

평가목적을 고려한 문제은행의 구축과 활용

- 교수·학습 지원을 위한

문제은행과 선발 및 자격 시험을 위한 문제은행 -

상 경 아(한국교육과정평가원 연구원)

양 명 희(군 산 대 학 교 교 수)

《 요약 》

이 연구에서는 평가의 목적에 따라 문제은행의 유형을 '교수·학습 지원을 위한 문제은행'과 '선발 및 자격 시험을 위한 문제은행'으로 개념적으로 구분하고, 각각의 문제은행이 어떻게 구축되며 활용될 수 있는지를 논의하였다. '교수·학습 지원을 위한 문제은행'은 교수·학습 과정에 유용한 평가 정보를 제공할 수 있도록 교육과정에 근거한 문항을 충분히 확보하고 교사나 학생들이 편리하게 활용할 수 있도록 하는데 강조점이 주어지며, 반면에 '선발 및 자격 시험을 위한 문제은행'은 시험의 목적에 부합되는 양질의 검사를 안정적, 효율적으로 개발하고 시험의 보안을 유지하는 것이 중요하게 고려된다. 이들 두 유형의 문제은행이 갖는 특징과 차이점들을 문항 개발, 문제은행 시스템 구축, 문제은행의 활용 측면으로 나누어 구체적으로 논의하였고, 이를 토대로 문제은행의 구축과 활용을 위한 몇 가지 제언과 함께 문제은행의 개발과 활용이 더욱 활성화될 필요가 있음을 지적하였다.

주제어 : 문제은행, 문제은행 유형, 교수·학습 지원을 위한 문제은행, 선발 및 자격 시험을 위한 문제은행

I. 서론

문항을 체계적으로 관리하고 활용하기 위해 고안된 문제은행은 1980년대 초반까지도 종이로 된 문항카드를 서류함이나 파일 형태로 저장하는 것이 일반적이었고 문항의 저장과 검색은 수작업을 통해 이루어졌다. 컴퓨터의 눈부신 발전은 문제은행의 전산화를 촉진하였는데 이를 통해 문제은행의 저장, 검색 기능은 획기적으로 확장되었다. 또한 단순히 문항을 저장하고 검색하는 기능에서 벗어나 검사 목적에 맞는 문항을 추출하여 검사를 제작하거나 컴퓨

터를 이용한 검사 혹은 개인별 능력수준에 따른 맞춤검사를 제공하는 기능들과 결부되면서, 문제은행은 성격과 활용 가능성에서 질적인 변화를 이루게 된다. 한편 문항반응이론의 등장은 문제은행의 발전에 또 다른 전기를 마련하였는데, 문항반응이론이 제안하고 있는 개인별 맞춤검사를 구현하기 위해서는 문제은행이 필수적으로 요구되었고 이로 인해 문제은행의 구축 방법, 적절한 규모, 효율적인 운영 방식 등에 대한 많은 연구를 촉발시켰다.

컴퓨터의 발전, 고속 통신망 구축, 문항반응이론과 같은 새로운 검사 이론의 등장은 문제은행을 단순히 문항을 저장하고 관리하는 차원에서 벗어나 검사를 개발하고 실시하는 새로운 패러다임으로 변모시키고 있다. 실제로 미국의 검사 전문 기관인 ETS는 우리나라의 대학수학능력시험과 성격이 유사한 SAT나 대학원생 선발 과정에서 활용되는 GRE, 의사자격시험인 MCAT 등의 시험을 문제은행을 토대로 운영하고 있다. 문제은행을 활용해 검사를 개발하는 것에 그치지 않고 문제은행을 토대로 한 컴퓨터 활용검사(computer based test)나 컴퓨터에 기반한 개인별 맞춤검사(computer adaptive test)를 진행함으로써 검사 실시와 운영 방식을 더욱 진보시켜 나가고 있는 것이다. 국내에서도 한국교육과정평가원에서 초·중등학교의 교사 및 학생들을 위한 문제은행을 구축하고 있는가 하면, 외무고시나 행정고시 등의 1차 시험인 공직적격성시험과 운전면허시험의 학과시험 등이 문제은행을 활용해 운영되고 있다. 또한 근래에는 우리나라에서 가장 민감한 시험 중의 하나인 대학수학능력시험에서도 문제은행을 통한 출제방식의 도입을 적극적으로 검토하고 있다.

문제은행은 이미 다양한 평가 영역에서 다양한 방식으로 활용되고 있으며 앞으로 그 활용 영역은 더욱 확대될 것이다. 그러나 아직까지 문제은행에 대한 실천과 연구들은 특수한 개별 문제은행의 구축과 활용의 문제에 천착했을 뿐, 규모나 형태, 운영방식 등에서 다양한 특성을 보이는 문제은행들을 유형화하고, 각각의 특성과 필요 요건 등을 명료화하는 이론화 작업이 부진하였다. 이러한 문제의식 하에 본 연구는 다양한 영역에서 여러 목적과 방식으로 운영되고 있는 문제은행을 그것이 활용되는 목적에 따라 ‘교수·학습 지원을 위한 문제은행’과 ‘선발 및 자격 시험을 위한 문제은행’으로 나누어 개념화해 보고, 각 문제은행의 특성은 무엇인가, 문제은행의 구축과 활용에 있어 어떠한 차이점이 존재하는가, 그리고 각 문제은행이 본연의 목적을 달성하기 위해 갖추어야 할 필요 요건은 무엇인가를 검토하고 논의해 보고자 한다.

문제은행의 유형을 구분하는 방식은 문제은행의 규모나 형태, 활용 방법 등 다양한 방식이 있을 수 있으나 본 연구에서 ‘교수·학습 지원을 위한 문제은행’과 ‘선발 및 자격 시험을 위한 문제은행’으로 구분한 것은, 평가의 목적에 따라 평가 방법과 형태를 달리하는 만큼 문제은행도 그것이 활용되는 목적에 따라 운영되는 형태와 방식이 달라져야 할 것이기 때문이다. 이러한 논의는 문제은행을 개발하는 데 있어 각각의 목적에 부합하는 적절한 형태가 무엇이며, 어떤 요소들이 고려되고 강조되어야 하는지, 혹은 이를 구현하기 위해 갖추어야 할

필수적인 조건이나 상황은 무엇인지에 대한 안내서의 역할을 할 것이며 동시에 그동안 미진하였던 문제은행에 대한 연구에 새로운 활력이 될 수 있을 것이다.

II . 문제은행

‘문제은행’에 대한 합의된 정의를 내리기는 어렵다. 문제은행에 대한 정의는 ‘검사 문항들을 모아 놓은 것’이란 매우 느슨한 정의에서부터, ‘상당히 많은 수의 문항들을 쉽게 찾아 쓸 수 있도록 모아 놓은 것(Millman & Arter, 1984; Ward & Murray-Ward, 1994)’, ‘검사 문항들의 집합체로서 각 문항의 내용과 측정학적 특성을 고려하여 색인화하고 조직화되어 있는 것(Choppin, 1981; Gronlund, 1998)’ 또는 ‘학습의 과정이나 결과를 능률적으로 평가하기 위해 양질의 문항을 제작, 정리, 보관하여 필요시에 사용할 수 있게 하는 일련의 절차 내지 체제(Booth, 1983; Wood, 1974)’라는 제한적인 정의에 이르기까지 다양하다. 또한 더 특수하게는 문제은행이란 반드시 전산화된 것으로서 컴퓨터를 이용하여 출제자가 필요한 문항을 선택, 편집, 인쇄하거나 경우에 따라 시험의 실시, 채점, 결과의 통계적 분석, 성적 통지까지 가능하도록 검사 문항을 분류, 저장, 관리하는 컴퓨터시스템(중앙교육평가원, 1986; Lippey, 1974)이라고 규정하거나, 문항들의 난이도 수준이 척도화된(calibrated) 것을 문제은행의 필수 요건으로 보는 견해(Anzaldua, 2002; Mead, 1981; Wright & Bell, 1984)도 있다.

사실상 문제은행에 대한 정의는 그것이 사용될 목적과 상황 그리고 이에 따른 문제은행의 형태와 운영방식에 따라 다르게 정의될 수 있다. 그러나 문제은행이란 그 규모나 형태에 관계없이 일반적으로 다음과 같은 공통적인 요소들을 포함한다고 할 수 있다.

첫째, 문제은행은 상당히 많은 수의 문항들로 구성된다. 여기서 ‘상당히 많다’는 것은 검사를 실시하는데 필요한 문항 수보다 적게는 수배에서 많게는 수십 배까지 더 많은 문항을 확보해야 함을 의미한다. 둘째, 문항들은 효율적으로 검색·활용될 수 있어야 한다. 이를 위해서는 문항들을 색인화하고 체계적으로 저장해야 하며, 원하는 문항을 쉽게 선택할 수 있도록 다양한 문항특성들(내용, 키워드, 난이도, 변별도 등)에 대한 정보를 함께 제공해야 한다. 셋째, 문제은행을 구축하고 유지·관리하기 위한 시스템이 전제되어야 한다. 문제은행은 고정적인 것이 아니라 끊임없이 새로운 문항을 추가하고 부적절하게 된 문항은 수정하거나 제거하는 유동적인 과정이라 볼 수 있다. 따라서 새로운 문항을 추가하거나 기존의 문항을 검토하고 갱신·폐기하는 일련의 작업을 일관성 있게 진행하기 위한 절차와 체계가 필요하다.

문제은행은 검사의 제작과 활용의 측면에서 많은 이점을 제공하지만 문제은행을 구축하는 과정에서 많은 시간, 노력, 비용이 요구된다. 그럼에도 일단 문제은행이 구축되면 다음과 같

은 장점이 있다(Anzaldúa, 2002; Millman & Arter, 1984; Rudner, 1998; Wright & Bell, 1984). 첫째, 문제은행은 검사 상황에 따라 적절한 수준과 내용의 문항들을 추출하여 비교적 짧은 시간 내에 검사를 쉽게 제작할 수 있도록 해준다. 둘째, 문제은행은 양질의 문항 및 검사를 개발할 수 있도록 해준다. 문제은행을 구축함에 있어서 다수의 문항개발자들이 문항 작성에 참여하므로 다양한 내용, 다양한 형식의 문항개발이 가능하며, 문항 출제 후에도 충분한 문항 검토와 보완을 거쳐 양질의 문항만을 문제은행에 저장하므로 결과적으로 문항과 검사의 질을 높일 수 있다. 셋째, 척도화된 문제은행의 경우, 피험자가 어떤 문항들로 구성된 검사를 치르든 피험자의 능력이 동일하게 추정된다. 척도화란 문항의 난이도와 피험자의 능력을 동시에 표현하는 척도 위에서 개별 문항의 위치를 규정하는 것으로, 이런 문제은행에 기초한 검사에서는 피험자의 능력 수준에 따른 맞춤검사가 가능하며 문제은행 내의 문항들로 구성된 어떤 검사들 간에도 검사점수의 직접적인 비교가 가능하다. 문제은행 내의 문항들을 척도화하기 위해서는 문항반응이론이 주로 사용된다. 넷째, 문제은행은 장기적인 관점에서 볼 때 검사 제작에 투입되는 시간과 비용을 절감할 수 있다.

Ⅲ. 문제은행의 유형 : 평가 목적에 따른 분류

문제은행은 그것의 규모나 형태, 조직방식 등에 따라 매우 다양한 모습으로 존재하지만 문제은행을 특성에 따라 분류하거나 유형화하려는 시도는 많지 않았다. Choppin(1979)은 국가 수준의 성취도 평가를 위한 문제은행과 교사용 문제은행을 구분하고 각각의 경우 문제은행을 활용하는 것의 이점과 문제은행이 갖추어야 할 형태 및 요건 등에 차이가 있음을 지적하였다. 한편 백순근(1997)은 문제은행의 유형을 크게 ① 다수의 다양한 형태의 문항들을 수합하여 각종 교육 및 심리검사를 개발한 후 검사를 필요로 하는 사람에게 일종의 완성된 형태의 '검사'를 제공하기 위한 문제은행과 ② 학교는 물론 가정이나 사회에서 행해지는 가르치고 배우는 과정에서 활용하고자 양질의 문항을 필요로 하는 사람에게 '문항'을 제공하기 위한 문제은행으로 구분한 바 있다. 즉, 문제은행을 활용하는 최종적인 형태가 완성된 검사인지, 아니면 문항인지에 따라 문제 은행을 구분하여 전자에 해당하는 것은 '검사은행'으로, 후자에 해당하는 것은 '문제은행'으로 개념적으로 분리하면서 교수·학습 활동 지원을 위한 문제은행의 구축의 필요성을 강조하였다.

이런 몇몇 시도에도 불구하고 아직까지 문제은행의 유형과 그에 따른 특성을 분석한 연구는 본격적으로 이루어지지 않았다. 따라서 본 연구에서는 문제은행을 그것을 활용하는 목적에 따라 '교수·학습 지원을 위한 문제은행'과 '선발 및 자격 시험을 위한 문제은행'의 두

유형으로 구분해보고, 문제은행을 구축하고 운영하는 과정을 중심으로 각각의 문제은행이 가지는 특징을 살펴보고자 한다.

1. 교수·학습 지원을 위한 문제은행

교수·학습 상황에서 학생들의 선수 학습 정도를 진단하거나, 수업 도중 학생의 이해도를 확인하기 위해, 혹은 일정한 교육활동이 완료된 후 성취도를 파악하기 위한 평가가 요구된다. 이러한 상황에서 문제은행은 평가와 관련된 교사의 업무부담을 대폭 감소시킬 수 있고, 동시에 양질의 평가문항을 제공하여 교사의 교수활동이나 학생의 학습활동을 촉진하고 개선하는데 기여할 수 있다(백순근, 1997).

교수·학습 지원을 목적으로 하는 문제은행은 교수·학습 과정의 질적인 변화에 궁극적인 목적이 있기 때문에 교수·학습과 평가 활동을 상호 통합할 수 있는 방향으로 문제은행을 구축하게 된다. 따라서 교수·학습 지원을 위한 문제은행은 다음과 같은 특성을 가지게 된다. 첫째, 문항은 교육과정을 토대로 성취기준별로 개발되어 문제은행에 저장된다. 선발이나 자격 시험에 활용되는 문항은 개인 간 능력 변별에 최대의 목적을 두기 때문에 문항의 난이도 등을 중요하게 고려하는 반면 교수·학습 지원을 위해 활용되는 문항은 학생들이 무엇을 성취하였고, 무엇을 성취하지 못하였는지에 대한 구체적인 정보 제공을 우선적으로 고려한다. 둘째 학습자의 흥미를 강조하여 이를 유발하고 유지하기 위한 방안을 모색한다. 교수·학습 상황의 문항은 난이도, 신뢰도, 타당도와 같은 문항의 측정학적 특성을 다소 희생하더라도 학습자의 흥미를 유도할 수 있는 방향으로 문항을 개발한다. 따라서 여러 가지 색상이나 그래프, 그림, 동영상 등을 적용하거나, 새로운 형식, 참신한 내용의 문항을 도입하고 강조한다. 셋째 사용자의 편의성을 지원하는 방식으로 시스템을 구축하고 확장한다. 예컨대 문제은행은 많이 활용될수록 바람직하므로 다수의 사용자가 접근할 수 있도록 고속통신망을 구축하고, 원하는 문항을 손쉽게 검색할 수 있도록 검색 기능을 강화하며, 선택한 문항을 자신의 수업에 적합하도록 수정할 수 있는 편집기능 등에 대한 다양한 장치와 전략을 강조한다. 이러한 노력은 모두 사용자의 편의성을 보장하기 위한 노력이며 궁극적으로는 교수·학습을 촉진시키는 데 목적이 있다.

교수·학습 활동을 지원하기 위한 문제은행의 예로는 ETS(Educational Testing Service)의 형성평가 문제은행(formative assessment item bank), 한국교육과정평가원에서 운영하는 평가시스템, 그 밖에 각국의 정부나 지역교육청 수준에서 제공하는 문제은행과 사설 기관에서 유료 또는 무료로 제공하는 교과별 문제은행 서비스 등 다양한 형태와 규모의 문제은행이 있다.

미국의 ETS(Educational Testing Service)는 형성평가 문제은행을 개발하여 초·중등학교 교사들이 수업에서 활용할 수 있도록 지원하고 있다. 문항은 구체적인 교수목표와 관련하여

학생들의 성취 정도를 파악하고 그 정보를 수업에 활용하는 것을 목적으로, 성취 기준에 근거해 개발된다. K-12 학년에 걸쳐 수학과 언어(읽기, 쓰기) 영역을 평가하는 문항들로 구성되어 있는데 문항은 교과 전문가들에 의해 제작되고 검토되었으며, 문항의 난이도 등에 대한 정보는 문제은행의 활용이 본격적으로 이루어지면서 데이터가 축적 되는 대로 전국 수준에서 혹은 주별로 규준에 대한 정보를 제공할 계획이다. 이 문제은행은 교수정보 관리시스템(IDMS: Instructional Data Management System)의 일부로 Test Builder라고 불리는 IDMS의 하위 모듈을 통해 제공된다. 이 모듈은 각 주에서 설정하고 있는 교수목표 혹은 성취기준과 그에 관련된 평가문항들을 보여주며, 주별 성취기준에 따른 평가도구를 작성해 주고, 검사지와 답안지를 출력할 수 있도록 해준다. 서비스는 유료로 제공되며 채점, 성적표 작성, 교육과정 안내, 수업 설계 등 다양한 기능들을 함께 제공하고 있다.

한국교육과정평가원은 1998년 국내 최초로 과학과 수학 교과에 대한 중학교 학력검사 시티켓(CT & CAT)을 개발하여(백순근 외, 1998) 1999년부터 인터넷을 통한 검사서비스를 시작하였다. 1999년에는 중학교 국어, 영어, 사회 교과의 평가 문항이 추가로 개발되었고(백순근 외, 1999), 2000년에는 중학교 2~3학년 5개 교과 평가문항이 포함되었다(채선희 외, 2000). 2001년에는 1998년부터 진행되어 이미 탑재된 총 2,000개의 문항에 대한 내용 검토가 시행되었고 시스템을 수정하고 보완하는 작업도 진행되었다. 동시에 중학교 1학년 국어, 수학, 영어, 사회, 과학 5개 교과에 대한 총 500문항을 개발·탑재하였다(채선희 외, 2001). 이 시티켓 시스템은 스스로 모의고사를 치러 보려는 학생을 대상으로 하고 있으며, 문항정보는 각 문항의 정답률과 변별도, 1모수 라쉬모형에 근거한 문항곤란도 지수, 문항적합도 등을 제시하고 있다. 4년 이상의 노력 끝에 탄생된 학력 검사 문제은행은 총 2,500개의 문항을 탑재하고 있는데 국내에서 컴퓨터 개별 적용 검사 시스템을 구현한 최초의 시도라는 점에서 커다란 의미가 있다. 그러나 2,500개의 문항은 교수·학습의 목적을 충족하기에는 제한적이고 미흡하였다.

본격적인 교수·학습용 문제은행은 한국교육과정평가원이 다시 2002년도부터 진행한 국민공통기본 교과의 학년별 문제은행 구축을 통해서이다. 시티켓이 학생을 대상으로 한 문제은행이었다면 새로 개발된 문제은행은 일선 교사들이 활용할 수 있도록 제공한다는 점에서 차별적인 특징을 지니고 있다. 2005년 현재 국민공통기본 10개 교과의 10개 학년의 16,000여 개 문항이 저장되어 있으며 매년 지속적으로 문항을 개발하여 추가로 탑재하고 있다. 문항들은 제7차 교육과정의 교과별 성취 기준을 근거로 개발되었는데, 새로운 교육과정의 도입으로 인한 교사들의 평가관련 활동을 지원하려는 목적으로 구축되었다. 문항검색은 교과별, 학년별, 단원별과 같은 교육과정별로도 가능하나 또한 키워드에 의한 검색도 가능하도록 디자인되어 있다. 이 문제은행은 교사의 등록된 아이디와 패스워드를 통해서 접속이 가능하다(조난심 외, 2002).

2. 선발 및 자격 시험을 위한 문제은행

선발이나 자격 인증을 목적으로 동일한 내용과 수준의 검사가 주기적으로 반복될 때, 이들 검사를 안정적이고 효율적으로 운영하기 위해 문제은행이 활용될 수 있다. 선발 및 자격 시험에서 문제은행을 활용하게 되면 다음과 같은 이점이 있다. 첫째, 시험의 내용 및 난이도를 일관성 있게 유지할 수 있다. 둘째, 사전에 충분한 검토와 보완을 거쳐 선별된 문항들을 사용해 검사를 제작함으로써 양질의 검사를 제작할 수 있다. 셋째, 장기적으로 검사 제작에 필요한 시간과 노력을 줄일 수 있다. 넷째, 검사 개발에 드는 시간과 노력을 줄임으로써 검사 실시 회수를 늘일 수 있다. 검사 실시 회수의 증가는 응시자들이 자신이 원하는 시기에 시험을 치를 수 있고 불가피한 사정으로 시험을 치르지 못하거나 시험에서 낮은 평가를 받은 경우 재시험을 치를 기회를 제공하는 등 응시자의 편의성을 제고할 수 있다.

선발 및 자격시험을 위한 문제은행을 효율적으로 활용하기 위해서는 다음과 같은 특성이 요구된다. 첫째, 문항의 난이도, 변별도 등 문항의 통계적 특성에 대한 정보가 필수적이다. 선발이나 자격 인증을 위한 평가에서는 검사가 응시자들의 능력 차를 적절히 변별해 주거나 혹은 합격/불합격을 나누는 분할점 부근에서 피험자들의 능력 수준을 세밀히 변별해 주어야 한다. 그러므로 검사 개발자는 응시자들의 능력 수준이나 분할점 등을 고려하여 검사 점수가 일정한 분포 형태나 분산을 갖도록 검사를 구성해야 하며 따라서 문항의 난이도나 변별도 등이 문항을 선정하는 중요한 기준이 된다. 둘째, 문제은행 및 완성된 검사에 대한 보안 유지가 무엇보다도 강조된다. 선발이나 자격 인증을 위한 평가는 평가결과가 응시자 개인에 대한 중요한 의사결정 자료로 활용되는 고부담시험이다. 따라서 응시자들은 평가에서 좋은 결과를 얻기 위해 평가에 대비한 연습을 하는 등 노력을 기울이는 한편 평가와 관련한 가급적 많은 정보를 수집하고자 하며 경우에 따라서는 불법적인 정보까지도 활용하고자 한다. 보안 문제는 문제은행의 활용방안을 좌우할 뿐만 아니라, 많은 경우 문항분석을 위한 예비 검사의 실시를 불가능하게 하는 등 문항의 질 관리 작업 내지 질 사전 보장 작업을 크게 제약하기도 한다.

문제은행은 실제로 선발이나 자격 인증을 위한 국내·외의 여러 시험에서 활용되고 있다. 국내의 경우 운전면허자격시험의 학과시험과 외무고시, 행정고시 등의 1차 시험인 공직적격성시험이 문제은행을 활용해 출제되고 있으며, 서울대학교 언어교육원에서 주관하는 영어능력인증시험(TEPS)도 문제은행 방식으로 운영되고 있다. 국외의 경우는 ETS에서 운영하는 여러 시험들(미국의 대학원입학시험인 GRE, 영어능력시험인 TOEFL, 의사자격시험인 MCAT, 교사자격시험인 PRAXIS 등)이 문제은행을 기반으로 검사를 제작하고 있으며, 나아가 문제은행을 기반으로 컴퓨터를 이용한 검사(CT: Computerized Test) 및 컴퓨터를 활용한 개인별 및

츄검사(CAT: Computerized Adaptive Test)의 형태로 운영되기도 한다.

이들 중 GRE(General Record Examination)에서 문제은행에 기반한 CAT를 도입 한 목적과 그 현황에 대해 간략히 소개하면 다음과 같다. GRE는 대학원 응시자들이 대학원에 입학할 수 있는 학력을 갖추었는지를 확인하고 대학원생들의 선발을 돕기 위한 시험이다. 이 시험은 학업에 필요한 일반적인 능력을 측정하는 검사(general test)와 특정 전공영역(심리학, 기계공학 등)에서의 성취를 측정하는 검사(subject tests)로 구성되는데, 이 중 GRE-일반검사는 분석적, 수리적, 언어적 추리력을 측정하며, 학부 성적과 같은 다른 예언변인들과 함께 사용될 때 대학원 입학생의 학업수행정도(대학원 첫 해의 성적)를 잘 예언하는 것으로 알려져 있다. GRE-일반검사는 1993년부터 지필검사와 함께 CAT 검사가 병행되고 있다. CAT 방식의 도입 목적은 기존의 선택형 문항으로만 구성된 검사에 서답형 문항 등 새로운 형식의 문항을 도입함으로써, 검사의 구인을 보다 개선된 방식으로 측정하고 현재 검사와는 다른 기능이나 능력들을 측정하는 새로운 검사를 가능하게 함으로써 궁극적으로 검사의 질을 향상시키는 것을 주된 목적으로 하였다. GRE-CAT가 도입된 후 5년이 지난 1997년에는 이미 약 50만 명이상의 응시자들이 CAT 검사를 받았으며, 75개의 CAT 문항풀이 개발되었다(Mills, 1999; Mills & Steffen, 2000).

IV. 유형별 문제은행의 구축과 활용 방안

앞서 우리는 문제은행을 활용하는 목적에 따라 ‘교수·학습 지원을 위한 문제은행’과 ‘선발 및 자격 시험을 위한 문제은행’을 구분하여 각각의 특징을 살펴보았다. 여기에서는 두 유형의 문제은행을 문항개발, 시스템 구축, 활용의 측면으로 나누어 각각의 특징을 구체적으로 정리해보고 점검해보고자 한다.

1. 문항 개발

문제은행을 구축하고 활용함에 있어서 핵심적인 사항 중 하나는 양질의 문항을 확보하는 것이다. 문제은행의 질은 그것을 구성하는 문항의 질에 달려 있다고 해도 과언이 아니다. 또한 문제은행은 다양한 내용, 수준, 형태를 포괄하는 충분히 많은 수의 문항들로 구성되어야 함을 앞서 논의한 바 있다. 많은 수의 다양한 문항을 확보한다는 것은 검사개발자 혹은 평가자가 원하는 내용, 형태, 통계적 특성에 합치하는 문항을 찾아낼 가능성이 더 높다는 것을 의미한다(Millman & Arter, 1984; 채선희, 1996). 그렇지만 문제은행을 평가 목적에 따라 교수·학습 지원을 위한 문제은행과 선발 및 자격 시험을 위한 문제은행으로 나누어 볼 때,

각각의 평가 목적을 충족하기 위해 요구되는 문항의 양호도 기준과 적정한 문항의 수는 다를 수 있다. 이러한 맥락에서 두 유형의 문제 은행에서 바라본 좋은 문항은 무엇이며, 각각은 얼마나 많은 수의 문항을 확보해야 하는가에 대한 문제를 제기해 볼 수 있겠다.

가. 교수·학습 지원을 위한 문제은행

교수·학습 지원을 위한 문제 은행이 갖추어야 할 가장 중요한 조건 중의 하나는 교수·학습 과정에 요구되고, 실제적으로 현장에서 활용도가 높은 문항들을 확보하는 것이다. 문항 개발은 다음과 같은 방향에서 이루어지는데 문제은행이 이러한 문항들로 구성된 경우 좋은 문제은행으로 분류된다.

첫째, 교육과정을 토대로 개발된 문항들로 구성되어야 한다. 교수·학습 지원을 위한 문제 은행은 교사들이 수업을 진행하는데 필요한 평가 활동을 지원하거나 혹은 학습자가 학습한 내용을 어느 정도 이해하였는지 스스로 점검하는데 초점을 두게 된다. 자신의 수업 혹은 학습 목표가 도달된 정도를 확인할 수 있으려면 문항들은 교육과정에 근거하여 개발되어야 한다. 따라서 교육과정에서 제시되는 성취 기준별로 문항을 개발하는 것이 바람직하며, 평가 결과도 각 성취 기준에 따라 그것의 달성 정도에 대한 정보를 제공하여 교수 및 학습 과정에 유용한 피드백을 제공할 수 있어야 한다. 교육과정에 따른 성취기준과 평가기준에 토대를 둬으로써 교육과정 타당도를 확보할 수 있게 된다.

둘째, 충분한 수의 문항이 개발되어 포함되어 있어야 한다(백순근 외, 1998). 평가활동은 수업의 전 국면에서 이루어질 수 있다. 따라서 수업의 어느 과정에서든 평가 활동이 가능하도록 다양한 층의 문항을 충분히 보유하고 있어야 한다. 예컨대, 진단평가에 활용되는 문항은 형성평가를 위한 문항과는 내용수준과 난이도에 있어서 차이가 있다. 성취도를 평가하고자 하는 문항은 진단을 목적으로 하는 문항과 내용과 형식면에서 또다른 차이를 보인다. 교수·학습 지원을 위한 문제은행을 활용하는 주체가 교사 혹은 학습자라는 점을 감안한다면 그들의 사용목적, 요구, 개인적 특성에 따라 필요로 하는 문항도 다양할 것이다.

셋째, 개발된 문항을 대상으로 좋은 문항을 선별하여 문항을 탑재하는 과정에서 가장 중요하게 검토되는 기준은 교육과정 타당도가 된다. 어떤 문항이 좋은 문항인지 혹은 어떤 문항을 최종적으로 문항 데이터베이스에 탑재할 것인지를 결정하는 데 있어 선발을 목적으로 하는 경우에는 개인 간 능력을 최대로 변별할 수 있는 난이도가 중요하지만 교수·학습을 목적으로 하는 경우에는 난이도 자체는 교육과정 타당도만큼 중요한 기준은 아니다. 문항이 측정하고 있는 내용이 교수의 목적에 부합하지 않다면 여러 능력 수준을 예리하게 변별하는 문항이라 할지라도 교수·학습의 과정에는 도움이 되지 못한다.

넷째, 학습자의 흥미를 유발할 수 있을 정도로 참신한 문항을 포함하고 있어야 하며, 학습자의 특성이 다양한 만큼 평가 문항도 다양화되어야 한다. 학습자의 흥미를 유발하기 위한

방법은 여러 가지가 있을 수 있다. 우선 문항 유형을 다양화할 수도 있고 만화나 그림, 사진, 듣기 등을 이용하여 문항을 구성해 보는 것도 한 가지 방법이다. 문항의 소재나 제재를 색다르게 하는 것도 좋은 대안이다. 교과서 이외에 실생활 속에서 소재를 찾아와 문항을 구성할 때 학습자의 능동적이고 적극적인 학습 참여를 기대할 수 있다(조난심 외, 2002). 이러한 문항은 비록 난이도와 변별도, 타당도와 같은 측정학적 기준에 있어서 다소 미흡하더라도 교수·학습의 과정을 촉진시킨다는 점에서 교수·학습 지원을 위한 문제은행에 꼭 필요한 요소라 볼 수 있다.

나. 선발 및 자격 시험을 위한 문제은행

선발 및 자격 시험을 위한 문제은행은 검사의 목적을 잘 달성할 수 있도록, 요구되는 검사특성(평가영역별 구성 비율, 적정한 점수 분포 등)을 충족시키는 검사를 효율적으로 개발할 수 있어야 한다. 이를 위해 문제은행은 평가 영역이나 난이도 수준별로 충분한 수의 문항을 확보해야 하며, 선발 및 자격 인증을 위한 검사에 활용되는 만큼 결함이 없는 양질의 문항들로 구성되어야 한다. 선발 및 자격 시험을 위한 문제은행에서 문항을 개발 혹은 수집하는 과정의 특징은 다음과 같다.

첫째, 문제은행의 구축에 앞서 검사의 목적, 대상, 내용, 형식 등 검사의 제반 특성을 명확히 규정하는 것이 필요하다. 검사의 제반 특성을 구체화한 것을 검사 설계도(test blueprint) 또는 검사의 상세화(test specification)라고 하는데, 이는 어떤 내용과 형식의 문항을 얼마나 개발할 것인가를 결정하는 기초가 된다. 이러한 검사 설계도는 일반적으로 검사에 관한 주요 의사결정을 내리기 위해 구성되는 검사 위원회에 의해 결정된다.

둘째, 평가 영역 및 난이도 수준별로 충분한 수의 문항이 확보되어야 한다. 보유한 문항의 수가 많다는 것은 동일한 문항이 자주 반복되어 사용되지 않는다는 것을 의미하고, 따라서 피험자들에게 문항이 노출되거나 피험자들이 답을 암기하는 등의 문제가 최소화됨을 의미한다. 또한 문항을 활용하는 검사개발자 혹은 평가자가 자신이 원하는 내용, 형태, 통계적 특성에 합치하는 문항을 찾아낼 가능성이 더 높아진다는 점에서도 다수의 문항을 확보하는 것이 바람직하다. 그러나 문제은행의 규모는 항상 그것을 구축하기 위한 비용 문제와 결부되어 있다. 특히 선발 및 자격 시험에서는 동일한 문항들이 계속해서 검사에 사용될 수 없으므로 검사의 시행 빈도 및 일 회의 검사에서 사용하는 문항의 수에 비례하여 지속적으로 문항을 개발하고 문제은행에 추가해야 한다. 따라서 문제은행의 적정 규모는 어느 수준이어야 하는가에 대한 논의가 끊임없이 계속되고 있다. 문제은행의 적정 규모는 학자들에 따라 대략 1회 검사에서 사용될 문항 수의 적게는 10배(Prosser, 1974)에서 많게는 50배(Davey & Nering, 1998)까지 제시되고 있다. 그러나 실제로 문제은행의 규모가 어느 정도여야 하는가는 검사의 실시 방법, 실시 빈도, 1회 검사에서 사용되는 문항의 수, 검사의 중요성이나 민감성 등에 따라 달라진다.

셋째, 일단 수집되거나 개발된 문항들은 문항의 양호도 검토 과정을 거쳐 양호도가 입증된 문항들만이 문제은행에 저장된다. 이때 문항의 양호도로서 중요하게 고려되는 측면들은 우선 문항이 평가 영역 및 평가하려는 행동 특성을 적절히 반영하는가, 문항의 내재적 결함은 없는가(정답이나 채점 기준이 명확한가, 성이나 문화, 인종에 따른 차별적 요소는 없는가 등) 하는 점이다. 다음으로 문항의 경험적 양호도를 분석하는데 이 과정에서 문항난이도, 변별도, 공정성 등 일반적인 문항 분석 절차에서 사용되는 요소들이 검토된다. 선발이나 자격 인증을 위한 시험은 응시자들의 능력 차이나 일정한 자격 요건의 충족 여부를 변별해 주어야 하므로 문항들은 적절한 변별력을 갖춰야 한다. 또한 피험자들의 선발, 자격 부여 등 중요한 의사결정과 관련된 시험인 만큼 성별이나 문화, 인종, 계층, 지역에 따른 편파성을 갖지 않는 공정한 문항들일 것이 요구된다.

넷째, 선발 및 자격 시험을 위한 문제은행에서는 문항의 개발과 검토 과정에서 문항이 외부로 유출되지 않도록 보안을 유지하는 것이 필요하며, 특히 문항 검토 및 수정을 마치고 문제은행에 최종 입력되는 문항들에 대해서는 더 엄격한 보안 유지가 요구된다. 이 때문에 문항 검토의 경우 내부 검토자들만을 활용하거나 또는 문항 검토를 지정된 공간에서 디스켓이나 카메라 등의 물품 반입을 통제한 가운데 실시하는 경우도 있다. 또한 문항에 대한 보안 유지 문제는 경우에 따라 사전 검사 및 문항 분석을 통한 문항의 질 확보 방안을 제약하기도 한다. 즉, 선발 및 자격 시험을 위한 문제은행에서는 검사 결과가 미치는 영향의 중요성으로 인해 정련되고 양호도가 입증된 문항만이 문제은행에 탑재될 것을 요구하는 반면, 동일한 이유로 문항의 양호도에 대한 검증에 제약을 받는 이중적인 특성을 보인다.

〈표 1〉 문항 개발에 있어서 유형별 문제은행의 특징

	교수·학습 지원을 위한 문제은행	선발 및 자격을 위한 문제은행
문항 개발 기준	<ul style="list-style-type: none"> 교육과정 및 교수목표. 상대적으로 중요한 비중을 차지하는 영역의 문항이 더 많이 개발되기는 하나 기본적으로 모든 영역을 고루 포함하여야 함. 	<ul style="list-style-type: none"> 검사의 목적, 평가 영역(내용, 행동) 및 영역별 비중, 문항 형태 등을 규정한 검사설계도
적정 규모	<ul style="list-style-type: none"> 교육 과정에서 제시하는 다양한 성취 기준, 평가 목적과 상황 등을 충족시킬 수 있는 충분한 문항이 필요 사용자의 관심과 흥미를 유지하려면 지속적으로 문항을 개발하고 탑재할 필요가 있음. 	<ul style="list-style-type: none"> 요구되는 검사 특성에 맞는 문항을 충분히 추출할 수 있고, 문항 보안의 문제를 해소하기 위해 충분한 문항 수 확보가 중요 일정 수준 이상 사용된 문항은 보안상의 이유로 폐기되거나 일정 기간 사용이 중단되므로, 지속적인 문항 개발과 문제은행의 갱신이 필요
문항 양호도 기준	<ul style="list-style-type: none"> 교육과정 타당도 학습자의 흥미를 유도할 수 있는 참신한 소재와 내용 	<ul style="list-style-type: none"> 문항의 난이도, 변별도, 편파성 정답 혹은 채점 기준의 명확성

2. 문제은행의 구축

문항 개발 및 검토 과정을 거쳐 양호도가 입증된 문항들만이 최종적으로 문제은행에 저장된다. 그러나 수많은 문항들을 관리하고 활용하기 위해서는 단순히 많은 문항들을 저장해 두는 것만으로는 충분치 않으며, 문항들을 체계적으로 분류하고 조직화해 줌으로써 사용자들이 간편하게 그리고 효율적으로 활용할 수 있도록 해야 한다.

사용자가 자신의 목적이나 필요에 맞는 문항을 효과적으로 검색하고 활용할 수 있기 위해서는 문항의 내용, 형태, 난이도 등에 대한 상세한 정보가 각 문항마다 수집되고 체계적으로 저장되어야 한다(부제율, 2002; 성태제, 2002; Veldkamp & van der Linden, 2000). 이처럼 각각의 문항에 대해 교수목표, 평가영역, 형태, 문항의 통계적 특성, 개발자나 검토자 등에 대한 정보를 체계적으로 기록한 것을 문항카드(문항 정보표)라 한다. 문항카드에 입력되는 정보들은 문항을 검색하고 활용함에 있어서 중요한 참조 기준이 되는 것들로, 문항 카드에 어떤 정보들을 어떻게 분류하여 입력할 것인가의 문제는 문제은행을 관리하고 활용함에 있어서 매우 핵심적인 사항이다. 따라서 문제은행의 구축 이전에 이에 대한 신중한 논의와 검토가 요구된다. 그렇다면 두 유형의 문제은행에서 어떤 문항 정보들이 특히 중요하게 고려되는지를 살펴보고 시스템을 구축하는 데 있어서 어떠한 차이점이 있는가를 살펴보고자 한다.

가. 교수·학습 지원을 위한 문제은행

문항에 대한 상세한 정보는 문항 분류를 위한 코드로서 작용하며 동시에 문제은행을 해독하는 열쇠가 된다. 또한 문항이 저장될 때 문항과 함께 저장되는 기타 정보들은 각 문항에 대한 교수·학습적으로 유의미한 여러 정보를 담게 된다. 각 문항은 위와 같은 문항카드에 포함되는 여러 정보와 일치할수록, 정보가 풍부할수록 활용도가 높아지게 된다. 학교 현장의 평가 활동을 지원하기 위해 한국교육과정평가원이 구축한 평가시스템의 수학과 문항정보표의 예를 통해 교수·학습 지원을 위한 문제은행에서 사용되는 문항 정보의 유형과 그 특성을 살펴보면 다음과 같다.

[그림1]의 정보표는 학년(단계), 영역, 중영역, 평가요소, 난이도, 내용수준, 평가유형, 행동요소, 성취기준, 평가기준을 구성요소로 포함하고 있다. 우선 각 문항은 제7차 교육과정에 규정하는 성취기준과 평가기준에 대한 정보를 담고 있는데 수학과와 경우 단계형 교육과정을 따르므로 ‘내용수준’에 대한 정보는 현장에서 활용하는 데 유용할 수 있다. ‘내용수준’은 수학 내용의 깊이를 표현한 용어로, 동일한 내용도 다루는 방식에 따라 깊이가 다를 수 있다. 이점에서 문항이 학생들이 풀기에 쉽거나 어려운 정도를 나타내는 난이도와 개념적으로 차이가 있다.

문항번호						
단계	7-가	영역	수와 연산	중영역	정수와 유리수	
평가요소		간단한 덧셈식 계산하기			난이도	어려움/보통/쉬움
					내용수준	심화/기본(상,중,하)/보충
평가 유형	문항 형태	선택형 / 단답형 / 서술형			행동요소	계산 / 이해 / 추론 / 문제해결 / 의사소통
	수행형	도전형 / 소양형 / 프로젝트형				
성취기준		정수와 유리수의 사칙계산을 할 수 있다.				
평가 기준	상	유리수의 사칙계산을 이용하여 여러 가지 문제를 해결할 수 있다.				
	중	유리수의 사칙계산을 할 수 있다.				
	하	정수의 사칙계산을 할 수 있다.				

(그림 1) 교수·학습 지원을 위한 문제은행의 문항정보표 예

한편 평가유형을 보면 이는 다시 문항형태와 수행형으로 나뉘어 진다. 개념적으로는 수행형 문항도 문항 형태의 범주에 들어가는 것이나 수행평가가 일반 평가와는 구분되기를 원하는 학교 현장의 요구 및 활용도에 염두를 둔 분류체계라 볼 수 있다(조난심 외, 2002).

이상에서 살펴본 바와 같이 교수·학습 지원을 위한 문제은행에서 문항정보는 엄격한 개념에 토대를 두기 보다는 학교 현장에서의 활용도를 염두에 두고 구성되고 진행된다. 이렇게 개발된 평가문항은 교사들이 더 유용하게 활용할 수 있도록 ‘문제’와 ‘풀이’는 물론 ‘오류 유형’ ‘별해’ ‘변형 가능한 부분’에 대한 정보를 추가적으로 담고 있다. 오류유형은 학생들에게 발생될 수 있는 오류에 대한 정보를 담고 있으며, 별해는 학생들의 창의적 사고를 유도하도록 정식 풀이는 아니지만 논리적으로 타당하고 충분히 등장할 수 있는 풀이이며 변형 가능한 부분은 교사가 문항을 변형하여 활용할 수 있는 부분에 대한 정보이다. 이들 정보는 모두 효과적인 교수·학습 과정에 필요한 정보들이며, 문항의 분류와 검색을 위한 코드들이다.

교수·학습 지원을 위한 문제은행을 구축하여 활용하는 경우 서버의 용량과 처리 속도는 시스템과 관련하여 가장 큰 문제라 볼 수 있다. 교수·학습을 지원하기 위한 문제은행에서는 모든 문항이 공개되며, 많이 활용될수록 좋은 문항이므로 이를 위한 다양한 장치와 노력을 하게 된다. 따라서 문제은행은 다수의 사용자가 동시에 접속할 수 있도록 서버의 용량을 크게 하고 시스템의 운영 처리 속도를 빠르게 하는데 많은 노력과 투자를 하게 된다. 많은 수의 문항을 개발하는 데도 적지 않은 물적·인적 자원과 시간, 노력이 요구되지만 특히 대규모의 서버와 빠른 통신 속도 등에 대한 요구는 교수·학습 지원을 위한 문제은행의 활용도를 결정하는 중요한 요소이므로 이에 대한 투자와 노력도 지속적으로 진행되어야 한다.

나. 선발 및 자격 시험을 위한 문제은행

선발 및 자격 시험을 위한 문제은행에서는 문항의 난이도, 변별도 등 문항의 통계적 특성

에 대한 정보가 중요한 의미를 갖는다. 선발이나 자격 인증을 위한 시험은 검사의 내용이나 난이도 수준에서 일관성을 유지해야 하며, 응시자들의 능력을 적절히 변별할 수 있어야 하고, 검사의 목적에 따라 응시자들의 점수 분포가 일정한 형태를 갖도록 조정할 필요가 있기 때문이다. 따라서 문항의 난이도, 변별도 등의 정보는 문항을 검색, 추출하는 과정에서 중요한 기준으로 사용될 뿐 아니라 최종적으로 특정 문항을 검사에 포함시킬 것인가 여부를 결정하는데 직접적인 영향을 준다.

그러나 선발이나 자격인증을 목적으로 하는 시험에서는 문항에 대한 높은 수준의 보안 유지가 요구되므로 사전 검사를 통해 문항의 난이도 등 통계적 특성에 대한 정보를 수집하지 못하는 경우도 있다. 이런 경우 대체로 문항 개발자나 검토자들이 문항의 난이도를 예측하여 예측난이도를 대신 사용하기도 한다. 그러나 예측난이도는 실제 난이도와는 차이가 있을 수 있으며, 문항에 따라서는 그 차이의 정도가 상당히 클 수도 있다. 또한 문항변별도나 편파성의 경우 문항 양호도에 대한 질적 검토만으로는 문항의 결함을 충분히 발견하지 못할 수 있으며, 결과적으로 변별력이 거의 없거나 편파성이 있는 문항이 검사에 사용될 가능성도 배제할 수 없다.

선발 및 자격 시험을 위한 문제은행에서는 문항의 사용 이력에 대한 정보도 중요하게 활용된다. 이 때 문항의 사용 이력 정보는 문항의 활용도란 의미보다는 문항의 노출 정도를 반영하는 지수로서 해석되고 활용된다. 즉, 한번 사용되었거나 혹은 일정한 회수 이상 사용된 문항은 폐기하거나 일정 기간 사용을 중단하는 것이다. 일반적으로 검사가 한 번 시행될 때 모든 응시자들에게 동일한 검사가 주어지는 경우 한 번 사용된 문항은 폐기하거나 일정 기간 이상 사용을 중단한다. 이에 비해 개인별 맞춤 검사 등 응시자에 따라 상이한 문항들로 구성된 검사가 제공되는 경우는 문항 사용 빈도나 비율이 일정 수준을 넘을 때 문항을 폐기 또는 사용 중지 한다.

한편 문제은행의 시스템을 구축하는 단계에서도 보안 유지는 중요한 고려사항이다. 선발 및 자격 시험을 위한 문제은행은 대체로 검사 운영 센터 등 독립된 공간에 독립된 시스템으로 구축된다. 문제은행 시스템이 설치된 공간은 일반적으로 외부인의 출입이 엄격히 통제되며, 온라인을 통한 문제은행 접속도 허용되지 않는다. 또한 문제은행을 상대적으로 규모가 작은 다수의 문제은행으로 분할하여 일정 간격을 두고 이들 문제은행을 순환하며 활용하는 방식을 취하기도 한다. 이런 전략은 모두 문항의 노출 정도를 효과적으로 통제하고 동시에 문제은행이 노출되는 사고가 발생하더라도 그 피해를 최소화하기 위한 방안이라고 볼 수 있다.

〈표 2〉 문항정보 및 시스템 구축에 있어서 유형별 문제은행의 특징

	교수·학습 지원을 위한 문제은행	선발 및 자격을 위한 문제은행
문항 정보	<ul style="list-style-type: none"> 문항정보표의 문항정보는 분류와 검색을 위한 도구로 엄밀한 학문적 분류보다는 현장에서 활용성이 높은 분류체계를 따름. 학생들이 자주 보이는 오류나 별해와 같은 상세한 정보 제공 	<ul style="list-style-type: none"> 문항의 난이도, 변별도 등 통계적 특성 문항의 사용 이력 : 문항의 노출 정도를 나타내는 지수로서 중요하게 관리함.
검색과 추출	<ul style="list-style-type: none"> 원하는 문항을 손쉽게 검색하고 추출하기 위한 다양한 기능이 활용도를 결정하는 관건 	<ul style="list-style-type: none"> 검사설계에 따라 평가영역, 난이도별 문항 수를 지정하고 그에 따라 문제은행으로부터 문항을 추출함.
사용자	<ul style="list-style-type: none"> 교사 및 학습자 	<ul style="list-style-type: none"> 문제은행 관리자 혹은 평가전문가
시스템 구축 시 고려점	<ul style="list-style-type: none"> 활용도를 높이기 위한 환경과 방안 마련이 우선적이므로 동시 접속자가 많은 경우 시스템 작동에 오류가 없도록 서버의 용량을 확대하고 운영시스템의 처리 속도를 가속화하는 노력이 필요 	<ul style="list-style-type: none"> 시험의 보안이 중요하므로 서버가 있는 곳의 출입을 통제하고 문제은행 접속도 통제하는 등의 보안유지 방안이 요구됨.

3. 문제은행의 활용

문제은행을 활용하는 방식은 문제은행으로부터 추출된 문항을 그대로 사용하는가 아니면 수정하거나 변형하여 사용하는가, 그리고 문제은행이 문항들을 검색하는 기능만을 제공하는가 아니면 문제은행 내에 자동으로 검사를 구성하거나 더 나아가 검사를 실시하는 기능이 결부되어 있는가에 따라 다양하다. 교수·학습 지원을 위한 문제은행과 선발 및 자격 시험을 위한 문제은행은 문제은행의 사용자와 활용 목적이 서로 다르므로, 그에 따라 문제은행을 활용하는 방식에서도 상당한 차이가 존재한다.

가. 교수·학습 지원을 위한 문제은행

교육과정은 실제 현장의 교수·학습 과정을 통해 구현되게 되는데 교수자의 특성, 학습자의 특성, 교수·학습이 이루어지는 상황, 환경적 요인 등에 의해 교수·학습 과정은 다양하게 영향을 받을 수 있다. 그러므로 수업에 따른 평가는 이러한 다양성과 독특성을 반영할 수 있어야 하며, 문제은행도 단순히 문항을 검색, 추출하는 기능만이 아니라 교사가 자신의 수업에 맞도록 문항을 변형하거나 수정할 수 있는 기능을 갖추고 있어야 한다. 검색을 통해 추출된 문항들을 편집할 수 있는 편집 기능을 구비하고 있을 때 문제은행 시스템은 보다 효율적이고 활용도도 높아진다.

문제은행에 검사를 실시하는 기능이 내제되어 있는 경우 검사 결과는 특정 영역의 교육 목표를 어느 정도 성취하고 있는지 그 달성도를 확인하는 과정이 된다. 따라서 평가 결과는

총점의 형태보다는 영역별로 제시되는 경우가 많다. 같은 맥락에서 오답하였을 경우에 오답을 한 문항의 정답이 무엇인지에 대한 정보와 더불어 문항이 다루는 교수·학습 내용으로 연결될 수 있도록 설계하는 것이 필요하다. 문항이 다루고 있는 학습 요소를 교수 자료와 바로 연결시키는 하이퍼링크 기능이 제공될 때, 평가와 수업이 동시적으로, 유기적으로 진행될 수 있으며 이는 교수·학습 지원을 위한 문제은행 만이 가질 수 있는 독특한 특성이 된다. 이러한 기능이 추가되려면 물론 서버의 용량이 충분히 커야하고 처리 속도 또한 충분히 빨라야 할 것이다.

나. 선발 및 자격 시험을 위한 문제은행

선발 및 자격 시험의 결과는 특정 영역에서 개인의 능력 수준이나 상대적인 우열을 나타내는 정보로서 활용된다. 검사결과는 응시자의 선발과정에 활용될 수 있는 형태의 검사점수로 통보되는데, 평가 영역별로 세분된 검사 점수가 제공되기도 하지만 일반적으로는 평가영역별 하위점수가 합산된 검사 총점이 중요한 의미를 갖는다. 자격 시험의 경우는 검사점수와 함께 일정한 분할점수에 의한 합격/불합격 등의 정보가 함께 제시된다.

한편 주기적으로 반복해서 실시되는 시험의 경우 검사 점수의 척도를 일관되게 유지해야 할 필요가 발생하기도 한다. 예를 들어 SAT 시험의 경우 일년에 7번 검사가 실시되는데, 서로 다른 시기에 시험을 치른 학생들의 점수가 대학에서의 신입생 선발 과정에서 선발자료로 함께 사용된다. 따라서 시험을 어느 시기에 보았는가에 따라 유리하거나 불리한 문제가 발생하지 않도록 매 시기 검사 점수들이 상호 비교 가능한 점수 척도를 유지해야 하며, 이를 위해서는 검사점수의 동등화가 필요하다.

〈표 3〉 활용의 측면에서 유형별 문제은행의 특징

	교수·학습 지원을 위한 문제은행	선발 및 자격을 위한 문제은행
문항 변형	<ul style="list-style-type: none"> 검색하고 추출한 문항을 자신의 수업에 맞게 변형할 수 있음. 따라서 편집 기능이 중요 	<ul style="list-style-type: none"> 문항은 검사에 바로 사용될 수 있는 형태, 수준이므로 특별한 변형이나 수정이 필요 없음.
평가 결과 활용	<ul style="list-style-type: none"> 학습자의 성취 정도는 총점뿐 아니라 영역별, 성취기준별로 나누어 자세히 통보됨. 오답을 한 경우 평가문항은 하이퍼링크를 통해 해당 교수 내용과 바로 연결되도록 디자인 	<ul style="list-style-type: none"> 개인의 능력 수준에 대한 정보이므로 검사 총점을 통해 선발이나 자격 인증에 활용될 수 있는 형태로 통보됨.

V. 맺음말

이 연구에서는 평가의 목적에 따라 문제은행을 크게 ‘교수·학습 지원을 위한 문제은행’과 ‘선발 및 자격 시험을 위한 문제은행’으로 개념적으로 구분한 다음 이를 토대로 각각의 문제은행이 어떻게 구축되고 활용되어야 하는지에 대한 검토와 논의를 진행하였다.

교수·학습 지원을 위한 문제은행에서는 교수·학습 과정에 유의미한 정보를 제공할 수 있는 양질의 문항을 충분히 확보하고 교사나 학생들이 그것을 편리하게 활용할 수 있도록 하는데 초점이 모아진다. 따라서 다양한 교수·학습 상황의 요구를 충족시킬 수 있도록 문항 수가 충분한지, 문항은 교육과정의 내용을 충분히 반영하고 있는지, 학습 흥미를 유발할 수 있는 참신한 문항이 포함되어 있는지, 원하는 문항을 손쉽게 검색하고 추출하도록 검색 시스템이 마련되어 있는지, 검사지로 바로 제작될 수 있는 템플릿이 존재하는지 등이 문제은행의 질과 활용도를 결정하는 평가 기준이 된다.

반면 선발 및 자격 시험을 위한 문제은행에서는 시험의 목적에 부합하는 양질의 검사를 안정적으로 개발하며, 문제은행 및 개발된 검사의 보안을 유지하는데 역점을 둔다. 그러므로 문제은행은 양호도가 입증된 양질의 문항들로만 구성되어야 하며 문항의 난이도, 변별도, 편파성 등의 통계적 특성들이 문항을 선정하는 과정에서 중요하게 고려된다. 또한 문제은행 및 개발된 검사에 대한 접근을 소수의 담당자들에게 제한하고, 문항의 사용 이력을 철저히 관리하여 일정 회수 이상 사용된 문항들은 폐기하는 등의 보안 유지 방안이 문제은행을 구축하고 운영함에 있어 중요한 고려사항이 된다.

이상의 논의를 토대로 문제은행의 구축과 활용에 관련하여 몇 가지 제언을 하고자 한다. 첫째, 문제은행을 개발할 때는 그것의 활용 목적에 따라 문항 개발 방향, 시스템 구축에 필요한 기능, 활용 방법, 예산 등에 대한 철저한 검토와 숙고의 과정을 거쳐야 한다. 하나의 문제은행이 두 가지 상이한 목적 즉, 교수·학습 지원과 선발 및 자격 인증을 동시에 충족시킬 수 없음을 앞서의 논의들이 분명하게 보여준다. 따라서 문제은행을 준비하고 설계하는 시점에서 그것의 목적이 무엇이며, 누구를 대상으로, 어떤 상황에서, 어떻게 활용할 것인지를 명료하게 할 때 효율적으로 개발하고 활용할 수 있을 것이다.

둘째, 교수·학습과정을 촉진시킬 수 있는 문제은행의 구축이 시급히 필요하다. 학교 교육의 질 향상을 위해서는 학생들의 능력을 정확히 파악하여 교수·학습 활동의 개선 정보를 추출하는 평가 활동이 필수적이다. 학생들이 어느 정도의 능력 수준에 도달해 있는지를 진단하고, 어떤 측면에서 왜 능력이 떨어지는지 혹은 높은지를 파악하고, 보충되어야 할 학습의 방향과 내용을 제시함과 동시에, 어떤 교수 활동을 강화해야 하는지 등의 개선 정보를 줄 수 있도록 문제은행을 개발하고 구축하려는 노력과 시도를 아끼지 말아야 한다. 외국 특

히 교육 선진국 중에는 교사의 교수·학습 활동을 지원하기 위한 평가 문항을 개발하여 웹 상에서 보급하고 있는 곳이 많다. 영국의 NGFL(National Grid For Learning), 뉴질랜드의 LM(Learning Media), 미국 플로리다의 BEACON 등이 그 대표적인 예이다. 이들 평가 시스템에는 매 차시, 주제 혹은 단원별로 교수·학습 자료와 평가 문항이 제시되어 교사가 편리하게 활용할 수 있는 체제를 이미 구축하고 있다.

셋째, 양질의 검사를 일관성 있게 안정적으로 시행할 수 있으며 장기적으로 검사개발에 필요한 시간과 비용을 효율적으로 절감할 수 있는 문제은행의 이점을 상기할 때, 대학수학능력시험이나 의·치학 입문시험, 교사 임용고시 등의 여러 시험에서 문제은행을 적극적으로 도입할 필요가 있다. 현재 이러한 시험들은 1년에 한 번 대규모의 합숙을 통해 검사를 개발하고 있으며, 문항개발자나 검토자들은 시험의 보안 유지를 위해 출제가 완료된 후에도 시험 당일까지 출제 장소에 묶여있게 됨으로써 많은 심리적 부담을 느끼고 활동에 제약을 받고 있다. 또한 이들을 관리하기 위한 행정적인 절차와 그에 따른 비용 또한 적지 않다. 따라서 이런 시험들을 문제은행 방식으로 전환한다면 현재 합숙 출제 방식이 안고 있는 많은 문제점들이 완화되고 양질의 검사를 보다 안정적이고 효율적으로 운영할 수 있을 것이다.

이 연구에서는 다양하게 존재하는 문제은행을 평가 목적에 비추어 교수·학습 지원을 위한 문제은행과 선발 및 자격 시험을 위한 문제은행의 두 유형으로 나누어 개념화하고, 각각의 특성과 차이점을 고찰해보았다. 이를 토대로 앞으로 국내·외의 문제은행 구축 및 활용에 대한 보다 심층적이고 구체적인 연구를 진행할 필요가 있다.

참 고 문 헌

- 박도순·배호순·김경성·남명호·구창현·유명규·박문기(1996). **대학수학능력시험의 문제은행식 출제·관리 방안에 관한 연구**. 대학수학능력시험 문제은행식 출제·관리방안 연구위원회.
- 박도순(1988). 문제은행의 운영과 중앙교육평가원의 역할. **문제은행의 운영현황과 발전방향**. 중앙교육평가원.
- 백순근(1997). 교수·학습활동을 돕기 위한 문제은행의 구축 및 운영 방안: 초중고교 문제은행을 중심으로. **교육평가연구**, 10(1), 105-127.
- 백순근·채선희·홍미영·임재훈·전은화(1998). **컴퓨터를 이용한 학력검사 개발 연구**. 한국교육과정평가원 연구보고 RRE98-4.
- 백순근·채선희·이춘식·부재울·전은화·김광주(1999). **컴퓨터를 이용한 학력 검사 지원체제 구축**. 한국교육과정평가원 연구보고 CRE 99-18.
- 부재울(2002). **컴퓨터활용검사**. 교육과학사.
- 성태제(2002). **현대교육평가**. 학지사.
- 조난심·양명희·조영미·이성흠·김수동(2002). **국민공통 기본 교과별 평가도구 개발연구·중등 수학과**. 한국교육과정평가원 연구보고 RRE 2002-7.
- 중앙교육평가원(1986). **문제은행 설치 운영에 관한 기초 연구**.
- 채선희(1996). 문제은행의 본질적 가치와 효율적 이용방안. **교육평가연구**, 9(2), 59-72. 한국교육평가연구회.
- 채선희·강운선·전은화·Linacre J.M(2000). **컴퓨터를 이용한 중학교 학력검사 개발 연구**. 한국교육과정평가원.
- 채선희·유준희·이미경·이선미(2001). **컴퓨터를 이용한 학력검사 개발 연구(II)**. 한국교육과정평가원 연구보고 RRE 2001-1.
- 한국교육과정평가원(2004). **대학수학능력시험 개선방안 연구**. 한국교육과정평가원. PRE-2004-9.
- Anzaldua, R. M. (2002). *Item Banks: What, Where, Why and How*. Paper presented at the annual meeting of Southwest Educational Research Association.
- Booth, J. A. G. (1983). *Computerization of Trade Examinations: A feasibility study*(Tech. Rep. No, B 83-1). Ottawa: Canada Employment and Immigration Commission, Trading Branch.
- Choppin, B. (1979). Testing the Question-The Rasch model and item banking In M. J. Raggett & C. P. Raggett (Eds.), *Assessment and Testing of Reading*. London: Ward Lock Educational.
- Davey, T. & Nering, M. (1998). *Controlling item exposure and maintaining item security*. A

- paper presented at the ETS colloquium, Computer-based testing: Building the foundation for future assessment.
- Gronlund, N. E. (1998). *Assessment of Student Achievement*(6th Ed.). Needham Heights, MA: Allyn and Bacon.
- Lipsey, G. (1974). Overview. In G. Lipsey (Ed.), *Computer-assisted test construction*(pp. 1-29). Englewood Cliffs, NJ: Educational Technology Publications.
- Mead, R. J. (1981). *Basic ideas in item banking*. Paper presented at the annual meeting of National Council on Measurement in Education. Los Angeles. (ED 208 029).
- Millman, J. & Arter, J. A. (1984). Issues in item banking, *Journal of Educational Measurement*, 21(4), 315-330.
- Mills, C. (1999). Development and introduction of a computer adaptive Graduate Record Examinations General Test. In F. Drasgow & J. Olson-Buchman (Eds.) *Innovations in computerized assessment*.
- Mills, C. & Steffen, M. (2000). The GRE computer adaptive test: operational issues. In W. van der Linden & C. Glas (Eds.), *Computerized adaptive testing: theory and practice*.
- Prosser, F. (1974). Item banking, In G. Leppy (Ed.). *Computer-assisted test construction*. Englewood Cliffs, NJ: Educational Technology Publications.
- Rudner, L. (1998). Item banking. *Practical assessment, research & evaluation*, 6(4).
- Umar, J. (1999). Item banking. In G. N. Masters & J. P. Keeves (Eds.), *Advances in Measurement in Educational Research and Assessment*.
- Veldkamp, B. P. & van der Linden, W. J. (2000). Designing item pools for computerized adaptive testing. In W. J. van der Linden & C. Glas (Eds.), *Computerized adaptive testing: theory and practice*.
- Ward, A. W, & Murray-Ward, M. (1994). Guidelines for the development of item banks. *Educational Measurement: Issues and Practice*, 13(1), 34-39.
- Wood, R. (1974). Question Banking. In H. G. Macintosh (Ed), *Techniques and problems of assessment*. London: Edward Arnold, 208-220.
- Wright, B. D. & Bell, S. R. (1984). Item banks: What, Why, How. *Journal of Educational Measurement*, 21(4), 331-345.

• 논문접수 : 2005년 10월 15일 / 수정본 접수 : 2005년 11월 14일 / 게재 승인 : 2005년 11월 28일

ABSTRACT

Construction and Utilization of Item Bank with Consideration of Evaluation Purpose

Kyong-Ah Sang(Research Fellow, Korea Institute of Curriculum & Evaluation)

Myong-Hee Yang(Professor, Kunsan National University)

With the introduction of computer technology and item response theory into testing practices, many item banks have been developed and are being used. This study examines two types of item banks, 'item bank for teaching and learning (IT-TL)' & 'item bank for selection and qualification(IT-SQ)', and discusses how each item bank is constructed and utilized.

The purpose of IT-TL is to facilitate effective teaching-learning process. Thus it is critical that the IT-TL is easy to use with diverse items that is based on the current curriculum. A user-friendly search function is very helpful in this context so that the user draw items, select and review them, then download items to user's computer for assembly. More time and resources should be committed to expanding the size and quality of item bank.

On the other hand, IT-SQ is basically designed to develop a valid and fair test in high-risk evaluation situations such as selection. Thus the descriptive statistics such as difficulty, discrimination, and reliability of test items are very important. IT-SQ provides substantial savings of time and energy over test development as well as security maintenance of test items. It is time to think about introducing item banking system in test situations such as college entrance examination in our country.

This study describes and illustrates two types of item bank. The most crucial step in developing an item bank is planning. This study explains some important points to consider at the planning stage of each item bank.

Key Words : item bank, item bank for teaching and learning, item bank for selection and qualification