

Rasch 모형을 적용한 AI 교육 콘텐츠 사용성 평가 검사 양호도 분석 연구

허 균[†][†]부경대학교(교수)

A Study on the Analysis of the AI Educational Contents Usability Test Appropriateness by Rasch Rating Scale Modeling

Gyun HEO[†][†]Pukyong National University(professor)

Abstract

The purpose of this study is to analyze the appropriateness of item, item difficulty, rating scale, and compare people's ability to item difficulty with the AI Educational Contents Usability Test(AI-ECUT) scale through applying Rasch Model based on the item response theory. The data are 105 students responding survey after attending to usability test. For data processing, frequency analysis, factor analysis, and Rasch model analysis are performed using jMetrk, Winsteps, and SPSS program. The results are as follows. First, it is found that the AI-ECUT scale is found to satisfy one-dimensional fitness. Second, it is found that appropriate to change the rating scale category from 5 points to 4 points Likert scale. Third, as a result of examining the item relevance of the AI-ECUT scale, the Q41 item is found to be inappropriate and needs to revise. Forth, there are lots of lower difficulty levels items compared to higher and middle difficulty levels items.

Key words : Rasch model, AI education contents, Usability test

I. 서론

최근 인공지능을 체계적으로 가르치려는 노력이 계속되어 오고 있다. 정부에서도 초·중등교육법 제23조 제2항에 기초를 두고 이러한 노력을 실행하고 있다. 구체적으로 교육부(Ministry of Education, 2020)에서는 ‘인공지능 기초’를 고등학교 전 학년에 시행하도록 계획하고 ‘인공지능의 이해’, ‘인공지능의 원리와 활용’, ‘데이터와 기계 학습’, ‘인공지능의 사회적 영향’ 등 4개의 영역

을 제시하고 있다. 이러한 배경의 한 예로 테슬라와 같은 혁신 기업이 인공지능 기술을 기반으로 자율주행을 현실화하고 전기차 보급을 확산하는 것을 들 수 있다. 현대차나 GM 등의 기존의 업체들도 현재는 사활을 걸고 추격하는 실정이다. 2016년 알파고가 이세돌 9단을 이기는 것이 전국민에게 충격을 주었다. 그 이후로 우리도 일상 생활에서 많은 AI 기술들이 확산되고 있다.

인공지능 교육 확산 노력 중 하나로 일반 학생들을 위한 AI 교육 콘텐츠 개발이 이뤄지고 있

[†] Corresponding author : 051-629-5970, gyunheo@pknu.ac.kr

* 이 논문은 2021년도 교육부의 재원으로 한국과학창의재단의 지원을 받아 수행된 성과물임.

다. 교육부와 한국과학창의재단 등과 같은 교육 관련 기관에서는 일반 초중고에서 인공지능이나 스마트 교육(Heo and Goo, 2017; Heo et al., 2017; Heo and Goo, 2018), 소프트웨어 교육을 활성화하기 다양한 콘텐츠 개발과 교육과정 개발을 지속적으로 추진하고 있다. 과학기술정보통신부, 교육부, 한국과학창의재단, 각 시도 교육청 등이 협력적 사업을 통해 전국의 500여개의 학교를 ‘AI 교육 선도학교’로 선정하여 운영하고 있다(The Korea Foundation for Science and Creativity, 2000).

정부의 지원으로 다양한 AI 교육용 콘텐츠가 개발되고 있음에도 불구하고 이를 적절히 평가할 수 있는 검사 도구는 부족한 실정이다. 또한 이러한 검사 도구가 제대로 되었는지 양호도 검사를 통해 검사 도구를 수정하고 보완하는 노력은 거의 부재한 실정이다. 기존의 다양한 형태의 사용성 평가와 관련된 연구들(Rha et al., 2003; Lee and Lee, 2004; Heo et al., 2015; Choi, 2018; Han and Heo, 2020; Jeon et al., 2020)이 있어 왔다. 특히, 교육용 모바일 앱 사용성 평가 요인 연구(Choi, 2018), 스마트러닝 앱 사용성 평가 방법 연구(Jeon et al., 2020), 웹기반 교육용 소프트웨어의 사용성 평가 기준 연구(Lee and Lee, 2004) 등은 AI 교육용 콘텐츠 사용성 평가 지표와 직접적인 연관이 있을 것으로 생각된다. 컴퓨터 기반의 콘텐츠가 인터넷이나 모바일 기기 기반으로 변화하고 그 내용에 있어서도 컴퓨팅 사고나 인공지능 등으로 다양화되고 있다. 내용의 다양화에도 불구하고 그 근본적인 검사 내용은 유사하기 때문에 검사 도구를 계속 수정하고 보완하여 발전시켜야 한다. 특히 검사 대상자가 늘어날수록 데이터를 축적하여 기준 정보가 쌓이게 되어 판정의 정확성이 더욱 높아질 수 있다. 유명한 많은 검사들이 많은 데이터를 기반으로 기준 정보를 바탕으로 정확한 진단이 이뤄질 수 있다. 하지만 사용성 평가 도구와 관련된 많은 연구들이 단발성으로 기준이나 요인을 찾는 수준에서의 연구로 그치고 있다. 또한 검사 도구나 지표를 만드는

연구가 이뤄지더라도 검사도구 개발로 끝나는 경우가 많아 양호도 검사를 통해 검사 도구를 수정하고 보완하여 검사도구를 발전시키려는 노력이 필요한 실정이다.

이에 본 연구에서는 델파이 연구를 통해 개발된 AI 교육 사용성 평가도구(Kim et al., 2022)를 기반으로 양호도 검사를 통해 검사도구를 수정하고 보완하고자 하였다. 이를 위해 측정도구를 제작하는 데 널리 활용되는 Rasch 모형을 활용하여 검사 양호도를 확인하고자 하였다. Rasch 모형은 문항반응이론(Item Response Theory)의 하나로 알려져 있으며, 측정도구의 양호도 검사와 함께 문항을 수정하고 보완하여 검사 도구를 발전시킬 수 있는 모형으로 알려져 있다. 구체적으로 문항 응답범주 수의 적절성, 문항적합도, 응답자 능력 점수와 문항난이도의 적절성 등을 확인할 수 있다. Rasch 모형을 활용한 양호도 검사를 위한 구체적인 연구문제는 다음과 같다.

- 첫째, AI 교육 사용성 검사 문항은 Rasch 모형 분석을 위한 일차원성 가정을 만족하는가?
- 둘째, AI 교육 사용성 검사 문항의 응답 범주 수는 적절한가?
- 셋째, AI 교육 사용성 검사 문항의 문항적합도는 적절한가?
- 넷째, AI 교육 사용성 검사 문항의 응답자 능력 점수에 따른 문항난이도는 어떠한가?

II. 연구 방법

1. 연구대상

초중고 학생 105명(남자 72명, 여자 33명)을 연구 대상으로 하였다. 이중 초등학생은 49명(46.7%)이었으며, 중학생은 26명(24.8%), 고등학생은 30명(28.6%) 이었다.

2. 측정도구

측정도구는 학생용 AI 교육 콘텐츠 사용성 평가 지표이다. 이 지표는 AI 교육콘텐츠의 사용성 평가를 위해 Kim et al.(2022)이 델파이 연구 방식을 통해 최종 51개의 문항을 개발하여 만든 사용성 평가 도구이다. 이 측정도구는 피검사자로 하여금 사전에 AI 교육 콘텐츠를 사용해 본 후 설문을 통해 본인이 느낀 정도를 표시하게 한다. 예를 들어, “학습 콘텐츠는 사용하기 편리하다”와 같은 문항에 대해 동의하는 정도를 5점 리커트 형태의 척도(1점: 아주 그렇지 않다. 3점: 보통이다, 5점, 아주 그렇다)로 응답하도록 되어 있다.

3. 연구 방법 및 절차

Rasch 모형을 활용한 양호도 검사를 위해 일차원성 평가, 문항 응답범주 수의 적절성, 문항적합도, 응답자 속성과 분포특성 분석과 같은 네 가지 절차를 따랐다.

첫째, 일차원성 평가를 하였다. Rasch 분석을 위해서는 일차원성 가정을 하고 있다. 요인분석의 주축분해법을 통해 변수 간의 상관을 통대로 공통문항에 대한 설명 값인 Eignenvalue를 기초로 스크리 도표를 활용하여 일차원성을 확인하였다.

둘째, 문항 응답범주 수의 적절성을 확인하였다. 본 연구에서는 평정척도모형(Rating Scale Model, 이후 RSM)을 선택하였다. Rasch 모델에서는 PCM(Partial Credit Model) 모형과 RSM 모형을 선택할 수 있다. PCM 모형은 부분점수모형으로 알려져 있으며, 인지검사에서 부분점수를 고려할 경우 많이 활용되는 모형이다. RSM모형은 다점척도 문항을 분석하기 위해 Rasch 모형을 일반화한 모형이다.

$p(x = 1) = \frac{\exp(\beta - \delta)}{1 + \exp(\beta - \delta)} \dots\dots\dots (a)$
$p_{jni} = \frac{\exp \Sigma(\beta - \delta_{ij})}{\Sigma \exp \Sigma(\beta - \delta_{ij})} \dots\dots\dots (b)$

수식 (a)는 2점 척도 문항에서 Rasch 모형을 적용한 것을 나타낸다. Rasch 모형은 학업성취도 검사에 맞춰 개발되어 능력(person ability, β)은 본 연구에서는 AI 교육콘텐츠에 대한 인식 정도를 나타낸다. 난이도(item difficulty, δ)는 문항에 대해 동의하기 어려운 정도로 생각해 볼 수 있다. $p(x=1)$ 은 응답 확률을 나타낸다.

수식 (b)는 j번째 응답범주를 개인 n의 기대되는 응답을 계산한다. 이 과정에서 RSM모형은 각 문항별 난이도(δ_{ij}) 대신 공통된 난이도(δ_j)를 추정한다(Hong and Jo, 2006). RSM은 Rasch 모형을 확장한 모형이다. 일반적으로 정의적 검사에서는 RSM 모형을 많이 활용하고 있다. 이에 분석 과정에서 두 모형을 비교하여 RSM 모형의 적합성을 확인하였다.

셋째, 문항 적합도와 적절성을 평가하여 단측관 AI 사용성 평가 문항을 구성하기 위해 척도수를 교정하였다. 이를 위해 Hong et al.(2006)의 절차에 따라 검사문항 응답 범주 수가 적절한지를 검증한다. 만약 응답 범주의 수가 적절하지 않으면 범주 수를 조절한다.

넷째, 응답자 능력 점수와 문항 난이도를 분석한다. Rasch 모형에서는 문항 난이도와 응답자 능력 모두 로짓으로 변환하여 직접 비교가 가능하다. 이를 통해서 문항 난이도가 응답자 수준에 적절한지 확인할 수 있다. 검사 도구의 특성에 따라 커트라인을 결정하는 방법으로 쓸수도 있다.

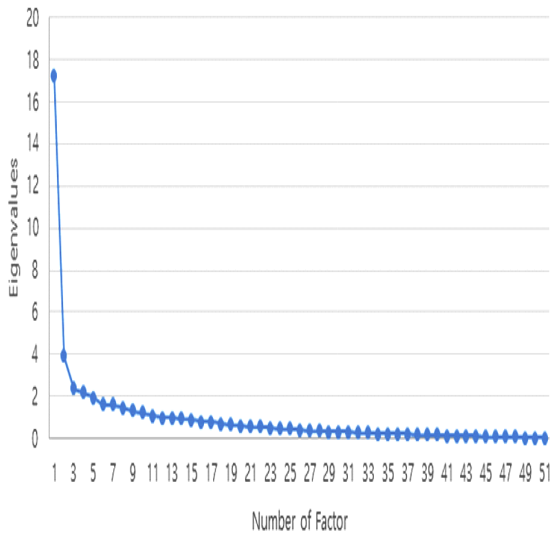
본 연구 분석을 위한 통계프로그램은 JMetrik, Winsteps, SPSS를 활용하였다.

Ⅲ. 연구 결과

1. 일차원성 평가

Rasch 모형 적용을 위해 선행요건인 일차원성(unidimensionality) 가정을 평가할 필요가 있다.

먼저, 스크리 도표를 통해 일차원성을 확인할 수 있다. [Fig. 1]은 첫 번째 아이겐 값에 비해 두 번째에서 상당한 폭으로 떨어지는 것을 시각적으로 보여준다. 이로부터 일차원성 가정이 성립한다고 가정할 수 있다. 첫 번째 요인의 아이겐 값이 나머지 값에 비해 특히 강할 때 일차원성이 성립하는 것으로 알려져 있다(DeMars, 2010).



[Fig. 1] Scree Plot

<Table 1>을 통해 변수 간의 상관을 통대로 공통문항에 대한 설명 값인 Eignenvalue 값이 1 요인에서 17.228로 이후 2 요인에서 3.915로 상당히 줄어드는 것을 확인할 수 있다. 이를 통해 일차원성을 확인할 수 있다.

2. 응답 범주 조정

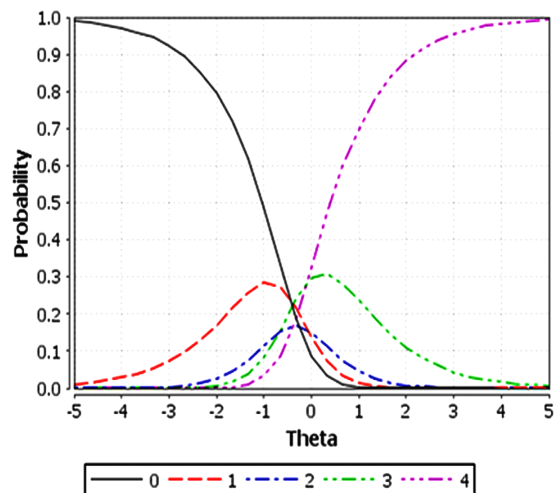
응답 범주는 조정이 필요한 것으로 나타났다.

<Table 1> Explained the Variance of Eigenvalue (Eigenvalue >1)

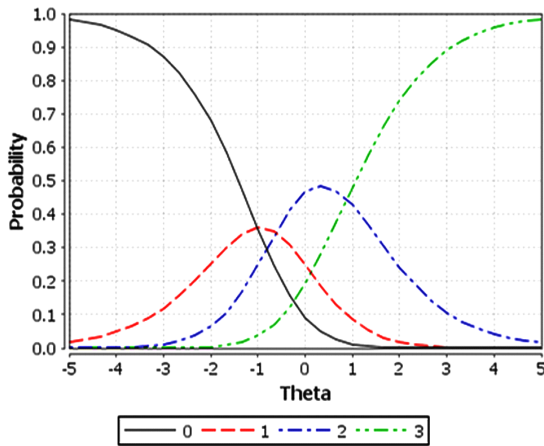
Factor	Total Initial Eigenvalue	Variance (%)	Accumuation (%)
1	17.228	33.781	33.781
2	3.915	7.676	41.457
3	2.380	4.667	46.123
4	2.168	4.251	50.374
5	1.973	3.869	54.243
6	1.614	3.165	57.408
7	1.596	3.130	60.538
8	1.460	2.862	63.400
9	1.314	2.576	65.976
10	1.213	2.378	68.354
11	1.067	2.092	70.446
12	1.001	1.962	72.408

응답 범주 분석과정에서 두 가지 문제를 발견하고 이를 해결하였다. 발견된 문제는 다음과 같다.

첫째, 척도에서 응답하지 않는 문항들이 있었다. 5점 리커르트 척도(Likert Scale)의 응답 범주를 따르고 있기 때문에 문항 응답 데이터가 없을 경우 모형 분석에서 오류가 나타났다.



[Fig. 2] 5-Likert Scale Response Category Probability Curve



[Fig. 3] 4-Likert Scale Response Category Probability Curve

둘째, 단계 모수의 비순서화(step disordering) 문제가 발생하였다. 특히 부분점수모형(PCM) 분석시 문항에 따라 비순서화 문항이 발생하여 이를 위한 조정이 필요한 사항이다. 그림은 단계모수 비순서화가 나타난 문항의 IRT그래프를 나타낸다.

두 가지 문제점을 해결을 위해 응답 점수 조정을 제안하였다. 구체적으로 5점 범주를 4점 리커트 척도로 응답범주를 조정하는 것이다. 응답 범주를 조정하고 평정척도모형(RSM)으로 문항을 분석할 경우 비순서화 문항이 발생하지 않았다. [Fig. 3]에서는 단계 모수에서 비순서화가 나타나지 않는 것을 시각적으로 확인할 수 있다.

<Table 2> The results of 4 Point Likert Category Statistics

Category	Threshold	Std. Err	Infit	Outfit
0				
1	-.76	.06	.97	1.03
2	-.37	.04	.87	.60
3	1.14	.03	.95	.97

<Table 2>의 지표값을 통해 4점 척도를 활용했을 때 한계값(Threshold)이 순차적으로 증가하고, 적합도 지수가 적절함을 확인할 수 있다.

3. 문항적합도

Rasch 모형은 문항적합도를 통해 원하는 AI 교육용 콘텐츠 사용성을 측정하지 못하는 문항을 교정하거나 제거할 수 있다.

<Table 3>은 51개의 척도 문항으로 이뤄진 AI 교육 사용성 평가 척도의 문항 난이도(Difficult)와 문항 적합도 지수 값(Infit/Outfit)을 확인할 수 있다. 일반적으로 외적 적합도 보다 내적 적합도가 더 심각한 문제를 야기할 수 있는 것으로 알려져 있으며, .7에서 1.3의 범위에 있을 때 문항적합도가 가장 적합한 상태로 알려져 있다(Bond and Fox, 2015). 여러 국내연구에서도 적합도 내적 적합도 기준을 .05에서 1.5내외로 설정(Shon, et al., 2021)하여 연구결과를 도출하고 있었다.

연구결과에서 문제가 될 수 있는 문항은 내적 적합도가 1.5를 초과하는 12, 13, 14, 22, 41번 다섯 문항이 있었다. 구체적인 문항 내용으로 12번은 “현재 학습자가 학습하고 있는 웹 브라우저 환경(크롬, 엣지, 익스플로러 등)에서 학습이 잘 이루어진다.”, 13번은 “PC, 노트북, 태블릿, 휴대폰 등에서 학습이 가능하다.”, 14번은 “검색기능이 있다.”, 22번은 “메뉴를 보면 콘텐츠에서 현재 내 위치를 알 수 있다.”로 구성되어 있다. 12, 13, 14, 22번 문항들은 내적 적합도 경계치에 있고 내용 타당서도 충분히 확보된 상황에서 수용할 수 있는 것으로 판단된다.

다만, 41번 문항은 내적 적합도가 3.38로 크게 벗어나 있어 척도에서 삭제나 수정이 요망된다. 구체적으로 41번 문항은 “학습내용이 인종, 성별, 종교, 나이, 경제적 수준 등에 따라 차별적 내용이 있다.”와 같이 복합적인 내용을 묻고 있어 삭제하여 50문항으로 구성하는 것이 타당할 것으로 생각된다.

<Table 3> AI Educational Usability Scale's Difficulty and Fitness

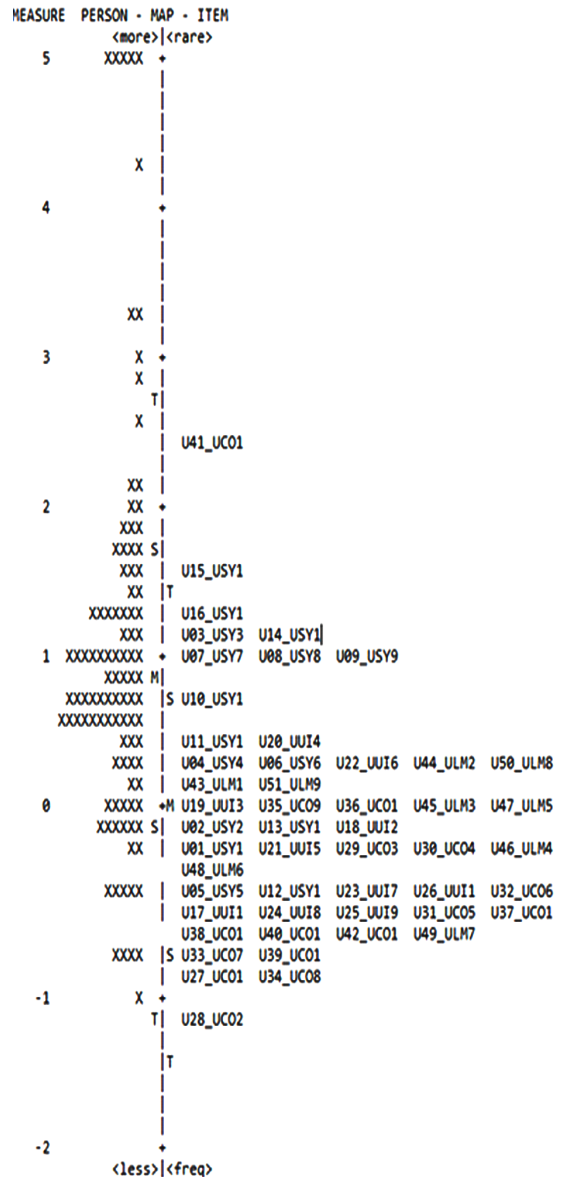
Item	Difficulty	Infit	Outfit	Item	Difficulty	Infit	Outfit
1	-0.21	0.76	0.73	26	-0.36	0.85	0.86
2	-0.13	1.03	0.92	27	-0.72	0.67	0.67
3	1.20	0.99	0.95	28	-1.01	0.76	0.75
4	0.25	1.08	1.08	29	-0.22	0.81	0.84
5	-0.42	1.22	1.28	30	-0.21	0.85	0.84
6	0.31	1.01	0.94	31	-0.56	0.59	0.55
7	1.04	1.01	0.95	32	-0.34	0.89	0.93
8	0.97	0.93	0.90	33	-0.70	0.70	0.69
9	1.01	1.20	1.22	34	-0.72	0.82	0.85
10	0.78	1.33	1.31	35	0.02	0.85	0.83
11	0.42	1.28	1.20	36	-0.01	0.87	0.81
12	-0.38	1.40	1.26	37	-0.50	0.72	0.67
13	-0.10	1.75	2.08	38	-0.54	0.80	0.73
14	1.17	1.50	1.44	39	-0.68	0.77	0.79
15	1.49	1.15	1.08	40	-0.52	0.67	0.66
16	1.31	1.10	1.00	41	0.36	3.38	9.78
17	-0.50	0.78	0.73	42	-0.50	0.64	0.58
18	-0.15	0.83	0.76	43	0.19	0.90	0.85
19	0.02	0.95	0.93	44	0.31	0.98	0.91
20	0.43	0.82	0.78	45	-0.01	0.68	0.62
21	-0.28	0.76	0.74	46	-0.22	1.16	1.31
22	0.30	1.53	2.10	47	0.07	0.94	0.91
23	-0.32	0.91	0.87	48	-0.17	0.76	0.79
24	-0.54	0.76	0.71	49	-0.52	0.74	0.73
25	-0.46	0.80	0.88	50	0.25	0.79	0.77
-	-	-	-	51	0.12	0.78	0.72

4. 응답자 능력 점수와 문항난이도 분석

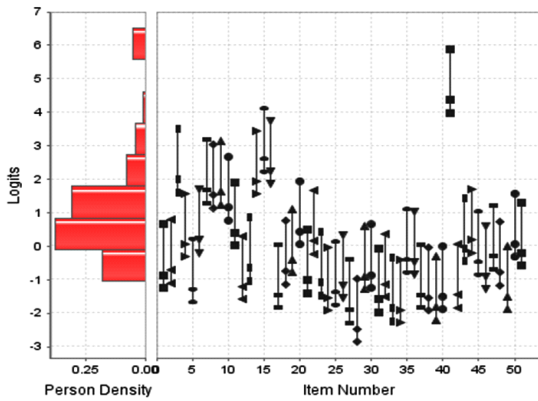
Rasch 모형에서는 문항의 난이도에 따른 응답자들의 능력 점수 분포와 대비해서 문항들의 난이도를 비교해서 확인할 수 있다. 이를 통해서

AI 교육 콘텐츠 사용성 평가에 응답한 학생들의 능력 점수와 함께 문항들의 난이도와의 관계를 파악할 수 있다.

분석결과 [Fig 4] 및 [Fig 5]와 같이 나타났다. 그림들은 문항지도의 형태로 시각화하고 있다.



[Fig. 4] Person Attribute and Item Difficulty Distribution



[Fig. 5] Person Density and Item Map

그래프의 좌측은 응답자의 능력치 속성 분포를 나타낸다. 상위로 갈수록 높은 능력을 가지고 있을 것으로 기대되는 사람들이다. 그리고 우측은 학생들의 능력치 범위를 측정할 수 있는 문항 난이도(곤란도) 등을 의미한다. 51개의 문항들이 응답범주에 맞춰 배치되어 있다.

연구결과를 통해서 상위에 있는 학생들에서 부터 초보 능력인 학생들까지 고르게 측정하는 것으로 나타났다. 특히 20번 문항 이후는 중간 이하의 능력들을 확인할 수 있는 문항들이 중복적으로 많이 배치되어 있는 것들을 확인할 수 있다.

IV. 결론 및 제언

본 연구는 선행 연구를 통해 만들어진 AI 교육 콘텐츠 사용성 평가 척도를 Rasch 모형을 적용하여 타당도의 증거와 함께 검사양호도 분석을 통한 개선 방향을 제시하는데 목적이 있다. 이 연구의 결과에 대한 시사점은 다음과 같다.

첫째, AI 교육 콘텐츠 사용성 평가 측정 도구는 일차원성을 만족하는 것으로 나타났다. 이로 부터 Rasch 모형을 적용하여 다양한 분석을 할 수 있는 기본적 가정이 만족되었다.

둘째, AI 교육 콘텐츠 사용성 평가 측정 도구

의 5점 척도 기반의 응답범주는 적합하지 않은 것으로 나타났다. 응답 결과에 따른 빈도와 적합도 지수를 고려하여 1번과 2번 범주를 수정하여 4점 척도로 수정한 경우 적합한 기준을 충족하였다. 이에 이후 연구에서는 4점 척도로 수정하여 AI 교육 콘텐츠 사용성 평가 척도를 활용해야 할 것을 제안한다.

셋째, 문항적합성 검사를 통해 AI 교육 콘텐츠 사용성 평가 척도 51개 문항 중 적합도가 적절하지 않는 문항에 대한 수정이나 삭제가 반영되어야 한다. 문항적합도는 적합성의 적절한 범위에 대한 기준(Bond and Fox, 2015)과 국내 여러 연구(Shon, et al., 2021)에서 반영되는 값 등을 종합해 문항 삭제나 수정이 요구된다. 구체적으로 41번 문항은 “학습내용이 인종, 성별, 종교, 나이, 경제적 수준 등에 따라 차별적 내용이 있다.”와 같이 복합적인 내용을 묻고 있어 삭제하여 구성하는 것이 타당할 것으로 생각된다.

넷째, AI 교육 콘텐츠 사용성 평가에 응답한 학생들의 능력 점수를 고르게 평가할 수 있는 문항난이도를 가진 문항개발이나 수정이 필요하다. AI 교육 콘텐츠 사용성 평가의 목적이 사용성에 문제가 없는지 확인하고 그 문제점을 개선하는데 있다. 따라서 설문 문항 개발에서 중간 이하의 능력에 문항들이 많은 것은 일정 부분 받아 들일 수 있다. 하지만 능력이 중간 이상인 사람들을 측정하기 위한 문항이 상대적으로 적어 이를 개선하기를 제안한다.

본 연구에서는 양호도 분석을 통한 AI 교육 콘텐츠 사용성 평가 지표의 개선 방향에 대해 살펴봐왔다. 추후 연구에서는 검사 데이터를 체계적으로 수집하여 표준 정보를 구축하여 판정의 정확성을 높이는 노력이 필요하다.

References

Bond, TG and Fox, CM(2015). Applying the Rasch Model: Fundamental Measurement in the Human

- Sciences (3rd ed.). Mahwah, NJ: L. Erlbaum.
- Choi EY(2018). Analysis of usability factors for Educational Mobile application, *Cartoon and Animation Studies*, 50, 453~478.
<http://dx.doi.org/10.7230/KOSCAS.2017.50.453>
- DeMars, C. (2010) *Item Response Theory: Understanding Statistics Measurement*. Oxford University Press, Oxford.
<https://doi.org/10.1093/acprof:oso/9780195377033.001.0001>
- Han SJ and Heo G(2020). Marine Education Web-Site Usability Analysis Case Study Busan Marine Education Guide, *The Journal of Fisheries and Marine Sciences*, 33(1), 91~98.
<https://doi.org/10.13000/JFMSE.2020.2.32.1.91>
- Heo G and Goo JM(2017). A Study on the Structural Equation Modeling of Smart Education between Fisheries Marine and General High School Students, *The Journal of Fisheries and Marine Sciences*, 29(6), 2,011~2,019.
<https://doi.org/10.13000/JFMSE.2017.29.6.2011>
- Heo G and Goo JM(2018). A Study on the Design and Development of Evaluation Based Learning System in the Field of Fisheries and Marine Education, *The Journal of Fisheries and Marine Sciences*, 30(6), 2,043~2,052.
<https://doi.org/10.13000/JFMSE.2018.12.30.6.2043>
- Heo G, Cho JH, Han SJ and Won HH(2015). A Study on the Design of Prototype for Developing Webpage of Fisheries and Marine Education, *The Journal of Fisheries and Marine Sciences*, 27(6), 1602~1609.
<https://doi.org/10.13000/JFMSE.2015.27.6.1602>
- Heo G, Goo JM and Han SJ(2017). A Meta-Analysis on the Effectiveness of Smart-Learning in the field of General Education and Fisheries & Marine Education, *The Journal of Fisheries and Marine Sciences*, 29(1), 128~136.
<https://doi.org/10.13000/JFMSE.2017.29.1.128>
- Hong SH, Jo YR(2006). Construction of a Short Version of the Dysfunctional Beliefs Test: An Application of Rasch Rating Scale Model, *The Korean Journal of Clinical Psychology*, 25(3), 865~880.
- Jeon HB, Kim HI, and Ko HY(2020). A Study on Efficiency Usability Evaluation Method for Smart Learning App, *The Journal of Image and Cultural Contents*, 19, 443~464.
<https://doi.org/10.24174/jicc.2020.02.19.443>
- Kim YY, Kim DG, Han SJ, Yun HJ, Park MH and Heo G(2022). A Delphi Study on the Development of Usability Evaluation Tools of AI Educational Contents for Students, *KSFME*, (decision in the process).
- Lee EW and Lee SJ(2004). A Study of Web-based Educational Software Usability Evaluation, *Studies on Constitutional Cases*, 17, 161~170.
- Ministry of Education (2020). Elementary and secondary school curriculum notice (No. 2020-236). 2020.9.11.
- Rha IJ, Heo G and Yu BM(2003). A Case Study of Usability Test for Developing User-centered WBI Contents, *The Journal of Korean Association of Computer Education*, 6(3), 109~117.
- Sohn WS et al(2021). Development and Validation of the Core Competency Scale for Primary Pre-service Teachers: Applications of Factor and Rasch Analyses, *Teacher Education Research*, 60(4), 683~698.
<https://doi.org/10.15812/ter.60.4.202112.683>
- The Korea Foundation for Science and Creativity (2020). Report on the results of AI education platform planning policy research, 29~32.

-
- Received : 24 February, 2022
 - Revised : 17 March, 2022
 - Accepted : 23 March, 2022