

## 라쉬 모델을 사용한 본초학 시험의 학업역량 분석 연구

채한<sup>1†</sup>, 이수진<sup>2†</sup>, 한창호<sup>3</sup>, 조영일<sup>4</sup>, 김형우<sup>1\*</sup>

<sup>1</sup>부산대학교 한의학과, <sup>2</sup>경성대학교 심리학과, <sup>3</sup>동국대학교 한의학과 내과학교실, <sup>4</sup>동국대학교 경찰행정학부

### Study on the Academic Competency Assessment of Herbology Test using Rasch Model

Han Chae<sup>1†</sup>, Soo Jin Lee<sup>2†</sup>, Chang-ho Han<sup>3</sup>, Young Il Cho<sup>4</sup>, Hyungwoo Kim<sup>1\*</sup>

<sup>1</sup>School of Korean Medicine, Pusan National University

<sup>2</sup>Department of Psychology, Kyungsung University

<sup>3</sup>Department of Internal Medicine, College of Korean Medicine, Dongguk University

<sup>4</sup>Division of Police Administration, College of Law and Police, Dongguk University

**Objectives:** There should be an objective analysis on the academic competency for incorporating Computer-based Test (CBT) in the education of traditional Korean medicine (TKM). However, the Item Response Theory (IRT) for analyzing latent competency has not been introduced for its difficulty in calculation, interpretation and utilization.

**Methods:** The current study analyzed responses of 390 students of 8 years to the herbology test with 14 items by utilizing Rasch model, and the characteristics of test and items were evaluated by using characteristic curve, information curve, difficulty, academic competency, and test score. The academic competency of the students across gender and years were presented with scale characteristic curve, Kernel density map, and Wright map, and examined based on T-test and ANOVA.

**Results:** The estimated item, test, and ability parameters based on Rasch model provided reliable information on academic competency, and organized insights on students, test and items not available with test score calculated by the summation of item scores. The test showed acceptable validity for analyzing academic competency, but some of items revealed difficulty parameters to be modified with Wright map. The gender difference was not distinctive, however the differences between test years were obvious with Kernel density map.

**Conclusion:** The current study analyzed the responses in the herbology test for measuring academic competency in the education of TKM using Rasch model, and structured analysis for competency-based Teaching in the e-learning era was suggested. It would provide the foundation for the learning analytics essential for self-directed learning and competency adaptive learning in TKM.

**Key Words** : Academic competency, Competency-based Test (CBT), Evidence-based Teaching (EBT), Herbology, Learning analytics, Rasch model

· Received : 31 March 2022      · Accepted : 6 May 2022

† H Chae and SJ Lee contributed equally to this work

· Correspondence to : Hyungwoo Kim

School of Korean Medicine, Pusan National University

49, Busandaehak-ro, Mulgeum-eup, Yangsan-si, Gyeongsangnam-do, 50610, South Korea

Tel : +82-51-510-8425, Email : kronos7@pusan.ac.kr

## 서론

컴퓨터 시험(Computer-Based Test, CBT)은 기존의 종이에 출력된 시험(Paper-Based Test, PBT)을 대신해서 컴퓨터나 태블릿과 같은 IT 기기를 도구로 사용하는 시험으로, 컴퓨터 및 정보통신기술(ICT)의 발전에 따라 급속도로 확산되고 있다<sup>1)</sup>. 기존의 PBT에 비하여 공정한 시험과 관리, 채점이 용이하면서, 피험자의 잠재된 역량을 효율적으로 측정할 수 있으며, 다양한 멀티미디어 자료를 사용할 수 있으면서도, 시간과 장소의 제약을 극복할 수 있다는 장점을 지닌다. 이에, 각종 어학 인증시험, 자격증, 수행 평가 및 수학능력 평가에 있어서 문제은행에 기반한 컴퓨터 적응형 검사(Computerized Adaptive Test, CAT)가 적극적으로 도입되고 있다<sup>1-3)</sup>.

의학교육 및 보건의료인 국가시험에 있어서, CBT는 해외의 의사와 간호사 자격시험에 이미 도입되었으며<sup>4,5)</sup>, 국내에서는 의사(2022년)를 시작으로 한의사와 치과의사(2023년)에도 도입될 예정이다<sup>6)</sup>. 보건의료인의 대학 교육에서부터 전문의 시험까지 다양한 분야에서 CBT의 도입과 활용에 대한 기반 확립 연구들이 지속적으로 진행되어 왔으며, COVID-19의 유행에 따라 각종 평가에 이러닝과 CBT의 도입이 폭발적으로 증가하였다.

그러나, 한의학 교육에 있어서는 대학 및 국가시험에서의 CBT 사용을 위한 시험 및 문항 분석에 대한 기초적인 연구가 제시되지 못하고 있는 실정이다. 이는 CBT 및 CAT의 개발과 운용에 기초가 되는 문항 반응이론(Item Response Theory, IRT)의 모형과 용어들이(Table 1) 비전공자들에게 상대적으로 생소하고 모형을 활용한 검사의 평가 절차 및 결과의 해석이 상대적으로 어렵기 때문이다<sup>7,8)</sup>.

오랫동안 시험의 기본적인 방법으로 사용되어 온 PBT의 통계적 기초를 제공한 전통적인 고전검사이론(Classical Test Theory, CTT)은, 함께 시험을 본 동료와 피험자의 점수(맞은 문항의 개수)를 토대로 학업

역량을 분석하고, 이를 기초로 검사 및 시험의 신뢰도 및 각종 타당도 지표를 제시한다<sup>9)</sup>. 이러한 지표들은 이해나 계산이 용이하여 오랫동안 사용되어 왔으나, 함께 시험을 본 동료와의 비교를 기준으로 한 상대평가에 기초하기 때문에, 피험자의 학업역량이나 능력 평가뿐만 아니라 문항의 특성이 시험에 참여한 집단에 따라서 달라진다는 한계를 지니고 있다. 예를 들어, 10문항으로 이루어진 시험에서 피험자가 6문제를 맞추었을 때, 동료들이 5문제를 맞추었다면 피험자의 학업 성취도 또는 학업역량이 높은 것으로 평가되지만, 만약 동료들이 7문제를 맞추었다면 피험자의 학업역량이 낮은 것으로 바뀌게 된다<sup>10)</sup>.

이와 달리, CBT에 활용되는 IRT는 시험을 구성하는 문항의 난이도와 함께 시험을 보는 동료의 점수들과는 독립적인 피험자의 고유하면서 변하지 않는 잠재 특성(latent trait) 또는 역량을 추정한다. 덧붙여서, 문항이나 동료가 달라져도 변하지 않는 학생 고유의 학업역량(academic competency)을 제시함에 있어서 계산에는 로지스틱 확률 함수를, 표현에는 특성곡선(characteristic curve)을 사용한다<sup>9,10)</sup> (Table 1).

IRT를 사용한 CBT는 문제은행 및 최신 ICT를 활용하여 교수자의 부담을 감소시킬 뿐만 아니라, 학생들의 불필요한 학업 부담을 줄여준다는 장점을 동시에 지니고 있다<sup>9)</sup>. 오랫동안 시험 장면에서 활용되어 온 CTT에 기초한 PBT는 학생들의 상대적인 채점점수를 기준으로 한 줄 세우기라는 상대평가에 주요하게 활용되기 때문에, 상대평가로부터 시작된 무한경쟁에 따른 높은 학업 스트레스는 의과대학생의 절반 이상에서 학업소진(academic burnout)을 경험하도록 한다<sup>11)</sup>. 그러나, IRT를 사용한 CBT는 의약계열 학생들의 학업역량에 대한 효율적인 절대평가로서<sup>9)</sup>, 불필요한 학업 스트레스는 줄이면서도 적절한 학업역량과 시간적 여유를 동시에 확보할 수 있도록 한다.

더군다나, 역량중심 교육(Competency-based Education, CBE)에서의 ‘역량’이 ‘특정 상황이나 직무를 효과적이고 우수하게 수행하기 위한 개개인의 내재적인 고

유 특성'으로 정의된다는 것을 고려한다면<sup>12-14)</sup>, 본 연구에서 활용하고자 하는 IRT는 개인의 고유하게 내재된 잠재 특성을 분석하는 통계 모형이다. 따라서 잠재 특성의 추정에 활용되는 IRT는 한의학 교육학에 시급히 적용되어야 할 근거기반 교육<sup>15-17)</sup>의 필수적인 연구 방법론이라고 볼 수 있다.

학업역량의 정확한 평가는 효과적인 교수법의 개발, 활용 및 평가에 필수적인 토대가 된다<sup>14,18)</sup>. IRT를 사용한 실시간 학업분석(learning analytics)은 차세대 교수법이 지향하는 학생들의 학업 성취도 및 개인별 학습 특성을 고려한 맞춤형(personalized), 학생들의 학습 속도나 학업 역량에 따라 계속적으로 변화해가는 적응형(adaptive) 교육의 토대가 된다. 따라서 IRT에 기초한 '평가 및 역량에 따른 맞춤형 피드백'은 새로운 의학 교육법인 플립드 러닝(flipped learning)과 객관구조화 진료시험(Objective Structured Clinical Examination, OSCE)의 핵심 과정이라고 할 수 있다<sup>10,19)</sup>. 시험은 교육의 결과로서 향상된 학업역량을 평가하는 과정인 동시에 자기주도 학습을 위한 새로운 학업의 시작이므로, 교육 내용과 교수법을 논하기에 앞서 시험과 평가에 대한 탄탄한 토대가 선행되어야 한다<sup>19-21)</sup>.

이에, 본 연구에서는 한의학 임상검사의 개발<sup>8,22-24)</sup> 및 문항분석(item analysis)에 IRT를 적용하였던 연구 성과<sup>10)</sup>를 토대로, 한의학 분야의 학업역량을 분석하기 위하여 본초학 기말고사에 사용된 객관식 문항을 포함한 시험에 대한 시험분석(test analysis)을 처음으로 진행하였다. 본 연구에 사용된 라쉬(Rasch) 모델은 Georg Rasch(1960)가 제시한 일반화된 1-모수 IRT 분석법으로서<sup>25,26)</sup>, 문항 모수(parameter)에 있어서 변별도(discrimination)를 1로 고정한 상태에서 문항의 난이도(difficulty)만을 추정하고, 이를 기준으로 피험자의 잠재 역량을 추정한다. 라쉬 모델은 계산이 비교적 용이하고 피험자나 분석 데이터의 크기가 적은 경우에도 추정 신뢰도를 유지할 수 있기에, 심리학, 교육학을 비롯한 다양한 보건의료 영역에서 지식,

심리, 행동, 정서, 능력, 임상 술기 등의 평가에 폭넓게 활용되고 있다<sup>8,25,27)</sup>.

본 연구에서는 본초학 시험 및 문항의 난이도 특성과 재학생의 학업역량을 특성곡선(characteristic curve)과 정보곡선(information curve)으로 제시하였다(Table 1)<sup>10)</sup>. 이와 함께, 본초학 학업역량의 시행연도별, 성별 특성을 스케일 특성곡선(scale characteristic curve), 커널 밀도지도(Kernel density map)와 라이트 지도(Wright map) 등을 사용하여 직관적으로 제시하였다<sup>8,25)</sup>. 또한, 성별, 시행연도별 채점점수와 학업역량의 차이를 T-test와 ANOVA를 사용하여 확인하였다.

라쉬 모델을 사용한 학업역량 분석은<sup>1,10,25)</sup>, 원하는 수준의 난이도 문항과 적절한 수준의 학업역량을 측정하는 시험을 개발하고 개선할 수 있기에, 학생 역량 맞춤형 수업과 평가를 손쉽게 기획하고 운영할 수 있도록 한다. 이와 함께, 비대면 시험에 있어서 객관식 문항의 유출 여부도 수치와 그림으로 확인할 수 있기에 평가 공정성을 보다 용이하게 유지할 수 있으며, 학생 학업역량의 시행연도별 변화를 통해 개정된 입시 전형과 교육 과정의 효과를 객관적으로 비교할 수 있다. 덧붙여서, 학업 역량 및 점수가 남아 및 연령대와 같은 특정 그룹에서 어떠한 차이를 보이는지 직관적으로 분석할 수 있도록 한다.

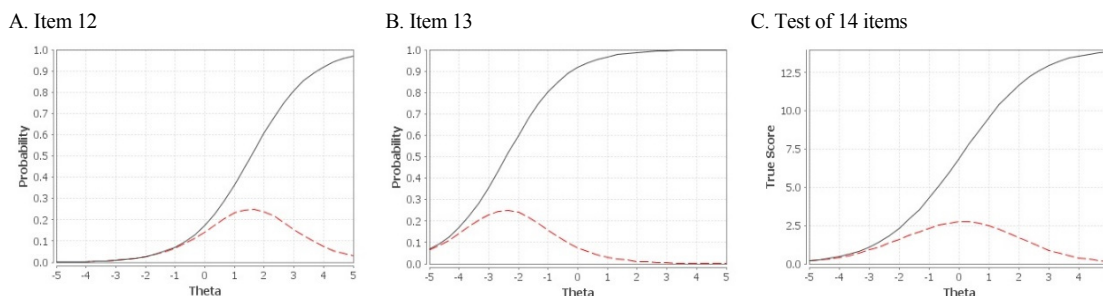
본 연구에 사용된 체계적인 시험분석은 한의학 교육에는 처음으로 활용되는 것으로 ICT를 사용한 근거기반교육(Evidence-based Teaching, EBT)에 기초 자료를 제공할 것이다. 또한, 후속 연구를 통해 한의사 면허 국가시험, 전문의 시험 및 한의학 기초종합평가와 임상 술기시험의 효율적 개발과 운영, 개선을 위한 한의학 교육학, 한의학교육 측정평가학의 토대를 마련할 수 있을 것이다<sup>14,18)</sup>.

## 연구방법

### 1. 연구 대상

**Table 1.** Description and Example of Item Response Theory Terms Used in Current Study

Term	Description	Example in current study
Subject	The entity with unique and latent ability or trait to be tested with clinical examination or academic evaluation.	390 students. Y axis of Figure 3 as a density. Left side of the Figure 4 as a density.
Ability, Theta ( $\theta$ ) or Competency	Unique and latent ability ( $\theta$ ) of a subject rated as a continuous variable with the range of -4 and +4. The academic competency is a major interest of Item Response Theory (IRT) in current study, and it can be used for the absolute evaluation of academic achievement.	Table 2 and Figure 2. X axis of Figure 1, 2 and 3.
$P(\theta)$ and Characteristic Curve	$P(\theta)$ is a logistic probability function made with item difficulty, discrimination and guessing parameters for estimating latent ability. The Characteristic Curve is an illustrative figure for intuitive understanding of $P(\theta)$ in IRT.	$P(\theta)$ for drawing ICC (Figure 1-A) of item and TCC (Figure 1-C) of test.
True Score or Expected Value	The score representing the latent ability of subject which is estimated using $P(\theta)$ .	Y axis of Figure 1-C. Y axis of Figure 2.
Test Score	The sum of item scores as for the relative evaluation and Classic Test Theory (CTT).	Table 2
Item	The basic unit consist of question and answer(s) for competency measure or test score.	14 items used in herbology test
Item Characteristic Curve (ICC)	A curve representing the probability ( $P(\theta)$ ) of getting correct answer corresponding to subject's ability in specific item.	Solid black line of Figure 1-A
Difficulty ( $\beta$ )	It is used for describing how difficulty an item is to achieve 0.5 probability of correct response at a given ability. The difficulty parameter of IRT is negatively correlated with that of CTT.	black marks on the right side of Figure 4.
Discrimination ( $\alpha$ )	It is the slope of the ICC at the point of 0.5 probability which measures the differential capability of an item. The Rasch model uses 1.0 as for the discrimination parameter.	Fixed as 1 in Rasch model
Item Information Curve (IIC)	It shows the amount of information yielded by the item at specific ability level. It can also be used for measuring the reliability of item.	Dashed red line of Figure 1-A
Test	Collected body of items for analyzing academic competency	Herbology test consist of 14 items
Test Characteristic Curve (TCC)	A curve representing estimated test score corresponding to subject's ability.	Solid black line of Figure 1-C
Test Information Curve (TIC)	It shows the amount of information yielded by the test at specific ability level, and it can also be used for measuring the reliability of test.	Dashed red line of Figure 1-C
Visualization of test analysis	A figure presenting results of IRT analysis for intuitively understanding.	
Scale Characteristic Curve	A figure showing estimated score corresponding to specific level of ability. It is useful for comparing estimated score corresponding to specific ability in several groups.	Figure 2
Kernel Density Map	A figure showing density of population corresponding to a specific ability. It is useful for intuitive understanding on distribution of subjects across ability levels.	Figure 3
Wright Map or item-person Map	A figure with two parts to illustrate whether the difficulty of item is proper for analyzing the ability of subjects; left box is for the distribution of subjects, and right box for difficulty parameter of items	Figure 4



**Fig. 1.** Characteristic Curve and Information Curve for (A and B) Items and (C) Test. The solid black line represents Characteristic Curve and the dashed red line the Information Curve of items and test.

본 연구에서는 한의학과 재학생(n=390명)의 본초학 시험(14문항)에 대한 문항별 응답과 채점 점수, 재학생의 성별 그리고 시험 시행연도(2011~2018)를 분석에 활용하였다. 본 연구는 사전에 기관 연구윤리위원회의 승인(2021\_148\_HR)을 받았으며, 피험자를 확인할 수 있는 인적 정보는 분석에 포함하지 않았다.

## 2. 문항과 시험 및 학업역량의 추정

### 1) 문항과 시험에서의 특성곡선

14개 문항들의 문항특성곡선(Item Characteristic Curve, ICC)과 난이도의 추정에는 라쉬(Rasch model) 모델이 사용되었고, 모수의 추정에는 결합 최대우도 추정(Joint Maximum Likelihood Estimation)이 활용되었다<sup>8)</sup>. 14 문항들의 문항특성곡선(ICC)은 역량별로 답을 맞출 확률을 나타내는 로지스틱 함수 곡선으로(Table 1), 각 문항에 대한 전체 피험자들의 응답치를 사용하여 계산하였다(Figure 1-A and 1-B).

IRT 문항 난이도( $\beta$ )는 해당 문항의 답을 맞출 확률로서 ICC상 피험자가 정답을 맞힐 확률이 0.5인 지점에서의 역량( $\theta$ )을 의미하는데, CTT에서의 난이도와는 부적상관을 지닌다<sup>10)</sup>. IRT 문항 변별도( $\alpha$ )는 학생의 역량을 구별해내는 능력으로서 ICC상 문항 난이도에 해당하는 점에서의 순간 변화율로, 본 연구에서 사용된 라쉬 모델에서는 이를 1로 고정한다<sup>26,28)</sup>.

시험(14개 문항으로 구성)에서의 채점점수(test score)

는 문항의 개별 점수를 모두 합한 것으로, CTT에서 검사 점수로 주요하게 활용된다<sup>19)</sup>. IRT로 추정하는 개개인의 진점수(true score)는 시험에서의 학업 역량 또는 학업 성취도에 따라 기대되는 점수로, 시험에 포함된 문항들의 문항특성곡선(ICC)을 모두 더한 검사 특성곡선(Test Characteristic Curve, TCC)을 사용하여 추정한다(Figure 1-C, Table 1).

### 2) 문항과 시험에서의 정보곡선

14개 문항들의 문항정보곡선(Item Information Curve, IIC)은 각 문항들의 역량 수준별 정확도로서<sup>8)</sup>, IRT에서의 문항 모수들을 사용하여 계산한다(Figure 1-A and 1-B, Table 1). 시험(14개 문항으로 구성)의 검사정보곡선(Test Information Curve, TIC)은 포함된 문항들의 문항정보곡선(ICC)을 모두 더한 값으로, 시험 또는 검사의 신뢰도 또는 각 역량수준별로 제공되는 정보의 양을 의미한다(Figure 1-C, Table 1).

자격증 시험과 같이 특정 점수를 기준으로 합격과 불합격을 결정하는 검사에서는 경계에 해당되는 역량 수준에서 정보의 양이 최대가 되어야 하며, 성격검사와 같이 비교를 위한 기준 참조검사에서는 넓은 역량 범위에서 고르게 높은 정보의 양이 유지되어야 한다<sup>19)</sup>.

### 3) 학생들의 학업역량과 문항의 난이도

학생들의 역량수준을 반영하는 학업역량과 채점점

수 그리고 난이도에 대한 분석에는 스케일 특성곡선, 커널 밀도지도와 라이트 지도 등이 사용되었으며, 이들에 대한 설명은 Table 1과 같다<sup>8,25)</sup>.

스케일 특성곡선(scale characteristic curve)은 각 문항 및 시험(14문항으로 구성)에서 학업역량에 따른 기대 점수를 추정하는 곡선으로<sup>8)</sup>, 가우시안 평활화(gaussian smoothing, 대역폭 1)를 사용하여 추정된 함수의 곡면을 매끄럽게 한다. 본 연구에서는 지면의 제약으로 인하여 대표적인 2개의 문항과 시험에 대한 스케일 특성곡선을 제시하였다.

커널 밀도지도(Kernel density map)는 다양한 역량 수준별 피험자의 빈도에 대한 확률 밀도함수를 추정하는 비모수 통계법으로<sup>8)</sup>, 본 연구에서는 성별 또는 시행연도에 따라 특정 수준의 역량을 지닌 피험자들의 빈도가 달라지는 것을 직관적으로 제시하기 위하여 사용되었다. 본 연구에서 추정된 밀도함수의 곡면을 매끄럽게 함에 가우시안 평활화가 사용되었다.

학생들의 본초학 채점점수와 학업역량의 성별, 시행연도별 그룹간 유의한 차이가 확인되었는데, 그룹간 차이를 기술 통계량으로 표현하기 위해서는 평균과 표준오차가 사용되었으며, 그룹내 역량수준별 피험자 분포를 직관적으로 이해하기 위해서는 커널 밀도지도가 사용되었다.

라이트 지도(Wright map) 또는 문항-피험자 지도(item-person map)는 하나의 동일한 로그 지표를 가진 그림에 문항의 난이도(오른쪽)와 피험자의 역량(왼쪽)을 동시에 제시하는 것이다<sup>25)</sup>. 라이트 지도는 피험자의 역량을 측정하기 위한 적절한 난이도를 지닌 문항들이 포함되어 있는지 확인하기 위하여 사용되었다.

### 3. 통계 분석

문항 모수로서의 문항 난이도와 피험자의 학업 역량 또는 진점수의 추정에는 라쉬 모델과 결합 최대우도 추정(Joint Maximum Likelihood Estimation)이 사용되었으며, 문항과 시험의 특성곡선(CC)과 정보곡선(IC)으로 제시하였다.

성별 및 시험 연도에 따른 채점점수 및 학업역량의 유의한 차이를 분석함에 있어서는 T-test와 ANOVA를 사용하였으며, 기술통계에는 평균과 표준오차가 사용되었다. ANOVA 분석에 있어서, 분산의 동질성을 Levene's Test를 사용하여 확인한 결과에 따라 사후 검정에 Bonferroni 또는 Dunnett's T3를 사용하였다. 성별, 시행연도별 추정 역량 수준별 피험자의 분포는 커널 밀도지도로 제시되었으며, 난이도별 문항과 추정 역량에서의 피험자 분포의 성별에 따른 차이는 라이트 지도로 제시되었다.

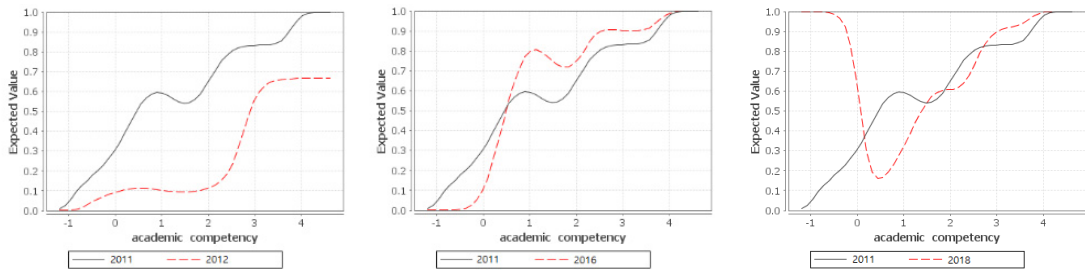
문항, 시험 및 학업역량의 모수 추정에는 jMetrik 4.1.1(Meyer JP, Charlottesville, VA)<sup>8)</sup>을 사용하였으며, 채점점수 및 학업역량의 성별 및 시행연도별 차이를 분석함에 있어서는 jamovi 2.2.1.(The jamovi project, <https://www.jamovi.org>)을 사용하였다. 통계 분석의 결과는 평균±표준오차 또는 빈도(%)로 표기하였으며, 통계적 유의성을 확인하기 위한 통계적 유의수준으로는  $p<0.05$ ,  $p<0.01$  및  $p<0.001$ 이 사용되었다.

## 결 과

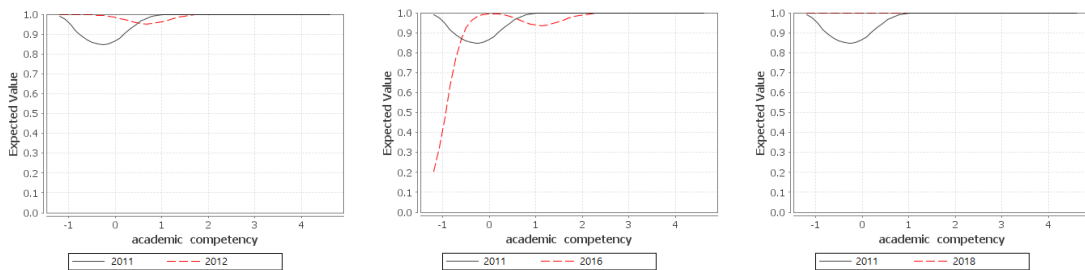
### 1. 문항과 시험에서의 특성곡선과 정보곡선

본초학 시험(14문항)에 대한 8년에 걸친 한의학과 재학생(n=390명)의 응답이 분석되었다. 14개 문항들의 난이도, 문항 특성곡선과 문항 정보곡선, 그리고 시험(14개 문항으로 구성)의 검사 특성곡선과 검사 정보곡선이 확인되었다. 다만 지면의 제한을 고려하여 대표적인 두개 문항의 특성을 제시하였는데, 높은 난이도의 12번(Figure 1-A)과 낮은 난이도의 13번(Figure 1-B) 문항만을 선택하여 문항 특성곡선과 정보곡선을 확인하였으며, 시험(14문항으로 구성된)의 검사 특성곡선과 검사 정보곡선(Figure 1-C) 또한 함께 확인하였다.

A. Item 12



B. Item 13



C. Test of 14 items

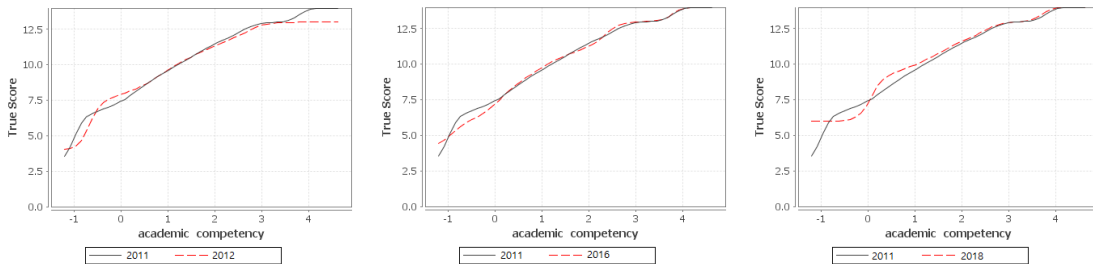


Fig. 2. Scale Characteristic Curve for (A) Item 12, (B) Item 13 and (C) Test in Years of 2012, 2016 and 2018 Compared to Those of 2011.

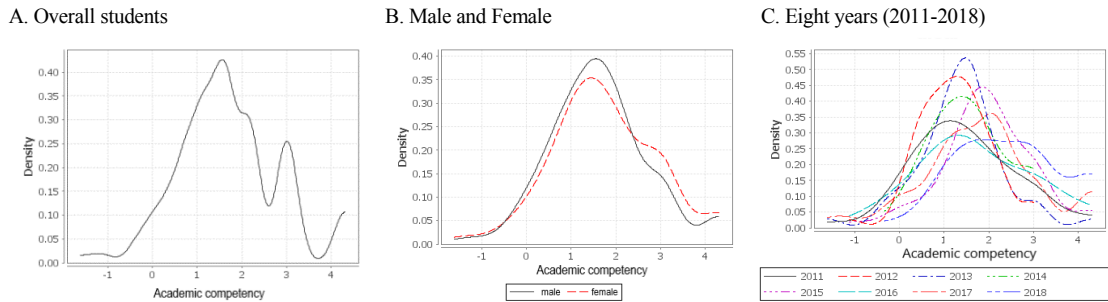
The solid black line is for specific test year of 2012, 2016 and 2018 and the dashed red line for the year 2011 as a reference.

2. 문항과 시험에 있어서 시행연도별 스케일 특성 곡선

처음 문항이 출제되었던 2011년을 기준으로 2012, 2016, 2018년에서의 2 문항(12번 및 13번, Figure 2-A and 2-B)과 시험(14문항으로 구성, Figure 2-C)의 스케일 특성곡선을 비교하였다. 분석 결과, 12번 및 13번 문항에서 시행연도에 따른 특성곡선의 큰 차이를 확인할 수는 있었으나, 시험(14개의 문항으로 구

성)의 특성곡선은 시행연도와 무관하게 일정한 것을 확인할 수 있었다.

이러한 결과는, 학업역량의 분석에 있어서 한 두개 문항만을 분석해서는 그 효용성과 타당도를 판단할 수 없으며, 다수의 문항으로 이루어진 시험으로 평가되어야 함을 의미한다. 또한, 12번 문항의 2011, 2012, 2016, 2018년도 스케일 특성곡선은 학업역량의 증가에 비례하여 기대점수도 함께 증가하는 프로파일을



**Fig. 3.** The Kernel Density Map of Representing Prevalence of Student in Response to Academic Competency in (A) Overall Students, (B) Male and Female Students, and (C) Eight (2011–2018) Years.

보였는데, 이는 해당문항이 학생들에게 유출되지 않았음을 의미한다.

### 3. 성별 및 시행연도별 채점점수 및 학업역량의 분포와 차이

전체학생(n=390), 남자(n=196)와 여자(n=194), 및 시행연도(2011~2018)에 있어서 학생들의 학업역량 (Table 2 및 Figure 3) 및 채점점수(Table 2)의 특성을 확인하였다(Table 1).

채점점수와 학업역량에 있어서 성별 및 시행연도에

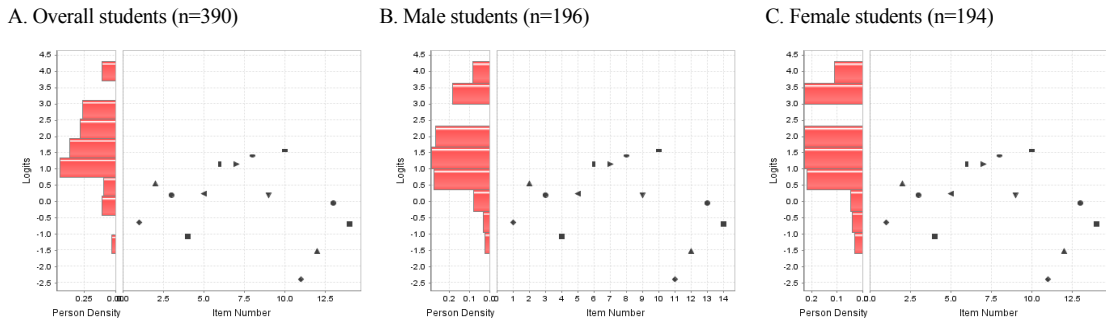
따른 차이를 T-test와 ANOVA를 사용하여 확인한 결과(Table 2), 성별에는 유의한 차이가 없었다. 그러나, 시행연도에 있어서는 유의한 차이를 확인할 수 있었는데, 2018년의 학업역량과 채점점수는 2011, 2012, 2013, 2014년보다 유의하게 높은 것을 확인할 수 있었으며, 2015년 또한 2012년보다 유의하게 높았다. 흥미롭게도 채점점수에서 확인된 2018년과 2016년의 차이가 학업역량에 있어서는 유의하지 않았는데, 이는 학업역량이 채점점수와는 다른 정보를 제공한다는 것을 의미한다.

**Table 2.** Test Score and Academic Competency According to Sex and Test Year

	Test Score (CTT)	Statistics	Academic competency (IRT)	Statistics
Total (n=390)	10.72±0.11		1.71±0.06	
Sex				
Male (n=196)	10.61±0.15	T=0.909, p=0.341	1.64±0.08	T=1.424, p=0.233
Female (n=194)	10.82±0.16		1.78±0.09	
Year				
2011 (n=44)	10.2±0.35	F=5.029, p<0.001 (2018>2011, 2012, 2013, 2014, 2016. 2015>2012)	1.46±0.18	F=6.376, p<0.001 (2018>2011, 2012, 2013, 2014. 2015>2012)
2012 (n=52)	9.98±0.25		1.25±0.11	
2013 (n=47)	10.15±0.3		1.36±0.14	
2014 (n=51)	10.61±0.24		1.57±0.12	
2015 (n=52)	11.29±0.27		1.97±0.15	
2016 (n=47)	10.64±0.35		1.73±0.19	
2017 (n=48)	10.81±0.36		1.81±0.19	
2018 (n=49)	11.98±0.25		2.49±0.17	

\* Bold represents significant difference found only in the Test Score.





**Fig. 4.** The Wright Map Showing Person Density in Relation to the Item Difficulty in (A) Overall, (B) Male and (C) Female Students.

The red box on the left is for person density and the black marks on the right for item difficulty.

2018년도 학업역량과 채점점수가 다른 시행연도보다 높게 나타난 이유는 학생의 분포가 학업역량 1.5~3에서 두텁게 모여 있고(Figure 3-C), 2018년도 스케일 특성곡선에서 낮은 학업역량의 기대점수가 높게 나타난 것(Figure 2-A)에서 기인한 것으로 보인다.

학업역량에 있어서 성별 및 시행연도에 따른 학생 분포 프로파일은 커널 밀도지도를 사용하여 분석하였다(Figure 3). CTT를 사용한 분석에서는 0~14점 사이의 11개 채점점수에서 불연속적 빈도만을 확인할 수 있지만, IRT를 사용한 분석은 -1.61에서 4.31에 이르는 연속적인 학업역량에서의 상대적 분포를 직관적으로 제시하였다. 학업역량에 따른 남학생과 여학생의 분포(Figure 3-B)는 유사하였으나, 학업역량별 학생분포는 시행연도에 따른 차이를 직관적으로 이해하기 쉽게 제시하였다(Figure 3-C).

#### 4. 학업역량 수준별 학생 분포와 난이도별 문항 분포

전체 학생 및 남학생과 여학생에 있어서, 학업역량 수준별 학생의 분포와 문항의 난이도 분포는 라이트 지도를 사용하여 제시되었다(Figure 4). 분석 결과, 남녀 성별에 따른 차이는 확인할 수 없었지만, 학생들의 학업역량 수준에 비하여 문항의 난이도가 비교적 낮은 것을 객관적으로 확인할 수 있었다.

이는, 낮은 수준의 학업역량을 지닌 학생에서의 높은 변별력을 목적으로 문항을 출제하였거나, 문항 출제 과정에서 재학생들의 학업역량을 너무 낮게 예상하였거나, 쉬운 객관식 문항을 사용하여 과락(총점 60점 이하)을 예방하는 방편으로 사용하였을 것임을 보여준다. 이와 함께, 시험을 개정할 때에는 11번, 12번 문항의 난이도가 -1.5보다 높게 수정되어야 함이 제시되었다.

### 고찰 및 결론

본 연구에서는, 라쉬 모델을 사용하여 본초학 교육 현장에서 8년간 활용되었던 필기시험에 대한 검사분석을 시행하였다. 문항과 시험의 특성곡선과 정보곡선, 학업역량과 채점점수, 그리고 스케일 특성곡선, 커널 밀도지도와 라이트 지도 등을 사용하여 학생들의 성별 및 응시연도에 따른 학업역량을 구체적으로 제시할 수 있었으며, 시험과 문항의 장단점과 개선 방안을 체계적으로 확인할 수 있었다.

첫째, 본 연구에서는 한의학계에서는 처음으로 IRT를 소개하면서 문항과 시험의 특성을 상세하게 분석하였다. 기존의 CTT를 활용한 경우에는 문항과 채점점수가 주는 단 2개(맞으면 1점, 틀리면 0점, 불연속 변수)의 정보만을 얻을 수 있지만<sup>2,10)</sup>, IRT를 사용한

경우에는 시험 난이도와 피험자 역량을 모든 수준의 역량(-4~+4, 연속형 변수)에서 거의 무한대의 정보를 확인할 수 있다(Figure 1-C, Table 1).

문항 및 시험의 스케일 특성곡선(Figure 2)을 통해 확인된 학업역량별 예상 점수는 문항과 시험의 개발과 수정에 토대와 가이드로 사용될 수 있다. 본 연구에서 문항의 스케일 특성곡선(Figure 2-A)이 매년 달라졌음에도 불구하고 시험의 스케일 특성곡선(Figure 2-C)은 일정하게 유지되었다. 이는, 본 연구에서 사용된 본초학 시험 일부 문항에 개선이 필요하지만 그럼에도 불구하고 시험은 적절한 신뢰성을 갖고 있었음을 확인하는 것으로, 시험의 타당성은 단순히 1-2개의 문항이 아니라 문항이 모인 시험으로 평가되어야 함을 재확인하는 것이다.

이와 함께, 스케일 특성곡선에서 예상점수가 학업역량에 비례하여 증가하고 있었기에 시험문항이 유출되는 부정행위는 없었음을 추정할 수 있었다. 문항이 유출되지 않은 문항에서는 학생들의 역량이 높아짐에 따라 예상 시험 점수도 정적 상관성을 보이면서 증가할 것(Figure 2-A 및 2-C)이다. 이에 반하여, 만약 문항이 유출되었다면, 학생들의 학업 역량이 높거나, 낮거나 또는 중간입에 따라 달라지지 않는 일정하게 높은 점수를 보였을 것이며, 이에 학업 역량과 예상 점수 사이에 유의하지 않는 또는 부적 상관성을 보였을 것이다.

둘째, 본 연구에 사용된 커널 밀도지도와 라이트 지도는 한의학 교육분석에 처음 적용된 것으로, 연령이나 응시연도와 같은 특정 그룹의 학업역량 특성을 체계적이고 직관적으로 제시한다<sup>8)</sup>. 본 연구에서 성별에 따른 학업역량의 차이는 확인할 수 없었으나 시행연도에 따른 변화는 확인할 수 있었다(Table 2 and Figure 3). 이와 같은 그룹간 역량의 차이는 채점점수(Table 2) 만으로도 확인할 수는 있으나, 그 이유에 대한 설명과 이해는 커널 밀도지도(Figure 3)로만 제시할 수 있었다. 이는 커널 밀도지도가 학업역량별 빈도를 사용하여 역량을 직관적으로 표현할 수 있기 때

문이다.

남학생과 여학생 사이에 유의한 역량 차이는 없었으나 학업역량 2.5 이상에서 여학생의 빈도가 약간 높은 것(Figure 3-B)을 알 수 있었는데, 이것이 통상적으로 여학생의 성적이 더 높았다고 생각되던 이유로 사료된다. 또한, 타 응시연도에 비하여 2018년도의 평균 점수가 높았던 것은(Table 2), 학생들의 학업역량이 1.0 정도 타 연도보다 높을 뿐만 아니라 학업역량 1.5~3.0 구간에 학생들이 폭넓게 분포하기 때문인 것으로 보인다(Figure 3-C).

재학생들의 학업역량은 상황에 따라 달라질 수 있는데, 이는 신입생 선발 전형의 변화에 따른 신입생의 기초 역량 변화, 학제 시스템이나 교육 내용의 변화, 교수법의 변경이나 개선, 새로운 교수요원의 투입 등으로 만들어질 수 있다<sup>10)</sup>. 기존에는 이러한 변화를 학생을 직접 교육하는 과정에서 막연하게 짐작하였으나, 본 연구에서는 커널 밀도지도도를 사용하여 구체적인 학업역량별 프로파일도 설명할 수 있음을 제시하였는데 의의가 있다고 하겠다.

라이트 지도(Figure 4)는 피험자의 분포와 개개 문항의 난이도를 함께 제시함으로써 문항의 효용성을 직관적으로 제시하는데, 본 연구에서는 학생들의 분포에 비해서 문항의 난이도가 비교적 낮은 것을 확인할 수 있었다. 이에, 학생들의 학업역량을 과도하게 낮게 추측하였음이 확인된 것으로, 학업역량이 -1.5보다 낮은 것으로 확인된 11번, 12번 문항이 적절한 난이도를 갖도록 수정하였다. 이러한 분석은 문제은행을 관리하고 출제할 때 필수적으로 고려되는 것으로, 심리검사의 개발 과정에서도 기초적인 타당화 지표로 사용되고 있다.

한의학 교육은 다음과 같은 단계를 거쳐 지속적으로 발전해 왔다<sup>17,29)</sup>. 첫번째 단계는 상한론, 황제내경, 의학입문, 동의보감 등의 원전에 대한 강독과 암기를 주된 교육법으로 사용하는 소규모 도제식 교육으로, 천여년 이상 활용되어 온 가장 오래된 방법이다. 두번째 단계는 1950년대 서양식 의료제도와 의학교육을

시스템을 위해 도입된 분과형 대형강의와 병원 실습 교육법으로, 현대적인 교육시스템에 맞추어 한의학 교육의 양과 질을 높이는 것을 목적으로 한다. 세번째는 ICT의 발전에 따라 새롭게 대두된 미래지향적 이러닝(e-Learning)으로<sup>1,9,10)</sup>, COVID19의 유행과 사회적 거리두기로 인해 갑작스레 확대된 원격 교육, 비대면 교육, 멀티미디어 활용교육, 역량 맞춤형 교육이다.

최근 한의학 학문 분과별 교육에 새롭게 도입되고 있는 역량중심 교육(CBE)<sup>30,32)</sup>은, 이러한 발전 과정 속에서 기존의 교육 시스템을 학생 맞춤형 역량개발 과정으로 재구조화(restructuring)하여 효율성과 수월성을 높이려는 교육방법론의 하나이다<sup>14,33)</sup>. 역량은 ‘특정 상황이나 직무를 효과적이고 우수하게 수행하기 위한 개인의 내재적인 고유 특성’으로 정의되는데<sup>12,13)</sup>, 통상적으로는 지식, 술기, 태도의 세가지로 구성된다.

한의학 교육 현장에서 역량중심 교육이 술기나 태도를 새롭게 도입하는 것으로 오해되기도 하지만, 실제로는 임상 능력의 효율적인 배양을 최종 목표로 학생들의 내재적인 능력(latent ability)을 지식, 술기, 태도의 세가지 분야로 나누고 최적의 교육을 통해 각 분야를 향상시키는 것을 의미한다<sup>34,35)</sup>.

이와 같은 효과적인 임상역량의 강화를 위해서는, 학업역량의 평가 또한 채점점수를 기준으로 한 기존의 줄세우기식 상대평가에서 벗어나<sup>9)</sup>, 학생 고유의 역량을 제대로 분석할 수 있는 IRT에 기초한 역량추정과 절대평가로 교육 시스템을 재구조화 해야 한다<sup>14,33)</sup>. 이러한 교육 시스템은 의약계에 만연한 학업소진을 줄이면서도 학업성취도를 효과적으로 향상시킬 수 있다는 큰 장점을 지니고 있지만, 분석 방법의 이해(Table 1)와 활용에 익숙하지 않았기에 제대로 활용되지 못하여 왔다<sup>1,10)</sup>.

본 연구는 이러한 체계적 학업역량 평가법을 한의학 교육 현장에 처음으로 적용하고자 하였다. 한의학 교육 현장에서 역량중심 교육이 모호한 구호에 불과하다는 저평가를 받아온 이유가 이러한 효율적 평가

시스템이 제 모습을 갖추지 못하였기 때문인데<sup>14,34,35)</sup>, 여타 교육현장에 이미 많은 이들이 사용하고 있는 라쉬 모델을 활용한 역량분석은 교육 과정의 시행과 평가, 분석과 개선의 되먹임 시스템을 온전하게 제시할 수 있을 것이다<sup>1,13)</sup>. 이를 위해서는 다음과 같은 한의학 교육학 연구가 필요할 것으로 사료된다.

첫째, 학업역량의 지속가능한 분석을 위해서는 교수자 및 학생을 대상으로 한 디지털 리터러시(digital literacy) 교육이 시급하다<sup>36)</sup>. 비대면 원격교육이 표준 교육과정에 포함되는 뉴노말 시대의 역량맞춤형 교육을 위해서는 문제 은행과 과정중심 교육을 안정적으로 운영하기 위한 학습관리 시스템(Learning Management System, LMS)의 활용이 필수적이다. 한의학 교육 현장을 ICT 기반 교육생태계로 재구조화하고 이를 학업 분석에 용이하게 활용하기 위해서는<sup>37)</sup>, 재학생을 대상으로 한 LMS 및 ICT 사용 교육과 함께 다양한 평가 기법의 개발과 활용을 위한 교수자 교육이 조속히 제공되어야 할 것이다.

둘째, 최근 한의학 교육현장에 CBT와 LMS가 급격하게 도입되고 있는데<sup>38)</sup>, ICT에 대한 이해의 부족은 도입 이후 여러가지 문제를 유발할 수 있다<sup>37)</sup>. LMS와 CBT 운영 체계가 너무 오래되어 새로운 역량 분석 모듈을 포함할 수 없거나, 학습기록저장소(LRS, Learning Record Store)로 활용할 수 없거나, 학습내용 및 학업결과의 공유를 위한 국제표준인 SCORM(Sharable Content Objective Reference Model) 또는 xAPI(eXperience Application Programming Interface)를 준수하지 않거나, 특정 회사 또는 플랫폼에 종속되어 CBT와 LMS, LRS의 지속적인 업데이트가 불가능하거나, 의약학 또는 한의학 교육의 특성을 반영하지 못하는 시스템을 구입하게 된다면, 향후 예상치 못한 교육현장의 혼란이나 예산의 낭비를 초래할 것이다. 이에, 본격적인 도입과 활용 이전에 한의학 교육학계 차원에서 한의과대학에서 공통으로 사용될 소프트웨어 및 하드웨어에 대한 가이드라인을 제시하여야 할 것이다.

셋째, 한의사 국가시험, 임상 술기시험, 한방 전문의 시험, 한의학 기초종합평가와 같은 자격시험에 IRT를 도입하기 위한 토대를 마련하기 위한 기반 연구가 진행되어야 할 것이다<sup>5,13,21</sup>. 이러한 자격시험은 매우 높은 수준의 안정성과 신뢰성, 공정함이 요구되는데, 선행연구가 부족한 상태에서 과목이나 문항 숫자의 임의 조정, 현장적용 연구가 충분치 못한 상태에서 CBT 시스템으로의 전환, 기존과 다른 문항 특성과 정보를 지닌 멀티미디어 문제의 도입 등이 급격하게 진행될 경우 타당도와 신뢰도에 예상치 못했던 문제를 유발할 수 있기 때문이다<sup>1,10</sup>.

본 연구에서는 라쉬 모델을 사용하여 한의과대학에서 본초학 학업역량을 평가하는 문항/시험의 특성을 분석하고, 학생들의 학업역량을 특성곡선, 정보곡선, 스케일 특성곡선, 커널 밀도지도, 라이트 지도 등을 사용하여 처음으로 제시하였다.

스케일 특성곡선을 통해 시험의 효용성과 타당도가 8년간 안정적으로 유지되어 왔으며, 문항들의 기밀성도 적절하게 유지되어 왔음을 확인할 수 있었다. 학업역량에 성별 차이가 없었으나 시행 연도에 있어서는 유의한 차이가 존재한다는 것을 커널 밀도지도와 ANOVA를 사용하여 확인할 수 있었으며, 다양한 학업역량 분석법이 현장에 적용될 수 있음이 제시되었다. 이와 함께, 문항 난이도와 학생들의 학업역량을 라이트 지도를 사용하여 비교함으로써 어떤 문항이 어떻게 수정되어야 하는지를 구체적으로 제시하였다.

이러한 미래 지향적인 학업역량 분석법은 한의학 교육학의 기본적인 도구로 활용되어야 하며, 이러닝 시대의 지속가능하고 역량을 중심으로 하는 미래지향적인 역량중심 한의학 교육, 한의학교육 측정평가학의 토대를 마련하는데 기여할 것이다.

### Acknowledgement

본 연구는 부산대학교의 연구비지원을 받았음 (This work was supported by a 2-Year Research Grant

of Pusan National University.)

### 참고문헌

1. Huh, S. (2014). Computer-based testing and construction of an item bank database for medical education in Korea. *Korean Med Educ Rev.* 16(1). 11-15. <https://doi.org/10.17496/kmer.2014.16.1.011>
2. Lee, C. K., Park, J. S., Lee, E., Lee, S. J., Park, E. S. & Park, Y. J. (2001). A comparative study of item analysis by item response theory based for initiating CAT (computer adaptive test) system. *Korean J Med Educ.* 13(1). 107-115.
3. Park, J. H., Son, J. Y. & Kim, S. (2012). Experiences with establishing and implementing learning management system and computer-based test system in medical college. *Korean J Med Educ.* 24(3). 213-222. <https://doi.org/10.3946/kjme.2012.24.3.213>
4. Im, E.-J., Lee, W.-K., Lee, Y.-C., Choe, B.-H., Chung, S.-K., Lee, T.-H., et al. (2008). Development of computer-based test (cbt) and student recognition survey on cbt. *Korean J Med Educ.* 20(2). 145-154. <https://doi.org/10.3946/kjme.2008.20.2.145>
5. Jeong, G.-H. & Yim, M. K. (2005). Applicability of item response theory to the Korean nurses' licensing examination. *J Educ Eval Health Prof.* 2(23-29). <https://doi.org/10.3352/jeehp.2005.2.1.23>
6. Korea Health Personnel Licensing Examination Institute. Computer based test. Available from: [https://www.kuksiwon.or.kr/cnt/c\\_2033/view.do?seq=11](https://www.kuksiwon.or.kr/cnt/c_2033/view.do?seq=11)
7. Choi, Y.-J. & Asiilkalkan, A. (2019). R package

- for item response theory analysis: Description and features. *Measurement: Interdisciplinary research and perspectives*. 17(3). 168-175. <https://doi.org/10.1080/15366367.2019.1586404>
8. Meyer, J. P. (2014). *Applied measurement with jmetrik*. New York: Routledge.
  9. Seong, T.-J. (2006). A propose for individualized comprehensive evaluation system. *J Curric Eval*. 9(2). 121-138. <https://doi.org/10.29221/jce.2006.9.2.121>
  10. Chae, H., Han, S. Y., Yang, G. & Kim, H. (2022). Study on the herbology test items in korean medicine education using item response theory. *Kor J Herbology*. 37(2). 13-21. <https://doi.org/10.6116/kjh.2022.37.2.13>.
  11. Kim, S. H., Han, S. Y., Kim, J. D., Choi, S., Lee, S. J., Lim, J. H., et al. (2015). Study on stress and burnout in medical education at the school of korean medicine. *J Oriental Neuropsychiatry*. 26(2). 103-116. <https://doi.org/10.7231/jon.2015.26.2.103>
  12. Short, E. c. (1985). The concept of competence: Its use and misuse in education. *Journal of Teacher Education*. 36(2). 2-6. <https://doi.org/10.1177/002248718503600202>
  13. Schauber, S. K., Hecht, M. & Nouns, Z. M. (2018). Why assessment in medical education needs a solid foundation in modern test theory. *Adv in Health Sci Educ*. 23(1). 217-232. <https://doi.org/10.1007/s10459-017-9771-4>
  14. Kim, Y. & Lim, C.-I. (2011). Competency-based medical education: Possibilities and limitations. *KMER*. 13(1). 13-23. <https://doi.org/10.17496/kmer.2011.13.1.013>
  15. Chae, H., Hwang, S., Kwon, Y., Baik, Y., Shin, S., Yang, G., et al. (2010). Study on the prerequisite chinese characters for education of traditional korean medicine. *J Physiol & Pathol Korean Med* 24(3). 373-379.
  16. Lee, Y., Kwak, M.-J., Jung, H., Ha, H.-y. & Chae, H. (2012). A study on the statistical methods used in kci listed journals of traditional korean medicine from 1999 to 2008. *Korea Journal of Oriental Medicine*. 18(2). 55-64.
  17. Lee, J., Han, J. H., Kim, M. S., Lee, H. S., Han, S. Y., Lee, S. J., et al. (2020). Teaching yin-yang biopsychology using the animation, “pororo the little penguin”. *Eur J Integr Med*. 33(10)1037. <https://doi.org/10.1016/j.eujim.2019.101037>
  18. Gray, A. C., Steel, A. & Adams, J. (2019). A critical integrative review of complementary medicine education research: Key issues and empirical gaps. *BMC Complement Altern Med*. 19(1). 73. <https://doi.org/10.1186/s12906-019-2466-z>
  19. De Champlain, A. F. (2010). A primer on classical test theory and item response theory for assessments in medical education. *Medical education*. 44(1). 109-117. <https://doi.org/10.1111/j.1365-2923.2009.03425.x>
  20. Schauber, S. K. & Hecht, M. (2020). How sure can we be that a student really failed? On the measurement precision of individual pass-fail decisions from the perspective of item response theory. *Medical Teacher*. 42(12). 1374-1384. <https://doi.org/10.1080/0142159X.2020.1811844>
  21. Lee, S. Y., Lee, Y. & Kim, M. K. (2018). Effectiveness of medical education assessment consortium clinical knowledge mock examination

- (2011-2016). Korean Med Educ Rev. 20(1). 20-31. <https://doi.org/10.17496/kmer.2018.20.1.20>
22. Lee, S., Lee, Y., Han, S. Y., Bae, N., Hwang, M., Lee, J., et al. (2020). Urinary function of the sasang type and cold-heat subgroup using the sasang urination inventory in korean hospital patients. *Evid Based Complement Alternat Med.* 2020(7313581). <https://doi.org/10.1155/2020/7313581>
23. Chae, H., Cho, Y. I. & Lee, S. J. (2021). The yin-yang personality from biopsychological perspective using revised sasang personality questionnaire. *Integr Med Res.* 10(1). 100455. <https://doi.org/10.1016/j.imr.2020.100455>
24. Lee, Y.-j., Lee, S., Kim, S.-h., Lee, J. & Chae, H. (2021). Study on the revision and clinical validation of the sasang digestive function inventory. *J.* 33(3). 54-71.
25. Park, J. C. & Kim, K. S. (2012). A comparison between discrimination indices and item-response theory using the rasch model in a clinical course written examination of a medical school. *Korean J Med Educ.* 24(1). 15-21. <https://doi.org/10.3946/kjme.2012.24.1.15>
26. Rasch, G. (1960). Probabilistic models for some intelligence and attainment tests. Danmarks Paedagogiske Institut.
27. Aksu, G., GÜZELLER, C. O. & Eser, M. T. (2019). Jmetrik: Classical test theory and item response theory data analysis software *Journal of Measurement and Evaluation in Education and Psychology.* 10(2). 165-178. <https://doi.org/10.21031/epod.483396>
28. Park, J. C. & Kim, K. S. (2012). A comparison between discrimination indices and item-response theory using the rasch model in a clinical course written examination of a medical school. *Korean J Med Educ.* 24(1). <https://doi.org/10.3946/kjme.2012.24.1.15>
29. Han, S. Y., Kim, H. Y., Lim, J. H., Cheon, J., Kwon, Y. K., Kim, H., et al. (2016). The past, present, and future of traditional medicine education in korea. *Integr Med Res.* 5(2). 73-82. <https://doi.org/10.1016/j.imr.2016.03.003>
30. Yu, J.-S. (2019). Recommendation for development of clinical skill contents in the competency-based sasang constitutional medicine education. *J Sasang Constitut Med.* 31(4). 1-8. <https://doi.org/10.7730/JSCM.2019.31.4.1>
31. Choi, J.-b. & Kim, Y.-j. (2018). A study on the education of medical classics through flipped learning. *J Korean Med Classics.* 31(2). 1-16. <https://doi.org/10.14369/jkmc.2018.31.2.001>
32. Kim, H.-S., Yang, U.-h. & Na, C.-s. (2018). A study on the perception of international healthcare development cooperation of korean medicine students for competency-based clinical herbology education. *Kor J Herbology.* 33(5). 39-46. <https://doi.org/10.6116/kjh.2018.33.5.39>
33. Hong, J. & Kang, Y. (2018). The implications of the case of medical education in north america on korean medicine education. *The Journal of Korean Medical History.* 31(2). 91-101. <https://doi.org/10.15521/JKMH.2018.31.2.091>
34. Harris, P., Snell, L., Talbot, M. & Harden, R. M. (2010). Competency-based medical education: Implications for undergraduate programs. *Med Teach.* 32(8). 646-650. <https://doi.org/10.3109/0142159x.2010.500703>
35. Lee, Y. B. (2013). The role of the concept of competence in korean outcome-based medical education. *Korean Med Educ Rev.* 15(3).

- 144-150.
36. Koschmann, T. (1995). Medical education and computer literacy: Learning about, through, and with computers. *Acad Med.* 70(9). 818-821.
37. Plch, L. (2020). Perception of technology-enhanced learning by medical students: An integrative review. *Medical science educator.* 30(4). 1707-1720. <https://doi.org/10.1007/s40670-020-01040-w>
38. Naciri, A., Radid, M., Kharbach, A. & Chems, G. (2021). E-learning in health professions education during the covid-19 pandemic: A systematic review. *Journal of educational*

*evaluation for health professions.* 18. 27.  
<https://doi.org/10.3352/jeehp.2021.18.27>

### ORCID

채 한 <http://orcid.org/0000-0002-8698-8229>  
이수진 <http://orcid.org/0000-0002-5543-7301>  
한창호 <http://orcid.org/0000-0003-3460-9263>  
조영일 <http://orcid.org/0000-0003-3235-7613>  
김형우 <http://orcid.org/0000-0002-1183-7434>