

# 공학계열 대학생 물리 기초학력평가 문항분석 및 성취수준 설정 사례연구

이금호\*·정혜경\*\*·†

\*한국기술교육대학교 데이터성과센터 연구교수

\*\*한국기술교육대학교 HRD학과 부교수

## A Case Study on Item Analysis and Standard Setting of the Physics Basic Ability Test for Engineering College Students

Lee, Keumho\*·Jung, Hyekyung\*\*·†

\*Research Professor, Center for Data & Outcomes, Korea University of Technology & Education

\*\*Associate Professor, Department of Human Resource Development, Korea University of Technology & Education

### ABSTRACT

This study is to examine the validity of assessing basic-level proficiency in physics among incoming engineering freshmen through item analysis and standard setting. For empirical analysis, we examined the physics subject taken by the freshman class of 2021 at K University, considering its significance for engineering students. In this study, we initially performed item analysis utilizing both classical test theory and item response theory. Subsequently, leveraging the item and test information, we employed a modified Angoff method and the Bookmark method for standard setting. Consequently, the difficulty level initially set during item development was found to be higher than the actual performance level exhibited by the students. This study highlights a discernible disparity between the expected university standard and the real proficiency level of incoming freshmen in terms of basic academic ability in physics. Based on these research findings, a comprehensive discussion on the fundamental academic competence of engineering students was conducted, underscoring the necessity for formulating a tailored learning approach leveraging the outcomes from the basic ability test.

**Keywords:** Engineering education, College students, Basic ability test, Physics, Standard setting

## 1. 서 론

우리나라 대부분의 대학에서는 대학 교육을 이수할 만한 학습능력을 판단하는 기준으로 고등학교 내신 성적, 대학수학능력평가가 성적 등을 활용하여 대학 신입생을 선발한다. 그러나 동일한 대학에 같은 기준으로 입학하였다고 해도 '신입생의 학습능력이 동일하다'라고는 할 수 없으며, 대학생의 기초학력 미달 문제는 대학 내 중요한 쟁점으로 부각되고 있다. 대학 교육의 대중화와 보편화에 따라 과거와 비교하여 학생들의 기초학력이 저하되거나 대학 수학을 위한 준비가 부족한 학생들의 대학 진학은 결과적으로 학습부진, 학사경고, 중도탈락 등으로 이어지고 있다(정미강·양유정, 2016).

2015 개정교육과정에서 문과와 이과가 통합되고, 고등학교 교과목 선택에 따라 기초 학습의 정도가 학생마다 다르게 되었다. 공학 전공 대학생에게 있어 수학이나 물리 등은 전공교과를 학습하기 전 필수로 이수해야 하는 기초학습에 해당한다(강유진·김지나, 2013). 그런데 7차 교육과정에 따른 선택형 교육과정, 2015 문·이과 통합교육 과정 등은 이공계 신입생의 수학과 기초과학 등 기초학력 저하의 주요 원인으로 제기되고 있으며, 기초학력의 저하가 더 심해질 것이라는 우려를 낳고 있다(노태완, 2008; 이영태·박성석, 2019). 물리 교과 학습은 공학 전공에서의 학업과 매우 밀접하기 때문에 물리학 기초의 부족으로 인한 전공과목 수강 어려움이 나타나고 있다(이종길, 2018). 실제로 어려운 교과목으로 인식되는 물리는 고등학교에서 선택 과목의 이수 비율이 급격히 감소하고 있다(변태진, 2020; 송영욱, 2019). 수능 과학탐구영역에서의 물리 과목 선택자 수가 급감하고 있다는 것에서 그 기피 현상이 심화하고

Received November 1, 2023; Revised November 24, 2023

Accepted November 27, 2023

† Corresponding Author: hjung@koreatech.ac.kr

©2023 Korean Society for Engineering Education. All rights reserved.

있음을 알 수 있다. 2008학년도 수능 응시자 중 물리 I을 선택한 응시자는 109,682명(57.6%), 물리 II는 19,597명(10.3%)이었으나(한국교육과정평가원, 2007.12.7.), 2023학년도 수능에서 물리학 I은 62,309명(14.4%), 물리학 II는 2,628명(6.1%)으로 나타나(한국교육과정평가원, 2022.12.8.), 응시생 수가 급감하였음을 알 수 있다. 이러한 배경 아래 대학 신입생의 기초학력 저하 문제, 특히 공학 전공 학생들의 수학 및 물리 기초학습 부족으로 인한 학습 곤란의 문제를 해결하기 위해 대학들은 다양한 노력을 기울이고 있다. 일부 대학에서는 고등학교 수준의 새로운 교과목을 개설하거나(문공주·남형주, 2018), 신입생 대상 기초학력평가를 실시하여 입학 초기에 적절한 교육적 조치를 취하려 하고 있다.

본 연구는 공학계열 학생들의 기초학력 증진 방안 마련에 앞서 기초학력을 진단하는 기초학력평가 자체의 중요성 및 질 관리 필요성에서 시작되었다. 여러 대학에서 재학생의 기초학력 제고 및 전공교육의 기초 향상을 위해 신입생을 대상으로 기초학력평가를 실시하고 있다. 기초학력평가는 학생들의 기초학력 수준을 파악하고, 수준별 교육을 통해 효과적인 학습성과 달성 도모를 목적으로 한다. 이는 전공교육의 수월성뿐 아니라 대학 학습을 시작하기 전 기초학력을 진단하여 적절한 교육을 제공하여 기초학력이 부족한 학생들에게 일어날 수 있는 학습부진이나 학사경고, 중도탈락 등을 미연에 예방하고자 평가 결과를 활용할 수 있다는 점에서 그 의의와 중요성이 있다. 대부분의 대학에서는 대학 자체적으로 문항을 개발하여 기초학력평가를 실시하고는 있으나, 문항 개발·분석에서의 전문성이나 결과 활용의 연계 측면에서 초·중등교육에서의 기초학력진단에 비해 부족함이 있다.

대학에서 기초학력평가를 실시하는 목적은 신입생의 기초학력 수준을 파악하여 수준이 낮은 학생에게는 보충학습을 제공하여 학습을 지원하기 위함이다. 또한 추적 조사를 통해 대학 학업성취도와와의 관련성을 분석하여 교육과정 개편 및 학생의 학력 증진을 위한 방안 등의 의사결정에 활용할 수 있다. 따라서 매년 다른 문항으로 시행되는 평가에서 문항의 관리의 매우 중요하다. 한편 대부분 대학에서의 기초학력평가는 문항의 난이도나 변별도 등에 대한 문항분석이나 기초학력 도달 여부를 판단하는 기준이 명확하게 설정되어 있지 않다. 대부분 원점수를 기반으로 일정한 점수를 기준으로 그 이상을 넘으면 기초학력이 있다고 판단하고 있는 실정이다. 그러나 매년 다른 문항으로 기초학력평가가 실시되고, 신입생의 기초학력 수준도 해마다 달라서 임의적인 원점수 컷(분할점수)을 기준으로 학생들의 기초학력을 도달 또는 미도달로 결정하는 것은 바람직하지 못하다.

본 연구는 K대학교 공학계열 신입생 대상의 물리 기초학력 평가의 사례를 통해 기초학력평가 문항에 대한 측정학적 분석 결과에 따라 성취수준 설정을 수행함으로써 대학생들의 기초학력 수준 도달 여부를 실증적으로 파악하고 수준별 교육을 제공하기 위한 활용 방안을 모색하는 데 목적이 있다. 성취수준 설정의 최종 목적은 성취수준에 대한 개념적 정의를 기반으로 수준을 결정하기 위한 분할점수를 도출하는 것으로(Cizek & Bunch, 2006), 본 연구에서는 기초학력평가의 결과를 활용하는 데 있어 타당한 기준을 마련하기 위한 것과 관련되어 있다. 본 연구 결과를 바탕으로 대학에서 신입생의 물리 능력에 대한 기초능력에 대해 전문가들의 의견을 수렴하고 물리 기초능력에 대해 실증적으로 파악해 봄으로써 대학에서 물리 교과 평가 데이터에 근거한 활용의 가능성 및 시사점을 제공하는데 연구의 의의가 있다. 나아가 본 연구는 대학교육에서의 평가측정학적 이론에 대한 이해 및 필요 인식이 상대적으로 낮은 현실을 직시하고 교육의 질 제고를 위한 대학생의 평가 결과의 체계적 활용가능성을 강조하고자 대학 교육 관계자들의 이해를 도모하기 위해 기초적인 이론 및 실증적인 증거를 제시하는 부가적인 목적도 포함하고 있다.

## II. 이론적 배경

### 1. 고전검사이론과 문항반응이론

문항분석은 검사의 각 문항이 본래 기능을 제대로 수행하고 있는가를 확인하고 검토해 보는 작업으로, 검사제작과정에서 문항분석을 실시하여 적절한 수준의 양호도를 보이는 문항을 사용하고 그렇지 못한 문항을 수정·보완하게 된다(이중승, 2009; Embretson & Reise, 2000). 그리고 문제은행을 구축하는 경우에 시스템에 적재할 문항을 선별하고 문항분석을 통해 얻어진 문항 정보를 저장하기 위해 활용하게 된다(황정규 외, 2016). 문항분석을 통해 산출되는 문항정보로는 문항의 난이도, 변별도, 추측도 등이 있다. 구체적인 개념을 정리하면, 문항 난이도는 검사문항의 어려운 정도를 나타내며, 변별도는 각 문항이 피험자의 능력을 얼마나 잘 구별해 주는가를 의미한다. 추측도는 피험자가 문항에 대한 지식 없이 추측해서 정답을 맞힐 비율을 의미한다. 문항분석은 고전검사이론과 문항반응이론으로 구분된다.

고전검사이론에 기반한 문항분석의 가장 큰 특징은 피험자의 응답정보를 활용하여 검사의 총점에 근거하여 문항의 특징을 산출하는 것으로, 관찰점수는 진점수와 오차점수의 합으로 정의된다. 이론적으로 진점수는 무한 반복 측정값의 평균이며,

진점수와 오차점수의 상관은 0이라는 가정에 따라 관찰점수의 분산은 진점수와 오차점수 각각의 분산의 합으로 정의할 수 있다(김석우, 2015). 고전검사이론에서 산출되는 문항정보 중 난이도는 전체 시험 응시자 중 정답을 맞힌 응시자 비율로 계산할 수 있다. 변별도는 총점이 낮은 집단과 높은 집단을 얼마나 잘 구별하는가를 나타내는데, 2가지 방법에 의해 산출이 가능하다. 검사 총점을 기준으로 상·하위집단으로 구분하여 두 집단을 얼마나 잘 구별하는가로 산출하거나 문항 정답 여부 즉 문항 점수와 총점 간의 상관을 산출하는 양분상관계수(point-biserial correlation)를 통해 산출 가능하다. 검사총점이 높은 사람들이 정답을 하고 낮은 점수를 얻은 사람들이 오답하는 경향이 강한 문항일수록 양분상관계수로 산출된 변별도는 높게 나타난다. 반면 음수값을 보이면 높은 검사점수를 받은 사람들이 오답을 하고 낮은 점수의 사람들이 정답을 하게 되는 경향이 강한 문항으로 변별력이 없는 좋지 않은 문항이다(이종승, 2009).

고전검사이론에 의한 문항분석은 비교적 계산이 쉬워 현장에서 적용하기 용이하다는 장점이 있다. 반면에 문항 난이도와 변별도는 응시자 집단에 종속되고 응시자 점수가 검사 특성에 종속된다는 점, 측정 오차가 검사 및 문항 난이도, 시험 응시자 능력에 상관없이 동일하다는 비현실적인 가정에 근거한다는 단점이 상존한다(Hambleton et al., 1991; 황정규 외, 2017). 즉, 검사의 난이도에 따라 피험자 능력 추정이 변화된다는 것으로, 쉬운 검사에서는 피험자 능력이 과대 추정되고, 어려운 검사에서는 피험자 능력이 과소 추정된다. 또한 같은 문항에 대해서도 응시자 집단의 능력 수준이 높을 때 문항분석 결과 해당 문항은 쉬운 문항으로 문항 정보가 산출될 것이고, 반대로 낮은 능력 수준 응시자들의 응답 결과에 기반하여 분석하면 해당 문항의 난이도는 높게 산출될 것이다. 변별도의 경우에도 피험자 집단의 특성에 따라 다르게 추정되는 문제점이 있다(성태제, 2009). 이러한 고전검사이론의 단점을 극복하기 위한 방법이 문항반응이론이라고 할 수 있다.

문항반응이론은 피험자가 어떤 문항에 대해 특정한 반응을 할 확률을 문항특성 및 능력의 함수로 표현하는 통계이론이다(van der Linden & Hambleton, 1997; Yen & Fitzpatrick, 2006). 고전검사이론은 검사총점에 의해 검사나 문항 수준에서 분석되기 때문에 응시자 집단의 능력 수준이 달라진다. 반면 문항반응이론은 문항 하나하나를 분석단위로 하여 개별 문항에 대한 응시자의 응답을 통계확률모형으로 표현하는 방식으로, 모형에 따라 추정 모수가 달라질 수 있다. 고전검사이론에 의한 난이도와 변별도는 피험자 집단의 성질에 따라 좌우되는 반면에, 문항반응이론에서는 문항모수치의 불변성과 능력모

수치의 불변성을 가정하게 된다(Hambleton et al., 1991).

문항반응이론에서는 일차원성(unidimensionality)과 지역독립성(local independence)이라는 중요한 두 가지 가정이 있다. 일차원성 가정은 하나의 검사도구가 하나의 특성을 측정한다는 것으로, 검사를 구성하는 모든 문항이 동일한 하나의 특성을 측정하여야 한다는 가정이다. 지역독립성 가정은 어떤 능력을 가진 피험자가 하나의 문항에 대한 응답은 다른 문항의 응답에는 전혀 영향을 주지 않는다는 가정으로, 피험자가 각 문항을 맞힐 확률이 상호 독립적이라는 의미이다(성태제, 2009). 문항반응모형은 문항모수 추정에 따라 난이도만을 산출하는 1모수 모형, 난이도와 변별도를 산출하는 2모수 모형, 난이도, 변별도에 추측도까지 산출하는 3모수 모형으로 구분된다. 문항모수는 문항특성곡선(item characteristics curve: ICC)으로 표현될 수 있으며, 문항난이도 모수는 피험자의 능력 모수와 동일한 척도로 표현된다. 문항반응이론에서의 난이도는 문항의 정답을 맞힐 확률이 .5에 해당하는 능력 수준을 의미하며, 일반적으로  $-2 \sim +2$  사이에 존재하게 된다. 값이 작을수록 쉬운 문항, 그 값이 클수록 어려운 문항으로 해석할 수 있다. 변별도는 문항의 정답확률 .5에서의 상위 능력 피험자와 하위 능력 피험자를 변별하는 정도를 의미한다. 일반적으로 변별도는  $0 \sim +2.0$ 의 범위에 있으며 값이 클수록 변별도가 높다고 해석하는데, 문항특성곡선에서 기울기가 클수록 변별도가 높은 문항으로 해석할 수 있다. 3모수 모형에서 추정되는 추측도는 능력이 전혀 없는 피험자가 문항의 답을 맞힐 확률을 의미하며, 추측도가 높으면 변별력이 떨어지는 문항으로 볼 수 있다.

문항반응이론의 가장 큰 특징은 분석의 단위가 문항으로 문항의 고유 정보를 문항모수 정보로 표현이 가능하며, 따라서 문항특성 불변성과 피험자 능력 불변성을 담보하는 것이다. 고전검사이론과 문항반응이론은 각각 고유한 특성과 장단점이 있다. 따라서 본 연구에서는 두 가지 방법을 이용하여 문항분석을 실시하고, 문항분석의 결과를 활용하여 기초학력 여부를 결정하는 정보로 활용하였다.

## 2. 수준설정

교육평가 방법은 평가 결과의 활용 및 해석적 측면에서 규준 참조평가와 준거참조평가로 구분될 수 있다. 규준참조평가는 학습자가 속해있는 집단 구성원의 학습성취도를 기준으로 하여 상대적인 위치를 비교하여 개인차를 변별하는 평가 방법이다. 준거참조평가는 어떤 절대적인 기준에 비추어 그 기준에 도달여부에 따라 학생의 성적을 평가하는 방법으로, 절대적 기준이란 교육목표 또는 학습성취도를 의미한다(이종승, 2009). 준

거참조평가는 절대적 기준, 즉 준거(standards)에 비추어 학생이 성취한 정도를 나타내주기 때문에 평가 결과가 교육적 피드백 정보로 활용될 수 있다는 장점이 있다(황정규 외, 2017). 학습능력이나 성취도에 대한 평가의 일차적 목적은 학생들이 교육과정에 명시된 학습목표를 어느 정도 성취했는가를 확인하는 것이라 할 수 있다. 따라서 평가 결과가 순위를 매겨 선발 등을 목적으로 하는 것이 아니라 학생들의 학습능력 수준을 파악하고 기초학습능력의 달성 여부를 알아보기 위한 것이 목적이라면 준거참조평가로 실시되는 것이 타당하다. 준거참조평가가 그 목적을 달성하고 그 결과가 제 기능을 하기 위해서는 성취수준(achievement levels)에 대해 타당성이 확보되어야 한다. 이를 위해 체계적이고 과학적인 수준설정(standard settings)의 절차가 뒷받침되어야 한다.

수준설정은 검사 결과에 대해 하나 이상의 분할점수(cut-off score)를 설정하여 피험자의 연속적인 검사 분포를 범주화하는 방법이다(임의진 외, 2019). 분할점수는 피험자의 성취수준을 구분하기 위한 점수를 의미한다. 이 분할점수를 산출하는 과정은 타당성을 확보하고 분류의 오류를 줄이기 위해 논리적이고 체계적이며 합리적인 방법에 따라 시행되어야 한다(임의진 외, 2019). 그 방법론으로 앵고프(Angoff) 방법과 북마크(Bookmark) 방법, 에벨(Ebel) 방법 등 다양하게 제시되었으며, 본 연구에서는 가장 대표적인 방법인 앵고프와 북마크 방법을 적용하였다.

앵고프 방법은 1971년 Angoff에 의해 소개되어 지금까지도 가장 널리 알려진 수준설정 방법으로, 개념적으로 정의된 성취수준에 대해 최소능력자(minimally competent person, MCP; 또는 minimally acceptable person, MAP) 집단을 정하고 이 최소능력자가 각 문항에 대한 정답 여부에 따라 1점 또는 0점을 부여하고 전체 문항의 합산 점수를 분할점수로 정하는 방법이다(김성숙 외, 2014). 이후 앵고프 방법은 수준을 설정하는 패널들이 검사지 문항별 정답을 맞힐 확률을 정하고 그 결과를 합산하여 분할점수를 산출하는 방법으로 변형되었다. 이 과정을 보통 3~5회의 라운드를 거쳐 패널들 간의 합의에 의해 최종 분할점수를 산출한다. 이를 수정된 앵고프(modified Angoff) 방법이라고 한다(Zieky, 2001). 패널들은 여러 차례의 라운드를 거치면서 각 라운드 사이에 다른 패널들이 평정한 분할점수, 성취수준집단별 비율 등의 결과를 제공받게 되며, 반복된 평정 과정에서 평정 결과에 대한 논의 및 합의를 통해 최종 분할점수를 산출하게 된다.

북마크 방법은 문항반응이론에 의해 도출된 문항분석 결과를 활용하게 된다(Skaggs & Tessema, 2001). 문항반응이론에 의해 분석된 문항난이도 순에 따라 문항을 배열한 순서화된 문

항집(ordered item book, OIB)을 활용하여 수준을 설정하게 된다. OIB는 가장 쉬운 문항부터 가장 어려운 문항 순으로 배열하여 구성된다. OIB에 분할지점을 표시(mark)하는 방식으로 분할점수를 설정하는 방식이기 때문에 패널들의 인지 작업을 단순화할 수 있으며, 성취수준이 여러 개일 때 비교적 용이하게 수준설정을 할 수 있다는 장점이 있다(장윤산·성태제, 2009; 임의진 외, 2019). 패널들은 쉬운 문항부터 순서대로 제시된 OIB를 앞에서부터 검토하면서 최소능력자 2/3 이상이 문항을 맞힐 확률 즉 응답확률(response probability, RP)이 .67인 마지막 문항을 보고하게 된다. 여기에 대응하는 척도점수( $\theta$ )를 산출하고 분할점수를 도출하게 된다.

앵고프 방법은 고전검사이론을 바탕으로 하고 있는 반면, 북마크 방법은 문항반응이론을 기반으로 하고 있다. 앵고프 방법은 모든 문항에 대해 최소능력자의 정답확률을 평정해야 하는 반면에 북마크 방법은 분할점수 수만큼만 문항을 선정하기 때문에 절차상 간편하다는 장점이 있다. 다만 북마크 방법은 문항이 적을 때 척도점수의 간격이 넓게 되어 정확한 원점수 산출이 어렵게 될 수 있다는 한계점도 있다.

### III. 연구방법

#### 1. 물리 기초학력평가의 구성

물리 기초학력평가의 문항 출제 범위는 고등학교 물리 교육 과정의 주요 내용으로, 이원분류표 상의 행동요소와 내용요소를 근거로 하여 출제된다. 물리 기초학력평가 문항 개발을 위한 행동요소는 Table 1에 제시된 바와 같이 계산(지식), 이해, 추론, 문제해결로 구분되며, 문항 개발 시 상위 행동요소로 갈수록 문항의 난이도가 올라가도록 하였다. 즉, 계산(지식)은

Table 1 Behavioral objectives in two-dimensional classification of physics basic ability test

분류	설명	난이도
계산 (지식)	<ul style="list-style-type: none"> <li>기본적인 물리 용어, 법칙, 이론 등 학습내용을 회상해서 푸는 문제</li> <li>기본 물리법칙을 계산하는 문제</li> </ul>	하
이해	<ul style="list-style-type: none"> <li>물리법칙을 수식으로 표현하여 설명하는 문제</li> <li>물리수식을 적용하여 실제 현상을 설명하는 문제</li> </ul>	하 또는 중
추론	<ul style="list-style-type: none"> <li>물리법칙 간의 연관성을 인지하여 주어진 현상을 증명하는 문제</li> <li>주어진 현상을 보고 물리법칙으로 표현하는 문제</li> </ul>	중 또는 상
문제 해결	<ul style="list-style-type: none"> <li>두 가지 이상의 물리 개념이나 법칙을 종합하여 푸는 문제</li> <li>실제 현상과 물리 이론의 연관성을 적용하여 푸는 문제</li> </ul>	상

하, 이해는 하 또는 중, 추론은 중 또는 상, 문제해결은 상의 난이도로 분류하여 출제하였다. 내용요소는 물리 교과와 주요 내용인 시공간과 우주, 물질과 전자기장, 정보와 통신, 에너지로 영역을 구분하였다. 난이도 하 5문항, 중 10문항, 상 5문항으로 총 20문항이 개발되었으며, 5지 선다형 객관식 문항으로 구성되었다.

## 2. 분석대상 및 자료수집

본 연구는 K대학교의 2021학년도 신입생 대상 기초학력평가 데이터를 활용하였으며, 2021년 2월 중 온라인 LMS(학습관리시스템)를 통해 실시되었다. 물리 기초학력평가는 공학계열 신입생만을 대상으로 하며, 총 655명이 응시하였다. 총 20문항에 대한 평가 결과를 100점 만점으로 환산하였을 때 평균은 47.3점, 표준편차는 22.2점이었으며, Fig. 1에 제시된 바와 같이 점수분포가 낮은 점수 쪽으로 편포된 분포였다.

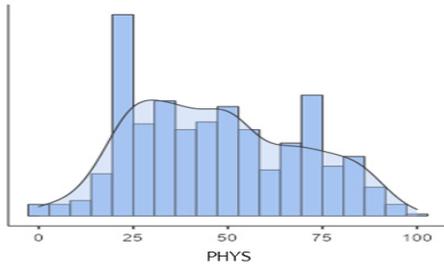


Fig. 1 2021 physics basic ability test score distribution

## 3. 문항분석 및 수준설정 절차

2021학년도 실시된 물리 기초학력평가 20문항에 대해 고전검사이론과 문항반응이론을 각각 적용하여 문항분석을 실시하였다. 고전검사이론에서는 난이도, 변별도, 답지별 반응 비율을 산출하였다. 난이도는 총 응시자 중 정답을 맞춘 응시자의 비율(정답율), 변별도를 알아보기 위해서는 총점과 각 응시자의 문항 점수의 상관계수를 산출하였다. 문항반응이론을 적용한 문항분석에서는 1모수(Rasch)와 2모수를 적용하였으며, 문항 분석 결과를 기반으로 수준설정 워크숍을 진행하였다. 수준설정 워크숍은 다음과 같이 진행되었다.

수준설정을 위한 전문가 패널은 총괄평가위원 1명, 교과평가위원 1명, 교과위원 5명으로 구성하였다. 총괄평가위원은 K대학교의 기초학력평가 연구를 진행한 교육학 박사가 맡았으며, 교과평가위원은 타대학의 평가전공 교수로 수준설정에 대한 이해도가 높은 교육학 전공 교수자 1명이 담당하였다. 교과위원은 K대학교에 교양교육과정에서 물리 교육을 담당하고 있는

전임교원 2명과 공학계열 전공교육과정을 담당하고 있는 전임교원 1명, 전문교육평가기관 연구위원으로서 물리교육을 전공한 박사 1명, 고등학교 물리교사 1명으로 구성하였다. 수준설정 워크숍은 2021년 12월 22일(수) 1일에 걸쳐 진행하였다. 앵고프 방법과 북마크 방법을 모두 적용하였으며, 각 교과위원이 평정한 결과에 대한 라운드별 평균, 중앙값, 표준편차, 최댓값, 최솟값, 도달과 미도달 학생의 비율을 확인할 수 있도록 하였다.

## IV. 분석결과

### 1. 문항분석 결과

#### 가. 고전검사이론 분석 결과

기존 K대학교에서 사용한 물리 기초학력평가에 대해 고전검사이론에 의한 문항분석 결과는 Table 2에 제시된 바와 같다. 음영으로 표시된 부분은 정답지를 의미한다. 물리의 경우 문항 정답률이 0.4 미만 6개 문항, 0.4~0.6 미만 6문항, 0.4~0.6 미만 4문항, 0.6~0.7 미만 8문항으로 나타났다. 0.8 이상의 매우 쉬운 문항은 없는 것으로 파악되었다. 한편, 변별도 측면에

Table 2 Analysis results of the classics test theory

문항	난이도	변별도	답지별 선택 비율 (%)				
			1	2	3	4	5
1	.315	.117	15.69	31.54	20.92	18.77	13.08
2	.605	.235	6.62	10.62	8.77	60.46	13.54
3	.620	.226	4.94	5.40	62.04	16.05	11.57
4	.650	.195	13.47	10.06	8.20	65.02	3.25
5	.601	.235	2.31	8.32	7.24	22.03	60.09
6	.354	.267	6.80	8.96	26.58	22.26	35.39
7	.408	.274	8.04	16.38	11.44	23.34	40.80
8	.390	.174	11.71	38.98	34.51	11.56	3.24
9	.632	.284	3.08	11.09	10.48	12.17	63.17
10	.523	.217	17.34	11.46	8.67	52.32	10.22
11	.626	.210	7.26	7.11	11.90	62.60	11.13
12	.381	.138	15.33	8.82	38.08	12.54	25.23
13	.348	.197	12.98	18.55	34.78	19.78	13.91
14	.608	.105	8.02	6.64	60.80	16.05	8.49
15	.404	.133	9.75	40.40	23.22	19.04	7.59
16	.488	.256	10.53	11.15	15.33	14.24	48.76
17	.413	.194	7.61	13.66	16.77	41.30	20.65
18	.602	.192	5.71	60.19	22.38	9.72	2.01
19	.283	.074	28.26	28.26	15.53	17.55	10.40
20	.318	.182	8.58	10.61	31.83	30.89	18.10

서 0.3 이상인 문항은 없는 것으로 나타났으며, 절반 정도의 문항이 0.2 미만으로 나타나 전반적으로 변별도가 낮은 것으로 확인되었다. 이러한 정보는 향후 대학의 물리 기초학력평가의 출제 교수진들에게 평가도구 개선을 위한 안내 정보로 활용될 수 있다.

나. 문항반응이론 분석 결과

문항반응이론 중 난이도만을 산출하는 1모수(Rasch)와 난이도와 변별도를 산출하는 2모수를 적용하여 물리 기초학력평가에 대한 문항분석을 한 결과 Table 3과 같다. Table 3에는 개발자가 문항 개발 시 평정한 난이도와 내용요소 및 행동요소가 제시되어 있다.

20개 문항 중 2모수 기준 난이도가 1이상인 경우 1번, 8번, 13번, 19번으로 Rasch에서는 19번이 가장 어려운 문항으로, 2모수에서는 8번이 가장 어려운 문항으로 추정되었다. 1번, 8번, 13번, 19번 문항에 대한 1모수와 2모수 분석 결과의 문항 특성곡선을 Fig. 2에 제시하였다. 2모수 추정과 1모수 모형에서 난이도 추정에 다소 차이가 나는 것은 2모수 모형의 변별도에 부분적으로 영향을 받을 수 있으며, 특히 8번(등속원운동, 에너지)의 경우 변별도가 전체 20문항 중에서 변별도가 가장 낮은 문항으로 파악되어 난이도는 높으나 변별력이 떨어지는 문항으로 나타났다.

Table 3에는 개발 당시 난이도가 제시되어 있는데, 개발 교수가 개발 당시 설정한 난이도와 실제 학생들의 수행 결과에서의 난이도가 차이를 보이는 문항이 발견되었다. 문항 1번은 개발 당시 난이도를 ‘하’로 설정하였으나, 고전검사이론에서의 난이도는 .315로 31% 정도의 학생만이 문제의 정답을 맞혔다. 문항반응이론에서의 난이도는 1모수는 .94, 2모수는 1.03으로 그 값이 높아 어려운 난이도의 문제로 볼 수 있다. 2~5번 문항도 개발 난이도는 ‘하’에 해당되는 문항이며, 정답률을 나타내는 고전검사이론에서의 난이도 결과를 보면 60~65% 정도의 학생이 정답을 맞혔다. 해당 문항들에 대한 문항반응이론에 의한 문항분석 결과, 2모수에서는 -.49 ~ -.82, 1모수에서는 -.51 ~ -.77의 난이도를 나타내었다. 개발 난이도 ‘중’에 해당되는 문항 6~15번 중 9번, 10번, 11번, 14번은 정답률이 50% 이상으로 나타났으며, 나머지 문항은 개발 당시 난이도보다 어려운 문항에 해당되는 것으로 문항분석결과 나타났다. 반면 개발 난이도 ‘상’에 해당되는 문항은 18번을 제외하고는 실제 난이도와 유사했다. 해당 문항에 대한 고전검사이론에 의한 난이도는 .318~.488, 1모수 .5~1.12, 2모수 .01~1.05의 난이도를 나타냈다. 특히 물리 기초학력평가에서 가장 어려웠던 문항은 19번으로, 고전검사이론 난이도는 .283으로 약 28%만이 정답

Table 3 Analysis results of the item response theory

문항	개발 난이도	내용 요소	행동 요소	2모수		
				1모수 난이도	변별도	난이도
1	하	전하와 전기장	계산	.94	.86	1.03
2	하	빛의 에너지	지식	-.54	1.14	-.49
3	하	전자기파의 특성	지식	-.61	.96	-.62
4	하	전자전이와 선스펙트럼	지식	-.77	.90	-.82
5	하	힘과 운동	지식	-.51	.80	-.59
6	중	소리의 성질	이해	.73	1.44	.56
7	중	광전효과	이해	.45	1.27	.37
8	중	등속원운동, 에너지	이해	.54	.29	1.56
9	중	등가속도운동	이해	-.67	1.62	-.51
10	중	에너지피이론	이해	-.13	.79	-.14
11	중	열역학법칙	이해	-.64	.91	-.67
12	중	전자기유도	이해	.59	.63	.83
13	중	운동량과 충격량	이해	.77	.63	1.09
14	중	파동의 성질	이해	-.55	1.18	-.49
15	중	역학적 평형	이해	.47	1.25	.39
16	상	직선전류에 의한 자기장	문제해결	.05	1.87	.01
17	상	전반사	문제해결	.42	.67	.57
18	상	운동량과 충격량	문제해결	-.52	1.46	-.41
19	상	역학적 에너지보존	문제해결	1.12	1.07	1.05
20	상	전류에 의한 자기장	문제해결	.92	.98	.92

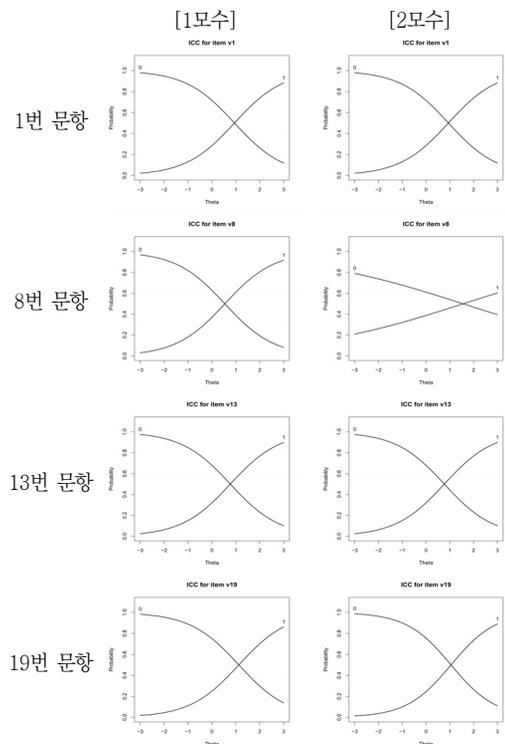


Fig. 2 Item characteristics curve of 4 items

을 맞췄다. 해당 문항은 역학적 에너지보존의 내용요소와 문제 해결 행동요소에 해당된다.

## 2. 수준설정 결과

### 가. 물리 기초학력개념 정의 및 최소능력자 특성

최소능력자에 대한 특성을 기술하기 위해 대학생의 기초학력의 일반적인 정의인 “대학교육에서 요구되는 기초적인 학습역량”과 K대학교 물리 교과 기초학력의 정의인 “대학에서의 전공기초 과목을 학습하는 데 필요한 물리 지식과 이를 활용하여 계산, 이해, 추론, 문제해결을 할 수 있는 능력”을 제시하였다. 각 교과위원은 대학생의 기초학력 정의를 논의하고, 대학에서의 물리 교육, 특히 공학계열에서의 물리 교육 필요성뿐 아니라 고등학교 물리 교육의 현황에 대한 논의를 통해 공학 전공 신입생에게 필요한 물리 교육 및 수준을 파악하고 K대학교에서 공학 전공을 학습하는 데 필요한 물리 기초학력의 수준을 다음과 같이 정의하였다.

- K대학교 물리 성취 수준 : 고등학교 교육과정(물리학 I)을 이수하여 역학, 전자기학, 파동 및 현대물리 등 기본 물리량과 물리법칙들을 이해하고, 성취도 수준 ‘중’을 도달할 수 있다.
- 상: 문제를 이해하고 정량적으로 풀 수 있다.
- 중: 문제를 이해하는 정도/물리법칙을 이해하고 도출할 수 있으나 때때로 문제를 풀지 못함(정형화된 간단한 수식이 필요한 경우 문제를 풀지만, 복잡하게 서술되거나, 복잡, 어려운 그래프에서 정보를 찾지 못해 풀지 못함).
- 하: 문제를 이해하지 못하는 수준

수준설정 교과위원들은 위와 같은 정의를 도출한 후 물리 기초학력평가에서의 기초학력 도달 여부를 판단하기 위한 최소능력자에 대한 특성을 다음과 같이 도출하여 기술하였다.

- 물리법칙을 이해하고 도출할 수 있으나 문제를 풀지 못함
- 기본 물리 개념 및 법칙을 이해하고 있으나 실제 적용 및 문제해결에 어려움이 있는 정도
- 문제에 적용된 기본적인 물리법칙을 찾을 수 있으나 수학적 해결책을 찾기 어려워하는 학생(기본적인 개념 및 수식을 알고 있지만 문제해결에 필요한 개념 및 수식을 판단할 수 없음)

### 나. 수정된 앵고프 방법 적용 결과

수정된 앵고프 방법 적용을 위해 교과위원들에게 실제 학생

들에게 제공된 20개 문항의 순서대로 되어 있는 문항집과 문항별 기대정답률을 기입할 수 있는 작성 시트를 제공하였다. 문항집에는 물리 기초학력평가 문항 개발을 위한 내용요소와 행동요소를 제시하여 각 문항의 특성을 파악할 수 있도록 하였다. 총 3라운드가 진행되는 동안 교과위원들은 문항별 기대정답률 시트를 활용하여 20개 문항이 차례대로 제시된 문항집을 검토하면서 최소능력자가 해당 문항을 맞힐 확률을 기입하였다. 5명의 교과위원이 기대정답률 시트에 작성하고 나면 진행요원은 해당 내용을 미리 준비된 엑셀에 입력하였다. 엑셀에는 각 교과위원의 문항별 평정결과를 입력하고, 평정결과와 문항별 평균값, 중앙값, 표준편차, 최댓값, 최솟값, 분할점수, 2021학년도 물리 기초학력평가 응시자 실제 점수 대비 도달 및 미도달 비율을 자동 계산되도록 준비하였다. 각 라운드마다 교과위원들은 다른 교과위원이 평정한 결과와 분할점수, 도달 및 미도달 비율을 확인하고 각자의 평정 이유와 기준을 설명하며 의견을 조율하는 논의과정을 거쳤다. 총괄평가위원과 교과평가위원이 이 과정을 진행하였으며, 이러한 과정을 거치면서 도달과 미도달을 구분 짓는 최소능력자에 대한 개념을 명세화하고, 궁극적으로 준거참조평가 기반 성취수준(도달/미도달)에 대한 조작적 정의를 패널과의 토론을 통해 확정 짓게 된다.

라운드별 분할점수를 원점수와 100점 만점 환산점수일 때의 중앙값과 평균값을 기준으로 살펴보았으며, 분할점수를 2021학년도 물리 기초학력평가 응시자 654명의 실제 점수에 대비했을 때의 도달과 미도달의 비율을 검토하였다. 그 결과는 Table 4와 같다.

Table 4 Standard setting results applying the Angoff method

구분	분할점수		도달/미도달 비율 _명(%)		
	원점수	100점 만점	도달	미도달	
중앙값	R1	8.4	42	355 (45.7%)	299 (54.3%)
	R2	6.6	33	460 (59.2%)	194 (40.8%)
	R3	7.0	35	460 (59.2%)	194 (40.8%)
평균값	R1	7.6	38	400 (51.5%)	254 (48.5%)
	R2	7.1	36	400 (51.5%)	254 (48.5%)
	R3	7.1	36	400 (51.5%)	254 (48.5%)

1라운드에서는 중앙값 기준 분할점수는 8.4점, 평균값은 7.6점으로 나타났다. 이때의 100점 만점 환산점수는 각각 42점과 38점이었다. 2라운드에서는 중앙값 기준 원점수 6.6점, 평균값 기준 7.1점이었으며, 100점 만점 환산점수는 중앙값은 33점, 평균값은 36점이었다. 3라운드에서는 중앙값 기준 7.0점, 평균값 기준 7.1점이었으며, 100점 만점 환산점수는 각각 35점과

36점이었다. 2021학년도 물리 기초학력평가 응시자 654명의 실제 점수에 대비했을 때 1라운드에서 중앙값을 활용하면 도달 355명(45.7%), 미도달 299명(54.3%)으로 나타났다. 평균값을 기준으로 했을 경우, 분할점수가 낮아져 도달하는 학생의 비율이 중앙값보다 높아졌다. 2라운드와 3라운드는 중앙값은 달랐으나 평균값은 동일하였으며, 도달과 미도달 학생의 수가 동일하게 나타났다. 중앙값을 기준으로 했을 경우에는 460명(59.2%)이 도달, 194명(40.8%)이 미도달이었다. 평균값이 기준일 때는 도달 400명(51.5%), 미도달 254명(48.5%)으로 3번의 라운드가 동일하였다.

다. 북마크 방법 적용 결과

앵고프 방법을 적용한 수준설정을 마친 후, OIB를 배부하고 북마크 방법을 적용한 수준설정을 진행하였다. 교과위원들은 쉬운 문항부터 어려운 문항 순으로 배열된 문항들을 살펴본 후, 최소능력자가 풀 수 있는 마지막 문항에 북마크를 하였다. 북마크 방법을 적용한 수준설정은 총 2라운드로 진행되었다. 1라운드에서 교과위원 3명은 8번 문항에 북마크를 했으며, 논의 과정을 거쳐 2라운드를 진행한 결과, 8번 문항으로 모든 교과위원이 북마크 하였다. 북마크 방법 적용 결과의 분할점수 및 도달과 미도달 비율은 Table 5와 같다.

Table 5 Bookmark method standard setting results for each round

구분	R1		R2		
	문항번호	척도점수	문항번호	척도점수	
교과위원1	8	-.07	8	-.07	
교과위원2	8	-.07	8	-.07	
교과위원3	8	-.07	8	-.07	
교과위원4	5	-.14	8	-.07	
교과위원5	11	.07	8	-.07	
기술 통계치	평균	8	-.06	8	-.07
	중앙값	8	-.07	8	-.07
	표준편차	2.12	.08	0	.00
	최댓값	11	.07	8	-.07
	최솟값	5	-.14	8	-.07

북마크 방법은 적은 문항 수인 검사에서 척도점수 간격이 넓어지기 때문에 정확한 원점수를 산출하는데 어려움이 있을 수 있다. 따라서 Table 6에는 원점수를 해당 척도점수에 가까운 원점수로 표시하였다. 1라운드에서 중앙값을 적용하면 원점수는 10점이었고, 기초학력 미도달 학생은 449명(57.8%)이었다. 2라운드에서도 동일한 결과가 나타났다. 평균값을 적용한 경우

에 1라운드에서 원점수는 9점이었으며, 2라운드 결과는 중앙값의 결과와 동일하였다.

Table 6 Standard setting results applying the Bookmark method

구분	분할점수		도달/미도달 비율_명(%)		
	척도점수	원점수	도달	미도달	
중앙값	R1	-.07	10	328 (42.2%)	449 (57.8%)
	R2	-.07	10	328 (42.2%)	449 (57.8%)
평균값	R1	-.06	9	325 (41.8%)	452 (58.2%)
	R2	-.07	10	328 (42.2%)	449 (57.8%)

앵고프 방법과 북마크 방법의 원점수와 도달 및 미도달 비율을 비교해 보면, 북마크 방법에서의 원점수가 더 높고 미도달 비율이 더 많게 나타났다. 각 방법의 마지막 라운드 결과를 보면, 앵고프 방법에서 중앙값 기준 원점수는 7점, 도달은 460명(59.2%)인 반면, 북마크 방법은 원점수 10점, 도달은 328명(42.2%)이었다.

V. 결론 및 논의

본 연구에서는 공학계열 신입생 대상 물리 기초학력평가 문항을 분석하고, 문항분석 결과에 기반하여 성취수준을 설정하였다. 이를 통해 대학에서 실시되고 있는 기초학력평가 실시 및 결과 활용을 위한 타당하고 실증적인 증거를 제시하는 데 주요 목적이 있다. 이를 위해 2021학년도 K대학교에서 실시된 물리 기초학력평가 20개 문항에 대해 고전검사이론과 문항반응이론을 적용하여 문항분석을 실시하였다. 고전검사이론과 문항반응이론에 의한 문항분석 결과, 난이도 측면에서는 1번, 8번, 13번, 19번이 공통으로 어려운 문항으로 나타났다. 변별도 분석 결과에서는 2가지 검사이론 결과 모두 8번 문항이 변별도가 떨어지는 문항으로 나타났으나, 고전검사이론에서는 물리 기초학력평가 문항 대부분이 변별도가 낮게 나타난 반면, 문항반응이론에서는 일부 문항을 제외하고는 대체로 양호하였다. 2가지 검사이론은 이론적 가정과 분석방법에서 차이가 있다. 고전검사이론은 검사총점에 근거하여 분석하기 때문에 응시자 집단의 특성에 따라 문항 고유 특성이 달라질 가능성이 있어 정확성이 결여될 수 있다는 단점이 있다. 반면 문항반응이론은 문항을 분석단위로 하므로 상황에 따라 문항의 특징이나 피험자의 능력점수가 변하지 않는 장점이 있다(Hambleton & Swaminathan, 1985). 고전검사이론이 분석방법이 쉽고 간단한 반면 문항반응이론을 이해하고 실제 분석에 적용하기에는 어려움이 있으나, 최근 범용 소프트웨어를 활용하여 결과산출이 용이하여 대학을 포함한 교육기관에서의 적용활용이 확

대될 것으로 기대된다.

개발 당시 작성된 난이도와 실제 학생들의 수행 결과에 대해 검사이론을 적용하여 문항분석한 결과를 비교하였을 때 난이도에서의 차이가 발견되었다. 개발 당시 매뉴얼에 따라 문항 개발 교수자가 평정한 난이도보다 실제 학생들이 수행한 결과의 난이도가 더 높게 나타났다. 즉, 물리 기초학력평가가 어려웠다는 것으로, 변별도 분석에서도 8번 문항의 경우 고난이도 이면서 변별력이 가장 낮아 기초학력진단을 위한 문항으로 적절한가를 검토할 필요가 있었다. 개발 당시 교수자의 난이도가 실제 학생들의 수행 난이도보다 높아 물리 기초학력평가는 어려운 시험으로 볼 수 있다. 그 원인 중 하나는 실제 학생들의 수행 수준과 대학에서 물리 교과를 담당하는 교수자의 학생 수준에 대한 판단의 차이에서 비롯된 것이다. 고등학교 물리 교육과정의 변화를 간과하고 문제를 출제할 때도 한 가지 원인으로 볼 수 있다. 2015 개정 교육과정 개편에 따라 물리2 교과를 배우지 않은 학생들이 있어 여기에 해당하는 내용의 문제를 못 푸는 학생이 많았다. 고등학교 교육과정을 고려해 볼 때 물리 기초학력평가의 경우 난이도를 조절할 필요가 있으며, 대학에서 기초학력평가 문항을 개발할 때 고등학교 교육과정에 대한 이해가 선행될 필요가 있다.

물리 기초학력 수준 도달 여부를 판단하는 타당하고 신뢰로운 기준을 제시하기 위해 수준설정을 위해 전문가로 구성된 수준설정 패널 집단과 워크숍을 진행하였다. 수정된 앵고프 방법을 적용하였을 때, 최종 분할점수는 20개 문항에 대해 원점수는 약 7점, 100점 만점 기준으로는 35~36점 정도에 해당하였다. 2021년 기준 K대학교 기초학력평가 대상자 중 기초학력 미도달은 중앙값 기준으로는 40.8%에 해당하였으며, 평균값을 기준으로 하면 48.5%의 학생이 해당하였다. 북마크를 적용하였을 때는 그 결과가 상이하였는데, 원점수 기준 분할점수가 10점으로 미도달 비율은 57.8%에 해당하였다. IRT 기반의 북마크 방법은 문항분석을 통해 산출된 문항의 난이도에 따라 쉬운 문항부터 어려운 문항 순으로 제시하고 최소능력자가 풀 수 있는 문항의 위치를 표시하도록 한다. 이때 척도점수를 활용하기 때문에 문항 수가 많아야 한다. 특히 도달과 미도달 2개로 수준을 구분하는 경우에는 난이도가 쉬운 문항이 많아야 하는데 그렇지 못한 경우에 분할점수와 비율이 달라질 수 있다. 반면 앵고프 방법은 각 문항에 대해 정답확률을 추정하기 때문에 분할점수가 문항 수에 영향을 크게 받지 않게 된다. 본 연구 결과에서 2가지 방법을 적용했을 때 원점수의 차이가 발생한 원인은 여기에 있다. 즉 본 연구에서 사용된 물리 기초학력평가는 20개의 문항으로 척도점수의 간격이 넓어 분할점수 산출에 있어 한계가 있었다. 이러한 연구 결과로

볼 때, K대학교와 같이 적은 문항으로 기초학력평가를 실시하는 경우에는 수준설정 방법으로 앵고프 방법을 적용하는 것이 더 적절한 것으로 판단된다. 물리 기초학력평가에 북마크 방법을 적용하기 위해서는 문항 수를 늘리고 난이도를 현재보다 낮게 출제해야 한다.

대학의 교육성과는 학생들에게 어떠한 학습을 제공하며 그 결과로 어떤 것을 성취했는가로 평가된다. 대학의 평가지표에는 학사경고자에 대한 조치가 포함되어 있어 각 대학에서는 학사경고자에 대한 다양한 사후 방안을 마련하고 중도 탈락률을 낮추기 위한 노력을 시도하고 있다(박덕희, 2017). 그러나 학습부진이 발생한 이후 사후 처방도 중요하겠지만, 사전에 이를 방지하는 방안 마련이 필요하다. 특히 공학계열은 수학 및 물리 등의 기초학습능력이 이후 전공학습에도 영향을 미치기 때문에 기초학력평가를 실시하고 그 결과에 따라 적절한 학습지원을 해주는 것이 선행되어야 할 것이다. 이를 위해서는 기초학력평가 도구 자체의 개발 및 질 관리가 담보되어야 하며, 이를 관리할 수 있는 인적·물적 지원이 확보되어야 한다. 이에 본 연구 결과를 바탕으로 대학교 신입생 대상 기초학력평가를 학생과 대학에 실제적으로 도움이 될 수 있도록 운영하기 위한 방안을 제안하면 다음과 같다.

첫째, 지속적인 문항 관리 및 분석을 위한 체계가 마련되어야 한다. K대학교에서는 매년 기초학력평가 문항을 자체적으로 개발하고 있으며 2020년 교내 자체 연구를 통해 문항 개발 매뉴얼을 작성하였다. 그러나 개발된 문항에 대해 평가 이후 문항분석 등 측정학적인 검토가 필요하다. 개발된 문항이 양호한 문항인지를 검토하고, 장기적인 관점에서 문제은행을 구축해 볼 수 있다. 문항의 개발과 문항분석, 수준설정 등은 전문성이 요구되는 분야이므로 지속적인 문항 관리 및 분석 체계의 마련을 위한 전문 기관과 연구원이 필요하다. 많은 대학이 데이터 기반 교육의 질 관리 체계를 마련하고 기관 연구(Institutional Research, IR)를 수행하고 있는데(배상훈·윤수경, 2016; 황영식, 2023), 기초학력평가도 데이터 및 통계 기반의 과학적인 의사결정 체계를 마련하고 그 정보를 대학 전체가 공유하는 정보의 환류로 이어지도록 구성할 필요가 있다.

둘째, 기초학력평가를 실시한 후 그 결과를 활용하기 위한 구체적인 방안 마련이 필요하다. 본 연구에서는 이를 위해 기초학력 수준에 대한 타당한 기준을 마련하기 위해 수준설정을 실시하였다. 본 연구에서의 결과 적용과 기초학력평가 결과의 활용에 대해서는 정책적인 의사결정이 필요하다. 특히 물리 기초학력평가에 대한 문항분석 결과, 상당히 어려운 수준의 도구로 나타났는데 이는 대학에서 요구되는 기초 물리 학습이 고등학교에서 이루어지지 않았다는 것에서 원인을 찾아볼 수 있다.

또한 공학계열 대학에서는 물리나 수학의 기초학력이 중요함에도 불구하고 고등학교 교육과정과 수능의 변화 등으로 인한 대학에서 요구되는 기초학습의 부족함을 해결할 방안 마련이 시급하다(노태완, 2008; 정미강 외, 2016). 기초학력의 저하는 이후 전공학습의 어려움으로 이어지므로 기초학력 미달 학생들에 대해 맞춤형 학습을 제공할 수 있도록 해야 한다(표용수·박준식, 2010).

셋째, 학생의 학습능력에 대한 정보인 기초학력평가 도구의 신뢰도와 타당도를 확보하고 이를 바탕으로 기초학습능력 수준을 측정하여 데이터 기반 교육의 질 관리 체계를 마련할 필요가 있다. 대학에서 교육연구의 혁신과 데이터 기반 교육성과 분석을 위해서는 학생들의 입학정보와 함께 기초학력평가 데이터의 관리-분석-피드백에 대한 환류 체계가 마련될 필요가 있다. 그리고 이 환류 체계는 맞춤형 학습 플랫폼 시스템에 활용될 수 있다. 문항분석, 내용요소별 하위점수분석, 학생의 능력분석, 전공별 분석, 입학년도별 추이 분석, 전공 성적과의 관련성 분석 등 다양한 연구 분석을 통해 학생들의 부족한 부분에 대한 선제적 대응과 우수 학생에 대한 교육의 수월성을 위한 지원을 제공하는 데이터 기반 맞춤형 교육 시스템을 개발하는 데 활용될 필요가 있다(이진숙 외, 2021; Montebello, 2018).

본 연구는 대학에서 보유한 기초학력평가 2차 데이터를 활용하여 교육측정·평가의 기본적인 이론을 바탕으로 검사 문항 분석 및 기준설정을 실제로 적용해 봄으로써 특정 대학 신입생의 물리 교과에 대한 기초학력 도달 정도를 실증적으로 분석해 보는 사례연구로, 2021년 1개년 데이터에 국한하여 분석, 논의된 한계점이 있으며 일회성의 검사 도구 및 응시자 집단 분석 결과를 바탕으로 일반화하기에는 어려움이 있다. 이에 향후 매년 신입생의 수준을 파악하고 연도별 비교해 볼 필요가 있으며, 이를 위해서는 매년 문항의 난이도가 유사한 수준으로 유지되어야 하는데, 본 연구 결과에서 나타난 것처럼 문항 개발자의 의도와 학생들의 실제 능력이 다를 수 있다. 따라서 출제 교수진 및 관계자들이 고등학교 교육과정에 대한 지식과 더불어 검사도구 제작 및 문항·검사 분석 결과에 대한 이해 또한 수반되어야 하며, 대학별 IR센터, 데이터센터, 성과인증센터 등에서 교육측정·평가 전문가 인력을 확보 또한 뒷받침되어야 할 것이다.

본 연구는 2021년 한국기술교육대학교 학부교육 질 제고 연구의 일환으로 수행된 “기초학력평가 문항분석 및 성취수준 설정”의 내용 일부를 발췌·수정·보완하였음

## 참고문헌

1. 강유진·김지나(2013). 물리학습에서의 인지적 신념과 동기 신념에 대한 공과대학 학생의 인식과 교수자의 기대 비교. *공학교육연구* 16(2), 50-57.
2. 김석우(2015). *교육평가의 이해*. 서울: 학지사.
3. 노태완(2008). 고등학교 제7차 교육과정을 이수한 공과대학 신입생의 수학 및 과학 과목에 대한 기초학력 평가 및 분석. 2008 공학교육학술대회 및 공과대학장협의회 워크샵 발표자료집, 431-434.
4. 문공주·남형주(2018). 고교 물리 선택 이수와 입시전형에 따른 공학계열 대학 신입생의 물리 학업성취도 비교-2017년 C 대학교 공과대학 신입생을 중심으로-. *학습지중심교과교육연구* 18(8), 155-169.
5. 박덕희(2017). 학업 저성취 대학생을 위한 일대일 학습공동체 프로그램 효과 분석. *학습지중심교과교육연구* 17, 375-402.
6. 배상훈·윤수경(2016). 한국대학에서 대학기관연구(Institutional Research) 도입 관련 쟁점과 시사점. *아시아교육연구* 17(2), 367-395.
7. 변태진(2020). 고교 과목 이수 및 수능 과목 선택과 대학 일반물리학 학업성취도와 관계 분석 연구: 지방대 공과대학 사례를 중심으로. *새물리* 70(4), 364-373.
8. 성태제(2009). *문항반응이론의 이해와 적용*. 서울: 교육과학사
9. 송영욱(2019). 대학 물리학 수업이 공과대학 학생들의 물리 이미지 및 인식에 미치는 영향. *새물리* 69(6), 637-648.
10. 이영태·박성석(2019). 공학교육의 기초학력 강화를 위한 AP (Advanced Placement)제도의 도입 및 활용방안, E-AP로 공학교육의 기초를 다지자. *공학교육동향* 26(3), 18-22.
11. 이종승(2009). *현대교육평가*. 서울: 교육과학사
12. 이진숙 외(2021). 대학의 AI 기반 맞춤형 강의 추천 시스템 개발 및 실제 적용 사례 연구: K 대학을 중심으로. *교육공학연구* 37(2), 267-307.
13. 임의진·전의성·이용원(2019). 유럽 언어 공통기준 연계를 통한 영어 독해 시험의 수준설정: 수정된 앙고프 방법과 북마크 방법 비교. *교육평가연구* 32(3), 523-548.
14. 장운선·성태제(2009). 문항반응이론에 기초한 준거설정 방법 비교. *교육평가연구* 22(3), 659-680.
15. 정미강·양유정(2016). 대학생의 기초학습능력 지원을 위한 교수-학습 참여자의 인식과 요구. *교육문화연구* 22(2), 101-126.
16. 표용수·박준식(2010). 대학수학 기초학력 부진학생을 위한 기초수학 지도 방안. *수학교육 논문집* 24(3), 525-541.
17. 한국교육과정평가원(2007.12.7). 2007학년도 대학수학능력시험 채점결과 보도자료.
18. 한국교육과정평가원(2022.12.8). 2023학년도 대학수학능력시험 채점 결과 보도자료.

19. 황영식(2023). 기관분석(IR)에 기반을 둔 한·미 대학정보공시 시스템 비교연구. *학습자중심교과교육연구* 23(7), 233-253.

20. 황정규 외(2016). *교육평가의 이해*. 서울: 학지사.

21. Cizek, G. J., & Bunch, M. B.(2006). *Standard setting: A guide to establishing and evaluating performance standards on tests*. California: Sage Publications.

22. Embretson, S. E., & Reise, S. P.(2000). *Item response theory for psychologists*. New Jersey: L. Erlbaum Associates.

23. Hambleton, R. K., & Swaminthan, H.(1985). *Item response theory: Principles and application*. Massachusetts: Kluwer-Nijhoff.

24. Hambleton, R. K., Swaminthan, H., & Jane Rogers, H. (1991). *Fundamentals of item response theory*. California: Sage Publications.

25. Montebello, M.(2018). *AI injected e-learning*. Cham, Switzerland: Springer International Publishing.

26. Skaggs, G., & Tessema, A.(2001). *Item disorderliness with the bookmark standard setting procedures*. Paper presented at the 2001 annual meeting of the National Council on Measurement in Education.

27. van der Linden, W. J., & Hambleton, R. K. (Eds.)(1997). *Handbook of modern item response theory*. New York: Springer-Verlag.

28. Yen, W. M., & Fitzpatrick, A. R.(2006). Item response theory. In R.L. Brennan (Ed.), *Educational measurement* (4th ed., pp.111-153). Westport, CT: Praeger Publishers.

29. Zieky, M. J.(2001). So much has changed: How the setting of cutscores has evolved since the 1980s. In G. J. Cizek (Ed.), *Setting performance standards: concepts, methods, and perspectives*(pp. 19-51). Mahwah, NJ: Lawrence Erlbaum Associates.



**이금호 (Lee, Keumho)**

2005년: 충남대학교 심리학과 학사  
 2010년: 동 대학원 교육학과 석사  
 2016년: 동 대학원 교육학과 박사  
 2020년~현재: 한국기술교육대학교 데이터성과센터 연구교수  
 관심분야: 검사개발, 교육평가, 고등교육  
 E-mail: april0401@koreatech.ac.kr



**정혜경 (Jung, Hyekyung)**

2000년: 연세대학교 교육학 학사  
 2002년: 동 대학원 교육학 석사  
 2010년: 미국 UCLA 교육학 박사  
 2020년~현재: 한국기술교육대학교 HRD학과 부교수  
 관심분야: 인과효과 추정, 기계학습 분석  
 E-mail: hjung@koreatech.ac.kr