

# 문항반응이론을 적용한 융합적 사고 및 문제해결 역량진단 도구의 병렬 단축형 개발 : H 대학교를 중심으로

유현주\*·남나라\*\*†

\*한국공학대학교 성과관리인증센터 연구교수

\*\*한국방송통신대학교 미래원격교육연구원 선임연구위원

## Development of Parallel Short Forms of the Convergent Thinking and Problem Solving Inventory Utilizing Item Response Theory : A Case Study of Students in H University

You, Hyunjoo·Nam, Na-Ra\*\*†

\*Research Professor, Center for Educational Evaluation & Improvement, Tech University of Korea

\*\*Senior Researcher, Institute of Future Distance Education, Korea National Open University

### ABSTRACT

The study was conducted to develop two parallel short forms for the Convergent thinking and Problem solving questionnaires which are part of H University's core competency diagnostic tools, based on Multi-Item Response Theory. Item responses of 2,580 students were analyzed using Graded Response Model(GRM) to determine item difficulty and discrimination of each item. The research results are as follows. Two parallel short tests were developed for the Convergent thinking questionnaire consisting of 12 items which were originally 17 items. Likewise, the Problem solving questionnaire, which originally consisted of 15 questions, was divided into two parallel short forms, each consisting of 9 items. The reliability of the shortened parallel tests was confirmed through internal consistency analysis, and their similarity to the original tests was established through correlation analysis. This study contributed to quality management of competency-based education and programs at H University by developing shortened tests. Based on the results, implications were presented as well as limitations and discussions.

**Keywords:** Item response theory, Core competency diagnostic tools, Parallel short forms, Convergent thinking, Problem solving competency

## I. 서 론

이론·지식보다 실천적 수행력으로서 '역량'을 중시하는 사회적 필요와 요구에도 불구하고, 대학교육과 산업체 요구 능력 간 '미스매치'의 문제는 오랜 시간 난제로서 지속되어 왔다(신현상 외, 2014). 이러한 교육적·사회적 요구에 부응하고자 각 대학에서는 그들 고유의 특성을 반영한 '핵심역량'을 설정하고, 역량 기반의 교육과정을 설계·운영하고 있다. 이와 더불어 고등교육 평가 시 대학 교육의 성과지표로서 학생의 역량 수준이 강조됨에 따라 그 성과를 진단하여 향후 개선에 활용하는 '역량 기반 교육 질 관리 선순환 체제'의 중요성이 점차 부각되어 왔다. 공

학계열 중심으로 이루어진 H 대학에서는 대학의 인재상 즉, 창의형·실천형·진취형 인재 양성을 위해 고유의 핵심역량 6개를 설정하고, 교과·비교과 전 영역의 교육과정을 핵심역량 기반으로 설계하여 운영해 오고 있다. 그리고 2019년부터 학생들의 핵심역량 수준을 진단하기 위한 97개 문항의 평가도구를 자체 개발·실행함으로써 그 결과를 대학 교육의 성과점검 및 교육과정으로의 환류·개선에 활용해 왔다. 이처럼 대학 핵심역량 진단 도구는 역량 기반 교육의 지속적 개선에 필요한 수단이며, 크게 2가지 측면에서 활용을 고려할 수 있다. 첫 번째로는, 주기적으로 학생들의 역량 향상도를 진단, 그 결과를 환류함으로써 부족 역량 확인 및 이를 개발하기 위한 교육과정 전반의 개선을 도모하는 것이다. 두 번째는 앞서 언급한 대학 단위의 주기적·전반적 활용 외, 프로그램 단위의 비정기적·개별적 활용을 들 수 있다. 다시 말해서, 프로그램에 대한 효과성 확인을 위해 진단 문항을 활용할 수 있다는 것이다. 이는 진단 문항 전체가 아닌 특정 역

Received March 29, 2023; Revised May 8, 2023

Accepted May 15, 2023

† Corresponding Author: naranam@gmail.com

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

량의 문항만을 선택·활용할 수 있다는 점에서 도구의 활용도를 극대화하는 방안이 되기도 한다.

특히 융합역량과 문제해결 역량의 경우, 지금까지 여러 선행연구(권재기·정미경, 2014; 김윤영·윤지영, 2021; 이민정, 2016; 조동현, 2016; 허지숙·황윤자, 2020)에서 공학계열 학생들이 함양해야 할 핵심역량으로 강조된 바 있다. H 대학 역시 4차 산업혁명 시대를 선도하는 공학 인재가 갖추어야 할 핵심역량으로 융합 사고 역량과 문제해결 역량을 설정하고, 이를 계발시키고자 다수의 프로그램을 설계·운영하고 있어 이들 프로그램의 질 관리를 위한 효과성 검증 도구의 필요성이 제기되는 상황이다. 이처럼 공학계열에서 핵심이 되는 융합·문제해결 역량계발을 위해 매년 가장 많은 프로그램이 개설되고 있지만, 대학 단위의 종합적 역량향상도 진단 및 환류만 이루어질 뿐 개별 프로그램의 질 관리는 다소 모호하고 주관적인 경향이 있다. 아울러, 이 같은 개별 프로그램의 효과성 확인이 적절히 이루어지려면 진단 도구에 대한 다음의 문제들이 개선되어야 할 것이다.

먼저 문항 수의 문제이다. 일반적으로 문항을 통한 효과성 확인은, 프로그램 참여 전과 후에 역량 수준을 진단하여 그 차이를 검증함으로써 알 수 있다. 그러나 진단 시 문항 수가 많아지면 응답자의 검사 피로도가 증가하고(Anastasi, 1976), 응답의 정확성과 성실성이 저하될 수 있어(김권현 외, 2015; 박종철 외, 2021) 주의해야 한다. 특히 프로그램 운영의 효율성 제고 및 참여 집중도 유지를 위해서는 단축형 문항을 통해 가능한 단시간 내에 사전·사후 검사를 마치는 것이 중요하다. H 대학의 핵심역량 진단 도구 내 융합 사고역량과 문제 해결역량의 문항 수는 총 32개로, 프로그램 전·후 실시하기에는 운영상 부담되는 부분이기도 하다. 다음으로는, 프로그램 참여 전·후 역량진단 시 동일 문항의 반복사용이다. 동일 문항을 반복 사용할 경우, 문항 암기(기억) 효과를 통제하기 어려워 검증 결과를 신뢰할 수 없게 된다. 따라서 암기 효과를 완화하고 결과의 신뢰를 높이기 위해서는 동일 능력을 측정하되 문항의 진술 내용은 다른 2개의 세트를 활용함이 적절하다.

이처럼 동일 능력을 측정하되 유사한 형식과 통계적 특성을 갖는 검사들을 동형 검사(equivalent form)(김화영·김현철, 2019)라고 하며, 유사한 여러 개의 검사라는 점에서 병렬형 검사(parallel or alternate form)라고도 한다. 이러한 병렬형 도구를 짧은 시간에 검사가 가능한 단축형으로 제작할 경우, 사전·사후 검사와 같이 반복측정이 필요한 상황에서도 사용할 수 있다(강연옥 외, 2000). 병렬형 검사 혹은 동형 검사 간에는 측정학적 요소 즉 문항 수·문항 난이도 및 변별도 등이 동일 혹은 유사해야 하며, 이를 위해서는 문항 분석을 통해 이러한 요소들을 확인하는 작업이 전제되어야 한다.

문항 분석의 근거가 되는 검사이론에는 크게 2가지가 있다. 첫째는 고전검사이론(classical test theory: CTT)으로 검사의 총점에 의해 분석하는 것이며, 둘째는 문항반응이론(item response theory: IRT)으로 총점이 아닌 문항 하나하나에 근거하여 분석하는 것이다(성태제, 2009). 문항반응이론의 경우 피험자의 잠재된 능력과 문항별 응답 간의 관계를 규명하므로(김향희·김수련, 2013), 문항 전체의 결과 즉 총점에 의해 분석하는 고전검사이론과 비교할 때 정확한 피험자 능력추정이 가능하다. 다시 말해서, 개인의 관찰 가능한 검사 문항 반응과 그 기저에 있는 관찰 불가능한 능력 간의 관계를 구체화 시키는 것이다(이종승, 2009). 따라서 문항반응이론에 의해서는 검사의 쉽고 어려운 정도에 상관없이 피험자의 능력을 안정적으로 추정할 수 있다(김석우, 2009).

단축형 검사 개발에 관한 선행 문헌의 경우, 이미 타당성이 입증된 성격검사나 심리검사에 대한 단축형 개발 및 타당화 연구(김세형·조동욱, 2020; 박종철 외, 2021; 배병훈 외, 2015; 서동기 외, 2019; 심예은 외, 2021; 정경미 외, 2020; 정문주 외, 2017)가 주를 이루고 있다. 그 외, 의학 분야에서 치료 전 신속한 진단을 목적으로 단축형 검사를 개발한 연구들(강연옥 외, 2000; 윤진웅 외, 2019)도 살펴볼 수 있다. 그러나 이러한 기존 연구들은 대부분 여러 개의 병렬 단축형이 아닌 1개의 단축형 도구 개발을 목적으로 하며, 그 내용 또한 수행 위주의 ‘역량’이 아닌 내적 속성을 측정하는 것이 대부분이다. 일부 병렬 단축형 개발 연구(강연옥 외, 2000)의 경우 특정 질환이 있는 환자를 대상으로 사용이 가능한 진단 도구이며, 교육 및 프로그램의 질 관리를 목적으로 하는 경우는 찾아보기 어렵다. 또한 선행 문헌 대부분은 문항반응이론보다는 고전검사이론에 근거하여 검사 내 요인구조의 타당성을 밝히는 경우가 다수였다. 일부 문항반응이론을 적용한 연구(김세형·조동욱, 2020; 김향희·김수련, 2013; 박종철 외, 2021; 배병훈 외, 2015)의 경우는 문항 난이도만을 고려한 경우가 많았다.

이에 본 연구에서는 공학계열 내 역량 기반 교육의 질 관리 차원에서, H 대학의 융합 사고역량과 문제 해결역량에 대한 기존 진단 도구를 활용도가 높은 병렬 단축형 도구로 분리·개발하고자 한다. 그리고 문항반응이론을 적용하여 문항별 난이도와 변별도를 동시에 고려함으로써, 단축형 검사 간 동형성을 높이고자 한다. 이와 같은 연구 목적 달성을 위한 연구 문제는 다음과 같다.

1. 융합사고역량 및 문제해결역량에 대한 H 대학의 병렬 단축형 검사 도구를 개발한다.
2. 개발된 병렬 단축형 검사의 타당성을 검증한다.

## II. 연구 방법

### 1. 분석 자료

#### 가. H 대학 핵심역량 진단 도구 구성

H 대학의 핵심역량 진단 도구는 최초 개발 이후 역량 및 문항에 대한 타당도 평가를 거쳐 6개 역량 및 18개 하위역량을 측정하는 총 97개 문항으로 구성된다. 그 중 본 연구에서 다루고자 하는 융합 사고역량과 문제해결역량에 대한 하위요인 구성은 아래 Table 1과 같으며, 문항 수는 총 32개이다. 신뢰도를 의미하는 알파값(Cronbach's  $\alpha$ )은 요인별 0.6 이상의 양호한 수준으로 확인되었다.

Table 1 Competencies to analyze

핵심역량	하위요인	문항 수	신뢰도
융합사고역량 (Convergent Thinking)	연결적 사고	5	.775
	아이디어 생성 능력	7	.909
	기업가 정신	5	.813
문제해결역량 (Problem Solving)	문제 인식 능력	5	.813
	비판적 사고 능력	5	.704
	창의 사고 실천	5	.674
전체		32	

#### 나. 분석 대상

이 연구에서 활용한 자료는 2020-2021학년도에 H 대학 핵심역량 진단 검사에 응시한 학생들의 응답 결과로(N=2,580), 남학생이 전체의 64.0%(1,650명), 여학생은 36.0%(930명)를 차지하고 있으며 1학년과 4학년의 응답 비율이 다소 높았다(1학년: 43.6%, 2학년: 14.7%, 3학년: 19.0%, 4학년 22.8%).

Table 2 Descriptives of data

구분	인원 수	구성 비율
연도	2020	1,593 61.7%
	2021	987 38.3%
학년	1	1,124 43.6%
	2	378 14.7%
	3	489 19.0%
	4	589 22.8%
성별	남	1,650 64.0%
	여	930 36.0%
전체	2,580	100.0%

### 2. 분석 방법

이 연구에서는 문항반응이론(Item response theory)을 활용하여 문항과 피험자의 특성을 추정하고, 해당 정보를 바탕으로 병렬 단축형 검사를 구성하였다. H 대학의 핵심역량 진단 도구에서는 각 문항에 대하여 5점 척도로 평정하므로(매우 그렇지 않다-그렇지 않다-보통이다-그렇다-매우 그렇다), 문항에 대한

반응이 세 개 이상인 다분 문항을 사용하여 피험자의 능력을 추정하는 다분 문항 반응이론을 기반으로 하였다(박정, 2001).

다분 문항을 위한 분석 모형은 등급반응모형(Graded Response Model), 부분점수모형(Partial Credit Model), 일반화점수모형(General Partial Credit Model) 등이 있으나, 이 연구에서는 인간의 정의적 행동 측정 시 주로 사용하는 Likert 척도에 의해 얻어진 값을 분석할 때 적절한 문항반응이론 모형의 하나인 등급반응모형을 활용하였다.

등급반응모형은 정답과 오답 사이의 중간 능력을 고려해서 피험자 반응을 상이한 등급을 갖는  $m$ (=반응범주-1)개 범주로 분류하여 점수화하는 것으로, Likert 척도로 측정된 정의적 특성을 추정할 경우 피험자 특성에 대해 다양하고 정확한 정보를 얻을 수 있다. 그뿐만 아니라, 문항들의 변별도와 각 범주별 위치모수를 통해 각 문항이 검사에 미치는 영향 등을 분석할 수 있어 문항 특성을 활용하여 병렬형 검사를 구성하려는 이 연구의 목적에 적합한 분석 방법이다(Koch, 1983; Muraki, 1990).

#### 가. 병렬 단축형 검사 개발 절차

이 연구에서는 H 대학의 융합사고역량과 문제해결역량 진단 도구의 병렬 단축형 검사를 구성하기 위하여 다음 절차를 거쳤다.

첫째, 전체 문항의 난이도와 변별도를 분석하였다. 문항 난이도( $b_m$ )란 각 문항 범주에 반응할 확률이 0.5인 지점의 능력값으로, 대체로  $\pm 2$  범위에 위치한다. 이 연구에서는  $m$ (=반응범주-1)개 범주에 생성된 각 단계난이도의 평균값을 산출하여 문항의 평균적인 난이도로 활용하였다. 또한 문항 변별도( $a$ )란 문항 난이도로 표현되는 능력 수준보다 낮은 능력의 피험자와 높은 능력의 피험자를 변별하는 정도를 나타내는 것으로, 일반적으로 0-2의 값을 취한다(이종승, 2009; Hambleton et al, 1991).

둘째, 각 역량의 하위요인별 문항 특성이 유사하게끔 2개의 병렬 단축형 검사를 구성하였다. 단, 하위영역 내 검사의 신뢰도를 고려하여 최소 3개 문항으로 구성하였으며, 문항 난이도는 유사하나 변별도에서 다소 차이가 큰 경우 등 하나의 단축형 검사에만 배정하기 어려울 시 두 검사 모두에 배정하는 가교 문항(공통 문항)으로 설정하였다.

셋째, 개발된 병렬 단축형 검사의 타당도를 검증하였다. 먼저

Table 3 Process for developing short forms

1단계: 문항 분석	⇒	2단계: 병렬 단축형 검사 구성	⇒	3단계: 병렬 단축형 검사 타당도 검증
문항 특성 분석 - 융합사고 17문항 - 문제해결 15문항		하위요인별 문항 특성이 유사하도록 병렬 단축형 검사 2개 세트 구성		병렬 단축형 검사 와 원 검사 간 신뢰도 점검 및 상관 분석

검사의 내적 일치도 계수를 점검하고, 원 검사 및 병렬 단축형 검사 간 상관 분석을 통해 병렬 단축형 검사의 적절성을 확인하였다. 구체적인 분석 과정을 정리하여 제시하면 다음과 같다.

나. 자료 분석

문항 특성 분석을 위하여 R 4.2.2 버전의 mirt 1.38.1 패키지를, 검사의 내적 일치도 계수 및 검사 간 상관 분석을 위하여 SPSS Window version 21.0을 활용하였다.

III. 연구 결과

1. 병렬 단축형 검사 개발

이 연구에서는 총 32개 문항(융합 사고역량 문항 17개, 문제해결역량 문항 15개)의 특성을 분석하여 병렬 단축형 검사 2개로 분리·개발하고자 하였다. 이때 문항 특성으로 문항 변별도와 문항 난이도를 산출하였으며, 융합사고역량 및 문제해결역량을 구성하는 하위요인별 문항 특성이 유사하도록 단축형 검사를 구성하였다. 이에 역량별로 구체적인 단축형 검사 개발 결과를 제시하고자 한다.

가. 융합사고역량

융합사고역량의 하위요인 - 연결적 사고, 아이디어 생성 능력, 기업가 정신 - 별로 문항 특성을 분석한 결과는 다음과 같다 (Table 4 참조). 먼저 연결적 사고를 측정하는 문항은 총 5개로, 3번과 5번, 그리고 1번과 2번의 문항 난이도 및 변별도가 유사

Table 4 Result of item analysis of Convergent thinking

하위 요인	문항	변별도 (a)	난이도 (b)	단계난이도			
				b <sub>1</sub>	b <sub>2</sub>	b <sub>3</sub>	b <sub>4</sub>
연결적 사고	1	1.589	2.932	7.058	4.448	1.96	-1.738
	2	1.431	2.750	6.862	3.954	1.649	-1.465
	3	1.723	1.538	6.445	2.806	-0.025	-3.076
	4	2.147	2.452	8.001	4.191	0.871	-3.256
	5	2.195	1.556	6.869	3.158	-0.159	-3.644
아이디어 생성 능력	1	2.513	0.791	5.946	1.949	-0.909	-3.822
	2	3.091	1.406	7.12	2.987	-0.383	-4.101
	3	1.934	2.019	6.362	3.615	0.834	-2.734
	4	2.49	0.466	5.482	1.931	-1.208	-4.342
	5	2.477	1.293	6.73	2.864	-0.538	-3.884
	6	2.535	1.077	6.25	2.557	-0.36	-4.141
	7	3.561	1.609	8.673	3.286	-0.558	-4.966
기업가 정신	1	1.758	0.654	4.207	1.281	-0.269	-2.603
	2	1.913	0.863	4.941	1.632	-0.385	-2.738
	3	1.563	0.617	4.139	1.436	-0.196	-2.913
	4	1.650	1.599	6.009	2.560	0.470	-2.644
	5	3.684	1.036	7.432	2.547	-0.636	-5.200

주: i = 문항 번호

하여 각 세트1과 세트2 검사에 배정하였으나, 4번 문항은 문항 난이도와 변별도가 동시에 높은 문항으로 검사 하나에만 배정하기 어려워 두 검사 모두에 배정하여 세트별 3문항씩 구성하였다. 다음으로 아이디어 생성 능력을 측정하는 문항은 총 7개로, 문항 난이도 및 변별도가 유사한 2번과 7번, 1번과 6번을 각 세트 1과 세트2 검사에 배정하였다. 3, 4, 5번 문항의 경우는 변별도가 대체로 유사하나 난이도 차이가 커 모두 공통문항으로 일괄 배정하여, 세트별 5문항씩 구성하였다.

마지막으로 기업가 정신을 측정하는 문항은 총 5개로, 문항 난이도와 변별도가 유사한 1번과 3번 문항을 두 세트에 나누어 배정하였다. 단, 2번과 5번 문항의 경우 난이도는 유사하나 변별도에서 다소 차이를 보였으며, 4번 역시 타 문항과 비교하여 난이도가 높아 2개 세트에 모두 배정되어 세트별 4문항씩 구성하였다. 이처럼 개발된 융합 사고역량의 병렬 단축형 검사는 원 검사 문항(17개)의 약 70% 수준(각 12개)으로 단축되었다(Table 5 참조).

Table 5 Items comprising two short forms for the Convergent thinking

하위 요인	세트/문항 개수	문항 번호							
		1	2	3	4	5	6	7	
연결적 사고	세트 1	3	1	2	3	4	5	6	7
	세트 2	3	1	2	3	4	5	6	7
아이디어 생성능력	세트 1	5	1	2	3	4	5	6	7
	세트 2	5	1	2	3	4	5	6	7
기업가 정신	세트 1	4	1	2	3	4	5	6	7
	세트 2	4	1	2	3	4	5	6	7
세트별 총 문항 수		12							

나. 문제해결역량

문제해결역량의 세 가지 하위요인 - 문제 인식 능력, 비판적 사고 능력, 창의 사고 실천 - 별로 문항 특성을 분석한 결과는 다음과 같다(Table 6 참조). 먼저 문제 인식 능력을 측정하는 문항은 총 5개로, 1번과 2번, 그리고 3번과 4번의 문항 난이도 및 변별도가 유사하여 각 세트1과 세트2 검사에 배정하였다. 그러나, 5번 문항의 경우 타 문항 대비 변별도가 다소 낮은 문항으로 단축형 검사 하나에만 배정하기 어려워 두 검사 모두에 배정하여 세트별 3문항씩 구성하였다. 다음으로 비판적 사고 능력을 측정하는 문항은 총 5개로, 문항 난이도 및 변별도가 유사한 1번과 3번, 5번과 4번을 각 세트1과 세트2에 배정하고 변별도가 다소 낮은 2번은 공통문항으로 구성하였다. 마지막으로 창의 사고 실천을 측정하는 문항은 총 5개로, 문항 난이도와 변별도가 유사한 1번과 2번, 4번과 5번을 각 세트에 나누어 배정하였다. 단, 3번 문항의 경우 문항 난이도가 타 문항들과 비교하여 이질적

로, 2개 세트에 모두 배정하여 세트별 3문항씩 구성하였다. 이처럼 개발된 문제 해결역량의 병렬 단축형 검사는 원 검사 문항 (15개)의 약 60% 수준(각 9개)으로 단축되었다(Table 7 참조).

**Table 6 Result of item analysis of Problem solving**

하위 요인	문항	변별도 (a)	난이도 (b)	단계난이도			
				$b_{i1}$	$b_{i2}$	$b_{i3}$	$b_{i4}$
문제 인식 능력	1	2.110	2.979	8.088	4.672	1.835	-2.680
	2	2.167	3.168	8.001	4.862	1.930	-2.121
	3	2.884	4.580	10.864	7.034	3.184	-2.762
	4	2.763	4.300	10.144	6.795	2.939	-2.677
	5	1.484	3.013	8.192	4.076	1.605	-1.823
비판적 사고 능력	1	1.767	2.989	8.172	4.509	1.494	-2.219
	2	1.053	2.032	5.949	2.778	0.917	-1.515
	3	1.478	2.798	7.722	4.258	1.492	-2.282
	4	1.893	2.074	7.125	3.697	0.651	-3.176
	5	1.509	2.007	6.158	3.413	0.838	-2.380
창의 사고 실천	1	0.973	1.471	4.740	2.291	0.585	-1.734
	2	0.994	1.938	6.356	2.925	0.756	-2.287
	3	1.726	0.674	5.507	1.933	-0.748	-3.996
	4	1.865	2.341	7.379	3.718	1.080	-2.813
	5	1.428	1.879	5.730	2.983	0.829	-2.027

주:  $i$  = 문항 번호

**Table 7 Items comprising two short forms for the Problem solving**

하위 요인	세트/문항 개수		문항 번호				
			1	2	3	4	5
문제 인식 능력	세트 1	3	■	■	■	■	■
	세트 2	3	■	■	■	■	■
비판적 사고 능력	세트 1	3	■	■	■	■	■
	세트 2	3	■	■	■	■	■
창의 사고 실천	세트 1	3	■	■	■	■	■
	세트 2	3	■	■	■	■	■
세트별 총 문항 수		9					

## 2. 병렬 단축형 검사 세트의 타당도 검증

이상과 같이 본 연구에서는 문항 특성에 근거하여 융합사고역량과 문제해결역량에 대한 병렬 단축형 검사 2개 세트를 구성하였다. 단축형 검사의 타당도를 검증하기 위하여 역량별 최종 구성된 문항들이 해당 속성을 일관되게 측정하는지 확인하고자 신뢰도 계수(Cronbach's  $\alpha$ )를 산출하였다. 아울러 응답자의 능력 추정치( $\theta$ )를 활용하여 원 검사와 단축형 검사 간 상관관계수 및 순위상관계수를 산출함으로써 3개 검사(원 검사, 단축형 검사 2개) 간 유사성을 확인하였다. 먼저 융합사고역량의 결과를 살펴보면 다음과 같다.

## 가. 융합사고역량

융합사고역량의 전체 신뢰도 및 세 하위영역별 신뢰도 분석 결과는 다음과 같다. 일반적으로 사회과학 연구 분야에서 Cronbach's  $\alpha$  값이 0.5-0.7이면 적정 수준, 0.7-0.9이면 높은 수준의 내적 일관성을 갖는다고 평가된다(Hinton et al., 2004). 이와 같은 기준에 비추어볼 때, 융합 사고역량의 세 하위요인을 측정하는 문항들을 문항 특성에 근거하여 나는 병렬 단축형 검사들의 내적 일관성은 양호하였다. 다만, 모든 영역에서 원 검사보다 신뢰도 계수 값이 다소 감소하였는데, 이는 원 검사와 비교해 줄어든 문항 수 때문으로 추정된다.

**Table 8 Cronbach's Alpha of competency tests for Convergent thinking**

구분	구분	신뢰도
융합사고 역량	원검사	.901
	세트 1	.860
	세트 2	.870
하위요인 1: 연결적 사고	원검사	.775
	세트 1	.667
	세트 2	.693
하위요인 2: 아이디어 생성능력	원검사	.909
	세트 1	.871
	세트 2	.874
하위요인 3: 기업가 정신	원검사	.813
	세트 1	.782
	세트 2	.790

다음으로 원 검사와 단축형 검사에서 산출된 응답자의 능력 추정치( $\theta$ )를 활용하여 3개 검사(원 검사, 단축형 검사 2개) 상관 분석을 수행하였다. 이때 스피어만(Spearman) 순위상관계수를 추가로 비교하여 원 검사로 측정하였을 때 학생들의 순위가 단축형 검사를 활용했을 때와 유사한지 점검하였다. 그 결과, 연결적 사고, 아이디어 생성 능력, 기업가 정신의 세 하위요인 모

**Table 9 Results of correlation analysis between original test and abbreviated tests for Convergent Thinking**

요인	구분	능력추정치		
		세트별	기존세트	세트_1
연결적 사고	상관계수	세트 1	0.933***	
		세트 2	0.946***	0.801***
	순위상관계수	세트 1	0.910***	
		세트 2	0.933***	0.762***
아이디어 생성 능력	상관계수	세트 1	0.971***	
		세트 2	0.976***	0.916***
	순위상관계수	세트 1	0.967***	
		세트 2	0.977***	0.901***
기업가 정신	상관계수	세트 1	0.987***	
		세트 2	0.990***	0.978***
	순위상관계수	세트 1	0.986***	
		세트 2	0.989***	0.978***

\*\*\* $p < .001$

두에서 단축형 검사와 원 검사 간 상관관계수 및 순위상관계수가 0.9 이상으로 매우 높았다( $p < .001$ ). 아울러 단축형 검사 간 상관관계수 및 순위상관계수 또한 최저값이 0.7 이상이었고, 대체로 0.9 내외의 값을 보여 문항 특성에 근거하여 배분된 단축형 검사와 원 검사 간 유사성을 확인하였다.

나. 문제해결역량

문제해결역량의 전체 신뢰도 및 세 하위영역별 신뢰도를 분석한 결과, 문항 특성에 근거하여 나는 병렬 단축형 검사들의 내적 일관성은 양호한 수준이었다. 다만, 모든 영역에서 원 검사보다 신뢰도 계수의 값이 다소 감소하였는데, 이는 문항 수에 민감한 신뢰도 계수의 특성 때문으로 추정된다(Nunnally & Bernstein, 1994; Streiner, 2003).

Table 10 Cronbach's Alpha of competency tests for Problem solving

구분	구분	신뢰도
문제해결 역량	원검사	.856
	세트 1	.758
	세트 2	.765
하위요인 1: 문제 인식 능력	원검사	.813
	세트 1	.683
	세트 2	.688
하위요인 2: 비판적 사고 능력	원검사	.704
	세트 1	.570
	세트 2	.585
하위요인 3: 창의 사고 실천	원검사	.674
	세트 1	.552
	세트 2	.529

다음으로 원 검사와 단축형 검사에서 산출된 응답자의 능력추정치 활용하여 3개 검사 간(원 검사, 단축형 검사 2개) 상관관계수 및 스피어만(Spearman) 순위상관계수를 비교하였다. 그 결

Table 11 Results of correlation analysis between original test and abbreviated tests for Problem solving

요인	구분	능력추정치		
		세트별	기준세트	세트_1
문제 인식 능력	상관계수	세트 1	.942***	
		세트 2	.947***	.827***
	순위상관계수	세트 1	.925***	
		세트 2	.933***	.788***
비판적 사고 능력	상관계수	세트 1	.878***	
		세트 2	.895***	.591***
	순위상관계수	세트 1	.865***	
		세트 2	.875***	.564***
창의 사고 실천	상관계수	세트 1	.931***	
		세트 2	.924***	.806***
	순위상관계수	세트 1	.924***	
		세트 2	.913***	.794***

\*\*\* $p < .001$

과, 문제 인식 능력, 비판적 사고 능력, 창의 사고 실천의 세 하위요인 모두에서 단축형 검사 두 세트와 원 검사 간 상관관계수 및 순위상관계수가 0.9 내외로 매우 높았다( $p < .001$ ). 아울러 단축형 검사 2개 간 상관관계수 및 순위상관계수를 살펴본 결과, 문제 인식 능력과 창의 사고 실천 요인에서는 0.8 내외, 비판적 사고 능력 요인에서는 0.5 이상의 값이 도출되었다. 상관관계수의 경우 0.5 이상이면 적절, 0.7 이상이면 강한 수준의 상관을 함의하므로(Akoglu, 2018), 문항 특성에 근거하여 배분된 단축형 검사와 원 검사 간 유사함을 확인하였다.

IV. 논의 및 결론

본 연구는 공학계열 중심인 H 대학의 역량 기반 교육 및 프로그램 질 관리 차원에서 핵심역량 진단 문항의 활용도를 높이고자, 융합사고역량 및 문제해결역량에 대한 기존 문항들을 문항반응이론에 근거하여 분석하고 이를 바탕으로 2개의 병렬 단축형 검사를 개발하였다. 그 결과, 총 32개 문항이었던 1개의 진단 도구는 21개의 간소화된 세트 2개로 분리되었고, 이를 통해 향후 다양한 영역에서 학생들의 역량 측정이 가능하게 되었다. 이제 연구 과정에서 도출된 몇 가지 한계점과 결론을 논하고자 한다.

먼저 신뢰도에 대한 한계를 들 수 있겠다. 32개의 기존 문항들을 2개 세트로 분리한 결과, 하위요인별 문항 개수가 감소한 것을 확인할 수 있었다. 즉, 기존 5개에서 최대 7개까지 구성되었던 요인별 문항이 분리 후에는 최소 3개에서 5개로 간소화된 것이다. 이는 '도구의 활용도 제고'라는 연구 목적에는 부합하나, 신뢰도 확보라는 또 다른 난제를 남겼다고 볼 수 있다. 신뢰도는 문항 수와 비례하는 것은 물론(Nunnally & Bernstein, 1994; Streiner, 2003), 적은 수의 문항으로 인간 속성을 측정할 때보다 많은 수의 문항으로 검사할 때 측정의 오차를 줄일 수 있다고 알려져 있다. 특히 신뢰도의 종류 중 Cronbach's  $\alpha$ 는 검사 도구의 신뢰도를 과소 추정하는 경향이 있기도 하므로(성태재, 2009), 2개 세트별 최종 문항 구성 후 신뢰도 계수가 0.6 미만으로 나타났던 요인에 대해서는 향후 다른 유형의 신뢰도 산출·비교 등 후속 검증과정이 필요할 것으로 예상된다.

다음은 공통문항에 관한 것이다. 개발된 병렬 단축형 세트 간에는 공통문항을 50% 이상 포함하는 하위요인(아이디어 생성 능력 등)들이 존재한다. 이는 세트 간 동등성 확보에는 유용하나, 문항 기억 효과 완화라는 또 다른 연구 목적에는 다소 반하는 부분이라고 볼 수 있다. 따라서 해당 하위요인별 문항의 추가 개발 및 Pool 형성·관리 등, 중복 완화를 위한 장·단기적 대안이 고려되어야 할 것이다.

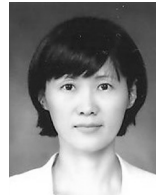
본 연구는 문항반응이론을 적용하여 공과대학생 대상의 병렬

단축형 역량진단 도구를 개발했다는 점에서 기존의 연구들과 차별성을 가지며, 역량 기반 교육의 질 관리를 목적으로 활용한다는 점에서 연구의 의의를 찾을 수 있겠다.

## 참고문헌

1. 강연옥·김향희·나덕렬(2000). 한국판 보스톤 이름대기 검사(K-BNT)의 병렬 단축형 개발. *대한신경과학회지*, 18(2), 144-150.
2. 권재기·정미경(2014). 공과대학생의 핵심역량 검사도구 개발 및 타당화. *교육방법연구*, 26(4), 687-716.
3. 김권현 외(2015). 설문지의 길이가 응답의 질에 미치는 영향. *조사연구*, 16(1), 1-48.
4. 김석우(2009). *교육평가의 이해*. 학지사.
5. 김세형·조동욱(2020). Rash 평정척도 모형을 적용한 한국과 미국 대학생 운동 동기검사지(EMI-2) 단축판 개발 및 비교. *한국체육과학회지*, 29(5), 1223-1239.
6. 김윤영·윤지영(2021). 공학계열 학생 핵심역량 진단도구 개발 및 타당화 연구. *공학교육연구*, 24(4), 3-20.
7. 김향희·김수련(2013). 문항반응이론을 적용한 한국판 보스톤 이름대기 검사 단축형(K-BNT-15) 개발. *한국콘텐츠학회논문지*, 13(12), 321-327.
8. 김화영·김현철(2019). 소규모 집단에 대한 동등화 방법 비교 : 모의실험 연구를 중심으로. *아시아교육연구*, 20(3), 797-829.
9. 박정(2001). *다분 문항반응이론 모형*. 교육과학사.
10. 박종철 외(2021). 한국판 어두운 성격 3요소 단축형의 요인구조와 타당화 지표. *한국심리학회지 산업 및 조직*, 34(3), 511-539.
11. 배병훈·이동귀·함경애(2015). Rash모형을 이용한 한국판 사회적 바람직성 척도 단축형(SDS-9)의 타당화. *상담학 연구*, 16(6), 177-197.
12. 서동기 외(2019). 단축형 심리검사 개발의 측정학적 방법과 타당화 : 한국형 역기능우울척도를 중심으로. *한국심리학회지 일반*, 38(1), 75-102.
13. 성태제(2009). *교육평가의 기초*. 학지사.
14. 신현상·이희원·김성근(2014). 교과기반의 학생 역량평가 연구. 서울과학기술대학교 공학교육혁신센터.
15. 심은은·최은실·정승철(2021). 성격 5요인검사 단축형(BFI-SF) 타당화 연구. *인지발달장애학회지*, 12(3), 105-134.
16. 윤진웅 외(2019). 단축형 스키마 척도의 신뢰도와 타당도. *신경정신의학*, 58(2), 138-145.
17. 이민정(2016). ACE사업 참여 대학의 핵심역량 및 실천 전략에 관한 연구. 박사학위논문. 경남대학교.
18. 이종승(2009). *현대교육평가*. 교육과학사.
19. 정경미·정승민·양윤정(2020). 한국판 청소년용 부모 양육스트

- 레스 검사 단축형의 타당화 연구. *한국심리학회지 일반*, 39(3), 335-357.
20. 정문주 외(2017). 한국판 5요인 마음챙김 단축형 척도 타당화 연구. *동의신경정신학회지*, 28(3), 207-216.
21. 조동현(2016). 공학계열 대학생의 핵심역량 도출과 평가지표 개발. 박사학위논문. 충남대학교.
22. 허지숙·황윤자(2020). 공학 융합역량에 대한 산업체와 대학생의 인식 비교분석. *공학교육연구*, 23(4), 3-13.
23. Anastasi, A.(1976). *Psychological testing* (4th ed.). NY: Macmillan.
24. Akoglu, H.(2018). User's guide to correlation coefficients. *Turk J Emerg Med*, 7:18(3), 91-93.
25. Hambleton, R.K., Swaminathan, H., & Rogers, H.J.(1991). *Fundamentals of Item Response Theory*. Newbury Park, CA: Sage.
26. Hinton, P.R., McMurray, I., & Brownlow, C.(2004). *SPSS Explained* (1st ed.). New York: Routledge.
27. Koch, W. R.(1983). Likert scaling using the graded response latent trait model. *Applied Psychological Measurement*, 7, 15-32.
28. Muraki, E.(1990). Fitting a Polytomous Item Response Model to Likert-Type Data. *Applied Psychological Measurement*, 14(1), 59-71.
29. Nunnally, J. & Bernstein, L.(1994). *Psychometric theory*. New York: McGraw-Hill Higher, INC.
30. Streiner, D. L.(2003) Starting at the Beginning: An Introduction to Coefficient Alpha and Internal Consistency. *Journal of Personality Assessment*, 80(1), 99-103.



**유현주 (You, Hyunjoo)**

2011년: 충남대학교 교육학 박사(교육심리학·교육과정)  
 2017년~현재: 한국공학대학교 성과관리인증센터 연구 교수  
 관심분야: 공학계열 역량평가 도구 개발, 대안적 평가 방법, 학습심리측정  
 E-mail: assess1003@tukorea.ac.kr



**남나라 (Nam, Na-Ra)**

2021년: 서울대학교 교육학 박사(교육 측정·평가)  
 2015년~현재: 한국방송통신대학교 미래원격교육연구원 선임연구위원  
 관심분야: 교육 분야 빅데이터, 프로세스 데이터, 인공지능을 활용한 교육평가  
 E-mail: naranam@gmail.com