

# 컴퓨팅 사고력 평가 도구로써 비버 챌린지 문항 분석: 문항반응이론을 기반으로

김은지<sup>†</sup> · 이태욱<sup>†</sup>

<sup>†</sup> 한국교원대학교 컴퓨터교육과

## Analyzing Beaver Challenge Questions as a Computing Computing Assessment Tool : Based on Item Response Theory

Eun\_Ji Kim<sup>†</sup> · Tae-Wuk Lee<sup>†</sup>

<sup>†</sup> Dept. of Computer Education, Korea National University of Education

### 요 약

본 연구에서는 컴퓨팅 사고력 평가도구로써 비버 챌린지 문항을 활용하기 위하여 문항반응이론을 통해 비버 챌린지 문항을 분석하고, 비버 챌린지에서 기존에 제시하는 난이도와 문항반응이론을 통한 난이도 간의 상관관계를 분석하였다. 분석 결과 비버 챌린지는 쉽고 변별력이 높은 검사로 나타났으나, 비버 챌린지에서 제시하는 난이도와 문항반응이론을 통한 난이도 간의 상관관계는 없었다. 난이도에 따라 가점과 감점이 이루어지는 비버 챌린지 채점 기준을 고려할 때 정확한 컴퓨팅 사고력 측정을 위해서는 난이도에 대한 수정 및 보완이 필요하다.

## 1. 서 론

2015개정 교육과정은 창의·융합형 인재 양성을 위한 국가·사회적 요구를 바탕으로 개발되었다. 2015개정 교육과정이 적용되기 전 소프트웨어 교육과정을 운영하는 안내서인 소프트웨어 교육 운영지침은 소프트웨어 교육이 컴퓨팅 사고력을 가진 창의·융합인재 양성을 추구함을 안내하고 있다[1][2]. 컴퓨팅 사고력은 컴퓨팅의 기본적인 개념과 원리를 기반으로 문제를 효율적으로 해결할 수 있는 사고 능력을 말한다. 운영지침은 소프트웨어 교육이 이루어진 후 타당한 평가 도구를 활용하여 컴퓨팅 사고력의 향상 정도를 측정하도록 권장하고 있다[1].

최근 컴퓨팅 사고력 평가 도구로써 비버 챌린지의 사용이 증가하고 있으나 문항의 적정성에 대한 연구는 아직 이루어지고 있지 않다. 따라서 본 연구에서는 비버 챌린지 문항을 문항반응이론으로 분석함으로써 비버 챌린지의 평가도구로써의 적정성을 검토해보고자 한다. 특히, 비버챌린지에서 제시하는 난이도와 우리나라 학교급, 학령 및 교육과정 편제에 따른 난이도가 일치하는지 확인하고자 한다.

비버 챌린지는 특별한 사전 지식이 없어도 도전이 가능한 과제로 구성되며 소프트웨어에 대한 동기를 유발하고 컴퓨팅 사고력을 평가하기 위한 목적으로 개발된 과제 해결 중심의 혁신적인 평가 모델이자 교육운동이다[3]. 비버 챌린지의 모든 과제들은 정보과학의 개념 및 컴퓨팅 사고의 요소를 포함하고 있다[4].

우리나라는 한국정보과학회의 주도로 2016년 비버 챌린지에 시범 참가하였고, 2017년 11월 한국 비버 챌린지 위원회를 통해 처음으로 정식 참가하였다. 과거에 출제된 문항은 공식 웹사이트에 공개되며 공개된 과제는 교육목적으로 사용할 수 있다[2].

비버 챌린지의 가장 일반적인 형태인 비버 대회는 대부분의 참가국에서 동시에 진행되며 15~25개의 과제를 과제당 3분 내에 해결하고 그 결과로 컴퓨팅 사고력을 평가하는 방식으로 진행된다. 채점은 각 과제의 난이도와 해결 정도에 따라 가점 또는 감점하는 방식으로 진행된다[3]. 이는 선다형 또는 단답형 과제의 부작용을 막고 주어진 시간을 효율적으로 사용하는 능력을 평가하기 위해 고안된 방식이다[2].

<표 1> 평가 방식

	정답	오답	무응답
쉬움	+6점	-2점	
보통	+9점	-3점	0점
어려움	+12점	-4점	

## 2. 이론적 배경

### 2.1 비버 챌린지

비버 대회에 참가하는 학생들은 연령에 따라 다른 과제를 부여받는다. 정웅열(2017)은 비버 챌린지를 우리나라의 학교급, 학령 및 정보 교육과정 편제에 맞도록 재편성하여 아래와 같이 재구성하였다[3].

**<표 2> 비버 챌린지의 초등 편제 재구성 예시**

연령	그룹명	우리나라 학력
6~8세	그룹 I (Pre-Primary)	초 1~2학년
8~10세	그룹 II (Primary)	초 3~4학년
10~12세	그룹III (Benjamins)	초 5~6학년
12~14세	그룹IV (Cadets)	중 1학년
14~17세	그룹V (Juniors)	중 2~3학년
17~19세	그룹VI (Seniors)	고 1~3학년

비버 챌린지의 내용 체계는 정보학 개념과 컴퓨팅사 고력 요소로 구분되며, 본 연구에서는 컴퓨팅 사고력에 초점을 맞추었다. 비버 챌린지에서의 컴퓨팅 사고력 요소는 다음과 같다.

**<표 3> 비버 챌린지의 컴퓨팅 사고력 요소**

요소	세부요소
추상화	- 불필요한 세부 사항 제거 문제의 핵심 요소 발견 시스템 표현
알고리즘적 사고	절차와 규칙의 관점에서 생각하기 알고리즘 실행하기 알고리즘 생성하기
문제 분석	문제 분할 구성 요소 및 조건 분할과 통합의 구조 추론
평가	- 최선의 해결책 찾기 자원의 유용성 목적 적합성 영부에 따른 의사결정
일반화	패턴 유사성 및 연결 관계 해결 방법의 일반화

## 2.2 문항반응이론

문항반응이론(Item Response Theory; IRT)이란 검사 문항의 고유한 속성을 나타내는 문항특성곡선에 의하여 문항을 분석하는 검사 이론이다[5].

문항반응이론은 문항의 특성으로서 문항난이도, 문항변별도, 문항추측도가 피험자 집단의 특성에 의하여 변하지 않는다. 고전검사이론의 경우 문항을 분석할 때 피험자 집단의 수준에 따라 문항의 난이도가 다르게 나타날 수 있는 반면, 문항반응이론에 의하여 문항을 분석하면 문항특성이 피험자 집단의 특성에 따라 변하지 않는다[6].

또한 피험자의 능력은 어떤 검사나 문항을 선택함으로써 변하는 것이 아니라 고유한 능력 수준이 있으므로, 고전

검사이론과 달리 어떤 피험자가 어려운 검사를 택하든 쉬운 검사를 택하는 능력 추정이 같게 나타난다[6].

### 2.3.1 문항난이도

문항이 어느 능력 수준에서 기능하는 가를 나타내는 지수로서, 문항의 어려운 정도를 알려준다. 문항난이도는 '매우 쉽다', '쉽다', '중간이다', '어렵다', '매우 어렵다'의 다섯 가지 수준으로 구분할 수 있다. 문항난이도의 언어적 표현에 대응되는 문항난이도 지수에 대한 명확한 범위는 설정되어 있지 않으며 <표 4>와 같이 분류할 수 있으나 절대적 기준은 아니다[6].

**<표 4> 언어적 표현에 의한 문항난이도 범위**

언어적 표현	문항난이도 지수
매우 쉽다	-2.0 이하
쉽다	-2.0 ~ -.5
중간이다	-.5 ~ +.5
어렵다	+.5 ~ +2.0
매우 어렵다	+2.0 이상

### 2.3.2 문항변별도

문항이 피험자를 능력에 따라 변별하는 능력에 따라 변별하는 정도를 나타내는 지수를 말한다. 문항변별도는 일반적으로 0에서 +2의 값을 가지며 높을수록 양질의 문항이다. 문항변별도가 음수 값을 지니면 검사에서 제외해야 한다. 문항변별도의 언어적 표현에 대응하는 문항변별도 지수의 범위를 엄밀하게 설정할 수 없으나 Baker(1985)는 <표 5>과 같이 설정하였다[6].

**<표 5> 언어적 표현에 의한 문항변별도 범위**

언어적 표현	로지스틱모형
없다	.00
거의 없다	.00 이상 ~ .34 미만
낮다	.35 이상 ~ .64 미만
적절하다	.65 이상 ~ 1.34 미만
높다	1.35이상 ~ 1.69 미만
매우 높다	1.70 이상
완벽하다	+∞

## 3. 연구방법

### 3.1 연구 대상

본 연구는 2017년 12월 경남 소재의 J초등학교 6학년을 대상으로 진행되었다. J초등학교는 소프트웨어 교육 선도학교로 지정되어 2년 동안 소프트웨어 교육을 실시하였기 때문에 본 연구에 참여한 6학년 학생들은 2년간 총 68시간의 소프트웨어 교육 경험이 있다. 응답자수는 96명이다.

### 3.2 연구 방법

정용열(2017)의 비버 챌린지 편제 재구성 예시에 따라 초 5~6학년에 해당하는 그룹 III을 선택하였으나, UK의 경우 그룹 III을 'Junior'로 지칭하므로, 본 연구에서는 UK의 Junior에 해당되는 문항을 사용했다는 점에서 차이가 있다.

비버 챌린지의 채점 방식에는 난이도에 의한 가점과 감점 방식이 있으나 본 연구에서는 문항의 난이도와 변별도를 구분하는데 중점을 두기 위하여 문항별로 정답여부만을 통해 분석하였다. 본 연구에서 사용한 비버 챌린지 문항의 난이도와 평가하고자 하는 컴퓨팅 사고력 요소는 <표 6>과 같다.

<표 6> 비버 챌린지 문항

문항 번호	문항	난이도	컴퓨팅 사고력
1	2015 beaver gates	하	추상화, 평가
2	2015 crane operating		알고리즘
3	2016 Maze		알고리즘, 분해
4	2015 setting the table	중	알고리즘, 평가
5	2014 village network		추상화, 알고리즘, 분해
6	2016 secret recipe		알고리즘, 분해
7	2015 cross country		평가, 알고리즘
8	2015 quick beaver code		추상화, 평가, 분해
9	2015 drawing patterns	상	알고리즘
10	2015 throw the dice		평가, 알고리즘
11	2014 loading lisas		평가, 분해
12	2016 Bottles	하	추상화, 평가
13	2016 tube system		알고리즘
14	2016 party guests		알고리즘, 분해
15	2016 soccer game		알고리즘, 평가
16	2014 toothbrushes		추상화, 알고리즘, 분해
17	2016 car trip		알고리즘, 분해
18	2015 trains		평가, 알고리즘
19	2015 geocaching	중	추상화, 평가, 분해
20	2016 concurrent directions		알고리즘
21	2016 magic potions		평가, 알고리즘
22	2016 bus stop		평가, 분해

### 3.2.3 분석방법

문항반응이론에 의한 문항 분석을 위하여 Bayesian 통계방법에 의하여 문항과 검사를 분석하는 BayesiAn(에버케이션, 2000)을 사용하였다. Bebras에서 제시하는 난이도와 문항반응이론에 의한 난이도 간의 상관관계를 알아보기 위하여 Spearman 상관분석을 사용하였으며, 검정을 위한 유의수준은 .01이하로 하였다. SPSS(Statistical Package for Social Science for Window) version 21.0을 사용하였다.

### 3.3 연구 결과

#### 3.3.1 문항반응이론에 의한 문항 분석

문항반응이론에서 사용하는 모형은 난이도, 변별도, 추측도 등 문항 속성 중 몇 개를 고려하느냐에 따라 1모수, 2모수, 3모수 모형으로 분류되는데[7], 본 연구에서는 난이도와 변별도만 고려하는 2모수 모형을 사용하였다.

가장 변별도가 높은 문항은 7번 1.253이고 가장 변별도가 낮은 문항은 3번 .555이다. 가장 어려운 문항은 5번 2.798이며 가장 쉬운 문항은 13번 -2.709이다. 문항의 평균 난이도는 -.173으로 쉬운 검사라 할 수 있고 평균 변별도는 .947으로서 변별도가 높은 검사이다.

<표 7> 문항반응이론에 의한 난이도와 변별도

문항 번호	난이도	변별도
1	-0.283	0.929
2	1.049	1.052
3	2.346	0.555
4	-2.447	1.168
5	2.798	0.59
6	1.504	1.17
7	-0.05	1.253
8	-1.012	0.937
9	0.354	1.149
10	-1.045	1.048
11	0.373	0.845
12	0.619	0.932
13	-2.709	0.786
14	-0.585	1.239
15	-0.756	0.795
16	-0.083	0.824
17	0.378	1.028
18	-0.622	1.015
19	-0.283	0.931
20	-1.967	0.738
21	-0.866	1.061
22	-0.513	0.783

<표 8> 문항반응이론에 의한 난이도

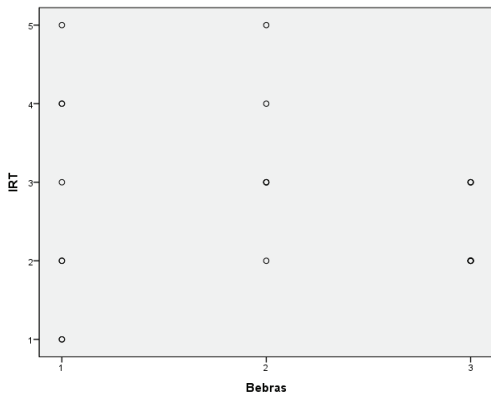
언어적 표현	지수	문항수	문항번호
매우 쉽다	-2.0 이하	2	4, 13
쉽다	-2.0 ~ -0.5	8	8, 10, 14, 15, 18, 20, 21, 22
중간이다	-0.5 ~ +0.5	7	1, 7, 9, 11, 16, 17, 19,
어렵다	+0.5 ~ +2.0	3	2, 6, 12
매우 어렵다	+2.0 이상	2	3, 5

<표 9> 문항반응이론에 의한 변별도

언어적 표현	지수	문항수	문항번호
낮다	.35 이상 ~ .64 미만	2	3, 5
적절하다	.65 이상 ~ 1.34 미만	20	1, 2, 4, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22

### 3.3.2 난이도 상관관계 분석

비버챌린지의 난이도와 문항반응이론에 의한 난이도 간의 상관관계는 -.132로 나타났으나 유의확률 .559로 상관 유의수준 .01 수준에서 유의하지 않았다. 이 관계는 양자 간의 산점도 <그림 1>에서도 확인할 수 있다.



## 4. 결론 및 논의

본 연구는 비버 챌린지 UK의 문항을 소프트웨어 교육경험이 있는 초등학교 6학년 96명을 대상으로 평가하고, 그 결과를 문항반응이론에 의해 분석함으로써 비버챌린지 문항의 난이도와 변별도에서의 적정성을 검토하였다.

문항 반응이론의 분석에 의하면 본 연구에서 사용한 비버 챌린지 문항에 변별도가 낮아 제외하거나 수정해야 할 문항이 없었다. 따라서 비버 챌린지 문항을 통한 컴퓨팅 사고력 측정이 적합하고 타당한 것으로 볼 수 있다.

하지만 비버 챌린지에서 제시하고 있는 난이도와 문항반응이론에 의한 난이도 간에는 상관관계가 없었다. 이는 본 연구에서 사용한 문항이 UK의 문항이기 때문으로 추측할 수 있다. 난이도는 전반적으로 쉬운 편이었기 때문에 정웅열(2017)이 제시한 것과 같이 학교급 및 학령에는 적절하나, 교육과정 편제에 의해 난이도 차이가 나는 것으로 보인다. 만약 기존의 채점 방식인 난이도에 따른 가점 및 감점 방식으로 채점한다면 우리나라에서의 컴퓨팅 사고력 평가 결과는 신뢰하기 어려울 것이다. 따라서 현행에서 컴퓨팅 사고력 평가의 신뢰도를 높여려면 가점 및 감점 없는 채점이 이루어져야 하고, 보다 신뢰로운 평가를 위해서 우리나라의 교육과정 편제에 따른 난이도 조정을 위한 연구가 필요하다.

또한, 문항반응이론에 의한 분석은 피험자 집단의 능력에 따라 문항 난이도와 변별도가 달라지지 않으며 검사 도구의 난이도에 따라 피험자 능력이 과소 또는 과대 추정되지 않으므로 본 연구의 결과가 의미 있지만, 본 연구의 대상이 초등학교 6학년이라는 한정된 학년이라는 점, 소프트웨어 교육 경험이 상대적으로 많은 학생들이라는 점에서 한계점을 가진다. 또한, 연구에 활용한 문제 또한 임의 선별되었다는 점 또한 한계점으로 볼 수 있다. 따라서 대상의 범위를 확대하여 추가적인 문항 분석이 필요하다.

## 참고 문헌

- [1] 교육부. (2015. 2). 소프트웨어 교육 운영 지침.
- [2] 정웅열, &이영준. (2017). 정보 교육에서 비버 챌린지(Bebras Challenge)의 활용 가능성과 향후 과제. 컴퓨터교육학회논문지 = The Journal of Korean association of computer education, 20(5), 1-14.
- [3] Bebras Challenge. (2017). <http://bebraschallenge.org>.
- [4] Dagiene, V., &Futschek, G. (2008). Bebras International Contest on Informatics and Computer Literacy: Criteria for Good Tasks. Lecture Notes in Computer Science, -(5090), 19-30.
- [5] 박도순, 권순달, 김석우, 김옥남, 김진규, 김형주, 남현우, 백순근, 양길석, 이기중, 이원석, 이현주, 조지민, 김명화, 김영애, 김종필, 김진하, 남명호, 박완성, 성태제, 어윤경, 이나현, 이종승, 장은아, &최대호. (2012). 교육평가. 과주: 교육과학사.
- [6] 성태제. (2004). 문항제작 및 분석의 이론과 실제. 서울: 학지사.