

## 수학과 평가의 특성과 문항개발

-고등학교 학업성취도평가를 중심으로-

나 지 영 (서강대학교)

정 순 영 (서강대학교)

### 1. 서론

평가는 학생들이 일정 시점까지 학습해야 할 교육목표를 어느 정도 달성하고 있는지를 측정하여, 그 결과에 근거하여 교육의 효과를 확인하고 나아가서는 교육의 계획 또는 정책의 결정 등에 활용하는 활동을 포괄적으로 의미하고 있다.

이러한 평가는 교사의 의식이나 교육환경의 수준에 상관없이 학습목표에 준거를 두는 소위 준거지향평가가 되어야 한다. 이는 표준 집단의 성적순서에 따라서 학생 개인의 성취도가 평가되는 상대평가에 대응하는 매우 중요한 평가의 과정이다.

그 동안 이 분야에 대한 많은 이론적 발전과 기술적 진보가 있어 왔으나 실제로 학교 현장에서 교사들이 준거지향평가를 이해하고 적용하기에는 거리감이 있어온 것이 사실이다.

이 연구에서는 이 준거지향평가의 이론적 배경을 수학과와 특성에 비추어 조명해 보고 그것을 위한 문항개발의 실제에 대해 연구하고자 한다.

우선 수학과에서는 무엇 즉 어떠한 목표를 측정해야 하는지를 여러 학자들의 이론을 통해 분석해보고 실제로 우리 나라의 개념에 맞는 수학교육의 행동목표를 소개하려고 한다. 또 각 행동목표의 특성을 연구하고, 실제로 문항개발을 하는 데 있어서 어떻게 이러한 행동목표를

측정할 수 있는지 예시문항과 함께 방법을 제시하려고 한다.

특히, 최근 수학교육에 있어서 가장 중요한 목표로 인식되고 문제 해결력(problem solving)에 중점을 두고 논의하려고 한다. 다음으로 수학과와 평가 문항의 개발에 있어서 기술적인 문제와 이론적 배경을 논하려고 한다. 즉, 준거지향평가를 위한 평가 문항 개발의 여러 단계를 소개하고 각 단계에서 수행해야 할 기술상의 원리들을 소개한다.

또, 후반부에서는 각 단계의 예로서 현행 고등학교 수학I의 전반부에 대한 내용 분석과 필수교육목표추출, 평가 목표의 설정과 이원분류표(또는 출제 계획표)의 작성 등의 예를 들고 있다.

이러한 과정을 수행하여 얻어낸 문항을 예비검사를 통해서 검증하는 과정을 실제로 96학년도 학업성취도평가(국립교육평가원 시행(1996))를 예로 들어 설명하였다. 여기서 문항을 어떻게 분석할 것인가에 대해 논한다. 즉, 문항분석의 수단과 분석의 결과를 어떻게 활용할 것인가에 대해 연구하였다. 또, 현장에서의 도움을 주기 위해 문항의 수정과 개선에 관한 과정을 실제 문항과 그 문항 분석 결과를 비교하여 제시하였다(제3장).

이 논문의 근본 목적은 고등학교 수학과 평가에 있어서 새로운 이론을 제시하는 것이 아니라 이론을 정리하여 교육 현장에서 교사들이 직접 수행할 수 있는 평가문항 개발의 한 알고리즘을 제시하는 데 있다.

## 2. 수학과 평가의 특성

### (1) 평가의 기초 이론

학업성취도의 평가란 일반적으로 학교에서 가르친 일정 기간동안의 교과목에 대한 학습내용과 그 수준을 확인하기 위해서 실시하는 시험(test)과 점수부여(grading) 및 그 해석이 동반되는 전반적인 과정을 의미한다. 이는 지능검사, 적성검사, 진로검사 등과는 달리 학교에서 이루어지고 있는 교육활동 전반에 대한 평가로 이해된다. 즉, 학업성취도 평가는 학생들이 얼마나 잘 배웠는가를 측정하는 것이 기본적인 핵심이지만, 학업성취도의 낮고 높음은 곧 학생의 책임으로 설명될 뿐만 아니라, 교사가 얼마나 잘 가르쳤는가, 또는 학교의 교육 행위가 얼마나 타당했는가를 설명해 주는 지표가 되기도 한다.

예컨대, 한 학생의 수학 성적이 낮고 그가 속한 학급의 수학 성적의 평균이 매우 낮다고 하자. 우리는 이 성취도의 결과를 보면서 이 학생의 주변에서 일어나고 있는 교육활동들을 생각한다.

왜 이 학생의 성적이 낮으며 어느 단원, 어떤 행동 능력에 취약한가? 왜 이 학급의 수학 성적이 낮으며 어떤 보완책이 필요한가? 담당교사에는 문제가 없는가? 담당교사가 맡고 있는 다른 반과의 성적 차이는 어떠한가? 다른 지역의 학생들의 성적과 비교해서는 어떠한가? 왜 학생들이 수학을 어렵다고 하는가?

이처럼 한 학생의 학업성취도의 수준은 곧 교육활동과 관련된 모든 것을 평가하는 기본 대상이 된다. 즉, 학업성취도 평가는 학생의 학습활동의 수준의 평가와 함께 그 개인, 교사, 학교, 교육청, 국가의 교육정책 등 교육활동 전반에 대한 평가의 출발점이 된다.

학업성취도에 관련된 평가는 목적과 용도에 따라 몇 가지 유형으로 구분해 볼 수 있다.

첫째는, 상대평가와 절대평가의 구분이다. 학

업성취도를 집단 내에서의 등위(等位: rank)에 관심이 있어서, 학생들을 서열화하는 데 평가의 용도가 있다면, 상대평가를 해야 한다. 공부 잘 하는 순서로 학생들을 분류하려는 목적이라면, 당연히 상대평가가 적격이다. 학기말에 학생 석차를 내려한다거나, 입학시험처럼 다수의 지원자 중에서 성적순으로 입학자를 뽑으려 한다면 상대평가는 필수적이다.

그러나 절대평가는 이와 다르다. 절대평가는 집단 속에서 개인의 위치를 알아내려는 것이 아니라, 학습한 내용의 몇 %를 어떻게 이해하고 있는가를 알고자 할 때 활용되는 평가이다.

물론 상대평가와 절대평가 둘 중 어느 것이 더 좋다 나쁘다는 말은 할 수는 없다. 평가의 목적과 용도에 따라 절대평가를 해야 할 때에 상대평가를 적용한다면, 그것은 문제가 될 수 있다. 우리 나라에서는 상대평가가 남용되어, 절대평가를 해야 할 상황에서조차 상대평가가 적용되어 온 경우가 많다.

둘째는, 형성평가와 종합평가이다. 형성평가는 학업성취도를 측정하되, 학습지도상의 문제를 발견하고 확인하기 위한 평가이다. 그러나 종합평가는 학습의 결과를 확인하고 평가하여 성적을 주기 위한 평가이다. 따라서 형성평가는 학급의 중간 단계에서 실시하는 것이 보통이고, 종합평가는 학습의 완료 시점에서 실시하게 된다. 설사 학습의 중간 단계에서 실시하는 것이라고 하여도, 그 평가의 결과가 최종 성적에 반영되는 것이라고 한다면, 엄밀한 의미에서 형성평가라고 보기 어렵다. 형성평가는 원칙상, 학습지도상의 문제를 발견하여, 학생의 학습 방법, 교사의 수업 지도 방법 등을 변화시키고자 하는 노력의 일환이어야 하기 때문이다.

셋째는, 개인평가와 집단평가이다. 개인평가란 학생의 학업성취도를 진단할 목적으로 시행되는 평가이지만, 집단평가는 개인보다는 집단의 평균 성취도에 관심을 갖는 평가이다. 예컨대, 남·녀 간의 학업성취도에 관심이 있거나, 학교간 지역간, 연령간, 국가간의 점수는 단순

히 집단의 평균치를 얻기 위한 자료로 활용될 뿐이다.

학교에서 교사가 학생의 성적을 준다거나 입학 선발과 관련된 시험에서는 개인평가가 주목적이 되지만, 연구 활동 등에서는 개인보다는 집단의 특성이 주된 관심이므로 집단평가의 개념이 적용되게 마련이다. 예컨대, 국가가 국민들의 기초 학력을 진단하고자할 때에는, 개인보다는 집단평가의 개념에 입각한 평가가 되므로, 해당 집단의 개인 전체를 평가할 필요는 없다. 따라서 그 중의 일부를 표본으로 추출하여 그 평균치를 살펴보게 되는 것이다.

이렇듯 학업성취도 평가에는 여러 유형이 있다. 어떤 유형의 평가를 할 것인지는 순전히 평가자의 목적과 용도에 따라 다르다. 따라서 목적과 용도에 부적합한 유형의 평가를 하지 않도록 하는 일이 중요하다.

## (2) 수학과 의 행동체계 및 평가 요소

학교에서 학생들에게 무엇을 어떻게 가르치고 평가할 것인가를 결정하기 위해서는 먼저 교과와 교육목표 및 학년목표를 근거로 하여 교과에서의 수업목표를 확인 진술하여야 한다. 이 때 기대되는 행동 또는 학습 결과의 유형에 따라 목표들을 분석하는 일이 필요하게 된다. Bloom과 Krathwohl의 교육목표 분류학(taxonomy of educational objectives)에서는 목표를 운동·기능 영역, 인지적 영역, 정의적 영역으로 분류하고 있다.

운동·기능 학습은 감각과 육체적 활동의 통합적인 조정을 수반하는 학습이다. 수학 학습에서의 운동·기능 학습은 자와 컴퍼스를 사용하여 도형을 그리거나 측정할 때와 같은 제한된 범위에서 이루어지며, 대부분의 학습은 인지적이다.

인지적 학습은 정신적 또는 지적 활동을 요하는 학습으로서 학습의 결과로 지식이 축적된다. 사실상, 수학에서 가르쳐지고 배우는 것

은 대부분 인지를 수반하며, 수학은 기본적으로 지적인 노력의 산물이라고 할 수 있다.

정의적 학습은 태도, 가치, 기호, 열심, 책임 등을 수반한다. 대부분의 학교 교사들은 수학 시간에 인지적 학습에 초점을 두는 경우가 많다. 그것은 아마도 수학 자체가 인지적 활동을 요하는 사고의 학문이라는 데서 그 원인을 찾을 수 있을지 모르나, 학교 수학은 인지적 학습에 못지 않게 정의적 학습이 중요함을 인식해야 한다. 그러나 여기서 주의해야 할 점은 인지적·정의적 학습은 어떠한 수학적 내용을 가르치고 배우더라도 항상 결합된 상태에서 이루어져야 한다는 사실이다. 따라서 학습목표는 수학적 내용에 따라 적절한 인지적·정의적 목표가 함께 설정되어야 한다.

특히, 인지적 학습 목표는 학생들이 각각의 학습 목표에 도달 하였는지의 여부를 보일 수 있도록 학생들이 할 수 있는 것을 나타내는 행동적 용어로 자세히 열거되어야 한다는 것을 행동주의 심리학과의 학자들(스키너, 베이거)은 주장한다.

예를 들어, 학습 목표를 '두 분수의 덧셈을 이해한다'라고 설정하였을 때 이 목표에 도달하였는지의 여부를 측정할 때 매우 어려움을 겪게 된다. 그 이유는

첫째, 두 분수란 어떠한 분수를 가리키고 있는가?

둘째, '덧셈'은 개념을 요구하는 것인가 아니면 원리인가?

셋째, '이해한다'란 무슨 뜻이며 어떻게 된 상태가 이해한 수준인가와 같은 애매한 표현 때문이다. 사실 '이해한다'라는 말은 정신적 활동을 뜻하는 용어이기 때문에 그 과정을 볼 수 없다. 우리는 다만 이해의 결과로서 산출되는 것만을 볼 수 있게 된다. 따라서 위와 같은 목표 진술은 '두 분수', '덧셈', '이해'의 수준을 명확히 측정할 수 있는 용어로 수정되어야 할 것이다.

인지적 학습 목표를 행동적 언어로 진술하는

것이 비교적 쉬운 일인 반면 학생들이 정의적 목표에 도달되었는지의 여부를 볼 수 있도록 행동을 미리 제시하는 것은 매우 어렵다. 예를 들면, 학생이 수학의 논리적 구조에 대한 아름다움을 느낄 수 있다는 것을 보기 위하여 학생들이 어떤 행동을 할 것을 기대할 수 있는가? 이 말은 정의적 목표가 잘 진술될 수 없다는 것을 뜻하거나 학생들이 정의적 목표에 도달할 수 없다는 것을 뜻하는 것은 아니다. 여기서 뜻하는 것은 다만 대부분의 교사들이 정의적 목표를 설정하는 것이 어렵고, 정의적 학습을 측정하는 것이 인지적 학습 목표를 설정하고 평가하는 것보다 어렵다는 것을 뜻하는 것이다.

현재 대입수학능력시험, 학업성취도평가 등 수학의 각종 평가에서 공통적으로 쓰이고 있는 평가의 틀은 Bloom의 분류 체계를 수학적 현실에 맞게 수정하여 다음과 같이 이원목표분류표라는 이름으로 사용되고 있다(<표 1>).

<표 1> 수학과목의 이원목표분류표

행동 내용 목표	계산 능력	이 해 능력	추론능력		문제해결력		정의적 영역
			추측 능력	증명 능력	내적 관련 성	외적 관련 성	
집합과 명제							
실수와 복소수							
유 리 식							
무 리 식							
방 정 식							
부 동 식							

이제 <표 1>에 나오는 행동요소에 대한 자세한 설명과 그 목표의 측정을 위한 예문을 들어보겠다.

가) 계산능력

계산능력은 판단이나 결정을 요구하지 않고, 이미 학습한 법칙을 기억하며 단순한 조작만을 통하여 답에 도달하는 알고리즘을 수행하는 능력이다. 주로 수와 식, 집합의 연산, 행렬의 계산, 표를 읽는 능력 등이 여기에 속한다. 따라서 계산능력을 묻는 문항은 다음과 같은 질문의 형태가 있을 수 있다.

- 주어진 수나 식을 간단히 하면?
- 연산 기호로 결합된 집합을 간단히 하면?
- 삼각함수표, 상용로그표를 이용하여 값을 구하면?
- 표준정규분포표를 이용하여 값을 구하면?
- 지수 또는 로그함수로 결합된 수와 식을 간단히 하면?
- 방정식의 해를 구하면?
- 2진법, 3진법 등을 10진법으로 바꾸면?

나) 이해력

이해력을 측정하기 위해서는 기본적인 수학적 개념, 원리, 법칙 등의 정확한 이해 여부를 묻게 된다. 따라서 적어도 중요한 수학적 원리를 이해하고 있다면 답에 도달할 수 있는 문항으로 표현되어야 한다. 예를 들어, 다항식을 주고 도함수를 구하라고 한다면 이는 단순한 계산능력을 묻는 것이 되고 그 함수가 증가하는 구간을 찾으라고 한다면 도함수 원리의 이해 수준을 묻는 문항이 된다. 실제로 이해능력을 측정하는 문항은 과거 우리 나라의 학력고사에 주로 출제되었던 문항들의 대부분이 그 예이다.

따라서 학습의 결과에 대한 보편적인 인지 정도를 묻게 되며 '문제 해결력'과 다른 점은 '이해력'을 묻는 문항은 한 가지의 원리, 법칙, 개념의 이해 수준에 대해 묻게 되지만, '문제 해결력'은 보통 두 가지 이상의 개념, 원리, 법칙 등을 동시에 재구성하여 전략을 수행할 수 있는 능력까지를 요구하게 된다. 이를 위해 다음과 같은 질문의 형태가 있을 수 있다.

- 다음 함수의 성질로 옳은 것은?

- 역행렬의 성질로 옳은 것은?
- 유리수, 무리수의 특성인 것은?
- 다음 상황을 그래프 또는 수식으로 바르게 표현한 것은?
- 다음 수학적 표현으로부터 말할 수 있는 것은?
- 다음 도형의 성질을 바르게 말한 것은?
- 어떤 성질을 이용하여 다음의 값을 구하면?
- 다음 급수 중 수렴하지 않는 것은?

다) 추론능력 : 추측능력

추측능력이란 흔히 귀납적 추론을 말한다. 몇 개의 주어진 사실 또는 상황으로부터 일반적 결과를 발견해내는 능력을 측정하게 된다. 따라서 주어진 상황을 면밀히 관찰하고 종합하여 일반적인 결과라고 추측되는 것을 유추해낸 다음, 실제로 이 결과가 옳은 것인지를 확인할 수 있어야 한다. 예를 들면 평면에 한 개의 직선을 그으면 평면이 둘로 분할되고, 평행하지 않도록 한 개를 더 그으면 평면은 네 개로 분할된다. 이와 같이 직선을 하나 첨가할 때마다 분할의 개수가 어떤 규칙을 따르는지 찾아내는 것이 귀납적 추론, 곧 추측능력의 전형적인 문제 유형이다. 따라서 추측능력의 평가 유형은 거의 수열과 관련되어 있다고 볼 수 있다.

일반적인 질문의 형태는 다음과 같은 것이 있을 수 있다.

- 몇 개의 주어진 항으로부터 수열의 일반항을 구하면?
- 주어진 현상을 관찰하여 얻어낸 법칙은?
- 반복되는 실험으로부터 무한히 계속된 실험의 결과는?
- 주어진 관계를 점화식으로 나타내면?
- 어떤 수들의 합을 무한히 더하면?
- 주어진 도형에 새로이 덧붙여 나갈 때 100 번째 도형의 모양은?

라) 추론능력 : 증명능력

증명능력이란 연역적 추론을 말한다. 삼단논법, 모순법, 간접증명법, 반례 들기, 수학적 귀

납법 등을 이용하여 주어진 명제를 증명할 수 있는가를 묻게 된다. 또 직접적인 증명 문제가 아니라 할지라도 증명과 거의 같은 연역적 추론 과정이 요구되는 문제로 표현될 수 있다. 증명능력을 측정하기 위한 문항은 원칙적으로 주관식(서답형) 형태로 주어져야 하나 객관식 문항으로만 물어야 하는 경우에는 주로 □ 안에 단어 또는 수식을 채워 문장을 완성하는 문항 형태로 제시될 수 있다.

증명 능력에 있어 나타나는 일반적인 물음 형태는 다음과 같은 것이 있다.

- 다음은 어떤 명제를 증명한 것이다. □ 안에 알맞은 것을 순서대로 적은 것은?
- 어떤 문제를 다음과 같이 풀었다. 틀린 곳은?
- 명제를 다음과 같이 증명하였다. 이 과정 중 처음으로 틀리기 시작한 곳은?
- 다음에 주어진 몇 가지 성질들을 이용하여 어떤 명제를 증명하려고 한다. 이 성질들을 어떤 순서로 적용해야 하는가?
- 다음 진술들로부터 범인을 찾으려면?
- 다음 명제 중에서 항상 참인 명제는?
- 다음에 주어진 성질을 이용하여 증명할 수 있는 것은?
- 어떤 선분 위에 10개의 점을 잡을 때, 이 점들 사이의 거리에 대하여 다음 중 바르게 진술한 것은? (비둘기집 원리)

마) 문제 해결력 : 내적 관련성

문제 해결력이란 수학의 어느 한 단위 또는 한 분야에 국한되지 않고 여러 가지 개념·원리·법칙 등이 복잡하게 얽혀 있는 상황으로부터 이를 분석하여 전략을 구성하여 답에 도달하는 능력을 말한다. 이 중에서 특히 내적 관련성이란 교과서 내의 또는 타 교과목이 아닌 수학이란 단일 교과목 내의 개념·원리·법칙 등의 상호 관련성을 묻게 된다.

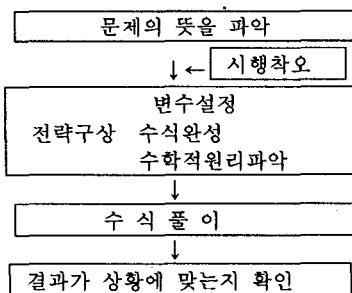
이해력을 측정하는 문항과의 차이는 이해력은 주로 한 두 개의 개념·원리·법칙 등의

간단한 적용으로 답에 도달할 수 있는 형태임에 반하여 문제 해결력의 내적 관련성은 적어도 두 개 이상의 개념·원리·법칙 등을 종합적으로 이용해야 하는 전략이 요구되는 문제 형태라 볼 수 있다. 따라서 고차적인 사고 능력이 요구되는 문제이어야 한다.

문제 해결력은 단순히 공식의 암기나 계산 기술만으로는 해결될 수 없어야 하고 다양한 문제 해결 전략을 사용하여 창의적으로 문제를 해결해 보는 경험을 충분히 가져야 비로소 해결될 수 있어야 한다. 그래프를 이용하거나 또는 몇 가지 경우로 나누어 보는 방법 등 다양한 전략을 요구하며 또한 여러 가지 시행 착오를 거치면서 해결 방안이 모색되어야 한다.

바) 문제 해결력 : 외적 관련성

문제 해결력에 있어서 외적 관련성이란 교과서에서 다루는 수학의 일반적인 내용과 다른 교과목과의 관련성 파악이 요구되는 통합 교과적 소재의 응용 문제를 해결하는 능력을 의미한다. 이른바 과거 우리 나라의 학력고사와 비교하여 현행 수학능력시험에서 가장 강조되고 있는 것이 바로 외적 관련성이다. 지식의 암기나 기술의 습득이라는 차원을 넘어 일상 생활 및 물리·화학·생물·경제·공학 등에서 접하게 되는 다양한 문제 상황 수학이라는 수단으로 해결하는 능력이다. 이러한 문제의 해결 과정은 대략 다음과 같다. (<그림 1>)



<그림 1>

외적 관련성을 측정하는 문항에 있어서 염두에 두어야 할 것은 비록 그 소재가 교과서 외적이고 통합적인 것이라 할지라도 평소에 쉽게 접할 수 있는 상황이라면 그것은 이미 이해 수준의 문항이 될 수도 있음에 유의해야 한다.

사) 정의적인 영역 ( affective domain )

어떤 외적인 현상이나 가치가 한 개인의 내적인 세계로 내면화되는 과정은 우선 그 현상이나 가치가 존재한다는 것을 인식하고 지각하게 되며, 다음에는 이것에 대해 주의를 기울여 반응하며, 나아가 단순히 관여하고 반응하는 수준을 넘어서 그러한 현상이나 활동을 가치롭게 여기며 적극적으로 일관성 있는 반응을 보이게 된다. 이 보다 더 내면화되는 단계에서 조직화된 행동은 더욱 다양화되고 구체화됨으로써 개인의 정의적 생활을 지배하게 된다.

이러한 정의적 행동의 성취 수준을 평가하는 것은 매우 어려운 일이며 상대평가에 있어서는 더욱 그렇다.

수학에 대한 흥미와 태도, 그리고 수학을 연구하려는 동기유발 등 정서란 보통 세 가지 요소를 갖고 있는데 첫째는 대상 즉 수학과, 둘째는 감정으로 이는 어떤 방향과 강도를 갖는다. 세번째 구성 요소는 그러한 방향성과 강도를 갖고 있는 감정에 따라 대상에 작용하는 경향성이다. 정의적인 결과 곧 수학에 대한 강력하고 적극적인 태도를 개발하고 유지하는 것은 수학교육의 중요한 목적이지만 인지적 영역에 비하여 매우 복잡하고 어려운 평가 상의 문제가 제기되며 아직까지 이에 대한 뚜렷한 연구 결과가 제시되어 있지는 못하다.

(i) 태도

수학에 대한 태도에는 여러 가지 유형이 있을 수 있지만, 다음과 같은 문항을 사용하여 수학을 얼마나 좋아하며 다른 교과에 비하여 수학을 얼마나 중요하다고 생각하고 있는지를 알아볼 수 있을 것이다.

- (1) 나는 수학책보다 이야기책을 좋아한다.
- (2) 나는 다른 것을 하기보다 수학을 공부하는 것을 좋아한다.
- (3) 나는 수학에 관한 문제를 공부하기보다 사회과목의 문제에 답하는 것을 더 좋아한다.
- (4) 나는 지리책을 읽기보다 분수의 뺄셈을 하는 것을 좋아한다.
- (5) 나는 수학을 가르치기보다 영어를 가르치는 것을 좋아한다.
- (6) 내가 좋아하지 않는 과목은 수학이다.
- (7) 나는 남들이 수학을 재미있다고 생각하는 이유를 이해할 수 없다.
- (8) 수학은 따분한 과목이다.
- (9) 수학은 재미있는 과목이다.
- (10) 수학은 이해하기가 힘들기 때문에 나는 다른 과목보다 좋아하지 않는다.
- (11) 나는 수학을 잘 하기 위하여 매우 열심히 공부한다.
- (12) 수학은 힘든 일이라기보다 일종의 게임이다.
- (13) 아무리 열심히 노력하여도 나는 수학을 이해할 수 없다.
- (14) 수학은 다른 과목보다도 이해하기가 훨씬 힘든 과목이다.
- (15) 수학은 너무나 어려운 내용이 많기 때문에 재미가 없다.
- (16) 나에게는 수학이 다른 과목보다도 쉽다.
- (17) 나는 “영이의 몸무게는 철수의 몸무게의 반이다. 순이의 몸무게는 영이의 몸무게의 반이다. 병철이의 몸무게는 영이의 몸무게의 반이다. 병철이의 몸무게는 24kg이다. 철수의 몸무게는 얼마인가?” 와 같은 문제보다 “ $359-574+6840-999-4693+97483=?$ ” 와 같은 문제를 좋아한다.
- (18) 나의 수학이 아주 실용적인 것은 아니라고 생각하고 있다.
- (19) 대부분의 수학은 너무 관념적인 것과 관련 되어 있어서 실제적으로 거의 쓸모가 없다.

- (ii) 흥미
  - (20) 나는 프로그래밍을 공부하고 싶다.
  - (21) 나는 수학자가 되고 싶다.
  - (22) 나는 더 깊이 있는 수학을 공부하고 싶다.
  - (23) 학교 밖에서 나는 수학을 이용하고 싶다.
  - (24) 나는 Gauss의 생애에 대하여 연구하고 싶다.
  - (25) 나는 과학을 공부하는데 수학을 이용하고 싶다.

- (iii) 동기유발
 

수학에 대하여 호의적으로 행동하게 하는 충동이나 욕구를 말하며, 이는 행동을 관찰함으로써 간접적으로 측정될 수도 있다.

  - (26) 나는 수학 문제에 대한 해답을 계산하는가?
  - (27) 내가 흥미 있는 수학 문제를 다루며 보낸 시간은 얼마나 되는가?
  - (28) 내가 이번 학기에 스스로 찾아 풀어 본 문제의 수는?
  - (29) 나는 수학 숙제를 다 해 오는가?

- (iv) 불안
 

수학에 대한 공포감이나 불안감은 아동을 자극하여 수학을 열심히 공부하게 하여 성적을 올리는 작용을 할 수도 있는 반면에, 불안감이 너무 크면 정상적인 수학의 학습이 방해될 수도 있다.

  - (30) 나는 수학 시험에 최선을 다한다.
  - (31) 나는 주로 숙제보다도 시험에서 최선을 다하여 좋은 수학 성적을 유지한다.
  - (32) 나는 수학 시험을 치르기 전에는 안정 부절 하지만 일단 시험이 시작되면 안정감을 되찾는다.
  - (33) 나는 내가 한 수학 숙제를 자랑하고 싶다.
  - (34) 나는 수학을 더욱 열심히 공부하고 싶다.
  - (35) 나는 수학을 더 잘 할 수 있었으면 좋겠다.
  - (36) 나는 수학 시간에 당황하지 않았으면 좋겠다.

- (37) 나는 수학 선생님이 내가 수학을 잘 못한다 는 느낌이 들도록 만들지 말아 주었으면 좋겠다.
- (38) 나는 수학 시간에 친구들 앞에 나가서 이야기하는데 어려움을 느낀다.
- (39) 나는 수학 공부 때문에 낙심하고 있다.

(v) 자아개념

수학을 배우는 한 학생으로서의 자기평가

- (40) 수학 시간에 친구들 앞에서 발표를 잘 했으면 한다.
- (41) 수학 숙제를 자랑하고 싶다.
- (42) 수학을 더 열심히 하고 싶다.
- (43) 내 자신이 수학을 잘 못한다고 느끼게 하지 않도록 해주시기를 선생님께 바란다.

(vi) 가치인식

수학의 성취도와 관련된 내용적 측면과 행동적 측면의 매우 미묘하고 복잡한 조합상태를 나타낸다.

- (44) 왜 수학을 연구하는가?
- (45) 수학이란 무엇인가?

(vii) 외적인 가치인식

- (46) 이번 한 주간에 수학을 어떻게 이용하였는가?
- (47) 수학을 나의 일에 어떻게 이용하였는가?
- (48) 대수는 건축가에게 어떻게 이용될 수 있는가?

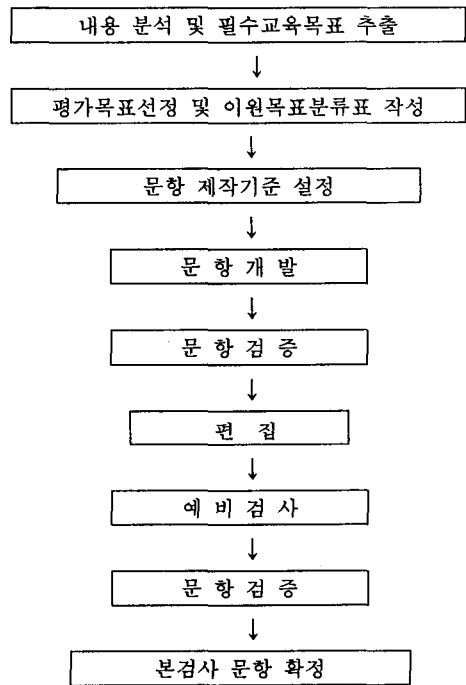
(viii) 내적인 가치인식

- (49) 수학에 대한 어떤 것을 가장 즐기는가?

3. 수학과 평가문항 개발의 실제

하나의 평가문항이 개발되기 위해서는 복잡한 절차를 거치게 된다. 곧 교육과정에 근거하여 학년별 교육내용을 분석하고 목표를 영역별 또는 단위별로 상세화 한 후 필수교육목표를

추출하고 이 필수교육목표 중에서 평가를 통해 측정하고자 하는 가장 핵심적인 평가목표를 선정한다. 이 때, 준거지향평가가 되기 위해서 이원목표 분류표를 만들어 상세한 문항 제작 기준이 설정되어야 한다. 이후 문항개발, 검증의 단계를 반복 또는 feed-back을 통하여 양질의 문항을 만들어 가는 것이다. 그 단계는 보통 다음과 같이 진행된다(<그림 2>).



<그림 2> 수학과 문항개발의 단계

(1) 내용 분석 및 필수교육목표의 추출

가) 내용 분석

교육목표를 진술할 때는 행동 영역과 내용 영역이 동시에 포함되지만, 종착 행동의 진술과 교과of 구체적인 내용을 분리해서 생각하는 것이 일반적이고 더욱 바람직한데, 이는 동일한 교과 내용에 대하여 상이한 행동 영역을 생각할 수 있기 때문이다.



여기에서는 목표 진술을 위한 교과 내용의 분석 방법을 논하고자 한다. 일반적으로 교과 내용은 단원의 장, 절, 항목을 근거로 좀 더 구체적으로 분류할 수 있다. 교과목의 내용을 학습이라는 관점에서 보면 많은 하위(구성)내용 또는 학습과제로 구성되어 있다. 이러한 내용 분석을 좀더 치밀하게 하려는 것이 과제 분석 방법이다. 이 방법은 검사의 목적이나 점수의 해석이 교육 목표의 도달 정도를 따지는 데 집중한 절대평가 방법에서 더욱 중요해진다. 한 단원의 학습을 효율적이고 성공적으로 이끌기 위해서 그 단원 학습을 통해서 학습자에게 주고자 하는 것이 무엇이며, 그 단원 학습이 종결되었을 때 학습자들이 무엇을 할 수 있게 되기를 기대하는지가 명백해야 한다. 교과 내용 속에는 목표(학습목표 또는 수업목표)를 달성하기 위하여 계열적으로 또는 순서적으로 성취되어야 할 학습 요소들이 있게 된다. 이들 학습 요소들을 추출해서 상호 위계적, 수평적 관계를 밝히는 작업이 곧 학습 과제 분석 또는 과제 분석이라 한다.

교과의 성질에 따라 위계성(계열성)이 뚜렷한 교과가 있는가 하면, 교과의 계열성이 명료하지 않는 경우도 있다. 그러나 본질적으로 계열화된 학습 과제로 구성되어 있는 것이 특징이다. 위계성이 뚜렷한 과목은 수학, 과학 등의 교과목이며, 국어, 사회 과목 등은 위계성이 모호하다. 이처럼 위계성이 모호한 과목은 학습 순서에 의한 과제 분석을 하기도 한다.

과제 분석이 철저히 이루어지게 되면, 학습에 수반하는 행동적 기능 요소들과 그 요소들 간의 종적, 횡적 관계 및 전체적인 과제에서 각 요소가 갖는 기능 등이 구체적으로 제시될 수 있다. 바른 과제 분석은 학습자가 과제를 성공적으로 수행할 때 밝게 되는 단계의 순위를 구체적으로 제시해 준다. 따라서, 과제 분석은 학습자가 그 목표를 달성했다는 것을 입증하는 데에 필요로 하는 일련의 학습 절차를 제시해주는 것이다.

과제 분석과 단원 계획을 구별하는 것이 중요하다. 단원 계획은 학습자가 무엇을 해야 할 것인가를 명료하게 하는 쪽에 관계하여 교사가 하고자 하는 일에 중점을 둔다. 단원 계획은 주제 및 과제의 개요를 밝히는 데 중요성을 둔다. 그러나 과제 분석은 교사가 할 일보다는 학습자가 할 일에 중점을 둔다. 과제 분석은 학습자가 습득하기를 기대하는 기능(skills)에 관한 청사진이라 할 수 있다. 과제 분석을 통해서 어떠한 교수 자료를 고안해서 써야 하며, 어떠한 교수법을 활용하는 것이 효과적인 것인지를 결정하게 된다.

과제분석에 관련된 질문은 다음과 같다.

(1) 이 목표를 성취하려면 어떤 기능과 지식이 필요한가?

(2) (과제의 주요한 요소들이 확인되면) 이 요소를 성취하기 위하여 어떤 하위 기능과 개념들이 요구되어지는가? 여기에서 말하는 하위 기능과 개념은 학습자의 출발점 행동이라 할 수 있다. 이 출발점 행동 능력에 일치하는 하위 기능과 개념까지 분석되면 과제 분석은 종결된 셈이다.

보통 과제의 분석에는 위계에 의한 학습 과제 분석과 학습 순서에 따른 학습 과제 분석으로 나눌 수 있는데, 수학과에 있어서는 내용의 계열성 또는 위계성이 분명하게 나타나 있으므로 위계에 의한 분석 방법을 적용한다. 이에 대해 좀더 구체적으로 알아보자.

Gagné에 따르면, 학습 내용을 “인간 활동의 어떤 영역에서 학습자가 나타낼 것을 기대하는 능력”으로 규정하며, 이 학습 내용을 더욱 세분화시켜서 구체적인(좁은) 내용 단위로 구조적으로 나타내려는 작업이 과제 분석이다. 교과목은 이러한 내용 단위들의 구조적 집합이다. Gagné와 Briggs가 개발한 과제 분석 절차는 다음과 같다.

첫째, 최종적 학습 목표가 무엇인지를 분석한다. 이 목표는 그 단원의 수업을 마쳤을 때 학습자들이 실제로 할 수 있기를 기대하는 것

으로서, 일반적으로 “단원 학습 목표”가 여기에 해당된다.

둘째, 상위 학습 요소의 확인 단계이다. 최종적 학습 목표까지 연결된 일련의 학습 요소를 밝히는 일이다.

셋째, 그 단원 학습의 기초가 되는 하위 학습 요소의 분석이다. 교육과정과 교과서의 해당 단원에 포함하는 핵심적인 학습 내용, 필히 학습해야 할 요소 등을 밝히는 일이다. 이 학습 요소들을 충분히 습득하게 되면 상위 학습 요소의 숙달에 도움이 된다. 그러나 충분 조건은 되지 않을 수 있음을 유의해야 한다.

넷째, 본 단원의 학습에 기초가 되는 선수 학습 요소의 분석이다. 이미 학습된 요소들 중 그 단원을 학습하기 위해 필히 숙달되었어야 할 학습 요소를 밝히는 일이다. 일반적으로 계열적인 위계가 명백한 교과목의 경우에는, 앞서 배운 단원의 핵심적인 요소들이 선수 학습요소로서 중요하다.

다섯째, 첫 단계로부터 넷째 단계까지 분석된 학습 요소들이 어떠한 순서로 학습되어야 효과적으로 최종 목표를 달성할 수 있는가를 정하는 일이다. 즉 이들 요소들의 학습 순서를 정하는 일이다. 하위 학습 요소들은 위계적인 구조에 따라 순서가 결정되지만, 동일한 수준의 요소들 간에도 학습 순서를 정한다.

그러나 위계적 과제 분석법은 다음과 같은 문제점을 갖고있다.

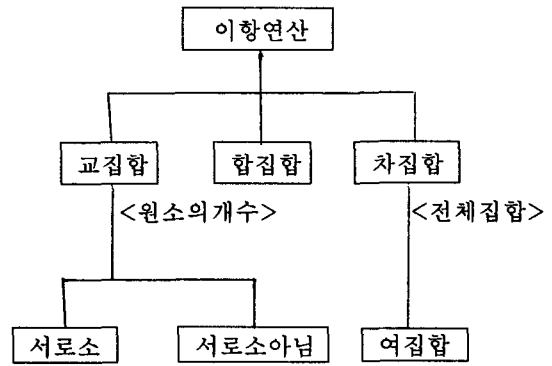
첫째, 행동에 대한 위계적 구조는 복잡하여 반드시 동일할 필요는 없다는 것이다. 분석을 통해 제안된 구조가 학습자들에게 타당할 수는 있지만 모든 학습자들에게 유일하게 타당하지는 않다.

둘째, 분석을 통해 제안된 위계의 구조나 계열이 완전한 것이라는 알 수 없다. 분석된 요소들은 가설적인 것이 많아서 좀 더 많은 경험적 자료에 의해 확인되어야 하는 것이다.

셋째, 누구나 학습 단계로 진입하게 되면 적당한 곳에서 시작을 하게 되고 또 적당한 단계

의 계열을 따르게 된다. 그러한 과정은 목시적일 수도 있고 직관적일 수도 있다. 실제로 어떤 이론에 따라 교과 내용을 구조화하는 것은 학습 행위와 이 학습 행위를 습득하는 계열을 표출하여 뚜렷하게 해주는 하나의 방법일 뿐이다.

다음은 중학교 수학에 있어서 “집합의 연산”(소단원)에 관한 과제 분석의 결과로 제안된 과제 분석표의 예이다 (<그림 3>).



<그림 3>

#### 나) 필수교육목표의 추출

이제 내용 분석이 이루어지고 나면 필수교육 목표를 추출하여야 한다. 여기에서는 이 필수교육 목표를 추출하는 방법에 대하여 알아보겠다.

Oliver는 교육목표들에 해당하는 목표의 기능을 다음과 같이 말하고 있다.

(1) 목표에 도달하는 과정 그 자체는 연수 방안(전략)이다.

(2) 초·중등 교육에 연속성을 유지시키는 것이 교육과정 구성자의 주된 목표이며, 목표를 조합해서 조직적으로 공식화하는 작업은 목표 설정을 도울 것이다.

(3) 목표 진술은 학습 활동의 선정에 있어서 그 기초를 제공해 줄 것이다.

(4) 목표는 평가에 관련 지워진다. 평가는 자료, 방법 및 활동들의 충분함과 적합성을 판단할 기준으로 주어져야 할 공식화된 목표를 필요로 할 것이다.

(5) 목표를 명백히 제시하는 일은 대중에게 학교를 설명하거나 대중이 기대하는 바가 무엇인가를 교육자들이 알도록 하는데 도움이 될 것이다.

(6) 폭넓고 포괄적인 교육목표들은 실제 교실의 장면으로 이행될 수 있다.

(7) 가르침은 학습자의 행동에 변화가 나타나기를 기대하게 된다.

(8) 심리학적 견해에 의하면 목표는 학습 동기를 주게 될 것이다.

Oliver의 제안은 교육과정에서 목표의 기능을 세목이고 행동적인 목표, 즉 수업·평가 목표의 입장에서 진술한 것이다. 교육 목표는 교과 내용의 선정, 조직, 학습 지도 및 평가 과정에 일관성 있게 반영되는 기준이 되어야 한다. 앞서 밝힌 바와 같이 교육과정 목표에 따른 교과서, 교사용 지도서 및 각종 교수-학습 자료를 철저히 분석하는 일은 필수적이다.

목표지향평가의 특성과 평가의 타당도 신장에 기여하는 것으로 알려진 영역 상세화의 원리를 적용하여 필수교육목표를 추출하는 방법을 생각해 보자.

(1) 교육과정에 근거하여 해당 교과목의 목표와 각 학년 목표를 분석한다.

(2) 교과서와 교사용 지도서에 의거하여 수업 목표와 내용을 적절한 기준에 따라 구체적으로 영역화한다. 이 때의 영역은 학년 목표, 단원 목표, 소단원 목표(필요에 따라 차시별 목표까지)의 수준을 의미한다. 영역 설정 과정에서 교육과정이나 교과서의 단원을 합리적이고 융통성 있게 통합하거나 분할할 수도 있다. 단원 내용의 성격이나 비중을 고려하여 한 단원이 보다 구체적인 2~3개의 영역으로 분할될 수도 있고, 몇 개의 단원이 하나의 영역으로 통합될 수도 있다. 이 때, 교과서와 교사용 지도서의 단원 목표, 차시별 목표 및 내용을 체계적으로 분석하여 통합 또는 분할해야 한다. 또한, 목표 진술의 여러 조건을 충분히 반영하여 설정한 평가목표(또는 수업목표)를 체계화해야 한

다. 내용 영역을 상세화할 때는 수업(평가) 목표의 상세화 방법과 비슷하지만, 평가에 소요되는 시간, 문항수 등에 따라 상세화 과정이 제한될 수 있다는 것도 고려해야 한다.

(3) 필수교육목표의 수정, 분류, 종합 정리 및 타당성 검토를 해야 한다. 핵심적인 목표와 내용이 무엇인지 재심의 하면서 중복 문제, 누락된 목표 여부 등을 철저히 검토한다. 교육목표의 타당성은 항상 평가 또는 비판에 의해 필요시 수정·보완·변경될 수 있어야 하고 또한 목표의 실현 가능성, 일관성 등도 중요하게 검토되어야 한다. 예를 들어 다음과 같은 사항들을 검토하여야 한다.

- 설정된 목표 달성에 요구되는 내용 선정과 조직에 무리가 없는가?
- 해당 학문의 특성에 비추어 보아 타당한 목표와 내용인가?
- 교육과정의 기본 원리나 입장에서 보아 타당한가?
- 교육 현장에서 얼마나 큰 비중을 두고 교수-학습되는가?

## (2) 평가목표의 설정과 이원분류표

학력검사에서 준거지향검사(절대평가)의 제작과정은 표준지향검사(상대평가)의 그것과 다르다. 일반적으로 준거지향검사 제작과정은 먼저 목표를 설정 또는 확인한 다음 그 목표의 행동을 분류하고, 목표의 영역을 상세화하고, 끝으로 영역별로 문항을 선정·편집한다. 참고로 준거지향검사와 목표 영역 상세화와의 관계에 대해 알아보자.

준거지향검사에 대한 개념 중에서 가장 보편적인 것은 구체적인 지식과 기능에 대한 정보를 제공하고 학습과제 혹은 학업성취도에 관해 직접적인 해석이 가능한 점수를 산출하는 검사로 규정하는 것일 것이다. 여기서 “준거”(criterion)란 학생이 수업을 통해서 성취하기를 기대하는 어떤 수준의 능력 또는 학습목표를

말한다. 다시 말하면 준거란 검사 점수가 아니라 학생이 성취해야 할 어떤 수준의 지식이나 기능으로서, 학습하기 전에 설정해 놓은 구체적인 학습목표 또는 과제라고 할 수 있다.

일반적으로 형성평가의 경우는 과제 분석에 따른 모든 구체적인 학습 목표를 다 측정한다. 그러나 학기말, 학년말 또는 몇 년 단위의 교육 성과를 알아보는 경우는 구체적인 학습 목표를 모두 다 측정할 수 없다. 따라서 이런 때는 학습과제 또는 목표 영역을 설정하고 그 영역을 대표하는 문항들을 선정하여 검사를 실시한다. 그리고 전체 검사의 결과뿐만 아니라, 영역별 즉 목표별 결과 해석도 하게 된다. 그래서 이런 검사를 특별히 영역지향검사(domain-referenced test)라고 한다. 이 영역지향검사는 준거지향검사의 일종으로 상세화된 과제 또는 목표 영역을 대표하는 문항들로 구성하는 검사다. 영역지향검사에서는 문항 모집단을 만들어 그 영역을 대표할 수 있는 문항들을 선정할 수 있도록 하기 위해 과제 또는 영역을 상세화하는 일이 중요하다.

검사에 대한 준거 지향적 해석의 성질과 타당성은 그 검사의 내용과 문항 선정의 상세화 정도에 좌우된다. 만약 문항 제작자가 단순히 우리 나라 삼국시대의 전쟁사에 대한 지식을 측정하는 문항 10개를 만들었다면, 그 검사의 결과에 대한 준거 지향적 해석이 불가능하다. 그것은 10개 문항이 그 영역을 잘 대표하였다고 볼수 없기 때문이다. 그 10개의 문항에서 하나도 답하지 못한 학생일지라도 우리 나라 삼국시대의 전쟁사에 대하여 어느 정도 알고 있을 수도 있다. 그런데 만약 “우리 나라 삼국시대의 전쟁사 중 중요한 전투에 대하여 학생들은 그 전쟁의 이름, 그 전투의 연도, 그 전투의 승전 장군의 이름을 제시할 수 있다” 라는 목표에 일치하도록 10개의 문항을 만들었다면, 이 검사는 문항에 의해서 표집된 과제 내용이 상세화되었기 때문에 그 결과점수에 대하여는 준거 지향적 해석이 가능하다. 그러나 여기서도

아직 준거 지향적 해석이 만족하게 될 만큼 그 내용과 문항 선정에 대하여 명확하지 못하다. 위의 상세화는 “중요 전투의 선정 기준은 무엇인가?”, “검사에 선정된 전투는 어떻게 선정되었는가?”, “채점에서 철자법은 볼 것인가?” 등의 질문에 대한 명확한 답을 줄만큼 더 상세하게 기술할 필요가 있다.

영역지향검사에서는 각 목표영역에서 선정된 문항들이 그 영역을 대표하는 것이 중요하다. 선정된 문항들이 그 영역을 효과적으로 대표할 수 있게 되기 위해서는 과제 혹은 영역을 정교하게 상세화해야 한다. 많은 목표들의 경우, 검사 문항들을 목표에 맞게 선정할 수 있도록 문항 모집단의 내용을 정교하게 상세화하는 것이 가능하다.

영역지향검사에서의 영역은 넓은 일반적 목표일 수도 있고 하나의 좁은 목표일 수도 있다. 비록 광범위한 목표일지라도 그 내용과 체계를 효과적으로 상세화하면 명확하게 목표를 표현할 수 있다. 여기서 주의할 것은 영역을 정의적 영역이나 독서 기능 영역과 같이 방대한 개념 영역으로 생각해서는 안 된다. 여기서의 영역은 훨씬 더 한정된 개념이다. 학년목표, 단원목표 또는 소단원 목표 등과 같은 수준이다.

잘 상세화된 과제 혹은 영역에서 문항들을 표집하여 만든 검사는 모두 영역 지향적인 검사다. “잘 상세화 되었다”는 말은 과제 혹은 목표 영역이 명확하게 규명 혹은 정의되었다는 뜻이다. 영역지향검사의 또 하나의 특징은 문항 선정시 영역 상세화에 일치하지 않는 문항들은 삭제되거나 표준지향검사에서는 달리 정답률이 100%인 문항들도 검사에서 삭제하지는 않는다는 것이다. 목표영역의 상세화에 들어가기 전에 하나 해야 할 간단한 일은 설정된 목표들의 행동분류이다. 목표가 어떤 유형의 행동인지 확인되어야 그 목표영역 상세화가 바람직하게 되고 나아가서 타당성 높은 문항을 제작할 수 있기 때문이다. 영역 상세화 또는 문항 상세화가 필요한 이유는 학년 목표는 어느 정도 구체적

으로 진술된다고 해도 비교적 일반적이기 때문이기도 하지만 이 목표가 구체적으로 상세화되어야 문항 작성이 용이해지고, 검사의 타당도가 보장되고, 검사 결과의 준거 지향적 해석 능력이 향상되기 때문이다.

상세화 방법은 여러 가지가 소개되고 있으나 여기서는 비교적 적용하기 쉽고 보편적이라고 보는 몇 가지만 소개하겠다. 하나의 영역 또는 목표 상세화 방법이 모든 과제 내지 목표에 다 적용될 수 있는 것은 아니다. 검사의 성격, 목표 또는 과제의 성질에 따라 방법이 다를 수 있다. 따라서 제시된 방법에 너무 구속되지 말고 검사의 질을 향상시킬 수 있는 새로운 방법도 찾아야 한다.

영역지향검사에서는 목표영역의 상세화란 그 목표가 어떤 조건에서 어떤 내용으로 어떻게 행동하는 것인가를 상세하게 분석하여 그 목표 행동을 구체적으로 규명하는 것이다. 다시 말하면 그 목표가 학습을 통해 달성되었다는 증거 행동을 상세히 규명하는 것이다. 이것은 그 목표를 측정할 문항의 성질을 상세히 분석해 놓은 것이다. 그래서 영역 상세화 즉 목표영역 상세화를 문항 상세화라고도 한다.

이제 고등학교 수학I의 전반부에서의 평가목표의 설정, 문항제작 기준, 그리고 평가의 가장 중요한 구조를 알 수 있는 출제계획표(이원목표분류표)를 예로 들겠다(<표 2>).

<표 2> 수학I 평가 목표

영역	평가 목표	행동 분류	문항 제작 기준
대수	1.행렬의 연산 법칙과 상등의 성질을 이용하여 행렬의 성분을 구할 수 있다.	계산	<ul style="list-style-type: none"> <li>행렬은 2차 정사각행렬로 제시한다.</li> <li>행렬의 상등과 덧셈·뺄셈, 실수배 및 곱셈 등의 연산을 포함하도록 한다.</li> <li>미지 성분은 2개 이내로 한다.</li> </ul>

영역	평가 목표	행동 분류	문항 제작 기준
대수	2.행렬의 연산법칙을 이해하고 이를 활용할 수 있다.	이해	<ul style="list-style-type: none"> <li>행렬의 곱셈이 수의 곱셈과 다른 성질을 포함하도록 한다.</li> <li>연산에 사용되는 행렬의 개수를 3개 이내로 한다.</li> <li>역행렬은 사용하지 않는다.</li> </ul>
	3.일상생활에서의 문제를 행렬을 사용하여 표현하고 이를 해결할 수 있다.	문제 해결   외적	<ul style="list-style-type: none"> <li>관광 코스, 교통망, 농도계산 등 일상생활의 소재를 택한다.</li> <li>단순한 표현 수준에 그치지 않고 간단한 연산을 포함한다.</li> <li>행렬의 크기는 3차 이하로 한다.</li> </ul>
	4.역행렬의 뜻과 존재 조건을 이해하고 단위원행렬, 행렬의 곱셈의 성질을 안다.	문제 해결   내적	<ul style="list-style-type: none"> <li>역행렬과 거듭제곱이 포함되도록 한다.</li> <li>행렬의 2차 정사각행렬을 제시한다.</li> <li>미지 행렬을 구할 때, 미지의 성분은 2개 이하로 한다.</li> </ul>
	5.역행렬의 성질을 이해하고 이를 연역적 추론에 활용할 수 있다.	증명	<ul style="list-style-type: none"> <li>완성형이 아닌 진위를 판단하는 문제로 참인 것은 간단한 증명으로 밝힐 수 있고 거짓인 것은 간단한 반례를 들 수 있는 것으로 한다.</li> <li>행렬의 수는 3개 이하로 한다.</li> <li>기본적 성질이 아닌 특정한 성질을 증명하도록 한다.</li> </ul>
	6.역행렬의 성질을 이용하여 행렬의 연산 및 상등에 관한 문제를 해결할 수 있다.	문제 해결   내적	<ul style="list-style-type: none"> <li>1~2개의 이차 정사각행렬을 제시한다.</li> <li>역행렬의 존재 여부를 판단하는 것을 포함시킨다.</li> <li>두 행렬의 복합적인 식으로 제시할 수도 있다.</li> </ul>
	7.특정한 조건을 가지는 연립 방정식의 미정계수를 구할 수 있다.	문제 해결   내적	<ul style="list-style-type: none"> <li>불능, 또는 부정의 경우의 연립방정식을 제시한다.</li> <li>미정계수는 2개 이하로 한다.</li> <li>근과 계수와의 관계를 이용하는 문제를 제시할 수도 있다.</li> </ul>

<표 3> 수학 I 출제 계획표

내용 영역	행동 수준	계산	이해	추론		문제해결		문항수	비율 (%)
				귀납, 유추, 추측	증명	수학내적 관련성	수학외적 관련성		
행렬		○						9	45
			○						
			○						
				○					
					○				
						○			
							○		
수열과 순서도		○						11	55
			○						
			○						
			○						
				○					
					○				
						○			
							○		
문항수		2	5	2	2	6	3	20	
비율 (%)		10	25	10	10	30	15		100

(3) 문항개발

타당성 검토를 거쳐서 최종적으로 평가목표가 선정되면 문항개발이 이루어진다. 문항은 상세화된 평가 목표를 잘 반영하고 그 영역을 잘 대표할 수 있어야 한다.

좋은 문항과 타당도 그리고 신뢰도 높은 검사 도구를 제작하기 위하여 다음과 같은 점이 매우 중요하다.

첫째, 교육목표와 교과 내용, 그리고 교과과

정에 대한 충분한 이해가 있어야 한다. 즉 학생이 무엇을 알아야 하고, 무엇을 잘 못 알고 있는지를 파악하여야 한다. 이럴 때 검사 도구의 타당도로서 교과 타당도와 교수 타당도를 증진시킬 수 있다.

둘째, 피험자 집단의 특성을 잘 알고 있어야 한다. 학습 발달 수준뿐만 아니라, 그들의 어휘 수준도 파악하여야 그 집단에 적절한 문항을 제작할 수 있다. 만약 질문에 어려운 단어를 사용한다면, 질문의 답을 알고 있다 하여도 질문에 사용된 단어의 의미를 모르므로 답이 틀리게 되는 경우가 있다. 이는 검사 도구의 신뢰도를 저하시키는 요인이 된다.

셋째, 측정 이론을 숙지하여야 한다. 문항의 종류와 각기 다른 문항 형태에 따른 특징과 장단점, 그리고 문항 제작 방법을 이해하여야 한다. 또한 어떤 문항이 좋은 문항인지를 숙지하여야 한다.

넷째, 문장력이 필요하다. 질문의 내용을 간결 명확하게 글로 표현하는 능력이 필요하다. 산만하고 긴 문제는 자칫하면 질문의 요지를 상실하게 하여 검사 도구의 신뢰도를 저하시키는 원인이 된다.

마지막으로 경험이 필요하다. 제작자가 위에서 설명한 조건들을 지니고 있다 하더라도 문항을 제작해 본 경험이 없으면 좋은 문항을 제작하기란 쉽지 않다. 문항을 제작·분석하고 수정·보완하는 경험을 통하여 새롭고 참신한 문항을 제작할 수 있다.

좋은 문항이 되기 위한 절대적 판단 기준은 없다. 그러나 좋은 문항이 되기 위해서는 몇 가지 고려할 점이 있다.

첫째, 문항이 타당성이 있어야 한다. 즉, 측정하고자 하는 내용을 정확하고 구체적으로 담고 있어야 한다.

둘째, 문항이 복잡성(complexity)을 지녀야 한다. 복잡성이란 의미는 질문의 내용이 단순 기억에 의한 사실보다는 고등 정신 기능으로서 분석, 종합, 평가의 능력을 측정할 수 있는 문

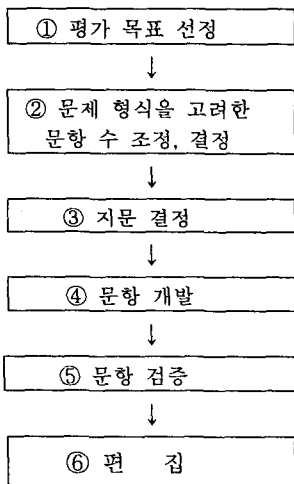
항이어야 한다.

셋째, 문항의 발전성을 들 수 있다. 질문이 열거된 단순 사실만을 질문하는 것이 아니라, 열거된 사실들을 요약하고, 일반화하고, 나아가 추상화시킬 수 있는 내용을 포함하여야 한다.

넷째, 문항의 참신성이다. 이는 내용적 측면이나 형식적 측면에서 기존에 존재하는 화석화된 형태의 문항이 아닌 문항을 의미한다.

여섯째, 문항의 난이도가 적절하여야 한다.

앞에서 언급한 바와 같이 하나의 평가문항이 개발되기 위해서는 복잡한 절차를 거치게 된다. 곧 교육과정에 근거하여 학년별 교육 내용을 분석하고 목표를 영역별 또는 단위별로 상세화한 후 필수교육목표를 선정, 이 목표를 평가할 수 있도록 문항을 제작해야 한다. 먼저 해당 단원의 교육과정 상의 내용 영역과 행동 영역을 분석·확인하여 평가 목표를 선정하는 것이 필요하다. 이후 문항 형식을 고려하여 문항 수를 조정 및 결정하고, 해당 목표를 측정할 수 있는 지문을 선정한 뒤 문항을 개발을 예비로 완성하게 된다. 물론 목표 이원분류표를 작성, 목표 체계를 뚜렷이 하고 문항 제작 기준을 작성한 후 해야할 일이다. 이러한 일련의 과정은 다음과 같이 흐름도로 요약된다(<그림 4>).



<그림 4>

#### (4) 문항의 분석

시험(test)은 여러 개의 문항으로 구성되기 때문에 시험의 질은 시험을 구성하는 각 문항의 질에 따라 결정된다. 이처럼 시험의 양호도는 각 문항의 양호도에 따라 결정되기 때문에 시험의 양호도를 알기 위해서 문항 분석은 필수적이다.

문항 분석은 첫째, 교수 효과를 평가하고 둘째, 각 문항의 성취 정도를 조사하여 개인 및 집단의 학습 수준을 높일 수 있다. 고전검사이론(성태제(역), 1991)에 의하면 문항 분석의 결과는 각 문항이 너무 쉬운지 혹은 너무 어려운지, 전체적으로 시험 점수가 높은 사람과 낮은 사람을 잘 구별해 주는 문항인지, 그리고 문항의 답지들이 제 기능을 발휘하고 있는지를 알려준다. 문항 분석의 결과에 따라 선별된 문항으로 구성된 시험은 문항 분석을 거치지 않은 시험보다 높은 신뢰도와 타당도를 가질 것이다. 일반적으로 문항의 양호도를 분석하기 위해서는 문항 정답률( $P_a$ ), 문항 변별도( $DI$ ), 답지 반응률( $R$ )을 산출한다. 이들에 대해 좀더 구체적으로 알아보자.

##### 가) 문항 정답률( $P_a$ )

문항 정답률은 어떤 문항에 정답을 한 응시자들의 백분율(%)로 표시된다. 문항 난이도나 문항 곤란도라고도 하는데, 이는 어떤 문항이 얼마나 쉽고, 어려운가의 정도를 말하며, 문항  $a$ 에 대한 정답을 한 학생의 비율인  $P_a$ 값이 크면 쉬운 문항이고 낮으면 어려운 문항이다.

검사의 정답률을 어떤 수준으로 하는 것이 적절한가는 검사의 목적과 성질에 따라서 다르나, 개인차의 변별을 목적으로 하는 규준지향검사(상대평가)는 검사 점수의 분산이 클수록 바람직하다. 모든 수준의 응시자를 타당하게 변별하기 위해서는 문항의 난이도를 넓게 퍼지게 하여 능력이 높은 수준의 피험자에게는 도전적인 문항이 되고, 낮은 수준의 피험자에게는 자

기의 능력을 나타낼 수 있는 문항이 되게 하여야 한다. 일반적으로 규준지향검사는 정답률이 20%~80%인 문항으로 구성되며, 그 사이에서 정규분포를 하도록 하는 것이 적절하다. 즉 문항의 정답률이 50%인 것이 가장 많고 차츰 적은 수의 문항이 20% 또는 80% 수준의 정답률을 나타내고 평균 정답률이 50% 정도가 되도록 하는 것이 바람직한 것으로 알려져 있다.

<표 4> 문항 정답률의 언어적 해석 기준

문항 정답률	언어적 해석
35%~85%	양 호
30%~34%, 86%~90%	판단 유보
29% 이하, 91% 이상	불 량

문항 정답률의 양호도를 판단하기 위해서 사용하는 기준은 <표 4>와 같으며, 객관식과 주관식의 문항 정답률을 산출하기 위해서는 다음과 같은 식을 사용한다.

(i) 객관적 문항 정답률(Pa)

$$Pa = \frac{R}{N} \times 100 (\%)$$

R : 정답자 수, N : 응시자 수

(ii) 주관식 문항 정답률(Pa)

논술형 문항의 경우에는 한 문항에 대하여 완전한 답안이 못할지라도 정답이나 모범답안에 어느 정도 접근했느냐에 따라 여러 단계의 중간 점수나 부분 점수를 줄 수 있기 때문에 다음과 같이 정답률을 계산한다.

$$Pa = \frac{\sum X}{N \times M_{\max}} \times 100 (\%)$$

$\sum X$  : 해당 문항에서의 응시자 득점 총계

$M_{\max}$  : 해당 문항의 배점(만점)

N : 응시자 수

나) 문항 변별도 (DI)

문항의 변별도는 각 문항이 그 검사가 측정하려는 특성 혹은 능력을 가지고 있는 사람과 그렇지 않은 사람을 구별하는 정도를 말한다. 한 문항이 높은 점수를 받을 피험자에게는 쉽고, 반면 낮은 점수를 받은 피험자에게는 어려우면 그 문항은 변별력이 높은 문항이다. 어떤 문항에 대하여 상위집단의 정답률과 하위집단의 정답률이 비슷하거나 하위집단의 정답률이 상위집단보다 높다면, 그 문항은 측정하려는 능력의 상·하위집단을 잘 변별하지 못하는 문항이다. 따라서 문항의 변별도는 검사의 총점이라는 준거에 비추어 본 문항의 타당도로 볼 수 있다.

음의 변별도 지수는 하위집단 피험자의 정답자 수가 상위집단 정답자 수보다 더 많을 때이다. 변별도 지수가 0에 가까운 경우는 상·하위집단의 정답자 수가 거의 같은 경우이거나 정답자가 거의 없는 경우이다. 문항 변별도 지수가 음이거나 0에 가까운 문항은 삭제되거나 수정되어야 한다.

양의 변별도 지수를 갖는 문항은 상위집단의 정답자 수가 하위집단의 정답자 수보다 많은 바람직한 문항이다. 좋은 문항의 변별도는 적어도 0.25이상은 되어야 한다.

<표 5> 문항 변별도 지수와 문항 평가 및 언어적 해석

문항 변별도 (DI)	문 항 평 가	언어적 해석
0.40 이상	매우 양호한 문항임.	양 호
0.30~0.39	양호한 문항으로 개선의 여지는 있음.	
0.20~0.29	경계선에 있는 문항으로 개선의 여지가 많음.	판단 유보
0.19 이하	양호하지 못한 문항으로 완전한 수정이나 삭제를 요함	불 량



문항 변별도 지수와 문항 평가 및 언어적 해석은 <표 5>와 같으며, 다음과 같은 식을 사용하여 문항 변별도를 산출할 수 있다.

(i) 객관식 문항 변별도(DI)

$$DI = \frac{R_H - R_L}{0.5N}$$

$R_H$  : 총점 상부집단 50%의 정답자 수  
 $R_L$  : 총점 하부집단 50%의 정답자 수  
 $N$  : 응시자 수

(ii) 주관식 문항 변별도(DI)

$$DI = \frac{R_H - R_L}{0.5N(X_{\max} - X_{\min})}$$

$R_H$  : 총점 상부집단 50%의 해당 문항에서의 득점 합계  
 $R_L$  : 총점 하부집단 50%의 해당 문항에서의 득점 합계  
 $N$  : 응시자 수  
 $X_{\max}$  : 해당 문항에서 받을 수 있는 최고 점수  
 $X_{\min}$  : 해당 문항에서 받을 수 있는 최저 점수

다) 답지 반응률 (R)

문항의 답지에 대한 반응의 분포 상태가 답지 반응 분포이다. 각 답지 반응률의 분포를 분석하여 오답지의 능률성을 조사하면 보다 매력적이고 제 능력을 가질 수 있는 오답지를 만드는 데 도움이 된다. 그리고 피험자의 그릇된 이해와 해석에 관한 진단이 가능하여 학습 지도에 도움이 될 것이다. 오답지가 제 구실을 하려면 매력이 있어야 하며, 오답지가 함정이 되어서는 안 된다. 그러기 위해서는 모든 오답지가 고르게 매력 있고 고른 답지 반응 분포를 나타내는 것이 좋다. 따라서 선다형 문항 제작 기술을 개발하기 위해서는 답지 반응 분포의 분석

은 필수적이다. 우리는 각 답지에 대한 답지 반응률을 산출하기 위해서 다음과 같은 식을 사용한다.

$$R = \frac{W_i}{N} \times 100(\%)$$

$W_i$  : 답지에 대한 응시자 수

$N$  : 응시자 수

라) 문항 분석의 실제

그 동안 언급한 문항 제작 과정을 통하여 제작된 '96 학업성취도평가(국립교육평가원, 1996)의 예비검사지는 ㉗, ㉘형 두 가지 형태로 편집되어 전국에서 표본으로 추출된 8개 학급 200여명을 대상으로 실시하였다. 그 결과 각 과목의 평균점과 표준편차는 <표 6>과 같았다.

<표 6> 각 과목의 평균점과 표준편차

과목	구분	문제 유형	평균	표준 편차	응시자 수
공 통 수 학		㉗형	58.71	17.91	197
		㉘형	62.11	19.33	199
수 학 I		㉗형	53.08	17.79	185
		㉘형	45.54	18.10	171
수 학 II		㉗형	61.03	19.36	184
		㉘형	53.05	18.13	195

또, 이들과 각 과목의 문항 분석의 통계는 <표 7>, <표 8>과 같았다.

<표 7> 공통수학 ㉗형 정답률 및 변별도

문항	정답	정답률	변별도	답지 반응률				
				1	2	3	4	5
1	4	83.8	0.26	2.0	12.7	0.5	83.8	1.0
2	2	29.4	0.12	9.1	29.4	7.1	6.1	47.2
3	2	37.1	0.41	9.1	37.1	20.8	19.3	13.7
4	5	40.6	0.17	39.1	9.1	9.6	1.5	40.6
!		!	!					
S1		45.7	0.49					
!		!	!					

\* S는 주관식 문항을 의미한다.

<표 8> 수학 I ㉔형 정답률 및 변별도

문항	정답	정답률	변별도	답지 반응률				
				1	2	3	4	5
1	1	96.8	0.18	96.8	2.2	0.5	0	0.5
2	4	80.0	0.37	0.5	8.6	3.8	80.0	7.0
3	2	62.7	0.32	4.3	62.7	17.3	10.3	4.9
4	5	62.2	0.38	14.1	9.2	9.2	5.4	62.2
⋮								
S1		73.5	0.52					
⋮								

예비검사의 문항 분석 결과(<표 7, 8, 9>)에 의하면 여러 가지 측면에서 예비검사 문항이 양호한 것으로 나타나 있음에도 불구하고 정답률이 20%이하 또는 80%이상인 문항과 변별도가 0.2 이하인 것이 나타나고 있다.

<표 9>

구분 과목	문항 번호	평균	문항 수	정답률						변별도	
				20% 이하		20% ~ 80%		80% 이상		0.2 이하	
				문항 수	%	문항 수	%	문항 수	%	문항 수	%
공통 수학	㉔	58.70	20	·	·	16	80	4	20	2	10
	㉕	62.11	20	·	·	16	80	4	20	1	5
수학 I	㉔	53.08	20	1	5	17	85	2	10	1	5
	㉕	45.54	20	3	15	16	80	1	5	·	·
수학 II	㉔	61.03	20	1	5	14	70	5	25	1	5
	㉕	53.05	20	1	5	16	80	3	15	1	5

㉔, ㉕형 두 세트를 통합하여 한 세트의 본 검사 문항을 확정하려고 하므로 다음과 같은 원칙으로 본 검사 문항을 선정한다.

- 우선 변별도가 0.2이하인 것은 모두 버리되 ㉔, ㉕형 모두가 0.2이하인 것은 대폭 수정한다.

- 정답률이 20%이하인 문항과 80%이상인

문항은 그 원인을 분석하여 주로 정답률이 높아지는 것으로 택하여 전체 정답률이 50~60%가 되도록 한다.

- 현장 교사에게 모델이 될 수 있는 문항, 평가목표의 달성에 필수적이라 생각되는 문항을 선제한다.

#### 4. 결 론

이 논문의 서두에서 언급한 바와 같이 이 논문의 목적은 수학과 평가의 새로운 이론을 제시하는 데 있지 않고 기존의 이론을 재검토하여 수학과 평가문항 개발의 이론적 배경을 제시하고 평가문항 개발에 있어서의 기술적 알고리즘을 제시하는 데 있다. 좋은 평가는 좋은 평가문항으로만 이루어지는 것이 아니다. 양질의

문항이 우선 개발되고 그것의 모임이 하나의 구조를 이루어 각각의 문항이 그 구조 속에서 주어진 역할을 하여야 한다는 것이 가장 중요한 점이다. 그 구조와 역할이 올바르게 파악되고 결정될 때 비로소 하나의 훌륭한 평가가 이루어 질 수 있다.

#### 참 고 문 헌

국립교육평가원 (1995). 고등학교 학업성취도평가 연구.  
 \_\_\_\_\_ (1996). 고등학교 학업성취도평가 연구.  
 \_\_\_\_\_ (1996). '96학년도 대학수학능력 시험 해설.  
 \_\_\_\_\_ (1991). 대학수학능력시험 2, 3, 4차 실험평가 결과 보고서, 비간행물.  
 \_\_\_\_\_ (1992). 대학수학능력시험 문항 개발에 관한 연구, 비간행물.  
 \_\_\_\_\_ (1996). 학업성취도 평가 우수 문항집.  
 박도순 (1996. 4.). 교육의 본질과 학업성취도, 교육진흥(봄호).

\_\_\_\_\_ (1976). 문항 작성 원리와 실제, 한국교육개발원.  
성태제 (역) (1991). 문항반응과 이론입문, 서울: 양서원.

## A Characteristic of the Assessment of School Mathematics and Development of a Items

- Focused on the Achievement Test for High School Mathematics -

**Na, Ji-Young**

Department of Mathematics, Sogang University Seoul 121-742, Korea

**Chung, Soon-Yeong**

Department of Mathematics, Sogang University Seoul 121-742, Korea;

email: sychung@ccs.sogang.ac.kr

The purpose of this paper is to investigate a characteristic of the assessment of high school math and to introduce an algorithm for school teachers to perform the achievement tests for math in classrooms.

Moreover, we introduce a system of objectives which will be appropriate for math in the situation of Korea and illustrate our approach via '96 High School Achievement Test held by Korea National Board of Educational Evaluation.