

# 선다형 적성검사에서의 선택지 위치, 선택지 지문 길이와 문항 진술 유형에 따른 문항 특성 차이 검증

한영석<sup>1\*</sup>, 김명소<sup>1</sup>  
<sup>1</sup>호서대학교 산업심리학과

## A study on the item characteristics differences of response position, response length, and question types of multiple-choice aptitude tests

Young Seok Han<sup>1\*</sup>, Myoung So Kim<sup>1</sup>

<sup>1</sup>Department of Industrial-Organizational Psychology, Hoseo University

**요약** 본 연구의 목적은 선다형 적성검사에서의 선택지 위치, 선택지 지문 길이 및 문항 진술 유형에 따른 문항 특성차이를 알아보는 데 있다. 이를 위해 H대학의 적성평가 수시 전형 사전검사로 개발된 적성평가 80문항을 전국 80개 고등학교에 재학 중인 고등학교 3학년생 3120명(인문계열-1650명, 자연계열-1467명)에게 실시하였다. 분석결과, 정답이 외곽치(1, 4번 선택지)보다 중앙치(2, 3번 선택지)에 위치할 경우 추측도가 높았고, 옳은 것을 고르는 문항보다 옳지 않은 것을 고르는 문항에서 변별도가 더 높았다. 또한 선택지의 길이가 길수록 선택지 선호도가 높은 것으로 나타났다. 마지막으로 선택형 적성검사에 대한 새로운 접근과 관련해서 본 연구의 의의와 추후 연구 방향을 논의하였다.

**Abstract** This study examined the difference in the item characteristics in multiple-choice aptitude tests focusing on the response position, response length and question types. A university aptitude test consisting of 80 questions was used for this study. The subjects were 3120 senior high school students from 80 schools nation-wide (liberal arts-1650, natural sciences-1467 patients). The results suggest that item prediction is higher for numbers 2 and 3 (located in the middle) than numbers 1 and 4. The item discrimination was higher for pick-the-'wrong'-items than pick-the-'right'-items. In addition, longer choices are preferred. The suggestions for future research are provided based on these findings.

**Key Words** : Response position, Response length, Question types, Aptitude tests

## 1. 서론

### 1.1 연구배경

2014년 대입 선발 장면에서 적성평가를 활용하는 대학은 총 28개교로 그 사용이 점차 확대되어 가고 있는 추세이고, 대기업의 82.5%가 적성검사를 이용하여 신입사원을 선발하고 있다[1]. 신입사원 선발과 대입에서 사용하는 적성검사의 대부분은 선다형 문항으로 구성되어 있고, 검사의 신뢰도와 타당도를 높이기 위한 기초 작업으로 심리검사 제작 시 일반적으로 활용하는 방법은 예비

조사를 통해 도출되는 다양한 문항분석 결과 값을 바탕으로 문항을 수정 및 보완하는 것이다. 그러나 선발과 같이 검사의 보안성 문제로 인해 문항분석을 위한 예비조사를 실시하기 어려운 경우, 기존의 문항분석 결과 값을 이용한 문항 수정은 현실적으로 가능하지 않다. 이와 같은 상황에서는 실제 문항개발과정에 도움을 줄 수 있는 문항개발 이슈들(예: 선택지의 위치, 길이, 문항 유형 등)에 대한 연구가 요망된다.

또한 최근에 Bruce B. F, Stephanie P, Lisa M. Edwards, J(2005)[2]는 지금까지 교육평가방법과 관련된

\*Corresponding Author : Young Seok Han(Hoseo Univ.)

Tel: +82-10-2239-8377 email: nicehan@hoseo.edu

Received January 28, 2014 Revised (1st April 10, 2014, 2nd May 20, 2014, 3rd June 3, 2014) Accepted June 12, 2014

20권의 책에서 제시한 선택형(선다형(Multiple-choice), 연결형(Matching), 진위형(True-false), 완성형(Completion)) 문항작성과 관련된 지침에 대한 문헌고찰 연구에서 가장 사용빈도가 높은 40개의 문항 작성 지침을 제시하였다. 그들은 이 가운데에 5개 지침만이 경험적으로 검증되었다고 주장한다. 즉, 지금까지 선다형 문항작성 지침과 관련하여서는 일부 지침만 객관적 자료를 기반으로 한 연구가 진행되어 왔고, 대부분의 작성지침은 개발자의 주관적 판단과 논리적 타당성에 기반하여 작성되었으며, 객관적이고 경험적인 자료를 기반으로 한 지침은 매우 제한적인 실정이다.

## 1.2 선다형 문항제작

선다형 문항제작과 관련된 지침은 첫째 질문을 작성하는 방법, 둘째 응답지를 작성하는 방법, 셋째 응답지를 배열하는 방법으로 구분할 수 있다. 먼저 문제의 질문 작성과 관련하여, Cohen(1994)[3]은 “질문의 형태 및 언어 제시문의 종류 요구되는 응답의 파악은 타당하고 신뢰할 만한 문항의 개발을 위한 매우 중요한 정보를 제공한다”라고 주장한다(전지현 2008에서 재인용[4]). Shohamy(1984)[5]는 그의 연구에서 읽기 능력을 평가할 때 선다형 질문이나 개방형 질문이냐에 따라 피험자의 읽기 점수에 큰 차이가 있음을 보고하였다. Chun(2005)[6]은 한국 대학생들의 질문 종류에 따른 영어 듣기과업 수행에 있어 차이를 알아보는 연구에서, 한국 대학생들은 구체적인 정보를 묻는 질문보다는 전체적인 내용을 파악하는 질문에 더 어려움을 느끼는 것으로 나타났다.

그러나 질문형식에 대한 연구는 대부분 진위형(true-false form), 선다형(multiple choice form) 유형과 같은 문항유형에 대한 연구가 주를 이루고 있고, 선다형 내에서 질문형식에 대한 연구는 매우 미흡한 실정이다.

선다형 문항의 질문형식과 관련하여 성태제(1996)[7]는 문항마다 질문의 내용이 하나의 사실을 문도록 단순 명쾌하게 구조화되어 있어야한다고 주장한다. 또한 틀린 답을 찾는 것보다 맞는 답을 찾게 하는 것이 보다 교육적이고, 부정문을 사용할 때 주의력이 산만한 피험자는 부주의로 답을 맞히지 못하는 경우가 발생할 수 있으므로, 문항의 질문형태는 가능하면 긍정문이어야 한다고 주장한다. 그러나 Bruce B. F 등(2005)[2]은 문항에 부정적인 용어를 사용하지 않는다는 지침에 대한 경험적 증거가 부족하고, 지침자체가 혼돈을 유발할 수 있으며, 애매한

지침이라고 언급하였다. 따라서 질문형식에서 긍정적 진술문과 부정적 진술문을 사용하는 것이 문항 특성에 어떠한 영향을 미치는 지를 살펴보는 것은 매우 의미있는 연구라고 판단된다.

응답지 작성과 관련한 원칙으로 성태제(1996)[7]은 가장 중요한 주의사항으로 그럴듯하고 매력적인 틀린 답지를 만드는 것을 제안한다. 이와 같은 이유는 매력적인 오답지를 만드는 것이 응시자들이 추측(guessing)에 의해서 답을 맞추는 것을 방지할 수 있기 때문이다[2]. 추가적으로 성태제(1996)[7]와 Bruce B. F 등(2005)[2]은 과제개발의 효율성 측면에서 가능하면 응답지를 짧게 만들 것을 제안하고, 추가적으로 응답지 길이를 비슷하게 하는 것이 바람직하다고 제안한다. 또한 답지들 간의 상호독립성을 확보해야 할 것을 제안한다.

그러나 성태제(1996)[7]는 이와 같은 주장을 하면서 왜 답지의 길이를 비슷하게 만들어야 하는지? 답지 길이가 차이가 나면 문항특성에 어떠한 영향을 미치는 지에 대한 구체적 설명을 제시하고 있지 못하고 있다.

응답지 작성과 관련한 경험적 연구로 윤은혜, 성태제(1998)[8]는 대학수학능력시험 문제 중 ‘모든 것이 정답’ 혹은 ‘정답 없음’ 답지가 가지는 문항특성을 분석한 결과, ‘모든 것이 정답’ 답지를 포함한 문항들은 어려운 문항으로 피험자가 추측에 의해 답을 맞힐 확률이 낮았으며, ‘정답 없음’ 답지가 오답인 문항은 문항추측도가 높아 검사의 행정적 기능을 높이지 못하는 결과를 발견하였다.

선다형 문항의 답지의 배열과 관련해 언급되는 원칙은 주로 다음의 두 가지이다[9,10]. 첫째, 정답이 특정번호에 편중되지 않고 위치가 고르게 분포하도록 한다. 이 원칙은 피험자들의 응답경향이 검사 결과에 미치는 영향을 최소화하기 위한 배려라고 볼 수 있다. 즉, 특정한 위치에 정답이 편중되지 않게 함으로써 특정 위치의 답지를 선호하는 응답경향이 있는 피험자에게 이익이나 불이익이 받지 않도록 하려는 것이다[11].

둘째, 과제개발의 효율성 측면에서 Bruce B. F 등(2005)[2]는 답지들 간에 논리적, 수리적 순서가 있을 때 그 순서에 따라 배열하고, 답지의 길이는 질문 보다는 짧아야 한다고 주장한다. 답지배열과 관련한 경험적 연구로 Yigal과 Maya(2003)[12]는 IRT분석을 통해 오답자들이 중앙에 위치한 선택지를 선택하는 경향이 있다는 것을 발견했다. 상경아, 양길식(2007)[13]는 정답이 외곽에

있는 것보다 중앙에 있는 경우 정답률이 더 높은 것을 발견하였다. 그러나 상경아, 양길석(2007)[13]이 선택지 재배열에 따라 달라지는 문항 특성을 발견했으나 IRT를 기반한 연구는 아니었다.

### 1.3 문항반응 이론

문항반응이론(Item response theory)은 검사의 총점에 의하여 문항을 분석하는 것이 아니라, 문항은 하나하나의 불변하는 고유한 속성을 지니고 있으므로, 그 속성을 나타내는 문항특성곡선에 의해서 문항을 분석하는 검사 이론이다[14].

최근 문항반응이론(IRT)을 활용한 문항분석에 대한 관심이 높아지고 있다. 문항반응이론은 피험자 응답 결과를 근거로 하여 각 문항에 대한 통계적 분석 방법을 이용하기 때문에, 검사와 관계없는 피험자의 능력추정(test-free person measurement)과 표본과 관계없는 문항 모수치의 추정(sample-free item calibration)이라는 특성이 있다는 점에서 효과적이다[14].

문항반응이론을 통해 도출되는 문항특성인 문항난이도, 문항변별도 및 문항추측도에 대한 해석은 아래와 같다. 문항난이도의 범위는 일반적으로 -2에서 +2사이에 존재하며, 난이도 정도를 언어적으로 표현하면 ‘매우 쉽다(-2.0 이하)’, ‘쉽다(-2.0 ~ -0.5)’, ‘중간이다(-0.5 ~ +0.5)’, ‘어렵다(+0.5 ~ +2.0)’, ‘매우 어렵다(+2.0 이상)’이다. 문항특성곡선의 기울기를 나타내는 문항변별도는 .00에서  $+\infty$  값을 가지며, 변별도를 언어적으로 표현하면 ‘없다(.00)’, ‘거의 없다(.01-.19)’, ‘낮다(.20-.38)’, ‘적절하다(.39-.79)’, ‘높다(.80-.99)’, ‘매우 높다(1.0 이상)’, ‘완벽하다( $+\infty$ )’이다. 문항의 답을 알지 못하는 피험자가 우연히 문항의 답을 맞힐 확률을 나타내는 문항추측도 오지선다형의 경우 일반적으로 .20이하가 된다. 이 기준을 넘는 문항은 문항추측도가 높은 문항이라 할 수 있다[7].

IRT분석과 관련한 다른 국내 연구들의 흐름을 보았을 때, 문항의 난이도와 변별도를 고려한 점수 조정, 성별 차별가능 문항 추출 등은 연구해왔지만 문항의 특성 비교 연구는 부족한 실정이며 문항제작과정에서 제기될 수 있는 이슈에 대해 문항반응이론에 기초한 경험적 연구는 매우 미흡하다.

#### 연구문제:

1. 4지선다에서 정답의 위치(1번, 4번 vs. 2번, 3번)에

따라 문항 특성(변별도, 난이도, 추측도)에 차이가 있는가?

2. 4지선다에서 질문의 유형(옳은 vs. 옳지 않은)에 따라 문항 특성(변별도, 난이도, 추측도)에 차이가 있는가?

3. 4지선다에서 선택지의 길이에 따라 선택지 선호도에 차이가 있는가?

## 2. 연구방법

### 2.1 연구대상 및 자료수집

국내 H대학의 적성평가 수시 전형 사전검사로 개발된 적성평가 80문항을 전국 80개 고등학교에 재학 중인 고등학교 3학년생 3120명(인문계열-1650명, 자연계열-1467명)에게 실시하였다. 적성평가 문체지는 계열에 따라 다르게 구성되었는데, 인문계열은 언어기초(20문제), 수열추리(10문제), 언어추리(10문제), 자료해석(10문제), 자연계열의 경우 수리기초(20문제), 수열추리(10문제), 언어추리(10문제), 도형추리(10문제)가 포함되었다. 각 검사는 4지선다형으로 구성하였으며 계열에 상관없이 총 50문항이 제시되었다. 연구문제 1에 사용된 문항은 언어기초, 수리기초, 수열추리, 언어추리, 자료해석, 도형추리 문항으로 총 80문항 중 76문항이 사용되었다. 연구문제 2에는 언어기초, 언어추리, 자료해석 문항에서 ‘옳은’과 ‘옳지 않은’ 지문이 사용된 36개 문항이 사용되었다. 연구문제 3에 사용된 문항은 언어기초와 자료해석 문항으로 총 30문항 중 17문항이며, 지문이 긴 선택지의 선호도를 정확하게 알기 위해 지문이 긴 선택지가 정답인 문제를 제거했다. 그리고 선택된 17문항에서 정답을 틀린 응답자만을 대상으로 나머지 3개의 선택지를 상대적으로 나누어(짧은 선택지=1 ~ 긴 선택지=3) 총 17문항의 선택지 51개에 대한 빈도를 합산하였다.

### 2.3 분석 방법

연구문제 1과 2를 검증하기 위해 BILOG-MG 3.0 프로그램을 사용하여 문항 특성(변별도, 난이도, 추측도)을 구하였고, SPSS 19.0을 사용하여 기술통계 및 차이검증을 실시하였다. 그리고 연구문제 3을 검증하기 위해 적합도 카이제곱검증을 실시하였다.

[Table 1] T-test analysis on Response position

item	M(SD)		t
	outside(N=34)	median(N=42)	
discrimination	1.058(0.510)	0.921(0.324)	1.365
item difficulty	0.570(2.032)	0.325(2.029)	.524
item guessing	0.174(0.051)	0.214(0.072)	-2.826**

\*p<.05, \*\*p<.01, \*\*\*p<.001

[Table 2] T-test analysis on Question types

item	M(SD)		t
	right(N=24)	wrong(N=12)	
discrimination	0.714(0.287)	0.964(0.348)	-2.287*
item difficulty	0.032(1.900)	1.308(1.560)	-1.962
item guessing	0.149(0.045)	0.155(0.037)	-0.405

\*p<.05, \*\*p<.01, \*\*\*p<.001

[Table 3] Chi-square test on Response length

	observed	expected	residual
	frequency	frequency(%)	
short	4553	5094(33.3%)	-541.0
middle	5178	5094(33.3%)	84.0
long	5551	5094(33.3%)	457.0
sum	15282(100%)		

$\chi^2 = 99.840^{***}$

### 3. 연구결과 및 해석

정답 위치에 따라 변별도, 난이도 그리고 추측도에 차이가 있는지 알아보기 위해 독립표본 t검증을 실시하였고, 결과는 표 1에 제시한 바와 같다. 그 결과 추측도에서 유의미한 차이가 나타났지만( $t = -2.826, p < .01$ ), 변별도와 난이도에서는 유의미한 차이가 없었다. 따라서 외곽치(1, 4번 선택지)보다 중앙치(2, 3번 선택지)에 답이 있을 경우, 정답을 모르는 피험자가 추측하여 문항의 답을 맞힐 수 있는 확률이 높다고 해석할 수 있다.

다음으로 연구문제 2를 확인하기 위해 문항 진술유형에 따라 차이분석을 실시한 결과는 표 2에 제시되어 있다. 표에 제시된 바와 같이, 문항 진술유형에 따라 변별도에서 유의미한 차이가 나타났기 때문에( $t = -2.474, p < .05$ ), 옳은 것은 고르는 문제보다 옳지 않은 것을 고르는 문제가 변별도가 더 높은 것을 알 수 있다. 한편, 난이도와 추측도에서는 진술유형에 따른 유의미한 차이가 나타나지 않았다.

연구문제 3을 검토하기 위해 적합도 카이제곱 분석을

실시한 결과, 선택지 길이에 따라 선택 빈도의 차이가 있는 것으로 밝혀졌다( $\chi^2 = 99.840, p < .001$ ). 이를 통해 정답을 틀린 피험자는 선택지 지문이 긴 문항을 비교적 더 많이 선택한다는 것을 알 수 있다.

### 4. 논의 및 결론

외곽치보다 중앙치의 추측도가 더 높게 나타난 본 연구결과는 정답이 외곽치(1번, 5번)에 위치할 때보다 중앙치(2번, 3번, 4번)에 위치할 때 정답률이 높아지는 경향이 있다는 상경아와 양길성(2007)[13]의 연구결과를 지지한다. 이와 같은 상황은 리커드 척도를 사용하는 연구에서 일반적으로 나타나는 중앙집중오류(central-tendency error)와 매우 일관된 현상이라고 볼 수 있다. 일반적으로 중앙집중오류는 리커드 척도와 같이 양이나 질의 정도를 평가하는 상황에서 척도의 중간지점에 평가가 집중되는 현상을 말한다[15].

비록 본 연구에서 사용한 ①,②,③,④라는 번호는 단지 구분을 나타내는 명목척도의 기능만을 가지고 있지만, 응답자들은 부여된 번호를 서열이나 등간으로 인식하였을 가능성이 있다. 따라서 응답자들은 추측상황에서 중앙에 편중된 응답한 것으로 추론된다. 그러나 이와 같은 주장을 좀 더 명확히 하기 위해서는 추후연구에서 번호를 부여하는 방식의 구분기호가 아닌 무의미 철자나 기호와 같이 서열의 개념으로 오해되지 않는 구분부호를 사용하여 이를 검증할 필요가 있다.

긍정형과 부정형 질문 유형에서 긍정형 질문보다 부정형 질문의 변별도가 상대적으로 높았다. 이와 같은 이유로 추론할 수 있는 것은 응답자들의 학습과정과 유사성으로 추론할 수 있다.

즉, 응답자들이 학습하는 과정을 통해서 추론해보면 응답자들은 옳은 지를 선택하는 과정은 기존에 응답자에게 저장되어있던 학습 및 교육내용과 일치하는 지를 확인하는 과정이지만 옳지 않는 내용을 검토하는 것은 기존의 학습과정과는 상이한 과정으로 기존 학습과정에서 옳지 않은 지를 확인하는 연역과정이 더해진다고 추론할 수 있다. 따라서 옳지 않은 것을 선택하는 과정이 더 복잡한 인지과정을 거친다고 추론할 수 있다.

추가적으로 인지과정 수준을 6단계로 나눈 개정된 Bloom의 분류법에 따라 옳은 답을 고르는 과정과 옳지

않은 답을 고르는 과정을 검토해 볼 수 있다. Anderson & Krathwoh(2001)[16]의 개정된 Bloom's Taxonomy는 인지과정을 기억(remember), 이해(understand), 적용(apply), 분석(analysis), 평가(evaluate), 창조(create) 수준에 따라 높은 인지과정으로 보고한다. 이들 내용으로 추론해 볼 때, 옳은 것을 선택하는 인지과정에서 상대적으로 기억(remember) 수준에서 많은 인지과정이 일어날 것으로 추론되고, 옳지 않은 것을 선택하는 과정은 기억 수준 보다는 상대적으로 이해(understand), 적용(apply) 수준에서 많은 인지과정이 일어난다고 추론할 수 있다.

이와 유사하게 Marcus와 Rips(1979)[17]은 인간의 연역추론과정에서도 사람들은 긍정논법에 비해 부정논법에서 논리적으로 타당한 결론을 잘 내리지 못하거나 논리적으로 타당한 결론을 잘 인정하지 않는다고 주장한다. 즉, 부정논법을 성공적으로 해결하기 위해서는 상대적으로 시간이 많이 걸리고 난이도도 올라가는 현상이 발생하는 것으로 해석할 수 있다

그러나 이와 같은 추론을 구체화하기 위해서는 추후 연구에서 부정논법 또는 옳지 않은 것을 질문하는 것이 문항의 난이도와 변별도에 차별적인 영향을 주는지 독립적인 영향을 주는 지를 파악해야할 것이다. 또한 보다 타당한 결과를 위해서 경험적 자료를 추가로 확보할 필요가 있다.

마지막으로 정답을 모르는 상황에서 지문이 긴 오답지를 선택한 연구결과는 긴 오답지가 응답자들에게 특출한 자극으로 인식된 결과로 해석된다. 즉, 과거 비슷한 유형의 시험에서 긴 지문의 보기가 정답인 사례에 대한 기억왜곡으로 인해 긴 지문이 정답으로 잘못 인식된 결과로 해석된다. 이와 같은 기억왜곡현상에 대해서 심리학에서는 특출성(salience)에 의한 상관착각(illusory correlation) 현상으로 설명한다. 특출성에 의한 상관착각 현상이란 특출한 자극일수록 사람들은 그와 관련된 행동이 더 많이 일어나고, 이에 대해 더 호의적이라고 지각하는 현상을 의미한다[18].

관련된 연구로 Sanbonmatsu, Shavitt, Sherman, Roskos-Ewoldsen(1987)[19]은 피험자들에게 5명의 행동진술문을 읽게 하고, 이들 중 한 명에게 주의를 두라고 함으로써 특출성을 조작하였다. 연구결과, 5명 모두에게 바람직한 진술문과 비바람직한 진술문을 동일하게 제공하였지만, 피험자들은 빈번한 행동과 특출한 대상 사이의 연합 정도를 과대추정하였다. 즉, 피험자들은 특정한

대상이 비특출한 대상보다 빈번한 행동을 보다 많이 수행하였다고 판단하였으며, 더불어 평가 또한 특출한 대상이 비특출한 대상보다 더 호의적인 것으로 평가하였다.

추가적으로 지문길이와 선택빈도와의 관련성을 확인하기 위해 1) 띄어쓰기를 제외한 문자의 수(이하 길이)를 측정하고, 2) 차이길이와 오답 선택 빈도를 이용하여 선행 상관분석을 실시하였다. 분석결과, 오답과 정답길이의 차이와 선택빈도의 상관값은  $-.083$ 으로 유의한 통계값을 보이지 않았다. 이와 같은 결과는 선택빈도의 차이가 지나치게 큰 분산에 기인한 것으로 판단된다. 즉, 오답과 정답의 길이 차이는 최소값  $-7$ 에서 최대값  $13$ 으로 평균  $2.59$ , 표준편차  $3.91$ 의 값을 보였지만, 선택빈도 수는 최소값  $49$ 에서 최대값  $881$ 로 평균  $299.6$ , 표준편차  $191.5$ 로 나타났다. 이에 따라 선택빈도가 아닌 빈도를 소, 중, 대 3점 척도로 변환하여 상관분석을 다시 실시하였다. 분석결과 상관값은  $.310$ 으로 유의수준  $.027$ 에서 유의한 상관값을 보였다. 추가적으로 회귀분석을 실시한 결과에서도 오답과 정답의 길이 차이가 선택빈도에 미치는 영향은 유의한 것으로 나타났다( $B=.065$ ,  $t=2.28$ ,  $p=.027$ ).

이와 같은 맥락으로 본 연구결과와 긴 지문의 보다 더 높은 선택빈도결과를 해석하면, 정답을 모르는 상황에서 긴 지문이 응답자들에게 특출한 자극으로 인식되었으며, 과거의 긴 지문이 정답인 경우에 대한 기억왜곡이 과대 지각되어 오답 중에서 긴 지문을 더 많이 선택한 것으로 판단된다. 그러나 이와 같은 추론을 명확히 하기 위해서는 실제로 긴 지문이 특출한 자극으로 인식되는지와 정답과의 연합정도에 대한 추가적인 연구가 필요하다고 판단된다.

이와 같은 연구결과는 선다형 평가문항 작성원칙에 대한 가이드와 경험적 자료를 제공하였다는 점에서 연구의 의미가 있다. 먼저 문항제작자들은 질문형식에 따라 난이도와 변별도를 다양화 할 수 있을 것이다. 둘째 문항제작자들은 선택지 구성에 있어서 선택지 중앙에 정답을 배치하는 것과 긴 지문 정답을 조절하여 문항 추측도를 줄일 수 있을 것이다. 즉, 선다형 문항을 구성함에 있어 선택지 지문의 길이를 가능한 동일한 수준으로 유지할 것을 제안하며, 정답이 가운데 위치하고 동시에 긴 지문을 사용하는 것을 병행하는 것을 지양할 것을 제안한다. 추가적으로 이와 같은 결과는 최근에 부정행위를 방지하기 위해, 답지 배열을 다르게 한 복수의 검사지를 사용하는 경우, 또는 문제은행에서 문항을 추출해 검사를 제작

하거나 검사 동등화를 위해 동일한 문항을 구성할 때 활용될 수 있을 것으로 판단된다.

마지막으로 선다형 문항제작 원칙에서 경험적 연구가 시도되지 않은 부분들에 대한 IRT를 이용하여 경험적 자료를 제공하고 시도한 점에서 본 연구의 추가적 의의를 찾을 수 있다.

추후 연구에서는 신입사원용 적성검사나 5지선다형으로 확대하여 이들 결과의 일반화 가능성 및 각각의 연구 문제에서 제시된 추가연구 문제를 구체화하여 검토해볼 필요가 있다.

## References

- [1] Newsis, 82.5% of large companies, which, aptitude tests, skills tests conducted, 2014 Available from: <http://www.newswire.co.kr/newsRead.php?no=432553>. accessed January., 01, 14.
- [2] Bruce B. F, Stephanie P, Lisa M. Edwards, J. (2005) Item-writing rules: Collective wisdom. *Teaching and Teacher Education* 21, 357 - 64  
DOI: <http://dx.doi.org/10.1016/j.tate.2005.01.008>
- [3] Cohen, A. (1994). *Assessing language ability in the classroom (2nd ed.)*. Boston, MA: Heinle & Heinle.
- [4] Ji-hyun, Chun. (2008). A study on item type and item difficulty: A case of qualifying exam for Business English. *The Korean Association of Sedretarial Studies*, 17(1), 141-156.
- [5] Shohamy, E. (1984). Does the testing method make a difference? The case of reading comprehension. *Language Testing*, 1, 199-215.  
DOI: <http://dx.doi.org/10.1177/026553228400100203>
- [6] Chun, H. (2005). The effects of contents, pretasks and question types on Korean college students' lecture listening comprehension. *Korean Journal of Applied Linguistics*, 21(1), 241-264.
- [7] Seong, Tae-Jae. (1996). Production and analysis of the theory and practice questions. Hakjisa.
- [8] Yoon, Hye-Kyong, Seong, Tae-Jae(1998). An analysis of the item and test characteristics of multiple-choice items including the 'all of the above' or 'none of the above' alternatives. *The Journal of Educational Research*36(1), 131-147.
- [9] Hwang Jung-gyu(1998). *School Learning and Assessment*. Seoul:Kyoyookbook.
- [10] Halady. T. M. & Downing, S. M.(1989). *A Taxonomy of multiple-choice item writing rules*. *Applied Measurement in Education*, 2(1), 37-50.  
DOI: [http://dx.doi.org/10.1207/s15324818ame0201\\_3](http://dx.doi.org/10.1207/s15324818ame0201_3)
- [11] Yun, Jee Hwan, Lee, Moonbok, & Park, Yong-Hyo. (2012). The influence of the language of directions, questions, and choices in practice CSAT on learners' test results. *Korean Journal of Applied Linguistics*, 28(1), 59-85.
- [12] Yigal A, Maya B. H(2003). Guess Where: The Position of Correct Answers in Multiple-Choice Test Items as a Psychometric Variable. *Journal of Educational Measurement*, 40(2), 109-128.  
DOI: <http://dx.doi.org/10.1111/j.1745-3984.2003.tb01099.x>
- [13] Sang, Kyong-Ah, Yang, Kil-Seok(2007). The effect of response position in a multiple-choice examination *Asian Journal of Education*. 8(1), 25-46.
- [14] Baker F. B. *Item Response Theory: Parameter Estimation Techniques*. New York: Marcel Dekker, 1992.
- [15] Bardo, J. W. Yeager, S. J. and Klingsporn M. J. (1982) Preliminary assessment of format-specific central tendency and leniency error in summated rating scales. *Perceptual and Motor Skills*: 54, 227-234.  
DOI: <http://dx.doi.org/10.2466/pms.1982.54.1.227>
- [16] Anderson, L. W., Karthwohl, D. R., Airasan, P. W., Crikbank, K. A., Mayer, R. E., Pintrich, P. R., Raths, J & Wittrock, M. C.(2001). *A taxonomy for learning, teaching, and assessment: A revision of Blooms's taxonomy of educational objectives*. 1st ed. Pearson Education, Inc.
- [17] Marcus, S.S., & Rips, L.J.(1979). Conditional reasoning. *Journal of Verbal Learning and Verbal Behavior*, 18, 199-223.  
DOI: [http://dx.doi.org/10.1016/S0022-5371\(79\)90127-0](http://dx.doi.org/10.1016/S0022-5371(79)90127-0)
- [18] Yongwon Sun, Youngseok Han. (2000). Effects of attention and judgemental encoding in individual and group targets on illusory correlation and performance evaluation. *Korean Journal of Industrial and Organizational Psychology*. 13(1). 1-21.
- [19] Sanbonmatsu, D. M., Shavitt, S. Sherman, S. J., & Roskos-Ewoldsen, Jr.(1987). Illusory correlation in the perception of performance by self or a salient other. *Journal of Experimental Social Psychology*, 23, 518-543.  
DOI: [http://dx.doi.org/10.1016/0022-1031\(87\)90019-9](http://dx.doi.org/10.1016/0022-1031(87)90019-9)

**김 명 소(Myoung-So Kim)**

[정회원]



- 1985년 2월 : The University of Texas at Austin(상담심리)
- 1991년 2월 : The University of Texas at Austin(심리측정)
- 1992년 3월 ~ 현재 : 호서대학교 산업심리학과 교수

<관심분야>

인사관리, 직무수행평가

---

**한 영 석(Young-Seok Han)**

[정회원]



- 1997년 8월 : 성균관대학교 심리학과 (산업심리학석사)
- 2004년 8월 : 성균관대학교 심리학과 (산업심리학박사)
- 2009년 3월 ~ 현재 : 호서대학교 산업심리학과 교수

<관심분야>

인사관리