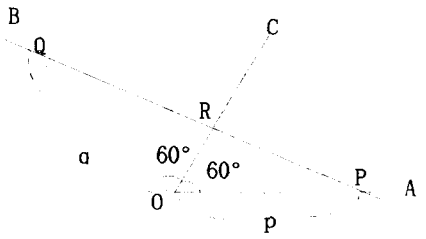
A가 직각 $\Leftrightarrow$ $a^2=b^2+c^2$

A가 둔각 $\Leftrightarrow$ $a^2>b^2+c^2$

11. $\angle AOB=120°$이고 $\angle AOB$의 이등분선을 OC라 하자. 임의의 직선이 반직선 OA, OB, OC와 만나는 점을 P, Q, R 이라 하고, $\overline{OP}=p$, $\overline{OQ}=q$, $\overline{OR}=r$ 라 할 때,

$\triangle OPR+\triangle OQR=\triangle OPQ$ 임을 이용하여 다음을 증명하여라.

$$\frac{1}{p}+\frac{1}{q}=\frac{1}{r}$$



12. 헤론의 공식

($S=\sqrt{s(s-a)(s-b)(s-c)}$, $2s=a+b+c$)을 증명하여라.