

일반화부분점수 모형에 의한 디자인역량 검사 특성 분석

이 대 용[†]
(북면초등학교)

An Item Characteristic Analysis of Competency Inventory for Designer via Generalized Partial Credit Mode

Dae-Yong LEE[†]
(Bukmyeon Primary school)

Abstract

This study was performed to analyze the item characteristics of competency inventory for designer (CID), which Gil (2011) developed for measurement of design competency. To accomplish the purpose of this study, general partial credit (GPC) model based on polytomous item response theory was applied. The findings were as follows: First, CID is a reliable instrument for measuring design competency. Second, most items of CID have low item discrimination and average item difficulty according to the GPC model. Especially, there are some problems to have low item discrimination in view of validation. To improve the goodness of CID, we will need to examine why CID has low item discrimination.

Key words : Competency inventory for designer(CID), General partial credit (GPC) model

I. 서 론

디자인 산업은 급격하게 성장하고 있다. 현재 디자인 산업 규모가 8조원을 넘어섰고, 이것은 세계 8위에 해당된다. 이런 추세는 전문 디자이너 양성에도 영향을 미쳐 한 해 디자인 전공자 졸업생 수가 25,000명 이상이다(KIDP, 2011).

그러나 영국의 디자인산업 규모와 비교할 때, 이러한 현황을 해결하기 위한 정책적 고려가 요구된다. 영국의 디자인산업 규모는 20조원이 넘는 데 비하여, 한해 디자인 전공자 배출은 19,000명 정도이다. 이점을 통하여 볼 때, 디자인 산업 규모가 대략 3배가 큰 영국에 비하여 우리나라의 디자인 전공자의 배출이 상당히 많다는 것을 알

수 있다. 산업 규모에 비하여 전공자 배출이 많아지게 되면, 디자인 전문 인력의 수요나 유용성은 낮게 된다. 결국 잉여 인력들은 디자인 전공 분야와 관련 없는 다른 직업을 구하거나, 실업자로 전락하게 될 것이다.

교육적인 관점에서 이러한 사태에 대한 대처 방안은 진로교육에서 찾을 수밖에 없을 것이다. 즉, 디자인을 전공하고자 하거나 전공을 하는 학생들에게 본인의 적성이 고려된 진로선택인지에 대한 근거 있는 정보를 제공해 주는 것이다. 이를 위해서는 디자인 적성을 검사할 수 있는 측정 도구가 필요하다.

현재 고무적으로 디자인 적성을 검사하기 위한 표준화 검사지가 개발되어 상용화되고 있다(Gil

[†] Corresponding author : 055-604-2636, daeyongc@daum.net

Im-Ju, 2006, 2014). 길임주 외의 디자인 진로 적성 검사(2006)는 2개의 하위 척도로 구성된다. 즉 디자인에 대한 기본적인 능력을 검사하는 기초소양 검사와 디자인의 전공분야를 판별해주는 전공소양 검사가 있다. 특히 기초소양 검사는 일반적인 교육을 받는 중등학교 학생들의 디자인 진로선택에 유용하게 활용될 수 있다. 물론 디자인을 전공하는 대학생들에게도 디자인 적성을 점검하는 데도 유용할 것이다.

최근 디자인 진로 적성 검사는 대학의 비실기 입시전형의 증가로 인하여 디자인 역량 검사로 개선되었다(Gil Im-Ju & Yi S, 2014). 비실기 전형은 통합적 문제해결력과 잠재적 디자인 적성을 중요시하기 때문에, 이에 대한 보다 정확한 정보가 요구된다. 그리하여 보다 문항의 개수가 보강된 디자인 역량 검사가 개발된 것이다.

디자인 역량 검사도 디자인 적성 검사와 같이 2개의 하위 척도로 구성되었다. 디자인 역량 검사는 그 개발 과정에서 신뢰도와 구인타당도에 대한 검토가 이루어졌다. 그러나 검사 도구의 문항에 대한 충분한 분석은 이루어지지 않았다.

교육측정학적 관점에서 보면, 검사를 개발하여 완성하는 과정에서 문항분석은 반드시 포함되어야 한다. 비록 타당하고 신뢰로운 검사 도구가 개발되었다 할지라도, 피험자의 반응을 근거로 문항분석이 면밀히 이루어 지지 않으면, 그 검사 결과는 유의미한 정보가 되기 어렵다. 즉 검사도구가 피험자의 특성을 얼마나 잘 변별해 주는지, 검사도구의 척도가 피험자의 다양한 반응을 반영하기에 적절한지 등을 파악할 수 있다면, 문항분석은 검사의 질을 향상시키는데 상당한 기여할 것이다(Koo Seul-Ki, 2011). 우리나라의 디자인산업의 과다인재 배출 상황을 고려한다면, 디자인 역량 검사에 대한 문항분석은 분명 요구된다.

문항특성을 탐색하기 위해서는 검사의 특성이 고려된 검사이론이 필요하다. 검사이론은 피험자의 반응을 고려하여 검사도구의 문항에 대한 변별도와 난이도, 추측도 또는 신뢰도를 확인하게

된다. 검사이론에는 피험자가 검사도구를 통하여 획득한 총점을 바탕으로 문항의 특성을 분석하고자 하는 고전검사이론과 검사도구의 문항과 피험자의 고유한 특성을 고려하여 문항의 특성을 파악하고자 하는 문항반응이론이 있다. 고전검사이론은 피험자의 특성에 따라 문항의 특성이 다르게 제시된다는 단점으로 인하여 최근에는 많이 사용되지 않는 편이다.

본 연구의 목적은 디자인 역량 검사 중 기초소양 검사(이하: 디자인 역량 검사)에 대한 문항특성을 탐색하는 것이다. 본 연구에서 기초소양 검사만 선정한 이유는 디자인 전공을 하고자 하는 학생들에게 그들의 디자인 적성에 대한 보다 정확하고 유용한 정보를 제공하는 데 있기 때문이다.

본 연구에서는 새롭게 개발된 디자인 역량 검사의 문항 분석을 위하여 문항반응이론 중 일반화부분점수 모형을 적용하였다. 이는 디자인 역량 검사가 Likert 척도를 적용한 다분 문항이고, 일반화부분점수 모형이 다분문항반응 모형들 중 가장 현실성이 많이 반영되었다고 평가받기 때문이다.

본 연구는 디자인 역량 검사의 결과를 해석하는 데 보다 엄밀한 정보를 제공하기 위한 문항분석을 통해서 검사 결과의 해석에 상당한 기여를 할 것이다.

본 연구의 목적을 달성하기 위하여 다음과 같은 연구문제를 설정하였다.

첫째, 디자인 역량 검사의 신뢰도는 어떠한가?

둘째, 일반화부분점수 모형에 의한 디자인 역량 검사의 문항특성은 어떠한가?

II. 이론적 배경

1. 일반화부분점수 모형

검사를 구성하는 각 문항에 초점을 두고, 문항의 고유한 특성에 의해 피험자의 잠재적 능력이나

문항의 특성을 추정하는 검사이론을 문항반응이론이라 한다(Korean Society of Educational Evaluation, 2004).

문항반응이론은 문항에서 피험자의 반응범주가 어떻게 구성되었느냐에 따라서 이분문항반응 모형과 다분문항반응 모형으로 나누어진다. 이분문항반응 모형은 피험자의 반응이 2개, 즉 정답과 오답으로 한정된 경우에 적용된다. 그러나 다분문항반응 모형은 피험자의 반응이 정답과 오답이 아닌, 3개 이상의 범주일 때에 적용된다. 이론적으로는 다분문항반응 모형에서 피험자가 선택하게 되는 각 반응범주는 이분문항반응모형의 이분 반응범주의 연장으로 이해될 수 있다.

다분문항반응 모형에서는 각 문항에 대한 피험자의 응답을 문항반응범주라 간주한다. 문항반응범주란 한 문항에 대하여 피험자가 반응할 수 있는 가능한 종류를 의미하는데, 연속적인 범주들을 연속적인 숫자로 점수화하여 나타낼 수가 있다(Korean Society of Educational Evaluation, 2004). 대표적인 예로는 각 중 심리검사지에서 주로 활용되는 Likert 척도가 있다.

다분문항반응 모형은 Samejima(1969)의 등급반응모형이 개발된 이후로 다양한 모형들이 개발되었다. 본 연구에서 디자인 역량 검사의 문항분석을 위하여 적용하고자 하는 다분문항반응 모형은 일반화부분점수 모형이다.

일반화부분점수 모형은 Master(1982)의 부분점수 모형을 확장하여 개발한 것이다. 부분점수모형은 이분문항에 적용되는 Rasch모형을 다분 문항에 적용할 수 있도록 변형하여 확장한 것이다. 그리하여 부분점수모형도 문항난이도만 추정하는 Rasch모형과 같이 문항범주별 단계 난이도만을 추정한다. 부분점수모형은 서술형 문항에서 부분 점수를 부여할 경우, 등급반응모형과는 달리 문항의 각 범주 간 서열화가 되도록 가정하지는 않는다는 장점이 있다. 그러나 부분점수모형은 모든 문항의 변별도를 고정하기 때문에, 이러한 제한은 모형의 현실 적용 가능성을 낮게 한다. 이

를 극복하기 위하여 개발된 모형이 일반화부분점수모형이다.

일반화부분점수모형은 문항특성을 추정할 시 각 문항의 변별도를 고려하면서 문항의 반응범주별 난이도를 서열화 하도록 가정하지 않기 때문에, 현실성이 많이 반영된 모형으로 평가된다.

다음 식 2.1은 일반부분점수모형에 대한 공식으로 능력이 θ 인 피험자가 문항 i 에서 반응범주 h 를 선택할 확률을 나타내는 것이다. 식 2.1에서 D 는 로지스틱 모형의 능력척도에 놓인 척도 상수를 의미하며, a_i 는 문항 변별도 모수, b_{iv} 는 문항의 범주난이도 모수 또는 단계 난이도 모수이다. 일반화부분점수모형에서 b_{iv} 는 두 개의 응답지 중에서 하나를 선택하는 경계선을 의미하므로 문항범주난이도라고 하기보다는 단계난이도라고 부르기도 한다(Park C, 2001).

$$P_{ih}(\theta) = \frac{\exp[\sum_{v=1}^h Da_i(\theta - b_{iv})]}{\sum_{c=1}^{m_i} \exp[\sum_{v=1}^c Da_i(\theta - b_{iv})]} \quad (2.1)$$

일반화부분점수모형에서의 모수추정은 주변최대우도추정법에 EM 알고리즘을 더한 MML-EM

(Marginal Maximum Likelihood-Expectation Maximization)방법에 의해서 구해진다. MML-EM 추정법은 PARSCALE 프로그램에 적용되었다(Muraki & Bock, 1998). 본 연구에서도 일반화부분점수모형의 모수추정을 위해서 PARSCALE 프로그램을 활용하였다.

2. 정보함수

정보함수는 문항반응이론에서 각 문항 또는 검사가 전체 능력 수준의 범위에서 피험자의 능력을 얼마나 정확하게 추정하는지를 알려주는 지표이다(Seong Tae-Je, 2001). 즉 정보함수는 고전검사이론에서의 신뢰도와 비슷한 개념으로서 문항의 적절성을 선별하여 좋은 검사를 만드는 준거

가 된다.

문항반응이론의 정보함수에는 문항정보함수와 검사정보함수가 있다. 문항정보함수는 각 문항이 제공하는 정보를 말하며, 검사정보함수는 검사가 제공하는 정보를 의미한다.

다분문항반응 모형의 문항정보를 얻기 위하여서는 다음 식 2.2와 같이 Donoghue(1994)의 공식을 활용한다.

$$I_i(\theta) = D^2 a_i^2 \left[\sum_{k=0}^{m_i} k^2 P_{ik}(\theta) - \left(\sum_{k=0}^{m_i} k P_{ik}(\theta) \right)^2 \right] \quad (2.2)$$

위의 식 2.2에서 k 값이 문항 i 의 부분점수를 의미한다면, $I_i(\theta)$ 는 능력 θ 에서 $D^2 a_i^2$ 과 문항 i 의 정답확률의 분산을 곱한 값을 의미한다. 즉 문항의 변별도를 고려하여 전체 능력범위 수준에서 어떤 피험자의 정답확률의 분산이 크면, 문항의 정보량이 많다는 것을 의미하게 된다. 피험자의 정답확률의 분산이 크기 위해서는 문항의 변별도와 적절한 난이도가 고려되어야 한다. 즉 변별도가 높고 적절한 난이도를 갖춘 문항은 피험자의 능력에 대한 더욱 정확한 정보를 제공하게 된다. 그리하여 문항특성이 고려된 문항분석을 위해서는 정보함수를 추정할 필요가 있는 것이다.

검사정보함수는 각 문항 정보함수의 합이다. 즉 문항반응이론에서는 문항간 지역독립성을 가정하므로, 각 문항의 정보함수의 합은 검사정보함수의 값에 해당된다. 검사정보함수는 다음 식 2.3과 같이 표현될 수 있다.

$$I(\theta) = \sum_{i=1}^m I_i(\theta) \quad (2.3)$$

이러한 검사정보함수는 디자인 역량 검사의 효과성을 평가하는 데 유용한 지표가 될 것이다.

피험자의 능력추정치의 표준오차는 능력추정치 분산의 제곱근의 역에 해당되어 다음 식 2.4와 같이 표현될 수 있다.

$$SE(\theta) = \frac{1}{\sqrt{I(\theta)}} \quad (2.4)$$

III. 연구 방법

1. 연구대상

본 연구는 전국 디자인 대학을 재학 중인 학생 1,205명을 대상으로 디자인역량 검사를 실시하였다(남학생: 322명(26.70%), 여학생 883명(73.30%); 1학년:321명(26.60%), 2학년: 358명(29.70%), 3학년: 351명(29.10%), 4학년:175명(14.50%)).

2. 검사도구

본 연구는 디자인 역량 검사(Gil Im-Ju & Yi S, 2014)에 대한 문항분석을 실시하는 것이다. 디자인 역량 검사는 대학의 디자인 학과에서 비실기 대입전형을 확대하는 것에 대하여 디자인 적성에 대한 검사결과의 타당성을 높이기 위해서 기존의 디자인 적성 검사(Gil Im-Ju & Yang S. Y, 2006)를 개선하여 개발된 것이다.

현재 디자인 역량 검사는 상용화되어 판매되기 때문에 검사의 전문을 소개하지는 못한다. 디자인 역량 검사의 문항 구성은 <Table 1>과 같다.

<Table 1> Items Construction of Competency Inventory for Designer(CID)

domain	factor	item	n
exp- ression	sketching	A3, A10, A14, A17, A27, A31	6
	formative	A7, A13, A21, A35	4
	aesthetic sensitivity	A5, A12, A24, A26, A38	5
creativity	software applicability	A4, A11, A18, A25, A32, A39	6
	imaginative	A6, A19, A20, A33, A34, A40, A44, A49, A50, A54, A55	11
	creativity	A2, A9, A16, A23, A30, A48	6
planning	planning	A1, A8, A15, A22, A28, A29, A36, A37, A42, A43, A45	11
	debating	A41, A46, A47, A51, A52, A53	6
sum			55

<Table 1>에서 보는 바와 같이 디자인 역량 검사는 총 55개의 문항으로 구성되어 있다. 하위 척도로 3개 영역에 8개의 하위 요인으로 구성되어 있다(하위요인: 소묘력, 공간표현력, 미적 감수성, 소프트웨어 사용력, 연상 전환력, 창의성, 기회력, 발표 및 토론 능력). 반응범주는 Likert 6점 척도로 되어 있다(1: 전혀아님~6: 매우그림).

3. 자료분석 절차

본 연구의 자료분석과 모수추정은 다음과 같은 절차로 수행되었다.

첫째, 문항반응이론은 검사의 일차원성과 지역 독립성을 가정한다. 이에 본 연구는 디자인 역량 검사에 대한 일차원성을 점검하였다. 이를 위하여 요인분석이 적용되었다.

둘째, 디자인 역량 검사의 신뢰도를 검토하기 위하여 고전검사이론의 알파계수와 문항반응이론의 정보함수를 살펴보았다.

셋째, 디자인 역량 검사에 대한 일반화부분점수 모형을 적용하여 문항특성을 살펴보았다. 본 연구가 디자인 역량 검사의 문항분석을 위하여 일반화부분점수모형을 적용하게 된 이유는 첫째, 디자인 역량 검사의 문항이 리커트 척도로 구성되었고, 둘째, 일반화부분점수 모형은 문항특성을 살피기 위해서 문항범주별 그 서열성에 대한 가정이 하지 않는다는 점이다. 즉 Likert 척도는 이론적으로 각 문항의 반응범주별 응답확률이 같다는 가정 하에 구성된다. 이에 대한 점검을 위하여서는 반응범주별 서열화를 가정하지 않는 일반부분점수모형을 적용하는 것이 적절하다고 여겨진다. 일반화부분점수 모형에 의한 문항의 모수추정을 위하여 PARSCALE 프로그램을 활용하였다.

마지막으로, 본 연구는 문항모수의 해석의 적절성을 위하여 추정된 정보함수를 그래프로 변환하여 살펴보았다(Masters, 1982; Muraki et al., 1998). 이를 위해서 IRTGRAPH 프로그램을 사용하였다.

IV. 연구 결과

1. 일차원성 검증

문항반응이론을 활용하여 문항분석을 하려면 우선 일차원성과 지역독립성이라는 두 가정이 만족되어야 한다. 본 연구는 디자인 역량 검사의 일차원성 가정을 검증하기 위해서 요인분석을 실시하였다. 요인추출은 최대우도법을 적용하였다.

<Table 2>에서 보는 바와 같이, 디자인 역량 검사의 제 1요인의 고윳값은 제 2요인의 것에 비하여 상당한 차이로 크다. 대체로 첫 번째 요인의 고윳값이 다른 요인의 고윳값에 비해 상대적으로 크다면 일차원성이 확보된 것으로 간주된다(Hattie, 1985). 또 제 1요인은 전체 검사점수의 분산(설명량)에 대하여 30.66퍼센트(%)를 설명하고 있다. Reckase 외(1988)는 첫 번째 요인이 전체 검사점수 분산의 20퍼센트(%) 정도를 설명할 수 있으면 문항반응이론을 활용할 있다고 말한다.

이에 본 디자인 역량 검사는 일차원성을 확보하였다고 여겨진다. 또한 일차원성 가정이 만족되면, 일반적으로 지역독립성 가정도 만족한다고 본다(Kang Tae-Hun et al., 2010).

<Table 2> Factor Explanation of CID(N=1,205)

factor	eigen-value	cumulative explanation(%)	factor	eigen-value	cumulative explanation(%)
1	16.86	30.66	5	1.68	3.05
2	3.65	6.63	6	1.52	2.76
3	2.73	4.95	7	1.30	2.36
4	1.94	3.53	8	1.07	1.95

2. 신뢰도 검증

가. 알파계수

디자인 역량 검사의 8개의 하위 요인별 신뢰도는 <Table 3>과 같다. 디자인 역량 검사의 전체 신뢰도(알파계수)는 .96이고, 하위 요인의 신뢰도의 범위는 .79에서 .89까지이다. 특히 이들 중 조

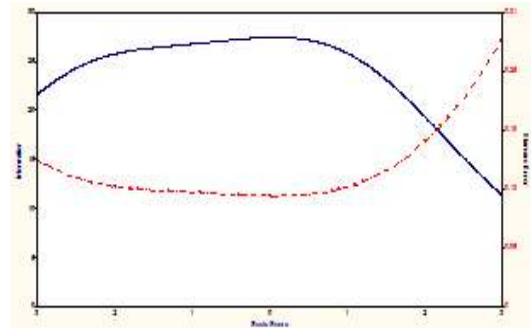
형력에 대한 신뢰도가 가장 낮았고, 기획력이 가장 높았다. 이를 종합하면, 디자인 역량 검사는 디자인 적성을 신뢰롭게 측정한다 할 것이다.

나. 정보함수

고전검사이론의 신뢰도와 같은 개념에 해당되는 검사정보함수는 검사가 능력 수준의 전체 범위에서 피험자의 능력을 얼마나 정확하게 추정하는가에 대한 정보를 알려주는 지표이다(Hambleton et al., 1991; Muraki, 1992). 즉 이 지표의 수치가 높으면 검사도구가 피험자의 능력을 정확히 추정한다는 것을 알 수 있다.

<Table 3>에는 디자인역량 검사의 하위척도별 검사정보함수가 제시되어 있다. 디자인역량의 하위척도인 소묘력의 검사정보의 범위는 1.06에서 2.58까지였다. 능력수준으로 볼 때, 0에서 검사정보가 가장 높았고 3에서 가장 낮았다. 다음 공간표현력의 검사정보는 .87에서 2.00까지 범위였다. 이를 능력수준으로 볼 때 0에서 검사정보가 가장 높았고, 3에서 가장 낮았다. 다음 미적감수성의 검사정보의 범위는 .98에서 2.55까지였다. 이를 능력수준으로 살펴보면 -2에서 가장 높았고, 3에서 가장 낮았다. 다음 소프트웨어 사용능력의 검사정보의 범위는 .87에서 1.37까지이다. 이를 능력수준에서 볼 때, 0에서 가장 높았고, 3에서 가장 낮게 나타났다. 다음 연상 및 전환능력의 검

사정보의 범위는 1.70에서 3.74까지였다. 이를 능력수준으로 살펴보면 0에서 가장 높았고, 3에서 가장 낮았다. 다음 창의성의 검사정보의 범위는 1.80에서 5.76까지였다. 이를 능력수준으로 살펴보면 0에서 가장 높았고, 3에서 가장 낮았다. 다음 발표 및 기획력의 검사정보의 범위는 2.92에서 6.88까지였다. 이를 능력수준으로 살펴보면 0에서 가장 높았고, 3에서 가장 낮았다. 마지막으로 토론능력의 검사정보의 범위는 1.07에서 2.71까지였다. 이를 능력수준으로 살펴보면 0에서 가장 높았고, 3에서 가장 낮았다.



[Fig. 1] Test information curve of CID

[Fig. 1]은 디자인 역량 검사의 전체 검사정보함수를 나타낸 것이다.

<Table 3> Cronbach α and Each Subtest Information function of CID(N=1,205)

subtest	ability level (θ)							Mean	item n	Cronbach α
	-3	-2	-1	0	1	2	3			
sketching	2.00	2.45	2.62	2.58	2.30	1.71	1.06	2.10	6	.83
formative	1.53	1.85	1.95	2.00	1.85	1.40	0.87	1.64	4	.79
aesthetic sensitivity	2.38	2.55	2.43	2.39	2.29	1.72	0.98	2.11	5	.79
software applicability	0.92	1.14	1.30	1.37	1.32	1.14	0.87	1.15	6	.82
imaginative	2.96	3.40	3.51	3.74	3.78	2.97	1.70	3.15	11	.87
creativity	4.73	5.52	5.74	5.76	5.05	3.37	1.80	4.57	6	.84
planning	5.75	6.81	6.67	6.88	6.80	5.24	2.92	5.87	11	.89
debating	1.38	2.04	2.54	2.71	2.43	1.75	1.07	1.99	6	.84
test information	21.65	25.76	26.76	27.43	25.82	19.3	11.27	22.57	55	.96

[Fig. 1]에서 실선은 검사정보함수이고, 점선은 검사정보의 역수로 피험자 능력치에 대한 추정오차를 나타내는 것이다. 이에서 보는 바와 같이, 전체 검사정보함수의 범위는 11.27에서 27.43이고, 능력수준 0에서 가장 높았고 3에서 가장 낮았다.

이를 종합하여 보면, 디자인 역량 검사의 하위 척도별 검사정보는 대체로 중간과 그 이하 능력수준에서 정보량이 많았고, 능력수준 1정도에서부터 급격히 낮았다.

3. 문항특성 분석

<Table 4>는 일반화부분점수 모형에 의해 추정된 디자인역량 검사의 문항모수들이다. 다음은 이들에 대한 문항분석 결과들이다.

가. 문항변별도

<Table 4>에서 보는 바와 같이 디자인 역량 검사에 대한 문항변별도의 범위는 .18에서 .72까지로 나타났다. 일반적으로 문항의 기울기 모수는 0에서 2의 범위 내의 수치로 나타난다. 문항의 변별도는 기울기 모수의 수치로 해석되고, 1을 기준으로 높고 낮음이 판단된다(Park C., 2001). 디자인역량 검사의 문항변별도의 평균은 .43정도이다. 추가하여 고전검사이론에 근거하여 총점과 문항들 간의 상관 정도를 살펴보면 .40에서 .64의 범위였다. 이 정도의 상관계수를 언어로 표현하면 상관이 있는 정도이다.

이를 종합하면, 디자인역량 검사의 문항변별도는 낮다고 할 수 있다.

나. 문항난이도

디자인 역량 검사에 대한 문항난이도는 <Table 4>에서 보는 바와 같이 -1.73에서 .44로 나타났다. 디자인 역량 검사의 문항난이도의 전체평균은 -.70정도이다. 문항난이도는 위치모수의 수치로 해석되고, 일반적으로 0을 기준으로 높고 낮음을 판단한다(Ayala et al., 2009). 이를 기준으로 볼 때, 디자인역량 검사의 문항난이도는 0을 기준으

로 비교적 낮으나 다양하게 분포되었다고 할 수 있다. 즉 문항난이도는 다소 낮게 형성되었다고 할 것이다. 문항변별도와 관련하여 문항난이도를 분석하면, 문항변별력은 문항의 난이도가 적절한 경우에 양호하게 나타난다. 그러나 본 연구의 경우에는 문항난이도도 낮으며, 문항변별도가 상당히 낮다.

다. 문항범주난이도

문항에 대한 평가는 문항변별도와 문항난이도 또는 문항범주난이도에 의해 실시된다. 문항난이도는 문항범주난이도의 중간점을 의미한다. Baker(1977)에 의하면 문항변별도가 높으며, 문항범주난이도의 간격이 유사한 문항이 좋은 문항으로 간주한다. 또 성태제(1998)에 의하면 질문이 간결하고 명확하며 부사를 사용하지 않는 것이 특징이라고 하였다. 즉 문항이 길면 주의집중하기 힘들고, 초점을 파악하기 어렵기 때문에 이해된다. 게다가 설문에 ‘매우’나 ‘절대’ 등과 같은 부사가 사용되면 양극단으로 치우치거나 양극단의 응답을 회피하는 경우가 있어 여러 범주들의 기능을 상실할 수 있다.

<Table 4>를 보면, 문항범주난이도의 간격이나 그 증감들이 제시되어 있다. 문항범주난이도의 간격은 일정하게 유지된 경우가 드물다. 또 문항범주난이도의 증감도 어떤 범주들 사이에서는 높은 범주가 낮은 범주보다 그 값이 작은 경우가 많았다. 이는 위에서 언급된 양호한 문항의 조건에서 상당 부분 위배되어서 실제로 그 문항범주가 그 기능을 상실하고 있다고 해석되어진다. 이를 개선하기 위해서는 문항의 내용 중 부사나 명확하지 않은 진술이 있는 지를 점검하여 수정할 필요가 있다고 여겨진다.

게다가 문항범주난이도 값이 -3을 초과한 단계가 많았다. 인간의 능력 특성이 -3에서 3사이에 분포되어 있다는 것에 비추어 보면, 이 단계는 파악이 되지 않는 경우에 해당된다. 즉 이 단계의 문항 범주는 실제로 기능하지 않는다고 해석

<Table 4> Item parameter estimate of CID(N=1,205)

subtest	item	Slope a_i	difficulty b_i	item category difficulty				
				b_{i1}	b_{i2}	b_{i3}	b_{i4}	b_{i5}
sketching	3	0.32	-1.35	-4.43	-1.08	3.26	2.27	2.63
	10	0.31	-0.66	-2.79	-1.33	1.93	2.67	2.52
	14	0.47	-0.58	-3.34	-1.57	2.63	3.16	3.00
	17	0.44	-0.80	-3.43	-1.45	2.63	3.01	2.53
	27	0.33	-0.79	-3.97	-1.32	3.37	2.99	2.47
	31	0.42	-0.88	-4.12	-1.41	3.53	3.09	2.40
formative	7	0.37	-0.48	-3.45	-1.40	3.04	2.98	2.67
	13	0.45	-0.40	-2.77	-2.07	2.33	3.72	2.87
	21	0.43	-0.77	-3.70	-1.43	3.17	2.93	2.37
	35	0.46	-0.83	-3.41	-1.38	2.77	2.68	2.28
aesthetic sensitivity	5	0.34	-1.14	-4.11	-1.73	2.93	3.28	3.26
	12	0.38	-1.01	-3.62	-1.75	2.51	3.36	2.94
	24	0.46	-1.12	-2.89	-2.21	1.46	3.87	2.96
	26	0.60	-0.84	-4.04	-1.37	3.27	3.10	2.69
	38	0.48	-0.86	-3.33	-1.60	2.27	3.28	2.78
software applicability	4	0.18	0.37	-2.41	-1.12	2.78	1.74	3.30
	11	0.26	-0.10	-2.71	-2.23	3.00	4.08	2.22
	18	0.26	-0.29	-3.58	-1.83	3.52	3.49	2.99
	25	0.24	0.76	-1.84	-1.67	3.27	2.87	1.76
	32	0.24	-0.60	-2.90	-1.82	2.70	3.07	2.06
	39	0.22	-0.98	-3.77	-2.24	2.93	4.15	2.84
imaginative	6	0.49	-0.74	-3.29	-1.30	2.38	2.75	2.75
	19	0.54	-1.13	-3.62	-1.54	2.39	3.06	2.71
	20	0.39	-0.81	-3.39	-1.46	2.81	3.04	2.01
	33	0.45	-1.02	-3.56	-1.28	2.77	2.69	1.94
	34	0.39	-0.89	-2.86	-1.42	1.98	2.73	2.05
	40	0.49	-1.11	-3.99	-1.65	2.86	3.13	3.07
	44	0.37	-1.28	-3.91	-1.16	2.65	2.51	2.40
	49	0.30	-1.16	-4.13	-1.21	3.30	2.84	1.89
	50	0.47	-1.14	-3.97	-1.35	3.00	2.88	2.31
	54	0.30	-1.65	-5.15	-1.39	3.77	3.02	2.73
	55	0.60	-0.70	-3.38	-1.53	2.92	3.12	2.12
creativity	2	0.47	-0.76	-4.59	-1.62	3.94	3.79	3.05
	9	0.52	-0.26	-3.10	-1.54	3.19	3.20	2.01
	16	0.65	-0.36	-3.52	-1.50	3.31	3.06	2.81
	23	0.48	-0.66	-3.31	-1.35	2.90	2.92	1.94
	30	0.45	-1.31	-4.00	-1.50	2.84	3.11	2.27
	48	0.67	-0.72	-3.37	-1.52	2.68	3.01	2.62
planning	1	0.36	-0.75	-3.46	-2.16	2.52	3.83	3.59
	8	0.50	-0.76	-3.63	-1.44	2.92	2.96	2.68
	22	0.55	-0.59	-3.73	-1.21	3.38	2.81	2.26
	15	0.63	-0.37	-2.92	-1.57	2.78	3.00	2.23
	28	0.71	-0.44	-2.99	-1.66	2.61	3.10	2.67
	29	0.53	-0.66	-3.45	-1.33	2.83	2.77	2.62
	36	0.61	-0.49	-2.87	-1.49	2.42	2.86	2.42
	37	0.72	-0.60	-3.19	-1.62	2.62	3.04	2.74
	42	0.42	-1.01	-3.82	-1.68	2.67	3.33	3.10
	43	0.54	-0.62	-3.01	-1.70	2.45	3.18	2.47
	45	0.35	-1.73	-5.54	-1.29	4.08	2.98	2.87
debating	41	0.32	-0.27	-2.51	-1.24	2.59	2.27	1.74
	46	0.34	-0.43	-2.47	-1.38	2.06	2.67	2.07
	47	0.37	-0.44	-2.74	-1.36	2.34	2.86	2.08
	51	0.30	0.44	-0.79	-1.25	1.03	2.45	1.68
	52	0.40	-0.20	-2.23	-1.36	2.15	2.53	1.96
	53	0.30	-0.64	-3.17	-1.12	2.52	2.45	2.32
<i>mean</i>		0.43	-0.70	-3.39	-1.51	2.78	3.01	2.50

된다. 이들은 그 척도 구성을 수정하거나 심지어는 인접 범주와 통합할 필요도 제기된다 할 것이다.

V. 논의 및 결론

본 연구의 목적은 디자인역량 검사에 대한 문항특성을 분석하는 데 있다. 이를 위하여 일반화 부분점수 모형을 적용하여 검사정보함수와 문항 특성 모수를 추정하였다. 다음은 본 연구의 결과를 정리하여 논의한 것이다.

첫째, 본 연구는 디자인역량 검사에 대한 정보의 신뢰성을 검증하기 위하여 검사정보함수를 추정하여 분석하였다. 분석 결과, 본 디자인역량 검사는 대체적으로 능력수준이 중간 정도인 대상들의 디자인 적성을 정확하게 추정하는 것으로 나타났다. 고전검사이론에 근거한 신뢰도 추정치인 알파계수는 본 평가지표가 신뢰롭다고 판단한 것에 비하여, 본 연구는 해당 평가지표가 제시하는 정보의 정확한 조건을 알게 되었다는 것은 검사 결과의 해석에 유의한 의미가 된다 할 것이다.

둘째, 본 연구는 일반화부분점수 모형에 의하여 디자인역량 검사에 대한 문항특성을 분석하였다. 문항특성 분석 결과, 본 디자인역량 검사의 문항들의 난이도는 낮고 다양하며 넓게 분포되었고, 문항들의 변별력도 상당히 낮았다. 실제로 양호한 검사 도구는 문항난이도가 적절히 분포되어 있으며, 높은 문항변별력을 갖추어야 한다(Baker, 1977; Koo Seul-Ki, 2011). 이에 의하면 디자인역량 검사는 문항변별도가 낮아서 양호한 검사도구의 조건을 충족하지 못한다 할 것이다.

디자인 역량 검사의 문항범주난이도를 살펴보면 그 간격이 일정하지 않고, 어떤 범주들에서는 높은 범주가 낮은 범주보다 난이도가 작은 경우가 있었다. 이는 양호한 문항의 조건에서 상당 부분 위배되어 실제로 그 문항범주가 그 기능을 상실하고 있다고 해석되어진다(Baker, 1977;

Seong Tae-Je, 1998). 이를 개선하기 위해서는 문항의 내용 중 부사나 명확하지 않은 진술이 있는지 점검하여 이를 수정할 필요가 있다고 여겨진다.

또 문항범주난이도 값이 -3을 초과한 단계에 대하여서는 인간의 능력 특성이 -3에서 3사이에 분포되어 있다는 것에 비추어 보면 이 단계는 그 척도 구성을 수정하거나 심지어는 인접 범주와 통합하는 할 필요도 제기된다.

디자인 역량 검사에 대한 문항특성 모수와 검사정보함수를 분석한 내용을 종합하면 다음과 같다. 디자인 역량 검사의 문항들은 그 난이도는 낮고 다양하게 분포되어 있지만, 문항변별도는 상당히 낮았다. 그리고 능력수준이 중간정도에 해당되는 피험자의 디자인 적성을 보다 정확하게 추정하는 검사지인 것을 판정된다. 또 디자인 역량 검사에서 표현된 부사나 명확하지 않은 진술은 수정이 요구된다.

본 연구는 대체로 양적 자료에 근거한 통계적 분석에 의하여 문항 분석을 시도하였다. 본 연구의 결과를 바탕으로 디자인 역량 검사에 대한 질적인 접근을 병행할 필요가 있다고 여겨진다.

References

- Andrich, D.(1988). The application of an unfolding model of the PIRT type to the measurement of attitude. *Applied Psychological Measurement*, 12(1), 33~51.
- Andrich, D.(1978). Application of a psychometric rating model to ordered categories which are scored with successive integers. *Applied Psychological Measurement*, 2(4), 581~594.
- Baker, F. B.(1977). Advances in item analysis. *Review of Educational Research*, 47, 151~178.
- De Ayala, R. J. · Dodd, B. G. & Koch, W. R. (1992). A comparison of the partial credit and graded response models in computerized adaptive testing. *Applied Measurement in Education*, 15(1), 17~34.

- Donoghue, J. R.(1994). An Empirical Examination of the IRT Information of Polytomously Scored Reading Items Under the Generalized Partial Credit Model. *Journal of Educational Measurement*, 31(4), 295~311.
- Gil, I. & Yi, S.(2014). Development of career aptitude test for designer(Extended Version). *Journal of Learner-Centered Curriculum and Instruction*, 14(2), 327~348.
- Gil, I. & Yang, S. Y.(2006). A Development of career aptitude scale for design majoring university students. *Archives of Design Research*, 63, 283~292.
- Hambleton, R. K. · Swaminathan, H. & Rogers, H. J.(1991). *Fundamentals of Item Response Theory*. Newbury Park, CA: Sage.
- Kang, T. · Cohen, A. S. & Sung, H. J.(2009). Model selection indices for polytomous items. *Applied Psychological Measurement*, 33(7), 499~518.
- Kim, S. H.(1998). *The theory and practice of polytomous item response model*. Seoul: Kyoyookbook Press.
- Koo, Seul-Ki(2011). *Item analysis of test for interpersonal competency by generalized partial credit model*. Ewah Woman's University.
- Master, G. N.(1982). A Rasch model for partial credit scoring. *Psychometrika*, 47(2), 149~174.
- Muraki, E.(1992). A generalized partial credit model: Application of an EM algorithm. *Applied Psychological Measurement*, 16(1), 59~176.
- Muraki, E. & Bock, R. D.(1998). PARSCALE: IRT item analysis and test scoring for rating scale data[Computer soft ware and manual]. Chicago: Scientific software.
- Park, C.(1999). Accuracy of estimates of item parameters in polytomous IRT models under test length, the number of categories in each item, sample size, and ability distribution. *Korean journal of educational research*, 12(1), 17~42.
- Park, C.(2001). An application of item response theory to an analysis of performance-base items. *Korean journal of educational research*, 32(2), 215~232.
- Samejima, F.(1969). Estimation of latent ability using a response pattern of graded score. *Psychometrika Monograph*, 17.
- Seong, Tae-Je(2001). *The understand and application of item response theory*. Seoul: Kyoyookbook Press.
- Seong, Tae-Je(1998). Item parameter estimation of the academic failure tolerance scale and person's trait estimation via the grade response model. *Korean journal of educational psychology*, 12(1), 203~218.

-
- Received : 03 September, 2015
 - Revised : 02 October, 2015
 - Accepted : 03 October, 2015