

수학과 평가 문항제작의 실제

전영주¹⁾

수학과 평가는 학생들의 수학적 사고력과 성취도를 측정하는데 있다. 또한 이를 분석·활용하여 수학교육의 질을 높이는데 그 목적이 있다. 이러한 평가는 좋은 문항제작에서부터 시작된다. 이에 본고는 문항제작 관련 이론을 고찰하고, 교육목표에 따른 예시 문항과 문항제작시의 유의점을 살펴봄, 계획 단계, 초안 문항제작 단계, 검토 단계, 최종 문항제작 단계 등 단계별 문항제작의 실재를 안내하고자 한다.

주요용어 : 수학과 평가, 수학과 문항제작

I. 서론

학교의 수학과 평가는 교수·학습 측면에서 두 가지 목적을 지닌다. 하나는 학생들이 교육 과정에 제시된 수학의 기본 개념·원리·법칙을 이해하고 있는지, 또 이를 적용하여 주어진 문제를 해결할 수 있는지를 알아보는 데 있으며, 다른 하나는 평가²⁾ 시행 자료를 분석·활용하여 수학교육의 질을 관리하고 높이는 것에 있다. 이와 같은 수학과 평가 목적을 달성하기 위한 제반사항으로 평가 틀과 평가 문항이 우선적으로 필요하다.

평가 틀은 좁은 의미로 평가도구의 개발 과정에서 고려해야 하는 항목에 대한 지침, 안내 및 준거가 되는 사항들을 말하며, 넓은 의미로는 평가의 전 과정에서 고려해야 할 항목에 대한 지침, 안내 및 준거가 되는 사항들을 말한다(황혜정 외, 1999). 평가 문항은 평가 틀을 기반으로 평가목표(why), 평가내용(what), 평가방법(how) 등의 요소를 담아 학생이 무엇을 알고 무엇을 모르거나 잘못 알고 있는지를 파악할 수 있는 측정도구라 할 수 있다. 따라서 좋은 문항은 평가의 성패를 좌우하는 준거가 되며, 좋은 문항제작은 성공적인 평가의 시발점이 될 수 있다.

이런 점에서 수학과 평가 문항제작 연구를 살펴보는 것은 의미가 있다. 그러나 최근의 수

1) 한국교육과정평가원 (whaljuro@kice.re.kr)

2) 평가와 비슷한 용어들이 아주 많다. 예를 들면 측정, 사정, 검사, 시험 또는 고사 등과 같은 낱말이다. 하지만 일반적으로 크게 두 가지 ‘assessment’ 와 ‘evaluation’ 으로 구분한다. ‘assessment’ 는 수학을 사용하는 능력과 수학적 성향에 관한 학생의 지식에 대한 정보를 수집하고 다양한 목적을 위해 수집한 정보로부터 결론을 도출하는 과정을 말하며, ‘evaluation’ 란 주의 깊은 조사와 판단에 기초한 어떤 것의 가치를 결정하거나 그 가치를 주는 가정을 말한다(NCTM, 1995). 여기서 말하는 평가의 의미는 ‘evaluation’ 이다.

학과 평가 문항제작 연구로는 수학적 지식 영역과 비교하여 평가 문항 제작에 필요한 지식의 영역이 무엇인지를 알아본 이석현(2010)의 ‘평가 문항 제작에 필요한 교사의 지식에 대한 연구’와 전국연합학력평가에서 출제되었던 문항 수정 과정을 소개한 신범영(2010)의 ‘바람직한 수학 평가 문항 제작에 대한 연구’ 정도이다. 하지만 이 연구들도 평가 문항을 제작하기 위해 교사가 갖추어야 할 소양 부분과 전국연합학력평가에 출제된 문항을 통해 문항 작성 요령과 오류를 피할 수 있는 방법을 살펴보는 정도이다.

이에, 본고 ‘수학과 문항제작의 실제’는 수학교사들에게 평가의 공정성과 객관성 등 평가 기능 제고는 물론, 수학과 교육과정의 내용과 수준을 고찰할 수 있는 기회를 제공해서 수학교사가 가르쳐야 할 교육과정에 대한 충분한 이해와 학생이 배워야 할 중요 내용을 점검할 수 있도록 현장에 직접적인 도움을 주고자 하였다. 이와 관련하여 먼저 수학과 문항제작의 이론을 고찰하고, 행동영역에 따른 예시 문항과 문항제작시의 유의점을 살펴보고, 계획 단계, 초안 문항제작 단계, 검토 단계, 최종 문항제작 단계 등 문항제작의 실재를 안내하고자 한다.

II. 이론 고찰

1. 내용영역과 행동영역

NCTM(1989)은 「Curriculum and Evaluation Standards for School Mathematics」에서 평가를 일반평가, 학생평가, 프로그램평가 세 부분으로 조직하고, 일반평가 3개 기준, 학생평가 7개 기준, 프로그램평가 4개 기준 등 모두 14개의 기준을 제시하였다.

특히 모든 평가는 교육과정 근거, 일관성, 평가의 목적에 맞는 다양한 평가 기법을 사용해야 한다는 일반평가 3개 기준과 수학적 힘, 문제해결, 의사소통, 추론, 수학적 개념, 수학적 절차, 수학적 성향 등 수학과 평가에서 평가되어야 할 수학적 지식에 대해 논하고 있는 학생평가 7개 기준은 교육목표를 내용영역과 행동영역으로 분류한 Bloom(1956)의 이원분류표(Specification Table of Educational Objectives)와 비교하여 상호 유사 개념으로 짝지을 수 있다. 예를 들면, NCTM의 일반평가 기준에서의 교육과정, 학생평가 기준을 Bloom의 내용영역과 행동영역으로 각각 대응시키는 것이다.

이렇듯 연구자마다 다루는 평가 기준이나 용어는 다르지만 이들의 개념이 유사한 것은 평가에서 다루고자 하는 근본이 같기 때문이다. 그 핵심에는 교과와 대단원, 중단원 등 평가하고자 하는 내용 영역과 각 문항의 내용을 측정할 때 어떤 인지능력을 측정할 것인가 하는 행동영역이 있다. 더욱이 Bloom(1956)에 의해 제안되어, 내용영역과 행동영역으로 교육목표를 작성하는 이원분류표(Specification Table of Educational Objectives)의 지속적인 사용은 현재까지 더 나은 평가기준을 정하지 못했기에 가능하다. 변창진 외(1998)의 학습목표의 성질을 구조적으로 파악할 수 있고, 문항 표집을 할 때도 최적으로 사용할 수 있는 것이 이원분류표라고 한 것에서 그 이유의 하나를 찾을 수 있다.

내용영역은 교육과정의 교수·학습목표를 따르고, 행동영역은 대체적으로 Bloom(1956)³⁾의

3) Bloom(1956)의 인지적 영역 교육목표분류학에 의한 세분화된 분류는 지식(Knowledge) 하위분류 12가지, 이해(Comprehension) 하위분류 3가지, 적용(Application), 분석(Analysis) 하위분류 3가지,

교육목표분류학에 근거하지만 교과 특성과 검사 목적에 따라 측정하고자 하는 영역이 다를 수 있다. 수학과는 측정하고자 하는 인지능력 수준을 한국교육과정평가원(이하평가원, 2005)이 제시한 계산, 이해, 추론, 문제해결로 분류하여 사용하고 있다. 우리의 대학수학능력시험(이하 수능)에 해당되는 중국의 고시(高考)에서는 연산능력(運算能力), 사유능력(思維能力), 실천능력(實踐能力), 창의적 사고능력(創新意識), 공간상상능력(空間想象能力) 등 5개의 능력으로 나누어 평가하고 있는데, 평가원의 인지능력과 비교하면 연산능력은 계산, 사유능력은 이해 및 추론, 실천능력과 창의적 사고능력은 문제해결과 비슷하다. 공간상상능력은 조건에 따라 정확한 도형을 그리는 능력과 상상한 도형을 직관적으로 표출할 수 있는 능력을 말하는 것으로 이해, 추론 및 문제해결에 조금씩 녹아들어 있다.(전영주, 2011).

2. 문항제작

문항제작을 할 때, 우선 고려해야 하는 것은 누구를 평가할 것인가 하는 평가 대상과 무엇을 평가할 것인가 하는 교육과정의 교수·학습목표이다. 평가 대상자를 고려하지 않을 경우 문항의 타당도와 문항의 적합도를 보장할 수 없게 되며, 교육과정의 교수·학습목표를 고려하지 않을 경우에는 문항제작을 위한 전략 수립 난항과 문항 자체에 대한 교육과정 위배 여부 논란으로 문항의 완성도를 떨어뜨릴 수 있기 때문이다. 여기에서 문항을 통해 평가 대상자에게서 얻고자 하는 정보와 얻은 정보를 어떻게 활용할 것인가 하는 후속 작업을 염두에 두어야 한다. 그러나 이러한 것들은 독립된 개체가 아니라 상호 유기적인 관계를 지니고 있어 복합적으로 이해하고 판단해야 한다.

성태제(2010)는 문항제작을 위해 교육목표와 교육내용, 피험자의 독해력과 어휘력 수준, 문항 유형 특성, 피험자에게 미칠 수 있는 영향 등을, Gronlund(1988)는 필요이상의 어려운 단어, 복잡한 문장구조, 모호한 문장, 두서없는 문장, 불분명한 그림, 혼란스런 발문 등이 고려대상이라고 주장하는 것은 전술 내용과 같은 맥락으로 이해할 수 있다. NCTM(1995)도 6가지 기준 - 평가는 모든 학생들이 알 필요가 있고 또 할 수 있는 수학을 반영해야 한다는 수학적기준(The mathematica standard), 학생들 스스로가 평가를 이용하여 독립적인 학습자가 되도록 수학학습을 촉진시키는 학습기준(The leaning standard), 각 학생들이 자신의 수학적 힘을 보여줄 수 있도록 다양한 평가 방법을 허용해야 한다는 공정성 기준(The equity standard), 평가 정보의 수집 과정과 평가 결과의 용도를 공개하는 개방성 기준(The openness standard), 수학 학습 방법에 대한 타당한 설명을 가능하게 해야 한다는 타당성 기준(The inferences standard), 평가 목적에 맞는 활동 계획을 수립하고, 수행 준거를 설정해야 한다는 일관성 기준(The coherence standard) - 을 제시하면서 새로운 평가 전략 시행의 필요함을 언급하고 있다.

좋은 문항은 이러한 문항제작시의 고려대상 및 기준을 전제로 두고, 새로운 아이디어를 찾도록 하는 문항의 참신성, 평가 대상자의 실력을 정확히 측정하기 위한 문항의 공정성, 평가 목표와 부합되는가의 문항의 일치성, 평가 대상을 고려하는 문항의 적합성, 평가 대상자를 변별하기 위한 문항의 난이도 등이 고려된 문항이다. 성태제(2010)는 좋은 문항의 조건을 측정 내용의 일치성, 문항 내용의 복잡성, 문항 내용의 요약성, 문항의 참신성, 문항 구조화,

종합(Synthesis) 하위분류 3가지, 평가(Evaluation) 하위분류 2가지로 되어 있다. Bloom은 교육목표분류에서 행동분류에 치중한 반면 내용분류는 소홀히 다루었다.

적절한 난이도, 학습동기 유발, 문항 검사 사용목적 부합, 측정오차 유발 방지, 문항 형식 제작지침 준수, 지필지침 준한 문항, 윤리적·도덕적 문제없는 문항, 특정 집단에 대한 유불리 방지 등을 언급하고 있다.

여기서 주의할 점은 교육과정과는 별개의 것으로 문항을 제작하는 우(愚)를 범하지 않은 것이다. NCTM(1995)은 제한된 자료를 주고 추론 하도록 하는 것, 수학적 지식을 증명하는 단일한 방법으로 학생들의 사고를 고정시키는 것, 수학 학습 기회를 제한하는 것 등을 평가할 때 배제해야 할 사항으로 지목하기도 하였다. 결국 좋은 수학 문항은 평가 대상자를 고려하고 교육과정에 따른 교수·학습 결과로서 그들의 수학적 사고력을 올바르게 측정 할 수 있는 문항일 것이다. 문항 유형은 크게 선택형(selection type item)과 서답형(supply type item)으로 구분한다. Mehrens & Lehmann(1975)은 선택형은 진위형, 선다형, 연결형으로, 서답형은 논술형, 단답형, 괄호형, 완성형으로 문항 형태를 구분하였다. 이 가운데 수학과 지필평가의 문항유형은 5지 선다형과 단답형의 문항 형태로 출제하는 비율이 가장 높으며, 수행평가에서는 서술형으로 출제하기도 한다.

III. 문항제작의 실제

1. 행동영역에 따른 예시 문항과 문항제작시의 유의점

평가에서 사용하는 각 문항은 학생들이 교수·학습 과정을 통해 어떤 내용을 어느 수준까지 알고 있는지를 알아보기 위해 내용영역과 행동영역으로 분류한다. 여기서 말하는 내용영역은 시험 출제범위에 해당되는 각 문항의 교과내용을 가리키며, 계산능력·이해능력·추론능력·문제해결능력 등의 행동영역은 각 문항의 내용을 측정할 때 어느 단계의 인지능력 수준을 측정할 것인가를 나타낸 것이다. 이렇듯 문항은 학생들의 내용영역과 행동영역의 교육 목표 도달점을 확인할 수 있다. 그러므로 학생들이 도달한 학습 성취도를 알아보기 위해서는 내용영역과 행동영역의 분류 구분을 이해하고 있어야 한다. 이를 돕기 위해 평가원(2005)에서 분류한 행동영역 계산, 이해, 추론, 문제해결을 예시 문항으로 두었다.

1) 예시 문항

(1) 계산능력

계산 문항은 기본적인 공식 또는 풀이 절차를 알고 있는지를 평가(평가원, 2005)하기 위한 문항으로 일반적으로 중하위권을 변별하는데 이용된다. 이와 동시에 수학 부진아 학생들에게는 수학에 대한 자신감을 불러일으켜 줄 수 있는 문항이기 때문에 수학교사들의 관심이 요구되는 문항이기도 하다. 따라서 계산문항은 하위권을 고려하면 난이도는 정답률 80% 정도가 적절한 수준이며, 복잡한 계산을 요구하지 않도록 제작하는 것이 바람직하다.

다음 예시 문항은 지수법칙 $(a^m)^n = a^{mn}$ 과 로그의 성질 $\log_a b^n = \frac{n}{m} \log_a b$ 을 이용하여 주어진 식을 간단히 할 수 있는지의 능력을 측정하기 위한 것으로, 보기 문항처럼 5지 선다형의 답지 배열은 관례적 순서나 논리적 순서에 따라 배열한다.

내용영역: 수학 I -지수와 로그	행동영역: 계산
$\frac{4^2}{2} \times 2 \log_3 \sqrt{3}$ 의 값은? ① 2 ② 4 ③ 6 ④ 8 ⑤ 10	

(2) 이해능력

이해능력은 주어진 문제 상황을 수학적으로 표현(수학적 용어, 기호, 식, 그래프, 표 등)· 적용하는 능력과 교과서에 나오는 기본 문제나 정형화된 응용문제를 해결하는 능력을 말한다(평가원, 2005). 이해 문항은 난이도 조정이 용이하기 때문에 출제의 폭을 어느 정도 자유롭게 할 수 있다. 따라서 중위권뿐만 아니라 상위권도 변별할 수 있는 문항을 제작할 수 있는 관계로 문항의 정답률이 나타나는 범위는 일반적으로 20~80% 정도로 매우 넓다.

다음 예시 문항은 수열 $\{a_n\}$ 이라는 용어, 그리고 수열의 합 기호 $\sum_{k=1}^n a_k$ 의 의미와 식

$\sum_{k=1}^n (a_k + b_k) = \sum_{k=1}^n a_k + \sum_{k=1}^n b_k$ 를 이해하고, 이를 적용할 수 있는 능력 측정을 평가목표로 하고 있다. 문제해결을 위한 조건을 문두에 주는 경우 예시 문항처럼 박스(box) 안에 (가), (나) 형태로 제시한다.

내용영역: 수학 I -수열	행동영역: 이해
수열 $\{a_n\}$ 은 다음 조건을 만족한다.	
(가) $a_1 = 1$ (나) $\sum_{k=1}^{10} a_k = 50, \sum_{k=2}^{10} (a_k + a_{k-1}) = 110$	
a_{11} 의 값은? ① 9 ② 10 ③ 11 ④ 12 ⑤ 13	

(3) 추론능력

추론능력은 발견적 추론능력과 연역적 추론능력으로 나눈다. 발견적 추론능력은 나열하기, 세어보기, 관찰, 유추 등을 통해 문제해결을 위한 핵심원리를 발견하는 능력을 평가한다. 연역적 추론능력은 수학적 개념·원리·법칙을 이용하여 참인 성질을 이끌어 내거나 주어진 명제의 참·거짓을 판별하는 능력을 평가한다(평가원, 2005).

일반적으로 발견적 추론능력은 보기 ㄱ, ㄴ, ㄷ이 주어지고 그 각각에 대해 참, 거짓을 판단하는 합답형 문항, 연역적 추론능력은 빈칸에 알맞은 수나 식을 찾는 완성형 문항으로 제작되어진다. 추론능력의 정답률은 대략 40~70% 정도로 나타난다. 다음 예시 문항은 역행렬의 뜻을 알고 계산할 수 있는지를 평가목표로 하고 있으며, 보기 ㄱ은 행렬 $A = \begin{pmatrix} 1 & a \\ 0 & 1 \end{pmatrix}$, $B = \begin{pmatrix} 1 & 0 \\ b & 1 \end{pmatrix}$ 를 주어진 식 $(A-B)^2$ 에 대입한 결과가 $(A-B)^2 = abE$ 가 되는지를 계산하면 확인할 수 있다. 보기 ㄷ은 참인 명제 ㄴ의 식 $A^{-1} = 2E - A$ 을 이용하여 $A + A^{-1} = 2E$ 임을 밝히면 된다.

합답형 <보기>의 개수는 보통 3개로 제작하고, 사용 문자는 ㄱ, ㄴ, ㄷ을 사용하며 상호연관성을 갖도록 한다. 보기 ㄱ, ㄴ, ㄷ의 난도(難度)는 낮음에서 높음 순으로(ㄱ: 낮음, ㄴ: 보통 ㄷ: 높음) 배열한다.

내용영역: 수학 I -행렬	행동영역: 발견적 추론
<p>0 이 아닌 두 실수 a, b에 대하여 두 행렬 A, B를 $A = \begin{pmatrix} 1 & a \\ 0 & 1 \end{pmatrix}$, $B = \begin{pmatrix} 1 & 0 \\ b & 1 \end{pmatrix}$이라 할 때, 옳은 것만을 <보기>에서 있는 대로 고른 것은? (단, E는 단위행렬이다.)</p>	
<p>—————<보 기>—————</p>	
<p>ㄱ. $(A - B)^2 = abE$ ㄴ. $A^{-1} = 2E - A$ ㄷ. $A + A^{-1} = B + B^{-1}$</p>	
<p>① ㄱ ② ㄴ ③ ㄷ ④ ㄱ, ㄴ ⑤ ㄴ, ㄷ</p>	
<p>[2012학년도 수능 6월 모의평가]</p>	

(4) 문제해결능력

문제해결능력은 두 가지 이상의 수학적 개념, 원리, 법칙이 사용된 문장에서 상호관련성을 파악하고 종합하여 문제를 해결할 수 있는가, 두 단계 이상의 사고과정을 거쳐서 문제를 해결하는 능력을 평가하는 내적문제해결과 실생활의 상황, 타 교과와 소재를 사용하여 제시된 문제에서 관련된 수학적 개념, 원리, 법칙 등을 파악하고 이를 적용하여 문제를 해결하는 능력을 평가하는 외적문제해결능력으로 나눌 수 있다(평가원, 2005). 특히, 수학적외적문제는 문제 상황이 현실에 반드시 부합되어야 하며, 행렬, 지수와 로그, 순열과 조합, 확률, 통계, 방정식과 부등식 영역 등에서 문항제작이 많이 이루어진다. 문제해결능력의 정답률은 5~40% 정도로 나타난다.

다음 예시 문항은 두 단계의 사고 과정을 거쳐야 한다. 첫 번째는 수열 $\{a_n\}$ 과 $\{b_n\}$ 을 이용하여 점 P_n 을 지나고 기울기가 a_n 인 직선의 방정식을 $y = a_n x + a_n b_n + b_n^2$ 로 나타내어야 하는 것이고, 두 번째는 수열 $\{b_n\}$ 의 계차수열이 $\{a_n\}$ 임을 이해하고 발문에 주어진 수열의

극한 $\lim_{n \rightarrow \infty} b_n$ 을 $\lim_{n \rightarrow \infty} b_n = 1 + \lim_{n \rightarrow \infty} \sum_{k=1}^{n-1} a_k$ 으로 구하는 것이다. 문항에서 사용하는 정의와 용어는 예시 문항에서의 ‘등비수열’, ‘교점’이라는 용어와 같이 교육과정의 정의와 용어를 따른다.

내용영역: 수학 I -수열의 극한	행동영역: 내적문제해결
첫째항이 12 이고 공비가 $\frac{1}{3}$ 인 등비수열 $\{a_n\}$ 에 대하여 수열 $\{b_n\}$ 을 다음 규칙에 따라 정한다.	
(가) $b_1 = 1$ (나) $n \geq 1$ 일 때, b_{n+1} 은 점 $P_n(-b_n, b_n^2)$ 을 지나고 기울기가 a_n 인 직선과 곡선 $y = x^2$ 의 교점 중에서 P_n 이 아닌 점의 x 좌표이다.	
$\lim_{n \rightarrow \infty} b_n$ 의 값을 구하시오.	
[2012학년도 수능 9월 모의평가]	

2) 문항제작시의 유의점

문항제작을 할 때 유의할 점은 다음과 같은 것들이 있다. 교육과정을 벗어나는 문항, 수학적으로 오류가 있는 문항, 출제의도가 명확하지 않은 문항, 반복되는 계산이나 많은 계산을 요구하는 문항, 불분명한 용어를 사용하는 문항 등이다. 이 가운데서도 교육과정 위배와 문항 오류는 가장 주의해야 할 점이다.

교육과정 위배
자연수 n 에 대하여 부등식 $x(x+2)(x-n)^n \leq 0$ 을 만족시키는 자연수 x 의 개수를 $f(n)$ 이라 할 때, $\sum_{n=1}^{10} f(n)$ 의 값을 구하시오.

위 문항은 부등식을 활용하여 수열의 합을 구할 수 있는지를 평가하고자 하였다. 그러나 현재 고등학교 교육과정에서는 ‘간단한 삼차부등식과 사차부등식을 풀 수 있다.’고 명시되어 있어, 문제에서 제시된 부등식 $x(x+2)(x-n)^n \leq 0$ 은 고등학교 교육과정의 범위를 벗어난다고 볼 수 있다.

오류 및 출제의도 불분명													
<p>오른쪽 상용로그표를 이용하여 구한 실수 $\left(\frac{10}{3}\right)^{20}$ 값의 범위로 옳은 것은?</p> <p>① $27 \times 10^9 < \left(\frac{10}{3}\right)^{20} < 28 \times 10^9$ ② $28 \times 10^9 < \left(\frac{10}{3}\right)^{20} < 29 \times 10^9$</p> <p>③ $29 \times 10^9 < \left(\frac{10}{3}\right)^{20} < 3 \times 10^{10}$ ④ $27 \times 10^{10} < \left(\frac{10}{3}\right)^{20} < 28 \times 10^{10}$</p> <p>⑤ $28 \times 10^{10} < \left(\frac{10}{3}\right)^{20} < 29 \times 10^{10}$</p>	<table border="1" style="width: 100%; border-collapse: collapse; text-align: center;"> <thead> <tr> <th style="padding: 2px 5px;">x</th> <th style="padding: 2px 5px;">$\log x$</th> </tr> </thead> <tbody> <tr> <td style="padding: 2px 5px;">2.6</td> <td style="padding: 2px 5px;">0.4150</td> </tr> <tr> <td style="padding: 2px 5px;">2.7</td> <td style="padding: 2px 5px;">0.4314</td> </tr> <tr> <td style="padding: 2px 5px;">2.8</td> <td style="padding: 2px 5px;">0.4772</td> </tr> <tr> <td style="padding: 2px 5px;">2.9</td> <td style="padding: 2px 5px;">0.4624</td> </tr> <tr> <td style="padding: 2px 5px;">3.0</td> <td style="padding: 2px 5px;">0.4771</td> </tr> </tbody> </table>	x	$\log x$	2.6	0.4150	2.7	0.4314	2.8	0.4772	2.9	0.4624	3.0	0.4771
x	$\log x$												
2.6	0.4150												
2.7	0.4314												
2.8	0.4772												
2.9	0.4624												
3.0	0.4771												

위 문항은 $\log\left(\frac{10}{3}\right)^{20} = 10.458$ 이므로 $\log(2.8 \times 10^{10}) < \log\left(\frac{10}{3}\right)^{20} < \log(2.9 \times 10^{10})$ 임을 이용하여 $28 \times 10^9 < \left(\frac{10}{3}\right)^{20} < 29 \times 10^9$ 이라는 것을 계산할 수 있는가를 묻고 있다. 그러나 $\left(\frac{10}{3}\right)^{20}$ 의 실제 값은 28679719907.9244로, 로그를 이용하여 얻은 근삿값의 범위와 오차가 10^8 정도나 발생한다. 따라서 $28 \times 10^9 < \left(\frac{10}{3}\right)^{20} < 29 \times 10^9$ 을 참값의 범위로 하기에는 무리가 있다. 즉, $\left(\frac{10}{3}\right)^{20}$ 의 값의 범위로 옳다고 하기 어렵다. 또한 $\log\left(\frac{10}{3}\right)^{20} = 10.458$ 이므로 $\left(\frac{10}{3}\right)^{20}$ 이 11자리의 실수라는 것을 밝히는 것은 지표의 성질을 묻고자 하는 출제의도가 명확하나 주어진 문항의 발문은 로그 계산을 하도록 하는 것이 평가 목표인지 부등식의 뜻을 묻고자 하는 것이 평가 목표인지가 불분명하다.

반복 계산
<p>함수 $f(x) = 2\sin x + \cos x$와 모든 실수에서 미분가능한 함수 g, h의 합성함수 $p(x) = (h \circ g \circ f)(x)$에 대하여 $x=0$에서의 $p(x)$의 미분계수가 $p'(x) = 12$일 때, $h'(g(1))g'(1)$의 값을 구하시오.</p>

위 문항의 평가 목표는 두 가지이다. 삼각함수의 미분 $f'(x) = 2\cos x - \sin x$ 을 구할 수 있는지와 합성함수의 미분 $p'(x) = h'(g(f(x))) \cdot g'(f(x)) \cdot f'(x)$ 를 구할 수 있는가이다. 그 중 합성함수의 미분 개념을 이해하고 계산을 할 수 있는가를 평가하기 위해, 같은 계산을 두 번 반복하도록 해서 문항을 복잡하게 구성하는 것보다는 합성함수를 $(g \circ f)(x)$ 로 주면 같은 개념을 알고 있는지를 확인할 수 있으면서도 계산량은 줄일 수가 있다. 이렇게 문항제작 과정에서 복잡한 계산과 계산량을 많게 하는 경향은 문항의 난도(難度)를 높이는 방법이라고 착각하는 것에서 일어난다.

불분명한 용어

빨간색과 파란색을 포함한 서로 다른 6개의 색을 모두 사용하여, 날개가 6개인 바람개비에 칠하려고 한다. 각 날개에는 한 가지 색만 칠하고, 빨간색과 파란색은 맞은편 날개에 칠한다고 하자. 바람개비를 회전하여 일치하는 경우는 모두 같은 것으로 정할 때, 서로 다른 경우는 모두 몇 가지인지 구하시오.

위 문항은 ‘맞은편’이라는 용어를 어떻게 해석하는가에 따라 정답이 여러 개 생길 수 있다. 예를 들면, 날개 1개를 정했을 때, 그 날개의 맞은편 날개의 수가 3개가 될 수 있다. 또한, 2가지 색으로 6개의 날개를 색칠할 수 있으므로 발생하는 경우의 수가 많아 계산하기도 쉽지 않다. 따라서 이와 같은 문제점을 해결하기 위해서는 ‘맞은편’을 ‘서로 맞은편’이라는 구체적인 용어 표현으로 바꾸고, 각 날개에는 한 가지 색만 칠한다는 조건을 주어야 한다. 이렇듯 불분명한 용어나 모호한 발문은 복답 발생과 출제의도와 다르게 해석될 수 있어 세밀한 주의가 필요하다.

2. 문항제작의 실제

문항은 계획 단계, 초안 문항제작 단계, 검토 단계, 최종 문항제작 단계를 거쳐 문항을 완성시킨다. 계획 단계는 문항제작 시 고려사항을 점검하고, 출제 계획(평균·표준편차와 난이도 등)을 세우고, 그에 따른 내용과 행동영역을 안배하며 평가의 방향을 결정하는 단계라 할 수 있다. 초안 문항제작 단계는 계획단계에서 작성한 이원목적분류표를 근거로 문항의 골격을 세우는 단계로 문항 구성과 발문을 고려하여 문항 유형(선택형 또는 서답형)을 결정한다. 수정 및 오류 검토 단계는 문항의 내용, 소재, 형식 등과 교육과정 위배 여부 등을 점검하는 단계로 문항이 이상이 있을 경우 문항을 수정·보완 및 교체한다. 이 단계에서 난이도 조정이 함께 이루어진다. 최종 문항제작 단계는 계획 단계, 초안 문항제작 단계, 오류 검토 및 수정 단계를 거치면서 완성된 문항의 최종 검토와 윤문 등을 거쳐 문항의 완성도를 높이는 단계이다. 문항이 완성되면 최종 문항카드⁴⁾를 작성한다.

이 절에서는 문항제작의 실제(시험시간 100분, 30문항 기준)를 1학년말 평가를 가상하여 단계별로 보여주고자 한다.

1) 계획단계

우선 평균과 표준편차의 목표를 설정(평균 50, 표준편차 34)하고, 설정된 목표를 달성하기 위한 구간별 문항 수를 정한다. 구간별 문항 수는 평가 전체의 난이도를 결정하게 되므로 평가대상자의 수준을 고려하며, 고르게 분포하도록 배치한다(<표 III-1>).

4) 문항카드의 별도의 양식은 없으나 일반적으로 과목명, 대단원·중단원명, 내용영역, 행동영역, 출제 의도, 문제 및 정답, 예상정답률 등을 기록하는 양식을 말한다. 이러한 문항카드는 초안 문항제작 단계에서 문항 개발과 함께 처음 작성하나 문항이 완성되면 그에 따른 최종 문항카드를 작성한다.

<표 III-1> 목표 평균과 표준편차에 따른 구간별 문항 수

시험	정답률(%)										평균	표준 편차
	0-10	10-20	20-30	30-40	40-50	50-60	60-70	70-80	80-90	90-100		
1학년말 평가	1	1	2	4	9	5	3	2	2	1	50.0	34.3

<표 III-1>을 기준으로 배점별 문항 수를 결정한다. 정답률 구간 40-50에서의 9문항을 3점 배점 4문항, 4점 배점 5문항으로 배점을 다르게 구분한 것은 두 가지 이유에서이다<표 III-2>. 하나는 40-45, 45-50으로 정답률 구간을 더욱 세밀하게 나눌 수 있으며, 다른 한 가지는 4점 배점 문항을 단답형 문항으로 출제할 수 있기 때문이다.

<표 III-2> 문항 정답률 구간과 배점별 문항 수

정답률 구간(%)	배점		
	2점	3점	4점
0-10			1
10-20			1
20-30			2
30-40			4
40-50		4	5
50-60		5	
60-70		3	
70-80		2	
80-90	2		
90-100	1		
계	3	14	13

그 다음은 <표 III-1>과 <표 III-2>를 근거로 이원분류표를 작성한다. 이원분류표 작성은 출제 계획에서 가장 중요한 작업으로 교육목표(내용영역, 행동영역)와 문항 수 등을 담는다. 내용영역은 출제내용이 되며, 행동영역은 측정하고자 하는 계산, 이해, 추론(연역적·귀납적), 문제해결(내적·외적)능력이다. 이 때, 내용영역과 행동영역이 특정영역에 치우치지 않도록 한다.

수학과 평가 문항제작의 실제

<표 III-3> 이원분류표

과목	내용영역	행동영역				합계	비율	배점	
		계산	이해	추론	문제해결				
수학 I	행렬과 그래프	행렬의 뜻				0	0.0	0	
		행렬의 덧셈과 뺄셈, 실수배	1*			1	3.3	2	
		행렬의 곱셈	1			1	3.3	3	
		역행렬	1	1**		2	6.7	7	
		연립일차방정식과 행렬		1		1	3.3	3	
		그래프와 행렬		1		1**	2	6.7	7
	지수함수와 로그함수	거듭제곱과 거듭제곱근	1*			1	3.3	2	
		지수의 확장				0	0.0	0	
		지수함수와 그 그래프		1	1**	2	6.7	7	
		지수방정식과 지수부등식		1		1	3.3	3	
		로그와 그 성질		1**		1	3.3	4	
		상용로그	1			1**	2	6.7	7
		로그함수와 그 그래프		1	1**	2	6.7	7	
	로그방정식과 로그부등식		1		1	3.3	3		
	수열	수열의 뜻				0	0.0	0	
		등차수열	1			1	3.3	3	
		등비수열	1			1	3.3	3	
		여러 가지 수열의 합		1		1**	2	6.7	7
		계차수열		1**		1	3.3	4	
		수학적 귀납법			1**	1	3.3	4	
		알고리즘과 순서도				1**	1	3.3	4
	수열의 극한	무한수열의 수렴과 발산		1**		1	3.3	4	
		극한값의 계산	1*			1	3.3	2	
		무한등비수열의 극한	1			1	3.3	3	
		무한급수의 수렴과 발산		1		1**	2	6.7	7
무한등비급수				1**	1	3.3	4		
계	문항 수	9	12	4	5	30	100	100	
	문항 비율	30.0	40.0	13.3	16.7				

※ * 문항은 2점 배점 문항이며, ** 문항은 4점 배점 문항을 표시함.

2) 초안 문항제작 단계

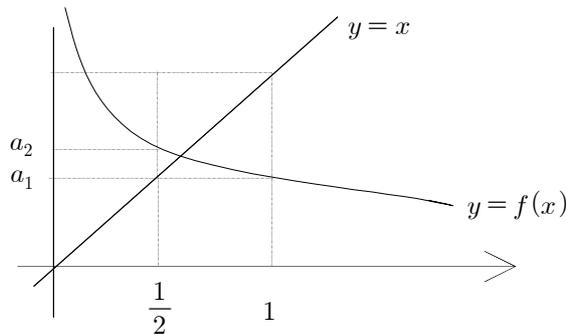
이제 평가목표에 부합하는 문제 상황을 결정하고, 문항의 체계 및 발문을 구상한 후 문항 초안을 제작한다. 이 때, 문항제작 상의 유의점을 고려하면서 창의적인 아이디어를 발휘하여 제작한다. 특히, 서술형 문항의 경우에는 채점 기준표와 모범 답안(유사답안 포함)을 함께 제작한다.

초안 문항제작 단계에서 주의해야 할 점은 다음과 같다: (1)학교 교육과정의 내용과 수준에 맞추어 수학적 사고력을 측정할 수 있는 문항을 출제하도록 한다. (2)수학적인 오류나 모호함이 없는 문항을 출제하도록 한다. (3)계산 능력, 이해 능력, 추론 능력, 문제해결 능력을 적절하게 평가할 수 있으며 평가 목표가 분명한 문제를 출제하도록 한다. (4)문항의 내용과 소재가 특정 영역에 지나치게 편중되지 않도록 고르게 출제하도록 한다. (5)실생활 및 타교과 소재나 상황을 이용할 때에는 현실에 부합되고, 타교과 지식에 적합한 문항을 출제하도록 한다. (6)단순 공식이나 지식의 암기 여부를 확인하는 문항의 출제는 지양한다. (7)문제의 표현은 학생들이 이해하기에 쉬우면서도 간결·분명·정확하게 하도록 한다. (8)교과서에 나오는 수학의 기본 개념이나 원리를 이해하고 있으면 풀 수 있는 문제를 출제하도록 하고, 교과서에 나오는 기본 공식이 아닌 특정 공식을 암기하지 못하면 풀 수 없는 문제의 출제는 지양한다. (9)지나치게 복잡한 계산 위주인 문제의 출제를 지양한다. (10)한 문제의 풀이에 지나치게 긴 시간이 소요되는 문항의 출제는 지양한다. (11)예상 평균 점수를 고려하여 문항의 난이도를 조절하여 출제하도록 한다(평가원, 2005).

다음은 <표 III-3>의 이원분류표를 근거로 작성한 초안문항의 예이다.

[초안문항] 지수함수 $f(x) = 2^{-x}$ 에 대하여 $a_n = f^n(1)$ 일 때, a_2, a_3, a_4, a_5 의 대소 관계를 옳게 나타낸 것은? (단, $f^1(x) = f(x)$ 이고 $f^{n+1}(x) = f(f^n(x)), n = 1, 2, 3, \dots$)

- ① $a_2 < a_3 < a_4 < a_5$ ② $a_5 < a_4 < a_3 < a_2$ ③ $a_2 < a_4 < a_3 < a_5$
 ④ $a_5 < a_3 < a_2 < a_4$ ⑤ $a_3 < a_5 < a_4 < a_2$



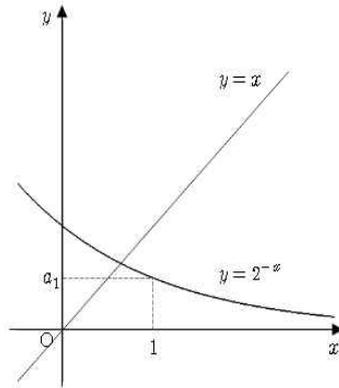
* 내용영역 : 지수함수와 그 그래프, 행동영역 : 이해

3) 오류 검토 및 수정 단계

검토 및 수정 단계는 일차적으로 제작된 초안문항에 대해 교육과정 및 교과서의 범위·수준 부합여부, 난이도 및 소요시간, 정답시비 가능성 여부, 단서 포함 여부, 물음과 답의 대응 관계, 편집 체계의 일관성 등을 검토 요목으로 선정하여, 문항 이상 유무를 점검하면서 발생되는 문항의 문제점을 수정·보완 및 교체한다. 이 단계는 많은 노력과 시간이 필요하며, 여기에서 얼마만큼의 보정 작업이 이루어졌는가에 따라 문항의 완성도가 결정된다. 특히, 출제자의 관점보다는 학생의 입장에서 검토하는 자세가 필요하며, 학생이 발문을 이해하고 출제자의 의도하는 방향으로 응답할 수 있는지, 채점 기준이 명료한지, 소요 시간이 적정한지 등을 중점 검토한다. 그리고 이 단계에서 수학교사들의 상호 검토와 문항 검토 항목을 체크리스트로 만들어 사용하는 것도 좋은 방법이다.

<수정문항1>은 [초안문항]의 함수열 표현 $a_n = f^n(1)$ 을 $a_1 = f(1)$ 으로 $f^{n+1}(x) = f(f^n(x))$ 은 $a_{n+1} = f(a_n)$ 으로 바꾼 것이다. 그러면서 함수의 그래프도 간단히 정리하였다. 함수열 표현을 바꾼 것은 고등학교 교육과정에서 다루지 않는 내용이기 때문이다.

<수정문항1> 지수함수 $f(x) = 2^{-x}$ 에 대하여 $a_1 = f(1)$, $a_{n+1} = f(a_n)$ 일 때, a_2, a_3, a_4, a_5 의 대소 관계를 옳게 나타낸 것은?



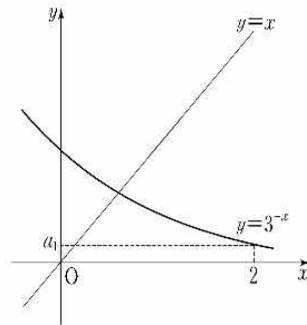
- ① $a_2 < a_3 < a_4 < a_5$ ② $a_5 < a_4 < a_3 < a_2$ ③ $a_2 < a_4 < a_3 < a_5$
 ④ $a_5 < a_3 < a_2 < a_4$ ⑤ $a_3 < a_5 < a_4 < a_2$

<수정문항1>의 지수함수 $f(x) = 2^{-x}$ 보다는 지수함수 $f(x) = 3^{-x}$ 가 학생의 입장에서 문제해결이 용이할 것으로 판단하여 <수정문항2>로 수정하였다.

<수정문항2> 지수함수 $f(x) = 3^{-x}$ 에 대하여

$$a_1 = f(2), a_{n+1} = f(a_n) \quad (n = 1, 2, 3, 4)$$

일 때, a_2, a_3, a_4, a_5 의 대소 관계를 옳게 나타낸 것은?



- ① $a_2 < a_3 < a_4 < a_5$ ② $a_5 < a_3 < a_4 < a_2$ ③ $a_2 < a_4 < a_3 < a_5$
 ④ $a_5 < a_3 < a_2 < a_4$ ⑤ $a_3 < a_5 < a_4 < a_2$

4) 최종 문항제작 단계

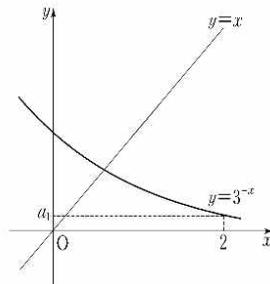
최종 문항제작 단계는 계획 단계, 초안 문항제작 단계, 수정 및 오류 검토 단계를 거치면서 어느 정도 완성된 문항의 최종 검토와 운문 등을 거쳐 문항의 완성도를 높이는 단계이다. 검사 문항의 수정·보완 작업을 마친 후에는 문항 편집 지침에 따라 최종 문항을 편집하고 인쇄하며, 문항 내용 정보를 문항카드에 기입한다.

<최종문항>은 <수정문항2>의 a_2, a_3, a_4, a_5 의 4개 항을 1개 줄여 a_2, a_3, a_4 3개 항으로 대소 관계를 찾도록 하였다. 문항을 수정한 것은 문제해결의 아이디어 발견이 중요한 것이기 때문에 개념을 이해한 뒤의 반복되는 문제풀이 계산량을 줄이기 위함이다.

<최종문항> 지수함수 $f(x) = 3^{-x}$ 에 대하여

$$a_1 = f(2), a_{n+1} = f(a_n) \quad (n = 1, 2, 3)$$

일 때, a_2, a_3, a_4 의 대소 관계를 옳게 나타낸 것은?



- ① $a_2 < a_3 < a_4$ ② $a_4 < a_3 < a_2$ ③ $a_2 < a_4 < a_3$
 ④ $a_3 < a_2 < a_4$ ⑤ $a_3 < a_4 < a_2$

IV. 결론 및 제언

평가는 학생들이 교실수업에서 이루어지는 교육과정의 정상적 이수를 통해 수학적 개념을 얼마나 이해하고, 또 이를 어떻게 적용할 수 있는가를 측정하는데 있다. 나아가, 상급학년에서 필요한 학습 내용을 보다 쉽게 이해하고 발전·심화 시킬 수 있도록 기초학습능력을 개발하고 실생활 속에서 부딪히는 여러 가지 난제들을 스스로 해결할 수 있는 수학적 힘을 배양하는 목적도 있다. 또한 교육목표 달성 진단과 학교학습을 위한 실제와 개선에 보다 적절한 정보를 제공한다(Wittrock, 1991)는 평가의 관리적 관점에서 보면 학교교육의 방향을 설정하는 지침이 되기도 한다. 그러므로 평가의 필수 항목인 좋은 문항은 교실수업과 학교교육의 성패를 좌우하는 요소라 할 수 있다.

본고는 이러한 평가의 시발점인 문항제작의 실재를 안내하였다. 이와 관련하여 이 장에서는 문제제작 능력을 향상시키기 위한 다음과 같은 내용을 제언하고자 한다.

첫째, 문항을 바라보는 안목을 기른다. 이를 위해 문항제작 관련 이론 및 저서들을 살펴보고, 학업성취도 문항이나 수능 문항 등 기 검증된 문항들을 분석해 본다. 문항제작관련 이론 및 저서를 통해서는 문항제작의 기초 배경 지식을 얻을 수 있으며, 문항 분석을 통해서는 문항의 타당성, 적합성, 난이도 등 문항제작을 위한 식견을 기를 수 있기 때문이다.

둘째, 아이디어를 내어 문항을 창작해 본다. 좋은 문항은 고민에서 나오는 것이 아니라 실천의 결과물로 산출되기 때문이다. 창작 문항이 많아질수록 풍부한 문항제작의 노하우도 쌓을 수 있다. 그렇지만 무엇보다 중요한 것은 문항제작의 절차를 계획하고 이를 실천하는 것은 평가자로서 갖추어야 할 기본 덕목이기 때문이다.

셋째, 문항은 교육과정안에서 제작한다. 교육과정은 교사에게는 가르쳐야 할 교육내용임과 동시에 학생에게는 알아야 할 중요한 내용이기 때문이다. 얼마 전 일부 중고교에서 상위 학년 교육과정에 있는 시험을 넘으로써 학생들이 사교육을 받을 수밖에 없어 학교 교육의 왜곡을 초래하였다는 보도가 있었다. 학교 교육은 교육과정에 의해 운영되므로 만일 평가가 교육과정을 따르지 않고 교육과정을 벗어나 과행적으로 운영된다면 학교 교육 자체의 근간이 흔들리게 되기 때문에 이 문제는 반드시 주의가 요구된다.

넷째, 평가자 스스로 자신의 문항제작 DB(Data Base)를 구축한다. 문항 DB는 각 문항의 히스토리(history)를 담고 있어 문항제작 과정에서 드러난 문제점을 사전 점검하여 향후 문항제작 과정에서 발생할 수 있는 동일한 문제점을 예방할 수 있기 때문이다.

다섯째, 유관기간의 문항제작 연수 확대 실시이다. 현장 교사들이 문항제작 관련 연수의 필요성을 인식하고 참여를 희망하지만 연수 인원의 제한, 방과 후 활동과 연수 기간의 중복 등 여러 가지 이유로 인해 현실적으로 참여할 수 있는 기회를 얻기가 힘들다는 토로를 하고 있다. 문항제작 연수는 현장 교사의 평가 전문성 향상과 학교 교육평가의 질 제고 측면에서 유관 기관은 이 문제에 대한 긍정적인 자세가 필요하다.

평가는 학습목표에 대한 학생의 발전을 모니터링 하는 것, 학습지도의 판단을 내리는 것, 어떤 시점에서 학생의 수학 성취에 대한 가치를 부여하는 것, 교육 프로그램에 대한 가치 판단을 하는 것이다(NCTM, 1995). 이러한 평가 작업의 기초는 좋은 문항으로부터이다. 그래서 문항제작은 중요하다. 그렇지만 문항제작이 쉬운 것은 아니다. 혹자는 문항제작을 산고(産苦)의 고통에 비유하기도 하는데 그만큼 극한의 작업을 이겨내야만 좋은 문항을 탄생시킬 수 있다는 뜻일 것이다. “문항제작은 예술이다.”라는 Embreston(1985)의 말은 이것을 잘 대변해 준다.

참고문헌

- 변창진, 최진승, 문수백, 김진규, 권대훈 (1998). 교육평가. 학지사
- 성태제 (2010). 현대교육평가. 학지사.
- 신범영 (2010). 바람직한 수학 평가 문항 제작에 대한 연구. 서울시립대 석사학위논문.
- 이석현 (2010). 평가 문항 제작에 필요한 교사의 지식에 대한 연구 : 중등학교 수학과 평가 문항 컨설팅 내용을 중심으로. 단국대학교 박사학위논문.
- 전영주 (2011). 중국 대학입학시험의 수학 평가내용 및 구성 고찰. 한국학교수학회 14(1), 90-99.
- 한국교육과정평가원(2005). 대학수학능력시험출제매뉴얼. 사단법인 교육진흥연구회.
- 황혜정, 최승현 (1999). 수학과 평가들에 관한 고찰. 대한수학교육학회<수학교육학연구>, 9(2), 459-471.
- Bloom, B. S. (1956). *The Taxonomy of Educational Objectives, Handbook I, Cognitive domain*. New York: David Mc. Kay.
- Embreston, S. E. (1985). Introduction to the problem of test design. In S. E. Embreston (ed.). *Test design: developments in psychology and psychometrics*. Florida: Academic Press, Inc.
- Gronlund, N. E. (1988). *How to construct achievement tests*, NJ: Prentice-Hall, Inc.
- Mehrens, W. A., & Lehmann, I. J. (1975). *Measurement and Evaluation in Education and Psychology*. NY: Holt, Rinehart, & Winston.
- The National of Council of Teachers of Mathematics (1989). *Curriculum and evaluation standards for school mathematics*, Reston, VA: Author.
- The National of Council of Teachers of Mathematics (1995). *Assessment standards for school mathematics*, Reston, VA: Author.
- Wittrock, M. C. (1991). Cognition and testing. In M. C. Wittrock, E. I. Baker(Eds) *Testing and Cognition*. New Jersey: Printice Hall, 1-4.

A Study on the Practice of Making up Questions on Schools Mathematics Tests

Jeon, Young-Ju⁵⁾

Abstract

Mathematics evaluation is aimed at measuring students' mathematical ability to think and achievement. Moreover, the quality of mathematics education at school is elevated by the test questions analysis and application process. And the starting point of evaluation is to make good test questions. This study is composed of four parts. First, theoretical background associated with making test questions is surveyed. Second, example questions according to the education goals and question-making cautions are presented. Third, four practical stages of making test questions which consist of designing, making the first draft of questions, verifying and making the final draft of questions, are illustrated.

Key Words : mathematics evaluation, mathematics test questions-making

5) Korea Institute of Curriculum and Evaluation (whaljuo@kice.re.kr)