



저작자표시-비영리-변경금지 2.0 대한민국

이용자는 아래의 조건을 따르는 경우에 한하여 자유롭게

- 이 저작물을 복제, 배포, 전송, 전시, 공연 및 방송할 수 있습니다.

다음과 같은 조건을 따라야 합니다:



저작자표시. 귀하는 원저작자를 표시하여야 합니다.



비영리. 귀하는 이 저작물을 영리 목적으로 이용할 수 없습니다.



변경금지. 귀하는 이 저작물을 개작, 변형 또는 가공할 수 없습니다.

- 귀하는, 이 저작물의 재이용이나 배포의 경우, 이 저작물에 적용된 이용허락조건을 명확하게 나타내어야 합니다.
- 저작권자로부터 별도의 허가를 받으면 이러한 조건들은 적용되지 않습니다.

저작권법에 따른 이용자의 권리는 위의 내용에 의하여 영향을 받지 않습니다.

이것은 [이용허락규약\(Legal Code\)](#)을 이해하기 쉽게 요약한 것입니다.

[Disclaimer](#)

조 영 일 교수 지도

석사학위 청구논문

개인합치도 지수 (person fit) 를
이용한 응답왜곡의 식별에 관한 연구

2020

성신여자대학교 대학원

심리학과

장 유 나

개인합치도 지수(person fit)를
이용한 응답 왜곡의 식별에 관한 연구

조 영 일 교수 지도

이 논문을 석사학위논문으로 제출함

2020년 6월

성신여자대학교 대학원

심리학과

장 유 나

인 준 서

장유나의 석사학위 논문으로 인준함

2020년 6월

심사위원장 진 경선 (인)
심사위원 박희경 (인)
심사위원 조영일 (인)

성신여자대학교 대학원

논문개요

본 연구는 Ferrando와 Anguiano-Carrasco(2013)가 제안한 새로운 개인 합치도 지수(person fit)의 효용성을 검증하는 것을 목적으로 하였다. 이를 위해 몬테카를로 시뮬레이션(Monte Carlo Simulation)을 이용하여 다양한 조건 하에서의 1종 오류(type 1 error)와 검정력(power)의 비율을 확인하였다. 본 연구에서 구성한 조건들은 다음과 같다. 1)검사의 길이(short, medium, long), 2)왜곡하기 쉬운 문항의 비율(20%, 50%), 3)피검자의 수(100명, 300명, 500명), 4)응답 왜곡 피검자들의 비율(honest:faker=1:1, 3:2, 4:1), 5)응답 왜곡 유형(faking good, faking bad) 6)응답 왜곡의 크기(Small, Large). 각 조건 하에서 1종 오류와 검정력의 비율을 확인한 결과, 검사 길이가 길어질수록 1종 오류가 감소하고 검정력이 증가하는 것으로 나타났으며, 왜곡 문항 비율 조건에서도 동일한 양상을 보이는 것으로 밝혀졌다. 또한 왜곡 응답자의 비율은 그 차이가 클수록 1종 오류가 감소하고 검정력이 증가하는 경향을 보였다. 피검자 수는 감소할수록 검정력이 증가했으나 피검자의 수가 100명 이하일 때에는 모델이 수렴되지 않는다는 문제를 보였다. 본 지수는 부정적 응답 왜곡에 더 효과적인 것으로 나타났으며, 왜곡의 크기가 클수록 왜곡 응답자들을 더 잘 식별하는 것으로 나타났다. 본 연구는 Ferrando와 Anguiano-Carrasco(2013)의 개인 합치도 지수를 시뮬레이션으로 검증했다는 의의를 지니지만, 절단값 k 를 1로 한정하여 상황 별 적절한 k 값을 탐색하지 못했으며 조건을 이분 문항으로 제한하였다는 한계를 지닌다.

주요어 : 개인 합치도 지수, 응답 왜곡, 몬테카를로 시뮬레이션

목 차

논문개요

I. 서론	1
1. 연구의 필요성 및 목적	1
II. 이론적 배경	8
1. theta(θ) shift	10
1) changing-persons paradigm	10
2) changing-items paradigm	11
2. 응답 왜곡 모델링	13
1) 1단계: 부분 동일성 모형 설정	13
2) 2단계: 개인 합치도 지수 산출	21
III. 연구문제 및 가설	25
연구문제1.	25
연구문제2.	26
IV. 연구 방법	27
1. 자료 생성	27
1) 독립변수	28
(1) 검사 길이	28
(2) 왜곡하기 쉬운 문항의 비율	28
(3) 피검자의 수	29

(4) 응답 왜곡 피검자들의 비율	29
(5) 응답 왜곡 유형	29
(6) 응답 왜곡 크기	30
2) 종속변수	31
(1) 제 1종 오류 비율(Type- I error)	31
(2) 검정력 (Power)	31
2. 자료 분석	32
V. 연구 결과	33
1. 자료의 특성에 따른 1종 오류 비율	34
1) 검사 길이에 따른 1종 오류 비율	34
2) 왜곡문항 비율에 따른 1종 오류 비율	39
3) 응답 왜곡 피검자들의 비율에 따른 1종 오류 비율	43
4) 피검자 수에 따른 1종 오류 비율	46
5) 응답 왜곡 유형에 따른 1종 오류 비율	50
6) 응답 왜곡 크기에 따른 1종 오류 비율	54
2. 자료의 특성에 따른 검정력 비율	58
1) 검사 길이에 따른 검정력	58
2) 왜곡문항 비율에 따른 검정력	62
3) 응답 왜곡 피검자들의 비율에 따른 검정력	66
4) 피검자 수에 따른 검정력	69
5) 응답 왜곡 유형에 따른 검정력	72
6) 응답 왜곡 크기에 따른 검정력	75

VI. 논의	79
1. 연구 결과 정리 및 논의	79
2. 종합 논의 및 연구의 제한점	86

참 고 문 헌

ABSTRACT

부 록

표 목 차

<표 1> 응답 왜곡 종류에 따른 응답 왜곡 크기	30
<표 2 > 긍정왜곡 조건에서의 검사 길이에 따른 1종 오류 비율	36
<표 3 > 부정왜곡 조건에서의 검사 길이에 따른 1종 오류 비율	37
<표 4> 긍정왜곡 조건에서의 왜곡 문항 비율에 따른 1종 오류 비율	40
<표 5> 부정왜곡 조건에서의 왜곡 문항 비율에 따른 1종 오류 비율	41
<표 6> 긍정왜곡 조건에서의 응답 왜곡 피검자 비율에 따른 1종 오류 비율	44
<표 7> 부정왜곡 조건에서의 응답 왜곡 피검자 비율에 따른 1종 오류 비율	45
<표 8> 긍정왜곡 조건에서의 전체 피검자 수에 따른 1종 오류 비율	47
<표 9> 부정왜곡 조건에서의 전체 피검자 수에 따른 1종 오류 비율	48
<표 10> 작은 크기의 왜곡 조건에서의 왜곡 유형에 따른 1종 오류 비율	51
<표 11> 큰 크기의 왜곡 조건에서의 왜곡 유형에 따른 1종 오류 비율	52
<표 12> 긍정 왜곡 조건에서의 응답 왜곡 크기에 따른 1종 오류 비율	55
<표 13> 부정 왜곡 조건에서의 응답 왜곡 크기에 따른 1종 오류 비율	56
<표 14> 긍정왜곡 조건에서의 검사 길이에 따른 검정력	59
<표 15> 부정왜곡 조건에서의 검사 길이에 따른 검정력	60
<표 16> 긍정왜곡 조건에서의 왜곡 문항 비율에 따른 검정력	63

<표 17> 부정왜곡 조건에서의 왜곡 문항 비율에 따른 검정력	64
<표 18> 긍정왜곡 조건에서의 응답 왜곡 피검자 비율에 따른 검정력	67
<표 19> 부정왜곡 조건에서의 응답 왜곡 피검자 비율에 따른 검정력	68
<표 20> 긍정왜곡 조건에서의 전체 피검자 수에 따른 검정력	70
<표 21> 부정왜곡 조건에서의 전체 피검자 수에 따른 검정력	71
<표 22> 작은 크기의 왜곡 조건에서의 왜곡 유형에 따른 검정력	73
<표 23> 큰 크기의 왜곡 조건에서의 왜곡 유형에 따른 검정력	74
<표 24> 긍정 왜곡 조건에서의 응답 왜곡 크기에 따른 1종 오류 비율	76
<표 25> 부정 왜곡 조건에서의 응답 왜곡 크기에 따른 1종 오류 비율	77

그림 목 차

<그림 1> 잠재 응답과 관찰 점수	14
<그림 2> 형태 동일성 모형	16
<그림 3> 부분 측정 단위 동일성 모형	17
<그림 4> 검사 길이에 따른 1종 오류 비율의 변화	38
<그림 5> 왜곡문항 비율에 따른 1종 오류 비율의 변화	42
<그림 6> 피검자 수에 따른 1종 오류 비율의 변화	49
<그림 7> 응답 왜곡 유형에 따른 1종 오류 비율의 변화	53
<그림 8> 응답 왜곡 유형에 따른 1종 오류 비율의 변화	57
<그림 9> 검사 길이에 따른 검정력의 변화	61
<그림 10> 왜곡문항 비율에 따른 검정력의 변화	65
<그림 11> 응답 왜곡 크기에 따른 검정력의 변화	78

I. 서론

1. 연구의 필요성 및 목적

1916년 Binet과 Simon이 최초의 지능검사를 개발한 이후로 다양한 형태의 자기보고식 검사들이 등장하면서, 심리학 전 분야에 걸쳐 자기보고식 검사들이 활발하게 사용되어 왔다. 특히 <한국심리학회지: 일반>에 2009년부터 2019년까지 발간된 435편의 논문들 중 270편(약 62%)이 모두 자기보고 형식을 활용한 연구임을 보았을 때, 자기보고식 검사는 가장 많이 사용되는 유구한 측정도구임을 알 수 있다. 그러나 자기보고식 검사는 피검자의 응답이 정직할 때에만 타당한 결과를 얻을 수 있다는 단점을 지닌다. 즉, 피검자가 응답을 왜곡했을 경우 검사의 타당도가 저해되어 결과를 신뢰할 수 없게 된다(Douglas, McDaniel, & Snell, 1996; Dunnette, McCartney, Carlson, & Kirchner, 1962; Ironson & Davis, 1979; Schmit & Ryan, 1992; Schmit, Ryan, Stierwalt, & Powell, 1995; Worthington & Schlottmann, 1986; Zickar와 Drasgow, 1996). 자기 보고식 검사에서 응답 왜곡(faking)에 대한 문제는 타당도에 대한 의구심을 야기하기 때문에 오랜 시간 동안 심리학 전반에서 중요한 문제로 다루어져 왔다(Morgeson et al., 2007; Ruch, 1942; Spencer, 1938).

응답 왜곡은 긍정적 응답 왜곡(faking good)과 부정적 응답 왜곡(faking bad)으로 구분할 수 있다. 긍정적 응답 왜곡(faking good)은 인사선발 등의 상황에서 특정 목적을 위하여 자신을 호의적으로 왜곡하는 것을 의미하며(이종구 & 한영석, 2008), 자신을 사회적으로 바람직한 방향으로 과장하는 것이라 정의할 수 있다(Paulhus, 1984). 반대로 부정적 응답 왜곡(faking bad)은 입대 등의 상황에서 특정 목적을 위해 타인에게 자신을 부

정적으로 왜곡하는 것을 의미한다(이종구 & 한영석, 2008). 이처럼 왜곡된 응답은 연구 결과의 신뢰도를 저해할 뿐만 아니라 선발과 진단 등의 장면에서도 선별을 어렵게 만든다는 문제를 야기한다(Paulhus, 1984). 때문에 응답 왜곡을 해결하기 위한 여러 노력이 있어왔으며, 이를 위한 방법은 크게 사전 예방적 접근방법(Proactive Approach)과 반응적인 접근방법(Reactive Approach) 두 가지로 분류할 수 있다(윤미리, 2015).

사전 예방적 접근방법(Proactive Approach)은 검사 시행 전 왜곡을 방지하기 위한 시도들을 총칭하는데, 대표적으로 응답 왜곡에 대해 사전에 경고하거나(Dwight & Donovan, 2003), 응답 왜곡에 보다 더 민감하다고 알려진 척도 형식(예, ipsative 형식¹⁾)을 사용하는 것 등이 있다(김명소, 이현주, 2006). 그러나 사전 경고 메시지를 사용하는 경우 그 효과성이 부정적 응답왜곡에서는 미미하며(지여운, 2013), ipsative 형식은 통계적인 분석이 어려울 뿐더러 그 해석이 제한적이라는 단점을 지닌다(Hicks, 1970; Johnson, Wood, & Blinkhorn, 1988).

반응적인 접근방법(Reactive Approach)은 검사 실시 후 수집한 응답들을 토대로 응답 왜곡을 탐지하는 방법으로, External 방식과 Internal 방식 두 가지가 존재한다. External 방식은 검사에 타당화 척도를 포함하여 응답 왜곡을 탐지하고 교정하는 방법이다. 이는 MMPI-2(K척도, L척도)와 16PF(Faking Good 척도, Faking Bad 척도) 등 주요 심리검사에서 사용되고 있는 방식으로, 해당 척도에 높은 바람직성 점수를 보인 피검자의 응답을 제외하거나 회귀 방정식을 이용하여 점수를 교정하는 목적으로 사용된다(Goffin & Christiansen, 2003). 그러나 다양한 학자들이 타당화 척도를

1) Cattell (1944)에 의해 명명된 심리측정 방식 중 하나로, 바람직성이 일치하는 2~4개의 진술문들을 짝지어 하나의 문항 세트를 구성한다. 이후 피검자들이 가장 선호하는 대안, 즉 자신과 가장 유사한 것으로 생각되는 대안을 선택하게 한다. 또한 가장 선호하는 대안과 가장 선호하지 않는 대안을 각 한 개씩 선택하게 하여 개인 내의 비교 경쟁을 통해 피검자가 선택한 대안의 속성에 점수를 부여한다(김명소, 이현주, 2006).

사용한 점수 교정에 대하여 다음과 같은 제한점들을 주장했다. 먼저 높은 사회적 바람직성 점수가 실제 응답 왜곡으로 인한 결과임을 확신할 수 없다. 몇몇 연구들에서 사회적 바람직성 척도에 긍정적으로 응답하도록 지도 받은 피검자들의 결과가 지도 없이 긍정 왜곡을 한 집단보다 정직하게 응답한 집단과 더 유사한 것으로 나타났다(Dwight & Alliger, 1997; Kroger & Turnbull, 1975). 이는 사회적 바람직성 척도가 응답왜곡을 제대로 탐지할 수 없음을 단적으로 보여주는 결과이다. 또한 응답왜곡 반응 패턴은 응답자의 목표에 따라 달라질 수 있는데(Dicken, 1959), 타당도 척도를 이용한 기존의 교정방식은 이에 대한 고려 없이 동일하게 적용되므로 교정된 점수가 적합하지 않을 가능성이 있다(김명소, 이현주, 2006). 실제로 선행연구들에 의하면 응답왜곡을 교정한 점수가 교정하지 않은 점수에 비해 타당도가 우수하지 않았다(Christiansen, Goffin, Johnston, & Rothstein, 1994; Ellingson, Sackett, & Hough, 1999; Ones, Viswesvaran, & Reiss, 1996).

마지막으로 Internal 방식에서는 검사로 수집된 응답 특성(response characteristics)을 이용하여 응답 왜곡을 탐지한다. Internal 방식 중 하나인 개인 합치도 분석(person-fit)은 정확하지 않은 측정에 영향을 주는 비전형적 반응 유형(aberrant response pattern)을 가진 검사자를 발견할 때 사용하는 기법이다(Karabatsos, 2003). 비전형적 반응 유형은 검사 동기 부족, 집중력 문제, 숙임수 등으로 인하여 전반적인 문항 반응 형태와 일관되지 않은 반응 유형을 일컫는다. 개인 합치도 지수는 이러한 비전형적 반응 유형을 탐지하기 위한 통계적 기법 중 하나로, 사회적 바람직성 혹은 응답 왜곡 행위를 탐지하는 데 사용되어 왔다(Zickar & Drasgow, 1996; Ferrando & Anguiano-Carrasco, 2013). 개인 합치도 분석은 측정 문항만을 사용하는 내적 준거 방법을 적용하고 있으므로 기존의 검사를 수정하

거나 문항을 추가하여야 할 필요가 없을 뿐더러 객관적인 판정 기준을 제공하기 때문에 해석이 용이하다(Ferrando & Chico, 2001). 또한, 검사 점수만으로는 밝힐 수 없는 개인의 특성까지 분석할 수 있으므로(Meijer & Sijtsma, 2001), 위 방법들의 단점을 보완할 수 있는 대안으로 사료된다.

개인 합치도 지수는 응답 패턴의 비교 대상이 추정된 문항반응이론(Item response theory, IRT) 모형에 준거 하는지에 따라 크게 모수적(parametric) 개인 합치도 지수와 비모수적(non-parametric) 개인 합치도 지수로 나눌 수 있다. 모수 개인 합치도 지수는 그를 산출하는 방법에 따라 세 가지로 분류 가능하다(Meijer & Sijtsma, 1995). 첫째는 모형으로 예측된 값과 실제 관찰된 값의 차이를 잔차를 이용하여 비교하는 방법이다. IRT 모형을 이용하여 피검자의 특성과 문항 모수와의 관계를 통해 특정 문항에 옳게 응답할 확률을 계산한다. 이때 잔차, 즉 예측된 반응과 실제 반응 간의 차이에 근거하여 피검사 반응의 적합 여부를 결정할 수 있다. 둘째는 우도 함수를 이용하는 것으로, 전형적인 반응에 대한 IRT 모형을 설정한 후 관찰 점수 패턴이 일어날 확률을 우도함수로 계산하는 방법이다. 이후 실제 피검자가 반응 패턴을 보일 확률을 계산함으로써 비전형적 반응으로 분류할 것인지를 판단한다. 셋째는 전형적인 응답에 대한 통계적 모형과 비전형적 응답 패턴에 대한 통계적 모형을 설정한 뒤, 우도비 통계치(likelihood ratio)를 계산하는 방법이다. 이를 통해 적절성 지수(appropriateness index)를 산출하여 피검자의 응답 패턴이 전형적인 반응과 비전형적인 반응 중 어느 쪽에 속하는지를 결정하게 된다. 반면 비모수 개인 합치도 지수는 응답 패턴을 추정된 IRT 모형이 아닌 Guttman perfect pattern과 비교하는 등 오로지 관찰된 점수들로만 계산한다는 점에서 모수 개인 합치도와 구분된다(Karabatos, 2003).

Ferrando와 Anguiano-Carrasco(2013)는 우도비 통계치를 통해 적절성

지수를 산출하는 세 번째 방법에 주목하였다. 이는 Zickar와 Drasgow(1996)의 이론을 기반으로 하여 발전시킨 것으로, 이들은 응답 왜곡의 메커니즘을 IRT의 관점에서 정의한 바 있다. 그에 따르면 응답 왜곡은 faking condition 하에서 일시적으로 피검자의 능력 수준이 상승하거나(긍정적 왜곡) 하락하여(부정적 왜곡) 검사 점수가 변화한 것으로, 일종의 ‘theta(θ) shift’ 현상이라 명명할 수 있다. 이로 인해 유발된 변화량은 두 가지 방법으로 모델링 할 수 있는데, 첫 번째는 문항 모수를 고정하고 변화를 능력 수준에서 모델링 하는 것이며(changing-persons paradigm), 두 번째는 능력 수준을 고정하고 변화를 문항 모수 수준에서 모델링 하는 것이다(changing-items paradigm). 즉, changing-persons paradigm의 경우 긍정 응답 왜곡 시 달라진 응답 패턴을 세타(θ)값이 증가한 것으로 모형화하는 반면, changing-items paradigm은 세타 값을 고정한 채 문항 모수(난이도 및 변별도) 값이 증가한 것으로 모형화 한다. Zickar와 Robie(1999)는 changing-persons paradigm을 이용하여 응답 왜곡을 식별하였으나 Ferrando와 Anguiano-Carrasco(2013)는 응답 왜곡 특유의 복잡성을 고려했을 때 changing-items paradigm이 더 적절할 것이라 지적했다. 이는 응답 왜곡이 문항들에 일정한 수준으로 발생하는 단순한 현상이라기보다는, 그로 인한 변화량이 문항마다 상이한 복잡한 현상이라는 논의에 기초한 것이다(Kuncel & Borneman, 2007). 따라서 Ferrando 와 Anguiano-Carrasco(2013)는 changing-items paradigm을 활용한 2단계 접근을 제안하였다. 이는 기존 절차에 구조방정식 모형(structural equation model, SEM)을 도입하여 조정을 보다 정교히 한 것으로, 개인 합치도 지수 산출의 효과성을 향상시키기 위함이다. 개인 합치도 지수의 최적성은 얼마나 정확하게 정직 응답 집단과 왜곡 응답 집단의 패턴을 모델링 하느냐에 달려있기 때문에 changing-items paradigm을 바탕으로 한 부분

동일성 모형의 설정은 이러한 정확성을 보다 향상시킬 것으로 예상 되었다. 이를 확인하기 위해 277명의 대학생을 대상으로 Eysenck Personality Questionnaire-Revised(EPQ-R; Eysenck et al., 1985) 척도를 사용하여 분석한 결과, 유의수준 5%에서 66%의 적중률(hit rate)을 보이는 것으로 나타났다. Ferrando와 Anguiano-Carrasco(2013)는 26%의 적중률을 보였던 Zickar와 Drasgow(1996)의 연구와 비교했을 때 해당 방법이 정확성을 보다 향상시켰음을 주장하였다. 그러나 이는 실제 데이터를 사용한 경험적 연구이므로 몇 가지 한계점을 지닌다. 먼저 피검자들에게 거짓으로 응답하도록 지시한 후 분석한 것이기 때문에 정확한 응답 왜곡자의 비율을 확신할 수 없다. 즉, 계산된 검정력이 정확한 결과인지 알 수 없으므로 시뮬레이션 연구를 통한 확인검증이 필요하다. 또한 연구에 사용된 척도들 역시 외향성(Extraversion) 23문항, 신경증(Neuroticism) 32문항, 정신증적 경향성(Psychoticism) 23문항으로 모두 중간 정도의 길이에 한정되어 있었다. 척도의 길이는 개인 합치도 지수의 식별 가능성에 중요한 영향을 미치는 중요 요인 중 하나이므로(Karabatsos, 2003; Li & Olejnik, 1997), Ferrando와 Anguiano-Carrasco(2013)의 새로운 방법이 짧거나 긴 척도에서도 효과적으로 작용하는지에 대한 검증이 필요하다. 이들은 또한 긍정적 응답 왜곡 조건만을 처치하였으므로 부정적 응답 왜곡 조건 하에서도 해당 지수가 효과성을 보이는지에 대한 확인이 요구된다. 마찬가지로, 피검자 수 역시 300명으로 제한되었으므로 피검자수가 300명 이상이거나 이하인 조건에서도 해당 지수를 적용할 수 있는지에 대한 확인이 필요하다.

따라서 본 연구는 Ferrando와 Anguiano-Carrasco(2013)가 고안한 새로운 개인 합치도 지수를 다양한 조건들 하에서 검증하여 그 효과성을 밝히고 일반화 가능성을 높이는 것을 목적으로 하였다. 때문에 검사 길이와 피검자의 수, 응답 왜곡 유형 조건에서 경험 연구로 확인되지 않았던 다양한

수준(예컨대 짧고 긴 문항과 300명 이상/이하의 피검자 수, 부정적 응답 왜곡)을 추가적으로 검증할 수 있도록 시뮬레이션 연구를 계획하였다. 또한 개인합치도 지수 연구에서 자주 다루어지는 조건들(왜곡하기 쉬운 문항의 비율, 응답 왜곡 피검자들의 비율, 응답 왜곡 크기)을 추가하여 보다 다양한 상황에서의 결과를 탐색하고자 하였다(강유진, 2012; Wang et al., 2010; 박찬호, 2009; Raju et al., 2009; Stark & Chernyshenko, 2006; Byrne & Stewart, 2006; Hancock et al.; 2000; Kaplan & George, 1995; Rojers et al., 1993). 위의 조건들은 특히 경험 연구에서는 조작하기 어려운 변인들이므로 시뮬레이션 연구를 통해 경험 연구의 한계를 보완할 수 있으리라 예상하였다. 따라서 시뮬레이션 자료를 활용하여 6개의 조건 하에서 1종 오류(Type-I error)와 통계적 검정력을 분석할 것이다.

II. 이론적 배경

Zickar와 Drasgow(1996)는 응답 왜곡을 검사 전체가 아닌 문항 수준에서 분석해야 함을 주장하였다. 다수의 선행연구들이 긍정적 응답 왜곡 시 정직하게 응답했을 때보다 전체 점수의 평균이 높음을 보였으나(Hough, 1998; Hough et al., 1990), 이러한 척도 수준의 결과는 응답 왜곡의 프로세스를 규명할 수 없다는 한계를 지니기 때문이다. 즉, 어떠한 문항이 응답 왜곡을 유발하는 것인지 알 수 없으므로 응답 왜곡의 과정을 밝힐 수 없다. 때문에 응답 왜곡을 문항 단위로 분석하려는 노력이 있어왔다. 대표적으로, Schmit와 Ryan(1993)은 확인적 요인 분석(Confirmatory factor analysis, CFA)을 이용하여 정직 응답 피검자와 응답 왜곡 동기가 있는 피검자(즉, 기업 지원자)의 성격 검사 응답 간 측정 동일성을 검증하였다. 그 결과, 두 집단 간 요인구조에 차이가 존재하는 것을 발견했다. 이후 다른 연구에서도 CFA를 이용하여 응답 왜곡 집단과 정직 응답 집단 간의 차이를 살펴본 결과, 잠재 변수의 수와 오차 분산, 잠재 변수들의 간의 관계에 차이가 있는 것으로 나타났다(Frei et al., 1997). 이러한 연구들은 응답 왜곡이 개인의 특성을 측정하는 것을 왜곡하는 현상임을 보였다는 것에 의의가 있으나, 여전히 문항 단위의 분석이 구체적으로 이루어지지 않았다는 점에서 한계를 지닌다.

문항반응이론(Item response theory)은 CFA를 활용한 측정 동일성 검증을 보완할 수 있는 방법 중 하나로 고려될 수 있다(Reise et al., 1993). IRT를 이용하여 문항 수준을 모델링할 시 다음과 같은 장점을 지닌다. 첫째, CFA에서 잠재변수와 지표변수 간의 관계는 선형으로 제한되는 반면(Waller et al., 1996), IRT는 비선형을 가정하기 때문에 데이터에 보다

적합한 모형을 설정할 수 있다. 둘째, IRT는 문항 단위 분석에 최적화된 통계적 기법으로, 문항 난이도와 변별도 등 CFA보다 더 많은 문항 정보를 제공한다. 이러한 이점을 바탕으로 Zickar와 Robie(1999)는 IRT를 응답 왜곡 분석에 적용하고자, 응답 왜곡을 차별기능문항(Differential item functioning, DIF)의 개념으로 설명하였다. DIF란 특정 문항이 집단에 따라 다른 응답 곡선을 가지는 현상으로, 피검자들의 능력수준이 같을지라도 각자가 속해있는 집단에 따라 특정 문항에 다른 응답 확률을 보이는 것을 의미한다(Camilli & Sheppard, 1994). Zickar와 Robie(1999)는 긍정 응답 왜곡을 한 피검자는 같은 능력 수준에서 정직하게 응답을 한 피검자보다 더 높은 응답 확률을 보일 것이라 예측했다. 즉, 응답 왜곡을 한 피검자와 정직하게 응답한 피검자가 서로 다른 문항 특성 곡선을 보일 것임을 가정하여 응답 왜곡의 프로세스를 탐색하고자 하였다. 이를 위하여 이들은 응답 왜곡을 능력 수준이 일시적으로 변화하는 현상, 즉 θ shift 현상이라 정의한 뒤 changing-items paradigm과 changing-persons paradigm 두 가지의 모델링 방법을 제시하였다. Zickar와 Robie(1999)는 특히 changing-persons paradigm을 이용하여 응답 왜곡을 분석할 수 있음을 보였는데, 실제로 그들의 연구에서 정직 응답 집단과 왜곡 응답 집단 간 능력 수준에 차이가 있을 뿐만 아니라 문항 곡선 또한 다른 형상을 보이는 것으로 나타났다. Ferrando와 Anguiano-Carrasco(2013)는 분석의 정확성을 보다 더 향상시키기 위해서 changing-items paradigm을 사용해야 함을 주장했다. 그에 따르면, 응답 왜곡은 문항마다 그로 인한 변화량이 다르게 나타나는 현상이므로(Kuncel & Borneman, 2007), 변화를 문항 수준에서 모델링해야 한다는 것이다.

1. theta(θ) shift

1) changing-persons paradigm

changing-persons paradigm은 응답 왜곡으로 인한 변화를 피검자의 능력 수준에서 모델링하는 것이다. IRT에서는 피검자 능력 불변성을 기본으로 가정하기 때문에 검사 도구와 목적에 따라 능력이 다르게 추정되지 않는다(성태제, 2001). 그러나 changing-persons paradigm에서는 응답 왜곡 시 개인이 시험 점수를 향상, 혹은 하락시키기 위해 자신의 본 능력 수준, 즉 true theta(θ)를 일시적으로 변화시키는 것으로 본다(Zickar와 Robie, 1999). 이 모델은 피검자가 문항에 그들의 true theta보다 더 높거나 낮은 theta를 가진 것처럼 응답할 수 있음을 가정한다. 예를 들어, $\theta = -1.5$ 인 피검자는 긍정적으로 응답을 왜곡할 시 $\theta = 0.0$ 인 것처럼 반응할 수 있다. Zickar와 Drasgow(1996)는 2모수 로지스틱 모델(two-parameter logistic model, 2PLM)을 이용하여 긍정 왜곡 반응을 모델링하였다. 이때 2PLM은 아래와 같은 형태를 지닌다.

$$P(u_i = 1 | \theta = t) = \frac{1}{1 + \exp[-1.7a_i(t - b_i)]'} \quad \text{식 (1.1)}$$

a_i 는 문항 i ($i = 1, \dots, n$)에 대한 변별도(discrimination or slope)이며, b_i 는 문항 i 에 대한 난이도(difficulty or location)이다. u_i 는 능력 수준이 θ 인 피검자의 문항 i 에 대한 응답을 의미한다. Zickar와 Drasgow(1996)는 긍정적으로 왜곡된 응답은 theta 수준이 +.50만큼 증가한 피검자가 응답한

결과라 가정한 뒤 이를 모델링하였다. 식 (1.2)는 이 가정을 식으로 표현한 것이다.

$$P(u_i = 1|\theta = t) = \frac{1}{1 + \exp\{-1.7a_i[(t + .50) - b_i]\}'} \quad \text{식 (1.2)}$$

이들의 연구에서 theta값이 +.25 상승되었을 경우와 +1.0 상승되었을 경우가 함께 검증되었으나 모두 +.50과 같은 것으로 나타났다(Zickar & Drasgow, 1996). 그러나 응답 왜곡이 발생한 모든 문항에서 theta값이 동일한 수준으로 상승했다는 가정은 응답 왜곡을 너무 단순화했다는 한계를 지닌다(Zickar & Robie, 1999).

2) changing-items paradigm

changing-items paradigm은 응답 왜곡으로 인한 변화를 문항 모수 수준에서 모델링하는 것이다. 응답을 왜곡하는 경우, 특정 보기를 선택했을 때의 결과를 다르게 기대하기 때문에 동일한 문항일지라도 정직하게 응답했을 때와는 다르게 인식할 가능성이 있다(Zickar & Robie, 1999). IRT에서 문항 모수 역시 불변하는 개념이나, 이 가정 하에서는 응답 왜곡 시 true theta는 고정된 상태로 문항 모수가 변화하게 된다. Ferrando와 Anguiano-Carrasco(2013)는 changing-items paradigm을 이용하여 응답 왜곡을 모델링 할 수 있음을 보였다. 모형은 다음과 같은 가정을 기본으로 한다. 첫째, 사용되는 척도는 이분 문항을 가정한다. 둘째, 모형은 정직 집단과 응답 왜곡집단 2그룹, 혹은 2시점을 전제로 한다. 이를 바탕으로 전

개된 2단계 모델링은 Zickar와 Robie(1999)의 모델에 changing-items paradigm을 적용하여 정교화한 것이다.

첫 번째 단계에서 두 개의 집단(혹은 시점)에 동일한 구조의 모형을 설정한 뒤 동일성 검증을 실시한다. 이 단계에서 수용 가능한 모형 합치도를 얻게 되면 모수들을 IRT의 형태로 변환시킨다. 변환된 모수들은 2단계에서 우도비를 기반으로 한 적절한 개인 합치도 지수를 산출하는 데 사용된다. 이후 계산된 개인 합치도 지수를 피검자의 응답과 비교하여 왜곡 응답자를 가릴 수 있다.

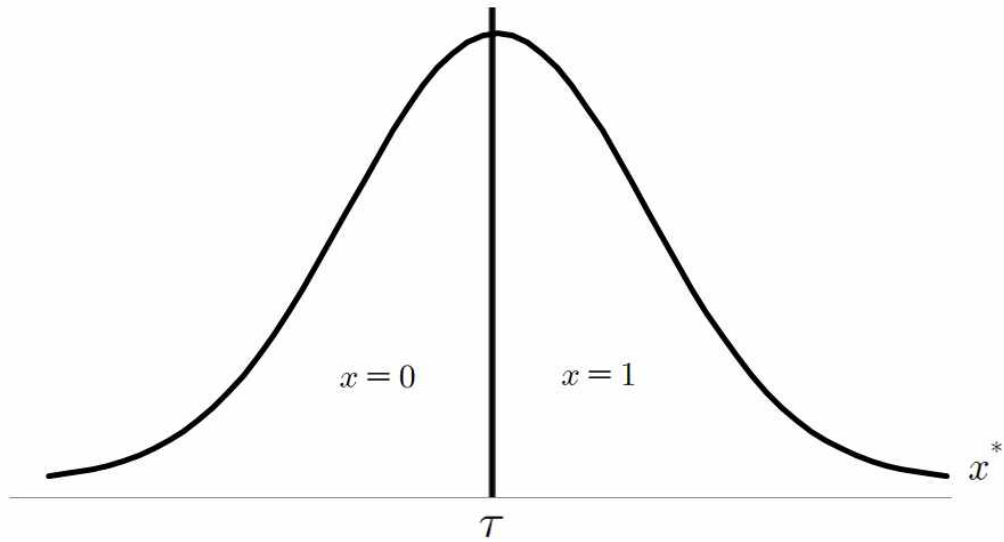
2. 응답 왜곡 모델링

Ferrando와 Anguiano-Carrasco(2013)의 모델링 절차는 아래와 같다. 검사는 능력 θ 를 측정하기 위한 n 개의 이분 문항으로 이루어져 있다고 가정한다. 또한 중립적인 조건에서 검사를 시행한 정직 응답 집단(집단1)과 응답을 왜곡하도록 지시한 응답 왜곡 집단(집단2) 총 두 개의 집단을 상정한다.

1) 1단계: 부분 동일성 모형 설정

집단1과 집단2에 동일한 측정 모형을 설정한다. 모형은 n 개의 문항에 대한 잠재 응답(underlying response) x^* 를 지표 변수로, 능력 θ 를 잠재 변수로 한다. x^* 는 관찰 점수 x 의 기저에 정규 분포 형태로 존재하는 연속형 변수라 가정하므로 관찰 점수 x 는 임계값(threshold) τ 에 대한 반응 결과로 가정한다(Muthén, 1984). 식 (2.1)은 이를 식으로 표현한 것이며, <그림 1>은 식 (2.1)을 도식화한 것이다.

$$\begin{aligned} x_{ij}^{(k)} &= 0 \text{ if } x_{ij}^{*(k)} < \tau_j^{(k)}, \\ x_{ij}^{(k)} &= 1 \text{ if } x_{ij}^{*(k)} \geq \tau_j^{(k)} \end{aligned} \quad \text{식 (2.1)}$$



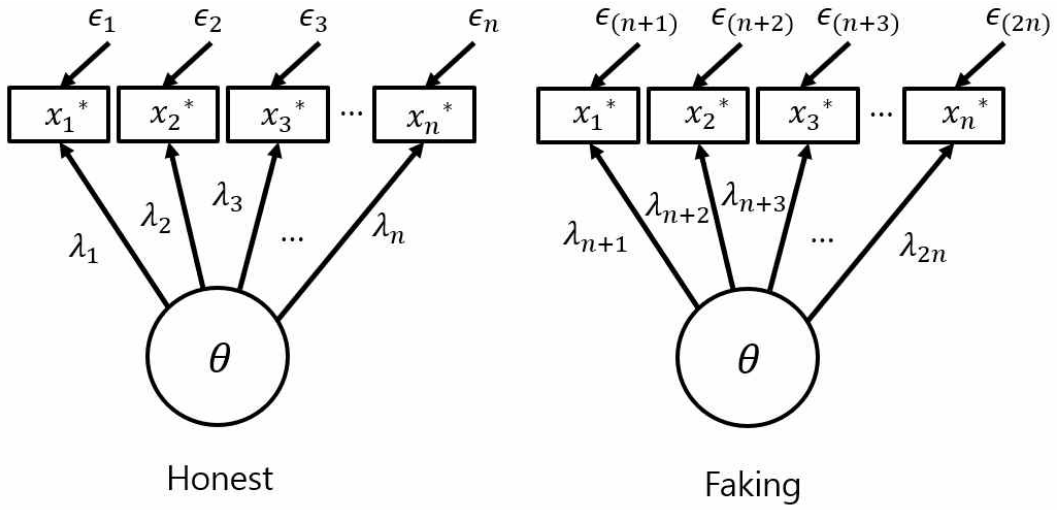
<그림 1> 잠재 응답과 관찰 점수

i 는 집단 k ($k=1$ 혹은 2)에 속하는 피검자를, j 는 문항을 의미한다. 이를 바탕으로 각 집단에 대한 측정 모형을 다음과 같이 설정한다.

$$x_{ij}^{*(k)} = \lambda_j^{(k)} \theta_i + \epsilon_{ij}^{(k)} \quad \text{식 (2.2)}$$

λ_j 는 요인 부하로, 잠재 응답 x^* 와 잠재 변수 θ 간의 관계를 의미한다. ϵ_j 는 θ 로 설명되지 않은 잔차이며, 평균값은 0이다. 잠재 변수 θ 는 평균을 0, 표준편차를 1로 고정하여 표준화 한 변수이다. 이때 집단1과 집단2의 θ 를 동일하게 고정하는데, 이는 두 집단의 true theta가 동일하다는 가정을 반영한 것이다. 때문에 θ 값은 고정된 상태로, 응답 왜곡으로 인한 변화는

문항 모수 수준에서 모델링 된다.



<그림 2> 형태 동일성 모형

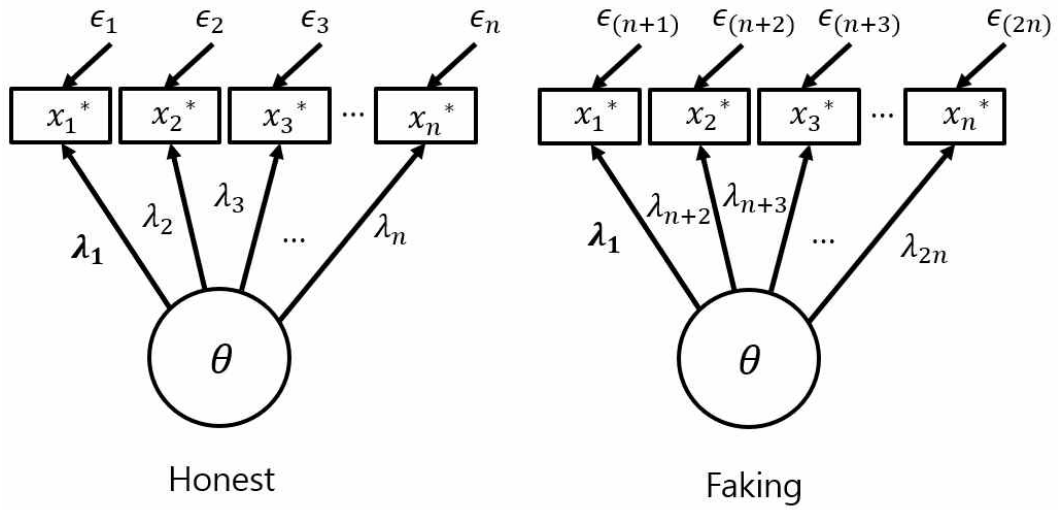
Honest

$$\begin{aligned}
 x_{i1}^* &= \lambda_1 \theta_i + \epsilon_{i1} \\
 x_{i2}^* &= \lambda_2 \theta_i + \epsilon_{i2} \\
 x_{i3}^* &= \lambda_3 \theta_i + \epsilon_{i3} \\
 &\vdots \\
 x_{in}^* &= \lambda_n \theta_i + \epsilon_{in}
 \end{aligned}$$

Faking

$$\begin{aligned}
 x_{i1}^* &= \lambda_{(n+1)} \theta_i + \epsilon_{i(n+1)} \\
 x_{i2}^* &= \lambda_{(n+2)} \theta_i + \epsilon_{i(n+2)} \\
 x_{i3}^* &= \lambda_{(n+3)} \theta_i + \epsilon_{i(n+3)} \\
 &\vdots \\
 x_{i4}^* &= \lambda_{(2n)} \theta_i + \epsilon_{i(2n)}
 \end{aligned}$$

식 (2.3)



<그림 3> 부분 측정 단위 동일성 모형

Honest

$$\begin{aligned}
 x_{i1}^* &= \lambda_1 \theta_i + \epsilon_{i1} \\
 x_{i2}^* &= \lambda_2 \theta_i + \epsilon_{i2} \\
 x_{i3}^* &= \lambda_3 \theta_i + \epsilon_{i3} \\
 &\vdots \\
 x_{in}^* &= \lambda_n \theta_i + \epsilon_{in}
 \end{aligned}$$

Faking

$$\begin{aligned}
 x_{i1}^* &= \lambda_1 \theta_i + \epsilon_{i(n+1)} \\
 x_{i2}^* &= \lambda_{(n+2)} \theta_i + \epsilon_{i(n+2)} \\
 x_{i3}^* &= \lambda_{(n+3)} \theta_i + \epsilon_{i(n+3)} \\
 &\vdots \\
 x_{i4}^* &= \lambda_{(2n)} \theta_i + \epsilon_{i(2n)}
 \end{aligned}$$

식 (2.4)

이후 각 집단에 설정된 모형을 다집단 분석의 동일성 검증 과정을 따라 추정한다. 구체적으로, 형태 동일성(configural invariance), 측정 단위 동일성(metric invariance), 측정 원점 동일성(scalar invariance), 분산 동일성(variance invariance) 순이다(이순목, 이찬순, 이현정, 여성칠, 2012). <그림 2>와 식(2.3)은 가장 첫 번째 단계인 형태 동일성 검증을 도식화한 것이다. 이 단계는 모든 단계 중 가장 제약이 없는 모형으로, 두 집단에 동일한 구조의 모형을 설정하나 집단 간 상응하는 모수가 같을 필요는 없다. 따라서 이 단계에서의 모형은 기저 모형(baseline model)처럼 취급된다. 측정 단위 동일성 모형은 <그림 3>과 같이 요인 부하 λ_j 가 동일하다는 제약을 더한 모형이다. 이전의 형태 동일성 모형에 제약을 가하여 추정해야 할 모수의 수를 줄였으므로 측정 단위 동일성 모형은 형태 동일성 모형에 위계적으로 내재 되어 있다. 따라서 카이제곱 차이 검정(chi-square test, x^2 test)을 사용하여 동일성 검증이 가능하다. x^2 차이 검정에서는 제약을 가하기 전의 모형과 제약을 가한 후의 모형 간 x^2 차이 값을 x^2 분포에서 검증하게 되는데, 이때 사용되는 x^2 분포는 두 모형의 자유도 차이 값을 따르게 된다. 검정 과정에서 사용되는 영가설은 ‘두 모형 간 적합도의 차이는 없다’이며, 이 경우 ‘측정 단위 동일성이 성립한다’이다. 만약 x^2 분포에서의 p 값이 .05 이상일 경우 유의수준 5%에서 영가설을 기각하는 데 실패하고 측정 단위 동일성이 성립한다는 결론을 내릴 수 있다. 이 단계에서는 순차적으로 λ_j 값을 동일하게 제약하여 Δx^2 을 확인한다. 이 과정에서 집단1과 집단2 간 다른 λ_j 값을 가지는 것으로 밝혀진 문항은 응답을 정직하게 하도록 지시했을 경우와 응답을 왜곡하도록 지시했을 경우의 답이 통계적으로 차이가 있다는 의미이므로 왜곡하기 쉬운 문항(fakeable item)이라 할 수 있다. 반대로, 두 집단 간 λ_j 값이 통계적으로 차이가 있지 않을 경우, 해당 문항은 왜곡하기 쉽지 않은 문항이라 결론 내릴 수 있다. 이러한 과정에서

동일성이 전반적으로 성립되는 것이 아닌, 부분적으로 성립되는 것이 필연적이기 때문에 1단계에서 설정되는 동일성 모형은 부분 동일성 모형 (partially invariant model)이라 일컫는다. 다음 단계인 측정 원점 동일성 모형 역시 이전 단계의 모형인 측정 단위 동일성 모형에 내재되어 있으므로 x^2 차이 검정을 실시한다. 이 단계에서는 두 집단의 평균 구조(mean structure)를 비교하여 τ_j 값을 추정하게 된다. 다음의 분산 동일성 검증 단계에서 역시 동일한 절차로 ϵ_j 의 분산 값을 추정할 수 있다.

동일성 검증으로 추정된 λ_j , τ_j , $\text{var}(\epsilon_j)$ 값은 변환식을 통해 2 모수 정규 오자이브 모형(two-parameter normal-ogive model, 2PNOM)을 따르는 문항 모수 값으로 치환 가능하다(McDonald, 1999). 식은 다음과 같다.

$$b_j^{(k)} = \frac{\tau_j^{(k)}}{\lambda_j^{(k)}} \quad \text{식 (2.5)}$$

$$a_j^{(k)} = \frac{\lambda_j^{(k)}}{\sqrt{\text{var}(\epsilon_j^{(k)})}} \quad \text{식 (2.6)}$$

식 (2.5)와 식 (2.6)을 통하여 응답 왜곡으로 발생한 변화가 문항 모수 수준에서 모델링되었음을 확인할 수 있다. 응답 왜곡에 영향을 받지 않은 문항은 임계값(τ_j)과 요인 부하량(λ_j)이 변화하지 않기 때문에 난이도 값(b_j) 역시 변화하지 않는다. 반면 응답 왜곡에 영향을 받은 문항은 임계값과 요인 부하량이 변화하기 때문에 다른 난이도 값을 갖게 된다. 예를 들어, 긍정 왜곡의 경우 임계값의 감소 혹은 요인 부하량의 증가로 인하여 난이도 값이 감소하게 되므로 정직하게 응답한 집단과 다른 난이도 값을 지니게 된

다.

계산된 난이도 값과 변별도 값(a_j)을 활용하여 해당 난이도 값과 변별도 값, θ 값에 따른 응답 확률을 계산할 수 있다. 식 (2.7)은 이를 2PNOM과 2PLM을 이용하여 계산한 것이다.

$$P_j(a_j^{(k)} b_j^{(k)} \theta) = \Phi(a_j^{(k)}(\theta - b_j^{(k)})) \approx \frac{\exp(Da_j^{(k)}(\theta - b_j^{(k)}))}{1 + \exp(Da_j^{(k)}(\theta - b_j^{(k)}))} \quad \text{식 (2.7)}$$

Φ 는 누적정규분포함수를 의미한다. 또한 $D=1.702$ 로, 로지스틱 모형으로 추정된 값을 정규오자이브 모형과 일치시키기 위한 상수이다(Harley, 1952). 식 (2.7)은 조건부 응답 확률(conditional response probability)이 θ 와 난이도 값의 차이에 따라 달라짐을 보이고 있다. 이에 따르면 난이도 값이 감소하는 것은 θ 값이 증가하는 것과 같은 맥락이다. 때문에 긍정 왜곡으로 인하여 증가된 θ 값은 난이도 값의 감소로 모델링 될 수 있다. 즉, 응답 왜곡으로 유발된 theta shift는 문항 모수 수준에서 모델링 가능하다.

요약하면, 1단계에서는 부분 동일성 모형의 검증을 통해 정직 응답 집단과 응답 왜곡 집단의 임계값과 요인 부하량, 잔차의 분산을 추정한다. 이 과정에서 두 집단 간 서로 일치하지 않는 값이 발견된다면 해당 값이 발생된 문항은 응답 왜곡에 영향을 받은 것으로 간주한다. 마찬가지로, 두 집단의 값이 일치한다면 해당 문항은 응답 왜곡에 영향을 받지 않은 것으로 여길 수 있다. 동일성 검증 과정에서 수용 가능한 모형 합치도가 산출되면 회귀 계수의 모수들을 식 (2.5)와 식 (2.6)을 통해 IRT 문항 모수로 변환한다. 일련의 과정들을 통해 얻은 문항 모수들은 다음 단계에서 사용되어진다.

SEM을 활용한 1단계 절차는 SEM의 모형 합치도 검증 과정을 거치기 때

문에 후에 정직한 집단과 응답 왜곡 집단의 응답 패턴 도출 시 정확성을 향상시켜준다는 이점을 지닌다. Levine & Drasgow(1998)는 2단계에서 사용될 적절성 지수 산출은 각 집단의 모형을 얼마나 정확하게 설정하고 문항 모수 값을 계산하느냐에 따라 좌우된다고 밝힌 바 있으므로 SEM을 활용한 교정 과정이 정확성에 긍정적인 영향을 미치리라 예측할 수 있다. 실제로 Ferrando와 Anguiano-Carrasco(2013)가 경험 연구를 실시한 결과, changing-persons paradigm을 사용한 기존 연구보다 적중률이 더 향상된 것으로 나타났다.

2) 2단계: 개인 합치도 지수 산출

Levine과 Drasgow(1988)는 네이만 피어슨 보조 정리(Neyman-Pearson Lemma)에 기반한 적절성 지수를 고안하였다. 이는 우도비 검정(likelihood ratio test)을 이용한 개인 합치도 분석으로, Zickar와 Drasgow(1996)는 2단계에서 이를 응답 왜곡의 식별에 적용하였다. 이들은 비전형적인 반응유형을 응답 왜곡자로 상정하여 탐지할 수 있도록 하였는데, 특히 changing-persons paradigm을 활용하여 응답 왜곡 패턴은 θ 값이 +.50만큼 상승한 것으로, 정직 응답 패턴은 θ 값이 변화하지 않는 것으로 모델링 하였다. 그러나 이는 응답 왜곡을 너무 단순화했다는 지적을 받았다(Ferrando & Anguiano-Carrasco, 2013; Zickar & Robie, 1999). 때문에 Ferrando와 Anguiano-Carrasco(2013)는 1단계에서 changing-items paradigm을 사용하여 응답 왜곡으로 인한 다양한 변화량이 SEM을 통해 모델링 될 수 있도록 하였다. 따라서 본 절차에서는 1단계에서 도출한 모형들을 기반으로 한 응답 왜곡 반응 패턴의 우도 함수와 정

직 응답 반응 패턴의 우도 함수를 계산하여 사용한다.

정직 응답 모형에서 능력 수준 θ 가 주어졌을 때 반응 패턴 x 의 우도 함수를 다음과 같이 계산할 수 있다.

$$L_H(x|a_H b_H \theta) = \prod_j^n P_j^{x_j}(a_H b_H \theta) [1 - P_j(a_H b_H \theta)]^{1-x_j} \quad \text{식 (2.8)}$$

이때 a_H 와 b_H 는 식 (2.5)와 식 (2.6)으로 도출된 문항 모수 값의 벡터 (vector)이며, P 는 식 (2.7)로 계산된 조건부 응답 확률이다. 식 (2.8)을 이용하여 비조건부 우도 함수(unconditional likelihood function)를 아래와 같이 계산할 수 있다. 이때 식 (2.9)의 적분을 통해 식 (2.8)의 우도 함수를 일정 수치로 근사시킬 수 있다.

$$P_H(x|a_H b_H) = \int_{-\infty}^{\infty} L_H(x|a_H b_H \theta) f(\theta) d\theta \quad \text{식 (2.9)}$$

일련의 계산 과정을 응답 왜곡 조건에서 반복하여 동일한 결과를 얻는 것이 가능하다. 식 (2.10)과 식 (2.11)은 각각 응답 왜곡 조건 하에서의 반응 패턴 우도 함수와 비조건부 우도 함수이다.

$$L_F(x|a_H b_H \theta) = \prod_j^n P_j^{x_j}(a_F b_F \theta) [1 - P_j(a_F b_F \theta)]^{1-x_j} \quad \text{식 (2.10)}$$

$$P_F(x|a_F b_F) = \int_{-\infty}^{\infty} L_F(x|a_F b_F \theta) f(\theta) d\theta \quad \text{식 (2.11)}$$

계산된 각 조건의 우도 함수를 이용하여 우도비(likelihood ratio, LR)를 구할 수 있다.

$$LR = \frac{P_F(x|a_F b_F)}{P_H(x|a_H b_H)} \quad \text{식 (2.12)}$$

응답 x 는 비로소 식 (2.13)의 결정 규칙(decision rule)을 통해 왜곡된 응답으로의 분류 여부가 가려지게 된다.

$$P_F(x|a_F b_F) \geq k P_H(x|a_H b_H) \quad \text{식 (2.13)}$$

이때 임계값(critical value) k 는 적중률(hit rate)과 오경보율(false alarm rate)을 결정하는 절단값(cut-off value)으로 해석할 수 있다. 이 경우 적중률은 왜곡된 응답을 한 피검자가 왜곡 응답자로 분류되는 비율을 의미하며, 오경보율은 정직하게 응답한 피검자가 왜곡 응답자로 잘못 분류되는 비율을 의미한다. 효과적인 k 값은 상황에 따라 다를 수 있으나, 사전 확률이 동일한 경우, 즉 왜곡 응답자와 정직 응답자의 비율이 같은 경우 k 값을 1로 설정하는 것이 가장 효과성이 크다(McNicol, 1972). 때문에 Ferrando와 Anguiano-Carrasco(2013)는 경험 연구에서 k 값을 1로 설정하여, 거짓으로 응답할 확률이 정직하게 응답할 확률보다 같거나 큰 피검자

를 왜곡 응답자로 분류하였다.

정리하면, 2단계에서는 1단계에서 설정한 왜곡 응답 모형과 정직 응답 모형을 바탕으로 각각의 우도 함수를 계산한다. 이를 바탕으로 우도비를 활용한 적절성 지수를 산출하며, 이 지수는 임계값 k 의 수치에 따라 응답을 왜곡된 응답으로 분류할지 정직한 응답으로 분류할지를 결정하는 데 사용된다.

Ⅲ. 연구문제 및 가설

연구문제1. 자료의 특성에 따라 제1종 오류(type-I error) 수준에 차이가 있는가?

가설 1-1. 검사 길이에 따라 제1종 오류 비율이 달라질 것이다.

가설 1-2. 왜곡하기 쉬운 문항의 비율에 따라 제1종 오류 비율이 달라질 것이다.

가설 1-3. 응답을 왜곡한 사람들의 비율에 따라 제1종 오류 비율이 달라질 것이다.

가설 1-4. 피검자의 수에 따라 제1종 오류 비율이 달라질 것이다.

가설 1-5. 응답 왜곡의 유형에 따라 제1종 오류 비율이 달라질 것이다.

가설 1-6. 응답 왜곡의 크기에 따라 제1종 오류 비율이 달라질 것이다.

연구문제2. 자료의 특성에 따라 통계적 검정력(power) 수준에 차이가 있는가?

가설 2-1. 검사 길이에 따라 검정력이 달라질 것이다.

가설 2-2. 왜곡하기 쉬운 문항의 비율에 따라 검정력이 달라질 것이다.

가설 2-3. 응답을 왜곡한 사람들의 비율에 따라 검정력이 달라질 것이다.

가설 2-4. 피검자의 수에 따라 검정력이 달라질 것이다.

가설 2-5. 응답 왜곡의 유형에 따라 검정력이 달라질 것이다.

가설 2-6. 응답 왜곡의 크기에 따라 검정력이 달라질 것이다.

IV. 연구 방법

1. 자료 생성 (data generation)

본 연구에서는 프로그램 R의 몬테카를로 시뮬레이션(Monte Carlo Simulation) 기능을 사용하여 자료를 생성하였다. 몬테카를로 시뮬레이션은 연구자의 목적에 따라 통제된 환경에서 확률적 분포를 따르는 자료를 무작위(random)로 생성하는 방법으로, 의사결정을 위한 일종의 모의실험이라 할 수 있다(Stephenson & Holbert, 2003).

본 연구는 changing-items paradigm을 활용한 적절성 지수가 다양한 조건 하에서 응답 왜곡을 식별할 수 있는지를 검증하는 것을 목적으로 한다. 때문에 dichotomous 2PLM을 따르는 자료를 다음과 같은 조건 하에 생성하였다. 1)검사의 길이(short, medium, long), 2)왜곡하기 쉬운 문항의 비율(20%, 50%), 3)피검자의 수(100명, 300명, 500명), 4)응답 왜곡 피검자들의 비율(honest:faker=1:1, 3:2, 4:1), 5)응답 왜곡 유형(Spuriously High, Spuriously Low) 6)응답 왜곡의 크기(Small, Large). 또한 능력모수 θ 값은 평균 0, 표준편차 1의 정규분포를 따르도록 생성하였으며 문항모수 a, b는 각각 최소 0.5, 최대 1.5 값을 지니는 연속균등분포를, 평균 1, 표준편차 0.25의 값을 지니는 정규분포를 따르도록 생성하였다. 이는 개인합치도 지수들의 효과성을 검증하기 위해 수행된 다수의 선행 연구들을 따른 것으로, 이전 연구들과 조건들을 동일하게 함으로써 검증 결과의 설득력을 높이기 위함이다. 이후 각 조건에서의 검정력과 제1종 오류 수준을 분석하고, 추가적으로 각 조건의 수렴율(convergence rate)을 확인하였다. 이 때 수렴율이란 전체 반복시행 중 모형 추정이 성공한 비율을 의

미하며, $1 - (NC^2)/\text{총 반복횟수}$ 로 계산하였다. 최종적으로 자료는 $3 \times 2 \times 3 \times 3 \times 2 \times 2$ 의 요인 설계를 바탕으로 생성되었으며, 주어진 조건을 각 500회 씩 반복하였다.

1) 독립변수

(1) 검사 길이

검사 길이는 개인 합치도 지수에 영향을 미치는 주요 변수 중 하나로, 검사 길이에 따라 개인 합치도 지수의 효용성이 달라지는 것이 일반적인 견해이다(Dimitrov & Smith, 2006; Karabatsos, 2003; Tendeiro & Meijer, 2014). Dimitrov & Smith(2006)와 Tendeiro & Meijer(2014)의 시뮬레이션 연구에서 조건은 10~40문항으로 구성되었으나 Ferrando와 Anguiano-Carrasco(2013)의 경험 연구에서 문항들이 20~30문항이었던 것을 감안했을 때, 다른 길이에서의 검증이 추가적으로 필요하므로 Karabatsos(2003)의 조건을 참고하여 short 조건 18문항, medium 조건 34문항, long 조건 66문항으로 구성하였다.

(2) 왜곡하기 쉬운 문항의 비율

기존의 DIF 연구에서는 전체 문항 중 10~40%를 DIF 문항으로 설정하였으며(Raju et al., 2009; Rojers et al., 1993; Wang et al., 2010), 50%까지 설정한 사례도 있기 때문에(박찬호, 2009) 본 연구에서는 왜곡하기 쉬운 문항의 비율을 20%, 50% 두 가지로 설정하였다.

2) Not Converged, 즉 모형이 수렴되지 않아 값을 추정할 수 없음을 의미한다.

(3) 피검자의 수

DIF문항 탐지에 구조방정식 모형을 이용한 강유진(2012)의 연구에서 전체 표본크기를 100, 300, 500으로 설정한 것을 따라 본 연구에서도 피검자 수를 동일하게 설정하였다.

(4) 응답 왜곡 피검자들의 비율

집단 간 표본크기의 차이가 결과에 영향을 주었다는 기존의 DIF 연구를 참고하여(Byrne & Stewart, 2006; González-Romá et al., 2006; Hancock et al., 2000; Kaplan & George, 1995), 참조집단(honest)과 관심집단(faker)의 비율을 1:1, 3:2, 4:1로 설정하였다. 이 때 1:1은 각 집단의 표본크기에 차이가 없는 경우이며 3:2는 표본크기의 차이가 작은 경우, 4:1은 차이가 큰 경우이다.

(5) 응답 왜곡 유형

선행 연구들을 참고하여, Spuriously High 조건과 Spuriously Low 조건으로 구성하였다(Sinharay, 2017; Tendeiro & Meijer, 2014). 이들은 각각 긍정적 응답 왜곡과 부정적 응답 왜곡으로 대입할 수 있다. 본 연구에서는 응답 왜곡을 문항 모수 수준에서 모델링하므로(changing-items paradigm), 응답 왜곡은 난이도 값(b_j)의 변화로 표현되어진다. 따라서 긍정적으로 응답 왜곡을 하는 것은 보다 쉬운 문항에 응답 하는 것과 같다고 할 수 있다. 이는 난이도 값(b_j)이 감소한 것과 같은 의미로, 식(2.5)에 따르면 난이도 값(b_j)은 임계값(τ_j)을 요인 부하량(λ_j)으로 나눈 값과 동일하

기 때문에 긍정적 응답 왜곡은 임계값(τ_j)이 감소하고 요인 부하량(λ_j)이 증가하는 현상이라 해석할 수 있다. 반대로 부정적 응답 왜곡은 난이도 값(b_j)이 증가하는 현상으로, 임계값(τ_j)이 증가하고 요인 부하량(λ_j) 감소한 결과라 할 수 있다. 그러므로 본 연구에서는 긍정적 응답왜곡을 임계값(τ_j)의 감소와 요인 부하량(λ_j)의 증가로, 부정적 응답왜곡을 임계값(τ_j)의 증가와 요인 부하량(λ_j)의 감소로 설정하였다.

(6) 응답 왜곡 크기

Stark와 Chernyshenko(2006)는 확인적 요인 분석(Confirmatory factor analysis)이 DIF문항을 효과적으로 변별할 수 있는지를 검증하기 위하여 시뮬레이션 연구를 수행한 바 있다. 본 연구는 그들의 절차를 따라 작은 왜곡 크기(Small) 조건에서는 요인 부하량(λ_j)을 0.15, 임계값(τ_j)을 0.25만큼 변화시켰으며 큰 왜곡 크기(Large) 조건에서는 요인 부하량(λ_j)을 0.4, 임계값(τ_j)을 0.5만큼 변화시켰다. 응답 왜곡의 종류에 따른 각 값들의 변화량을 아래 표에 정리하였다.

<표 1> 응답 왜곡 종류에 따른 응답 왜곡 크기

		왜곡 크기	
		Small	Large
왜곡 종류	긍정왜곡	λ_j	+0.15
		τ_j	-0.25
	부정왜곡	λ_j	-0.15
		τ_j	+0.25

2) 종속변수

(1) 제 1종 오류 비율(Type-I error)

제 1종 오류는 영가설이 참임에도 불구하고 영가설을 기각하는 오류로, 가설 검정 시 발생하는 대표적인 착오 중 하나이다(성태제, 2014). 본 연구에서는 응답을 왜곡하지 않았음에도 응답 왜곡자로 분류되는 경우를 일컬으며, 이는 정직 응답자 중 응답 왜곡자로 분류되는 횟수의 비율로 표현된다.

(2) 검정력 (Power)

검정력이란 참이 아닌 영가설을 기각하는 확률로, 본 연구에서는 응답 왜곡자를 응답 왜곡자로 분류하는 경우를 의미한다. 따라서 응답 왜곡자 중 최종적으로 응답 왜곡자라 정확히 분류된 횟수의 비율로 계산하였다.

2. 자료 분석

1단계에서 생성된 자료를 바탕으로 모형의 적합성을 평가하기 위해 합치도 지수를 확인하였다. 합치도 지수는 통계적 합치도 지수인 χ^2 값과, 실용적 합치도 지수인 RMSEA(Root Mean Square Error of Approximation), CFI(Comparative Fit Index), TLI(Tucker and Lewis Index)값을 사용하였다. 2단계에서는 모형에서 추정된 요인 부하량(λ_j)과 임계값(τ_j), 잔차의 분산(ϵ_j)값을 사용하여 정직 응답 집단과 왜곡 응답 집단 각각의 문항 난이도(b_j)값과 문항 변별도(a_j)값을 계산하였다. 이후, 계산된 문항 모수 값과 기존에 생성한 응답 데이터를 바탕으로 하여 개인별 응답 확률을 구하였고, 이를 통해 각 집단에서의 우도 함수를 산출하여 적절성 지수를 계산하였다. 그 후 계산된 적절성 지수로 가상의 피검자들을 응답 왜곡 집단과 정직 집단으로 분류한 뒤, 데이터 생성 시 피검자가 원래 속해있던 집단(이하 true 집단)과 비교하였다. 마지막으로, 적절성 지수를 통해 배정된 집단과 true 집단 값이 일치하는지 여부를 통해 오경보율과 검정력을 확인하였다. 위 과정을 각 자료 특성 조건 아래에서 동일하게 시행하였으며, 최종적으로 각 조건 아래에서의 1종 오류와 검정력의 변화를 관찰하였다. 추가적으로 1단계에서 설정한 모형이 자료에 수렴된 정도를 확인하기 위해 전체 반복횟수 중 모형이 수렴된 횟수의 비율을 확인하였다. 분석에는 프로그램 R과 Mplus 8.0을 사용하였다.

V. 연구 결과

응답 왜곡을 문항 모수 수준에서 모델링하는 방식이 다양한 조건 하에서 응답 왜곡을 효과적으로 탐지하는지를 확인해보고자 하였다. 이를 위해 몬테카를로 시뮬레이션 기능을 이용하여 데이터를 생성한 뒤 각 조건에서의 1종 오류와 검정력의 비율을 구하였다. 또한 그 과정에서 모형 합치도와 수렴율을 통해 모형의 적합성을 확인하였다. 본 연구에서의 주된 관심사는 1종 오류 비율과 검정력이므로 조건 별 합치도 지수와 수렴율을 <부록1>에 제시하였다.

모형 합치도를 살펴본 결과, 대부분의 조건 하에서 모형 합치도는 양호한 수준 이상인 것으로 나타났다. 모든 조건에서 χ^2 값은 통계적으로 유의하지 않은 것으로 나타났는데, 이는 표본의 공분산 행렬과 추정치로 재생산된 공분산 행렬이 동일하다는 가설을 기각하는 데 실패했음을 의미한다. 그러므로 모형이 자료에 적합하다는 결론을 내릴 수 있다. 또한 CFI와 TLI의 값은 표본 수가 100명이었던 경우를 제외하고 모두 0.90 이상인 것으로 나타났다. Hu와 Bentler(1999)에 따르면 이는 양호한 모형이라 해석 가능한 수치이다. 표본 수가 100명이었던 조건들에서는 CFI와 TLI의 값이 0.90을 넘지 않았으나 최소 0.89로, 0.90에 가까운 수치였다. 마지막으로 RMSEA 값은 모든 조건에서 0.05 이하인 것으로 나타났는데, 이는 적합도가 매우 좋음을 의미한다 (Browne & Cudeck, 1993).

수렴율의 경우 대부분의 조건에서 80% 이상이었으나, 표본이 100명인 조건에서는 모두 70% 이하였다. 긍정적 응답왜곡의 경우 문항 수가 많아질수록 낮은 수렴율을 보였으며, 특히 긍정적 응답왜곡의 크기가 큰 경우에는 왜곡하기 쉬운 문항의 비율이 50%일 때 300명 이상의 큰 표본 수에서도 낮은 수렴율을 보고했다. 따라서 해당 조건에서의 1종 오류와 검정력 해석에 주의가 필

요하다.

1. 자료의 특성에 따른 1종 오류 비율

6개의 조건(검사 길이, 왜곡하기 쉬운 문항의 비율, 피검자의 수, 응답 왜곡 피검자의 비율, 응답 왜곡 유형, 응답 왜곡 크기) 하에서 1종 오류 비율이 변화하는 양상을 살펴보았다. 각 독립변수들의 개별적인 효과를 살펴보기 위해 분석의 주가 되는 독립변수 외의 조건들은 동일하게 하여 통제하였다.

1) 검사 길이에 따른 1종 오류 비율

검사 길이에 따른 1종 오류의 비율을 <표 2>와 <표 3>, <그림 4>에 제시하였다. 그 결과, 모든 조건에서 검사 길이가 증가할수록 1종 오류의 비율이 감소하는 것으로 나타났다. 수렴율이 80%이상이었던 조건들 중 변화량이 가장 큰 조건은 전체 피검자 500명의 부정 왜곡 조건에서 응답 왜곡 피검자의 비율이 4:1이고 왜곡 문항의 비율이 20%이며, 왜곡의 크기가 클 때였으며, 그 수치는 0.167이었다. 변화량이 가장 작았던 조건은 전체 피검자 수가 500명이고 응답 왜곡 피검자의 비율이 1:1이며 왜곡 문항의 비율이 20%, 왜곡의 크기가 작은 긍정 왜곡 조건으로, 그 수치는 0.062였다. <그림 4>는 위의 두 조건을 기술한 것으로, Biggest는 변화량이 가장 컸던 경우를, Smallest는 가장 작았던 경우를 의미한다. 변화량에 차이가 있으나 검사 길이가 길어질수록 1종 오류가 낮아지는 양상을 동일하게 보임을 알 수 있다.

추가적으로 수렴율을 살펴보았을 때, 전체 피검자 수가 100명인 조건에서는 수렴율이 모두 80% 미만이었으며, 긍정 왜곡 조건의 경우 검사 길이가 길어질수록 수렴율 또한 80%미만으로 낮아졌다. 특히 긍정 왜곡 조건에서 왜

곡 크기가 큰 경우, 검사 길이가 중간 길이 이상일 때 일관되게 낮은 수렴율을 보였다.

<표 2 > 긍정왜곡 조건에서의 검사 길이에 따른 1종 오류 비율

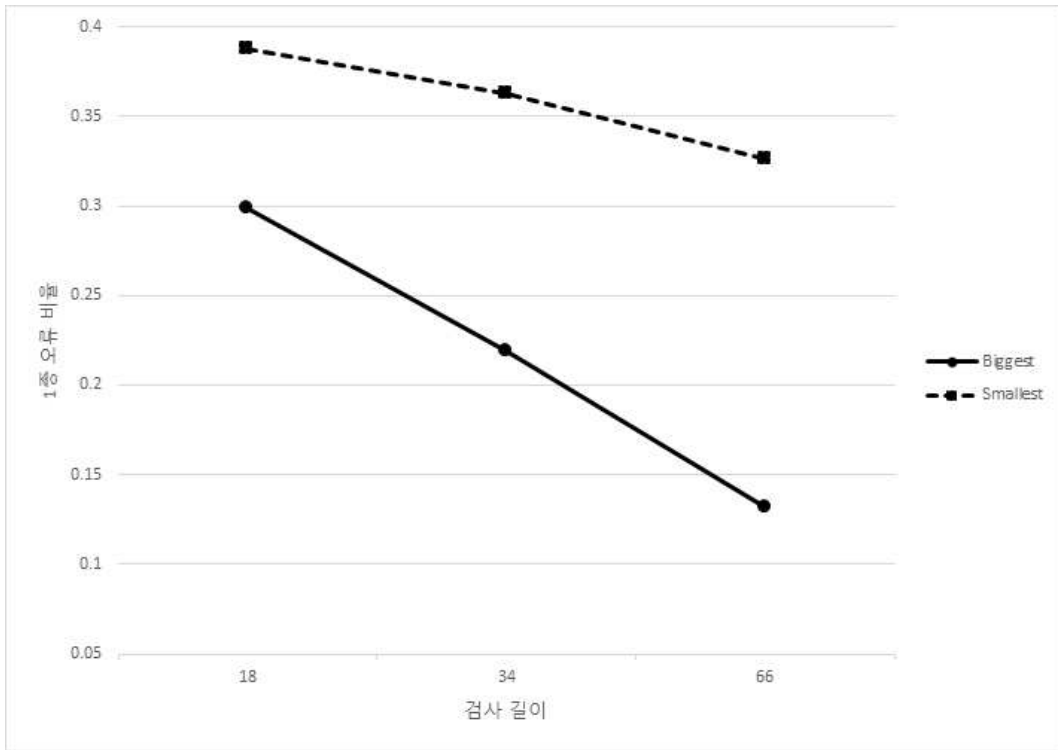
전체 표본 크기	검사 길이	집단 간 표본 크기의 비율											
		1:1				3:2				4:1			
		왜곡문항 20%		왜곡문항 50%		왜곡문항 20%		왜곡문항 50%		왜곡문항 20%		왜곡문항 50%	
		응답왜곡 크기		응답왜곡 크기		응답왜곡 크기		응답왜곡 크기		응답왜곡 크기		응답왜곡 크기	
		Small	Large	Small	Large	Small	Large	Small	Large	Small	Large	Small	Large
100	18	0.278	0.252	0.245	0.198	0.289	0.265	0.255	0.215	0.267	0.258	0.237	0.215
	34	0.240	0.224	0.194	0.154	0.236	0.212	0.202	0.146	0.188	0.183	0.163	0.138
	66	0.205	0.152	0.155	0.145	0.194	0.161	0.139	0.063	0.115	0.106	0.074	0.067
300	18	0.370	0.332	0.326	0.259	0.372	0.334	0.335	0.268	0.353	0.330	0.329	0.273
	34	0.338	0.290	0.291	0.209	0.335	0.294	0.293	0.215	0.297	0.272	0.269	0.204
	66	0.295	0.241	0.232	0.149	0.288	0.239	0.232	0.162	0.229	0.194	0.202	0.116
500	18	0.388	0.341	0.346	0.269	0.390	0.346	0.350	0.273	0.390	0.354	0.356	0.290
	34	0.363	0.308	0.307	0.209	0.360	0.309	0.305	0.208	0.342	0.301	0.300	0.222
	66	0.326	0.262	0.255	0.167	0.318	0.258	0.251	0.153	0.275	0.232	0.233	0.151

주. 수렴율이 80%미만인 조건은 색칠하여 표시하였음

<표 3 > 부정왜곡 조건에서의 검사 길이에 따른 1종 오류 비율

전체 표본 크기	검사 길이	집단 간 표본 크기의 비율											
		1:1				3:2				4:1			
		왜곡문항 20%		왜곡문항 50%		왜곡문항 20%		왜곡문항 50%		왜곡문항 20%		왜곡문항 50%	
		응답왜곡 크기		응답왜곡 크기		응답왜곡 크기		응답왜곡 크기		응답왜곡 크기		응답왜곡 크기	
		Small	Large	Small	Large	Small	Large	Small	Large	Small	Large	Small	Large
100	18	0.279	0.233	0.253	0.174	0.282	0.235	0.250	0.173	0.251	0.208	0.220	0.157
	34	0.242	0.175	0.194	0.108	0.230	0.165	0.180	0.091	0.180	0.142	0.140	0.086
	66	0.197	0.114	0.135	0.058	0.171	0.103	0.114	0.043	0.098	0.073	0.065	0.036
300	18	0.361	0.300	0.315	0.210	0.356	0.296	0.317	0.209	0.334	0.269	0.300	0.190
	34	0.321	0.228	0.259	0.129	0.314	0.222	0.254	0.127	0.273	0.189	0.224	0.104
	66	0.273	0.154	0.198	0.071	0.263	0.146	0.188	0.064	0.199	0.107	0.143	0.044
500	18	0.379	0.320	0.335	0.221	0.379	0.317	0.336	0.221	0.365	0.299	0.324	0.208
	34	0.342	0.245	0.279	0.138	0.338	0.241	0.275	0.136	0.314	0.219	0.255	0.119
	66	0.298	0.168	0.215	0.075	0.287	0.162	0.206	0.072	0.244	0.132	0.175	0.057

주. 수렴율이 80%미만인 조건은 색칠하여 표시하였음



<그림 4> 검사 길이에 따른 1종 오류 비율의 변화

2) 왜곡문항 비율에 따른 1종 오류 비율

왜곡문항 비율에 따른 1종 오류 비율은 <표 4>, <표 5>와 같다. 분석 결과, 모든 조건에서 왜곡문항의 비율이 높아질수록 1종 오류 비율이 낮아지는 경향을 보였다. 수렴율이 80%이상으로 관찰된 조건들 중 가장 큰 변화량을 보인 조건은 전체 피검자 500명의 부정왜곡 조건에서 응답 왜곡 피검자의 비율이 1:1일 때, 검사 길이가 34문항이며 응답 왜곡의 크기가 큰 경우였다. 해당 조건에서의 1종 오류 변화량은 0.107이었다. 반면 가장 작은 변화량을 보인 조건은 전체 피검자 300명의 긍정왜곡 조건에서 응답 왜곡 피검자의 비율이 3:2일 때, 검사 길이가 18문항이며 응답 왜곡의 크기가 작은 경우로, 해당 조건에서의 변화량은 0.024이었다. 위 조건들을 각각 Biggest와 Smallest로 명명하여 <그림 5>에서 제시하였다. 감소폭에 차이가 존재하지만 두 조건 모두 왜곡문항 비율이 20%일 때보다 50%일 때 1종 오류 비율이 더 낮음을 알 수 있다.

그러나 전체 피검자 수가 100명인 조건에서는 수렴율이 모두 80% 미만인 것으로 나타났다. 긍정 왜곡 조건의 경우 검사 길이가 길어질수록, 왜곡 문항의 비율이 커질수록 수렴율 또한 80%미만으로 낮아졌다. 또한 긍정 왜곡 조건에서는 응답왜곡의 크기가 클 때 왜곡 문항의 비율이 50%인 경우 일관되게 낮은 수렴율을 보였다.

<표 4> 긍정왜곡 조건에서의 왜곡 문항 비율에 따른 1종 오류 비율

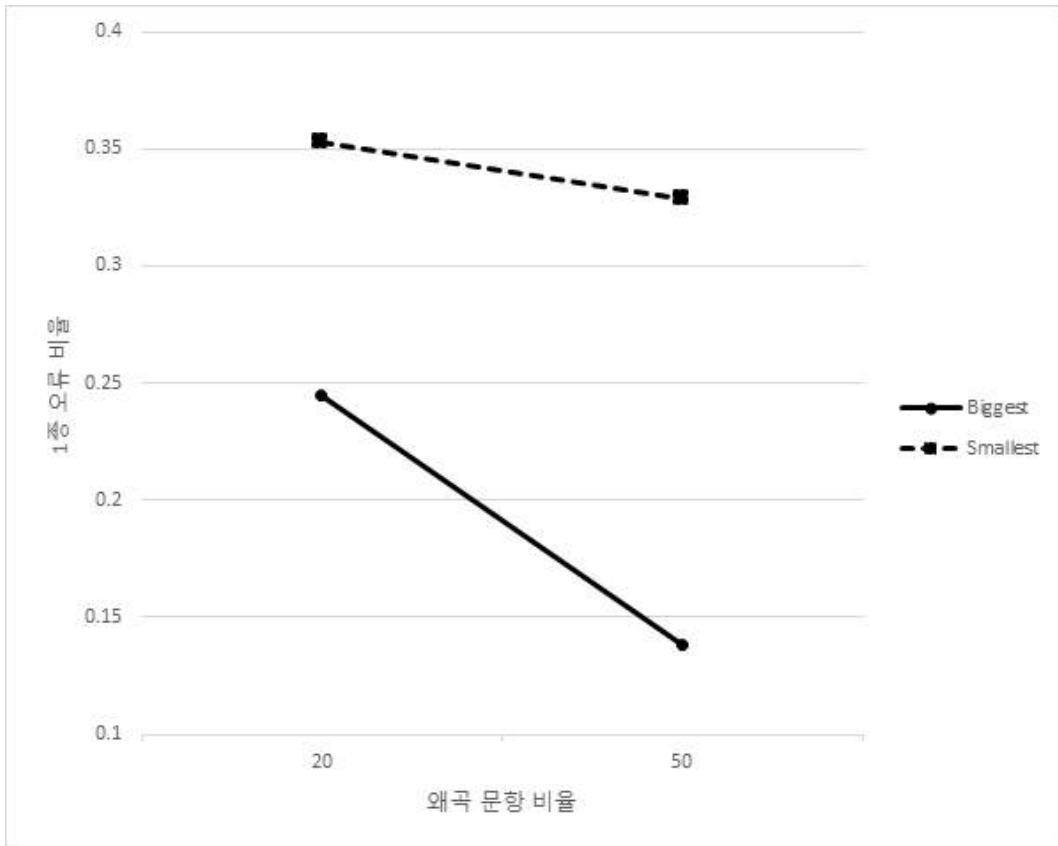
전체 표본 크기		왜곡 문항 비율		집단 간 표본 크기의 비율																	
				1:1				3:2				4:1									
				18문항		34문항		66문항		18문항		34문항		66문항		18문항		34문항		66문항	
				응답왜곡 크기		응답왜곡 크기		응답왜곡 크기		응답왜곡 크기		응답왜곡 크기		응답왜곡 크기		응답왜곡 크기		응답왜곡 크기		응답왜곡 크기	
		Small	Large	Small	Large	Small	Large	Small	Large	Small	Large	Small	Large	Small	Large	Small	Large	Small	Large		
100	20%	0.278	0.252	0.240	0.224	0.205	0.152	0.289	0.265	0.236	0.212	0.194	0.161	0.267	0.258	0.188	0.183	0.115	0.106		
	50%	0.245	0.198	0.194	0.154	0.155	0.145	0.255	0.215	0.202	0.146	0.139	0.063	0.237	0.215	0.163	0.138	0.074	0.067		
300	20%	0.370	0.332	0.338	0.290	0.295	0.241	0.372	0.334	0.335	0.294	0.288	0.239	0.353	0.330	0.297	0.272	0.229	0.194		
	50%	0.326	0.259	0.291	0.209	0.232	0.149	0.335	0.268	0.293	0.215	0.232	0.162	0.329	0.273	0.269	0.204	0.202	0.116		
500	20%	0.388	0.341	0.363	0.308	0.326	0.262	0.390	0.346	0.360	0.309	0.318	0.258	0.390	0.354	0.342	0.301	0.275	0.232		
	50%	0.346	0.269	0.307	0.209	0.255	0.167	0.350	0.273	0.305	0.208	0.251	0.153	0.356	0.290	0.300	0.222	0.233	0.151		

주. 수렴율이 80%미만인 조건은 색칠하여 표시하였음

<표 5> 부정왜곡 조건에서의 왜곡 문항 비율에 따른 1종 오류 비율

전체 표본 크기		왜곡 문항 비율		집단 간 표본 크기의 비율																	
				1:1				3:2				4:1									
				18문항		34문항		66문항		18문항		34문항		66문항		18문항		34문항		66문항	
				응답왜곡 크기		응답왜곡 크기		응답왜곡 크기		응답왜곡 크기		응답왜곡 크기		응답왜곡 크기		응답왜곡 크기		응답왜곡 크기		응답왜곡 크기	
		Small	Large	Small	Large	Small	Large	Small	Large	Small	Large	Small	Large	Small	Large	Small	Large	Small	Large		
100	20%	0.279	0.233	0.242	0.175	0.197	0.114	0.282	0.235	0.230	0.165	0.171	0.103	0.251	0.208	0.180	0.142	0.098	0.073		
	50%	0.253	0.174	0.194	0.108	0.135	0.058	0.250	0.173	0.180	0.091	0.114	0.043	0.220	0.157	0.140	0.086	0.065	0.036		
300	20%	0.361	0.300	0.321	0.228	0.273	0.154	0.356	0.296	0.314	0.222	0.263	0.146	0.334	0.269	0.273	0.189	0.199	0.107		
	50%	0.315	0.210	0.259	0.129	0.198	0.071	0.317	0.209	0.254	0.127	0.188	0.064	0.300	0.190	0.224	0.104	0.143	0.044		
500	20%	0.379	0.320	0.342	0.245	0.298	0.168	0.379	0.317	0.338	0.241	0.287	0.162	0.365	0.299	0.314	0.219	0.244	0.132		
	50%	0.335	0.221	0.279	0.138	0.215	0.075	0.336	0.221	0.275	0.136	0.206	0.072	0.324	0.208	0.255	0.119	0.175	0.057		

주. 수렴율이 80%미만인 조건은 색칠하여 표시하였음



<그림 5> 외국문항 비율에 따른 1종 오유 비율의 변화

3) 응답 왜곡 피검자들의 비율에 따른 1종 오류 비율

응답 왜곡 피검자들의 비율(이하 faker 비율)에 따른 1종 오류의 비율을 살펴본 결과는 <표 6>, <표 7>과 같다. 분석 결과, faker 비율이 변화할 때마다 1종 오류의 비율이 변화하였으나 그 양상은 조건마다 비일관적이었다. 먼저, faker 비율 조건을 제외한 나머지 72개 조건 중 43개의 조건에서 faker의 비율이 작아질수록 1종 오류의 비율이 감소하는 경향이 나타났다. 43개 조건들에는 부정 왜곡 조건에서의 34문항, 66문항 조건 전체와 긍정 왜곡 조건에서의 34문항, 66문항 조건 일부가 포함되었다. 또 다른 18개의 조건들에서는 faker 비율이 3:2일 때 1종 오류가 가장 높거나, faker 비율이 1:1일 때와 3:2일 때의 값이 동일하게 가장 높았다. 위 조건들은 이전의 43개 조건들과 비교했을 때, faker 비율이 4:1인 경우 1종 오류가 가장 낮았다는 공통점이 있다. 반면 나머지 11개의 조건에서는 faker의 비율이 증가할수록 1종 오류 비율도 함께 증가하는 양상을 보였다. 이러한 경향은 긍정 왜곡 조건에서의 18문항 조건들 일부에서 관찰되었으며, 이 중 절반 이상은 80% 미만의 수렴율을 보인 조건들이었다.

추가적으로, 전체 피검자 수가 100명인 조건에서 수렴율은 모두 80%를 넘지 못했으며, 긍정 왜곡의 경우 왜곡 문항이 50%일 때 응답 왜곡의 크기가 크다면 일관되게 낮은 수렴율을 보였다. 특히 검사 길이가 66문항으로 클 경우 대다수의 조건들에서 faker의 비율에 관계없이 수렴율이 낮았다.

<표 6> 긍정왜곡 조건에서의 응답 왜곡 피검자 비율에 따른 1종 오류 비율

전체 표본 크기	집단 간 표본크기 비율	전체 문항 수											
		18문항				34문항				66문항			
		왜곡문항 20%		왜곡문항 50%		왜곡문항 20%		왜곡문항 50%		왜곡문항 20%		왜곡문항 50%	
		응답왜곡 크기		응답왜곡 크기		응답왜곡 크기		응답왜곡 크기		응답왜곡 크기		응답왜곡 크기	
		Small	Large	Small	Large	Small	Large	Small	Large	Small	Large	Small	Large
100	1:1	0.278	0.252	0.245	0.198	0.240	0.224	0.194	0.154	0.205	0.152	0.155	0.145
	3:2	0.289	0.265	0.255	0.215	0.236	0.212	0.202	0.146	0.194	0.161	0.139	0.063
	4:1	0.267	0.258	0.237	0.215	0.188	0.183	0.163	0.138	0.115	0.106	0.074	0.067
300	1:1	0.370	0.332	0.326	0.259	0.338	0.290	0.291	0.209	0.295	0.241	0.232	0.149
	3:2	0.372	0.334	0.335	0.268	0.335	0.294	0.293	0.215	0.288	0.239	0.232	0.162
	4:1	0.353	0.330	0.329	0.273	0.297	0.272	0.269	0.204	0.229	0.194	0.202	0.116
500	1:1	0.388	0.341	0.346	0.269	0.363	0.308	0.307	0.209	0.326	0.262	0.255	0.167
	3:2	0.390	0.346	0.350	0.273	0.360	0.309	0.305	0.208	0.318	0.258	0.251	0.153
	4:1	0.390	0.354	0.356	0.290	0.342	0.301	0.300	0.222	0.275	0.232	0.233	0.151

주. 수렴율이 80%미만인 조건은 색칠하여 표시하였음

<표 7> 부정왜곡 조건에서의 응답 왜곡 피검자 비율에 따른 1종 오류 비율

전체 표본 크기	집단 간 표본크기 비율	전체 문항 수											
		18문항				34문항				66문항			
		왜곡문항 20%		왜곡문항 50%		왜곡문항 20%		왜곡문항 50%		왜곡문항 20%		왜곡문항 50%	
		응답왜곡 크기		응답왜곡 크기		응답왜곡 크기		응답왜곡 크기		응답왜곡 크기		응답왜곡 크기	
		Small	Large	Small	Large	Small	Large	Small	Large	Small	Large	Small	Large
100	1:1	0.279	0.233	0.253	0.174	0.242	0.175	0.194	0.108	0.197	0.114	0.135	0.058
	3:2	0.282	0.235	0.250	0.173	0.230	0.165	0.180	0.091	0.171	0.103	0.114	0.043
	4:1	0.251	0.208	0.220	0.157	0.180	0.142	0.140	0.086	0.098	0.073	0.065	0.036
300	1:1	0.361	0.300	0.315	0.210	0.321	0.228	0.259	0.129	0.273	0.154	0.198	0.071
	3:2	0.356	0.296	0.317	0.209	0.314	0.222	0.254	0.127	0.263	0.146	0.188	0.064
	4:1	0.334	0.269	0.300	0.190	0.273	0.189	0.224	0.104	0.199	0.107	0.143	0.044
500	1:1	0.379	0.320	0.335	0.221	0.342	0.245	0.279	0.138	0.298	0.168	0.215	0.075
	3:2	0.379	0.317	0.336	0.221	0.338	0.241	0.275	0.136	0.287	0.162	0.206	0.072
	4:1	0.365	0.299	0.324	0.208	0.314	0.219	0.255	0.119	0.244	0.132	0.175	0.057

주. 수렴율이 80%미만인 조건은 색칠하여 표시하였음

4) 피검자 수에 따른 1종 오류 비율

피검자 수에 따른 1종 오류의 비율을 <표 8>과 <표 9>, <그림 6>에 제시하였다. 그 결과, 긍정 왜곡 조건에서 faker 비율이 3:2, 전체 문항 수 66문항, 왜곡문항 50%, 응답 왜곡의 크기가 큰 조건 한 개를 제외하고 모든 조건 하에서 피검자 수가 증가할수록 1종 오류의 비율이 증가하는 것으로 나타났다. 그러나 전체 피검자 수가 100명인 조건에서는 수렴율이 모두 80% 미만이었고, 긍정 왜곡 조건의 경우 검사 길이가 길어질수록 수렴율 또한 80%미만으로 낮아졌다. 특히 긍정 왜곡 조건에서 왜곡 크기가 큰 경우, 검사 길이가 중간 길이 이상일 때 일관되게 낮은 수렴율을 보였다. 수렴율이 80%미만이었던 조건 모두를 포함해서 살펴보았을 때, 변화량이 가장 큰 조건은 긍정 왜곡 조건에서 faker의 비율이 4:1이고 전체 66문항 중 왜곡 문항의 비율이 20%이며, 왜곡의 크기가 작을 때였으며, 그 수치는 0.16이었다. 변화량이 가장 작았던 조건은 부정 왜곡 조건에서 faker의 비율이 1:1이고 전체 66문항 중 왜곡 문항의 비율이 50%이며, 왜곡의 크기가 큰 조건으로, 그 정도는 0.017이었다. <그림 6>은 위의 두 조건을 기술한 것으로, Biggest는 변화량이 가장 컸던 경우를, Smallest는 가장 작았던 경우를 의미한다. 두 조건은 증가 폭에 차이가 존재하나, 피검자 수가 증가할수록 1종 오류의 비율도 증가하는 공통된 양상을 보이고 있다.

<표 8> 긍정왜곡 조건에서의 전체 피검자 수에 따른 1종 오류 비율

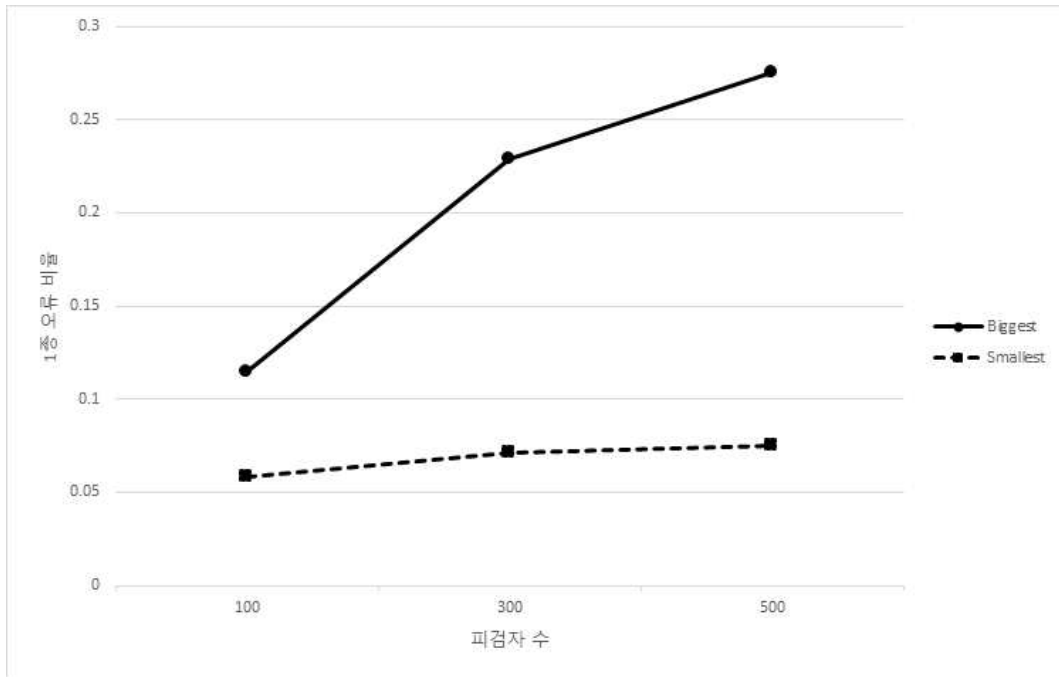
집단 간 표본크기 비율	전체 표본 크기	전체 문항 수											
		18문항				34문항				66문항			
		왜곡문항 20%		왜곡문항 50%		왜곡문항 20%		왜곡문항 50%		왜곡문항 20%		왜곡문항 50%	
		응답왜곡 크기	응답왜곡 크기	응답왜곡 크기	응답왜곡 크기	응답왜곡 크기	응답왜곡 크기	응답왜곡 크기	응답왜곡 크기	응답왜곡 크기	응답왜곡 크기	응답왜곡 크기	응답왜곡 크기
		Small	Large	Small	Large	Small	Large	Small	Large	Small	Large	Small	Large
1:1	100	0.278	0.252	0.245	0.198	0.194	0.154	0.194	0.154	0.205	0.152	0.155	0.145
	300	0.370	0.332	0.326	0.259	0.291	0.209	0.291	0.209	0.295	0.241	0.232	0.149
	500	0.388	0.341	0.346	0.269	0.307	0.209	0.307	0.209	0.326	0.262	0.255	0.167
3:2	100	0.289	0.265	0.255	0.215	0.236	0.212	0.202	0.146	0.194	0.161	0.139	0.063
	300	0.372	0.334	0.335	0.268	0.335	0.294	0.293	0.215	0.288	0.239	0.232	0.162
	500	0.390	0.346	0.350	0.273	0.360	0.309	0.305	0.208	0.318	0.258	0.251	0.153
4:1	100	0.267	0.258	0.237	0.215	0.188	0.183	0.163	0.138	0.115	0.106	0.074	0.067
	300	0.353	0.330	0.329	0.273	0.297	0.272	0.269	0.204	0.229	0.194	0.202	0.116
	500	0.390	0.354	0.356	0.290	0.342	0.301	0.300	0.222	0.275	0.232	0.233	0.151

주. 수렴율이 80%미만인 조건은 색칠하여 표시하였음

<표 9> 부정왜곡 조건에서의 전체 피검자 수에 따른 1종 오류 비율

집단 간 표본크기 비율	전체 표본 크기	전체 문항 수											
		18문항				34문항				66문항			
		왜곡문항 20%		왜곡문항 50%		왜곡문항 20%		왜곡문항 50%		왜곡문항 20%		왜곡문항 50%	
		응답왜곡 크기		응답왜곡 크기		응답왜곡 크기		응답왜곡 크기		응답왜곡 크기		응답왜곡 크기	
		Small	Large	Small	Large	Small	Large	Small	Large	Small	Large	Small	Large
1:1	100	0.279	0.233	0.253	0.174	0.242	0.175	0.194	0.108	0.197	0.114	0.135	0.058
	300	0.361	0.300	0.315	0.210	0.321	0.228	0.259	0.129	0.273	0.154	0.198	0.071
	500	0.379	0.320	0.335	0.221	0.342	0.245	0.279	0.138	0.298	0.168	0.215	0.075
3:2	100	0.282	0.235	0.250	0.173	0.230	0.165	0.180	0.091	0.171	0.103	0.114	0.043
	300	0.356	0.296	0.317	0.209	0.314	0.222	0.254	0.127	0.263	0.146	0.188	0.064
	500	0.379	0.317	0.336	0.221	0.338	0.241	0.275	0.136	0.287	0.162	0.206	0.072
4:1	100	0.251	0.208	0.220	0.157	0.180	0.142	0.140	0.086	0.098	0.073	0.065	0.036
	300	0.334	0.269	0.300	0.190	0.273	0.189	0.224	0.104	0.199	0.107	0.143	0.044
	500	0.365	0.299	0.324	0.208	0.314	0.219	0.255	0.119	0.244	0.132	0.175	0.057

주. 수렴율이 80%미만인 조건은 색칠하여 표시하였음



<그림 6> 피검자 수에 따른 1종 오류 비율의 변화

5) 응답 왜곡 유형에 따른 1종 오류 비율

응답 왜곡 유형에 따른 1종 오류 비율은 <표 10>, <표 11>과 같다. 4개의 조건을 제외하고 모든 조건에서 부정 왜곡의 1종 오류 비율이 더 낮은 것으로 관찰되었다. 긍정 왜곡의 1종 오류 비율이 더 높은 것으로 나타난 조건들은 모두 100명의 피검자 수 조건들로, 수렴율은 80%미만이였다. 수렴율이 80% 이상으로 관찰된 조건들 중 가장 큰 변화량을 보인 조건은 피검자 수 300명의 큰 크기 왜곡 조건에서 faker 비율이 4:1일 때 18문항 중 왜곡 문항이 50%였던 경우와, 동일 조건에서 34문항 중 왜곡 문항이 20%였던 경우였다. 둘 모두 변화량이 0.083으로 가장 높았다. 반면 가장 작은 변화량을 보인 조건은 피검자 수 300명의 작은 크기의 왜곡 조건에서 faker 비율이 1:1일 때 18문항 중 왜곡 문항이 20%였던 경우로, 그 수치는 0.009였다. 위 조건들을 각각 Biggest와 Smallest로 명명하여 <그림 7>에 제시하였다. 감소폭에 차이가 존재하지만 모두 부정 왜곡의 경우가 긍정 왜곡의 경우보다 1종 오류 비율이 더 낮음을 알 수 있다.

그러나 전체 피검자 수가 100명인 조건에서는 수렴율이 모두 80% 미만인 것으로 나타났다. 긍정 왜곡의 경우 왜곡의 크기가 작을 때, 문항 수가 많을수록, 왜곡 문항의 비율이 높을수록 수렴율이 낮았다. 또한 긍정 왜곡은 왜곡의 크기가 클 경우 faker 비율이 크다면 피검자 수에 관계없이 낮은 수렴율을 보였다. 또한 긍정 왜곡은 검사 길이가 중간 길이 이상의 조건 하에서도 일관되게 낮은 수렴율을 보였다. 반면 부정 왜곡의 경우 왜곡 크기가 크고 faker의 비율이 작을 때, 문항의 길이가 긴 조건에서 수렴율이 낮은 것으로 나타났다.

<표 10> 작은 크기의 왜곡 조건에서의 왜곡 유형에 따른 1종 오류 비율

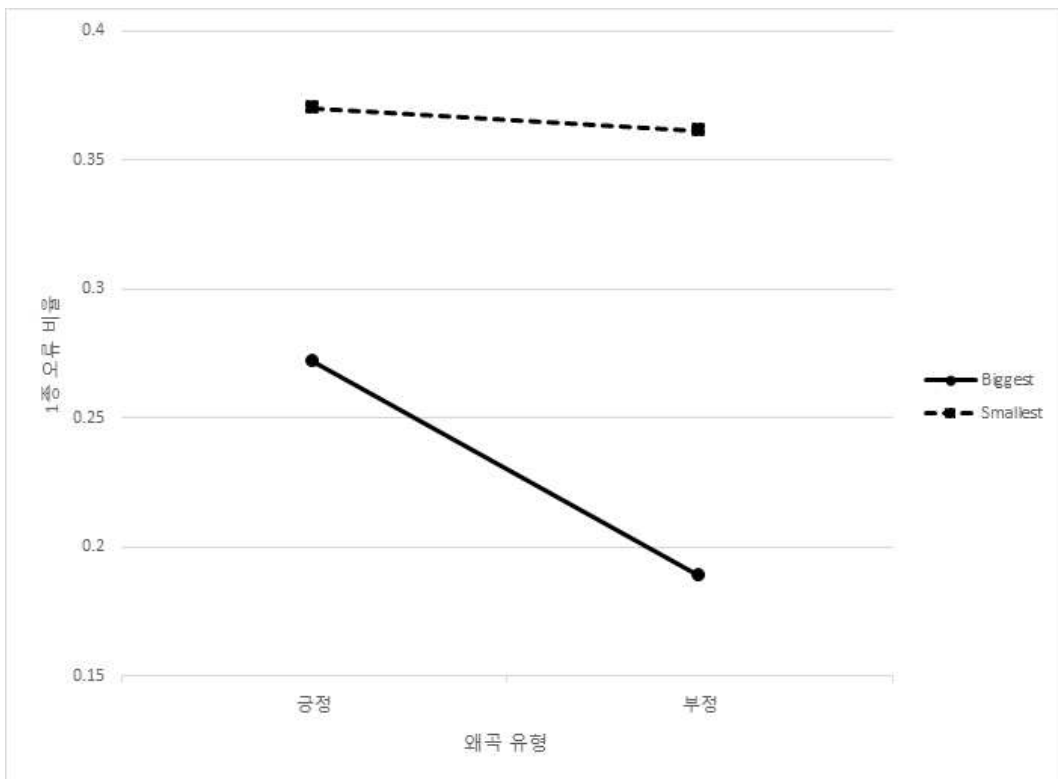
전체 표본 크기	왜곡 유형	집단 간 표본 크기 비율																			
		1:1						3:2						4:1							
		18문항		34문항		66문항		18문항		34문항		66문항		18문항		34문항		66문항			
		왜곡문항		왜곡문항		왜곡문항		왜곡문항		왜곡문항		왜곡문항		왜곡문항		왜곡문항		왜곡문항			
		20%	50%	20%	50%	20%	50%	20%	50%	20%	50%	20%	50%	20%	50%	20%	50%	20%	50%		
100	긍정	0.278	0.245	0.240	0.194	0.205	0.155	0.289	0.255	0.236	0.202	0.194	0.139	0.267	0.237	0.188	0.163	0.115	0.074		
	부정	0.279	0.253	0.242	0.194	0.197	0.135	0.282	0.250	0.230	0.180	0.171	0.114	0.251	0.220	0.180	0.140	0.098	0.065		
300	긍정	0.370	0.326	0.338	0.291	0.295	0.232	0.372	0.335	0.335	0.293	0.288	0.232	0.353	0.329	0.297	0.269	0.229	0.202		
	부정	0.361	0.315	0.321	0.259	0.273	0.198	0.356	0.317	0.314	0.254	0.263	0.188	0.334	0.300	0.273	0.224	0.199	0.143		
500	긍정	0.388	0.346	0.363	0.307	0.326	0.255	0.390	0.350	0.360	0.305	0.318	0.251	0.390	0.356	0.342	0.300	0.275	0.233		
	부정	0.379	0.335	0.342	0.279	0.298	0.215	0.379	0.336	0.338	0.275	0.287	0.206	0.365	0.324	0.314	0.255	0.244	0.175		

주. 수렴율이 80%미만인 조건은 색칠하여 표시하였음

<표 11> 큰 크기의 왜곡 조건에서의 왜곡 유형에 따른 1종 오류 비율

전체 표본 크기		집단 간 표본 크기 비율																			
		1:1						3:2						4:1							
		18문항		34문항		66문항		18문항		34문항		66문항		18문항		34문항		66문항			
		왜곡문항		왜곡문항		왜곡문항		왜곡문항		왜곡문항		왜곡문항		왜곡문항		왜곡문항		왜곡문항			
		20%	50%	20%	50%	20%	50%	20%	50%	20%	50%	20%	50%	20%	50%	20%	50%	20%	50%		
100	긍정	0.252	0.198	0.224	0.154	0.152	0.145	0.265	0.215	0.212	0.146	0.161	0.063	0.258	0.215	0.183	0.138	0.106	0.067		
	부정	0.233	0.174	0.175	0.108	0.114	0.058	0.235	0.173	0.165	0.091	0.103	0.043	0.208	0.157	0.142	0.086	0.073	0.036		
300	긍정	0.332	0.259	0.290	0.209	0.241	0.149	0.334	0.268	0.294	0.215	0.239	0.162	0.330	0.273	0.272	0.204	0.194	0.116		
	부정	0.300	0.210	0.228	0.129	0.154	0.071	0.296	0.209	0.222	0.127	0.146	0.064	0.269	0.190	0.189	0.104	0.107	0.044		
500	긍정	0.341	0.269	0.308	0.209	0.262	0.167	0.346	0.273	0.309	0.208	0.258	0.153	0.354	0.290	0.301	0.222	0.232	0.151		
	부정	0.320	0.221	0.245	0.138	0.168	0.075	0.317	0.221	0.241	0.136	0.162	0.072	0.299	0.208	0.219	0.119	0.132	0.057		

주. 수렴율이 80%미만인 조건은 색칠하여 표시하였음



<그림 7> 응답 왜곡 유형에 따른 1종 오류 비율의 변화

6) 응답 왜곡 크기에 따른 1종 오류 비율

응답 왜곡 크기에 따른 1종 오류의 비율을 <표 12>와 <표 13>에 제시하였다. 그 결과, 모든 조건 하에서 응답 왜곡의 크기가 커질수록 1종 오류의 비율이 감소하는 것으로 나타났다. 수렴율이 80%이상인 조건들 중 변화량이 가장 큰 조건은 피검자 수 500명의 부정 왜곡 조건에서 faker의 비율이 1:1이고 전체 34문항 중 왜곡 문항의 비율이 50%인 조건이었다. 이 때 변화량은 0.141이었다. 반대로 변화량이 가장 작았던 조건은 피검자 수 300명의 긍정 왜곡 조건에서 faker의 비율이 4:1이고 전체 18문항 중 왜곡 문항의 비율이 20%였던 조건으로, 그 수치는 0.023이었다. <그림 8>은 위의 두 조건을 기술한 것으로, Biggest는 변화량이 가장 컸던 경우를, Smallest는 가장 작았던 경우를 의미한다. 두 조건 모두 응답 왜곡 크기에 따라 1종 오류 비율이 감소하는 양상을 보이고 있음을 확인할 수 있다.

추가적으로 수렴율을 살펴보았을 때, 전체 피검자 수가 100명인 조건에서는 수렴율이 모두 80% 미만인 것으로 나타났다. 특히 긍정 왜곡 조건의 경우 66문항 중 왜곡 문항 비율이 50%일 때 일관되게 낮은 수렴율을 보고했다. 또한 긍정 왜곡 조건에서 응답 왜곡의 크기가 클 경우 18문항 중 20%가 왜곡 문항이었던 조건을 제외하고 대부분의 조건에서 수렴율이 낮은 것으로 나타났다.

<표 12> 긍정 왜곡 조건에서의 응답 왜곡 크기에 따른 1종 오류 비율

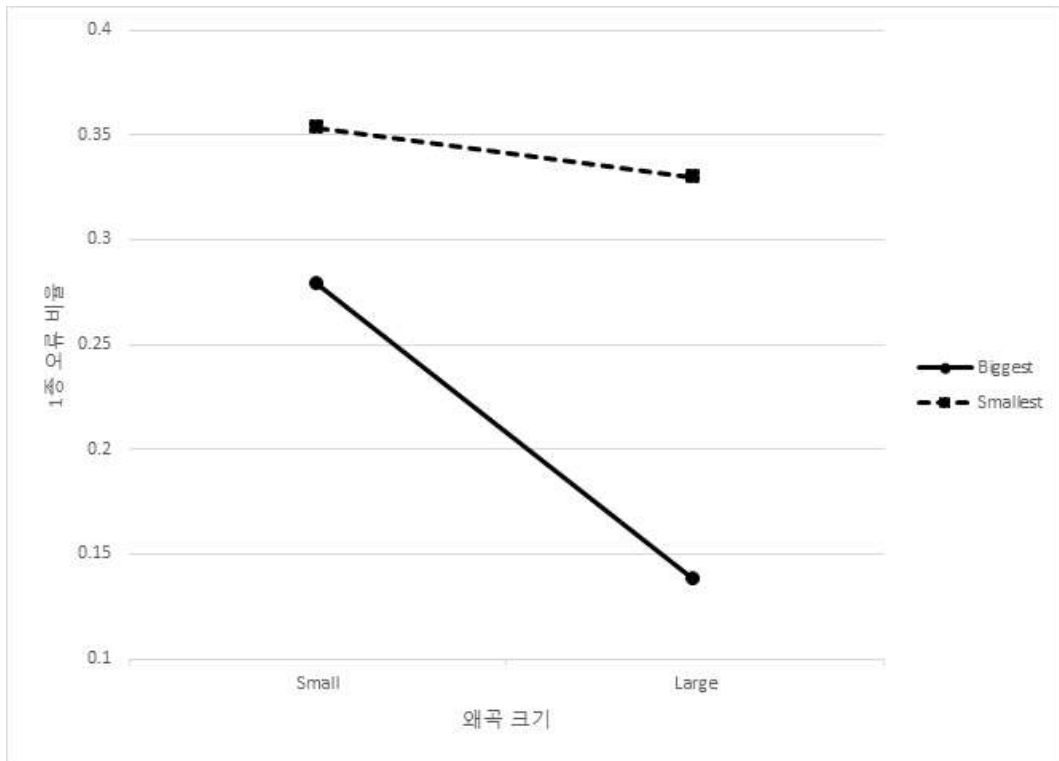
전체 표본 크기		응답 왜곡 크기		집단 간 표본 크기 비율																	
				1:1				3:2				4:1									
				18문항		34문항		66문항		18문항		34문항		66문항		18문항		34문항		66문항	
				왜곡문항		왜곡문항		왜곡문항		왜곡문항		왜곡문항		왜곡문항		왜곡문항		왜곡문항		왜곡문항	
				20%	50%	20%	50%	20%	50%	20%	50%	20%	50%	20%	50%	20%	50%	20%	50%		
100	Small	0.278	0.245	0.240	0.194	0.205	0.155	0.289	0.255	0.236	0.202	0.194	0.139	0.267	0.237	0.188	0.163	0.115	0.074		
	Large	0.252	0.198	0.224	0.154	0.152	0.145	0.265	0.215	0.212	0.146	0.161	0.063	0.258	0.215	0.183	0.138	0.106	0.067		
300	Small	0.370	0.326	0.338	0.291	0.295	0.232	0.372	0.335	0.335	0.293	0.288	0.232	0.353	0.329	0.297	0.269	0.229	0.202		
	Large	0.332	0.259	0.290	0.209	0.241	0.149	0.334	0.268	0.294	0.215	0.239	0.162	0.330	0.273	0.272	0.204	0.194	0.116		
500	Small	0.388	0.346	0.363	0.307	0.326	0.255	0.390	0.350	0.360	0.305	0.318	0.251	0.390	0.356	0.342	0.300	0.275	0.233		
	Large	0.341	0.269	0.308	0.209	0.262	0.167	0.346	0.273	0.309	0.208	0.258	0.153	0.354	0.290	0.301	0.222	0.232	0.151		

주. 수렴율이 80%미만인 조건은 색칠하여 표시하였음

<표 13> 부정 왜곡 조건에서의 응답 왜곡 크기에 따른 1종 오류 비율

전체 표본 크기		응답 왜곡 크기		집단 간 표본 크기 비율																	
				1:1				3:2				4:1									
				18문항		34문항		66문항		18문항		34문항		66문항		18문항		34문항		66문항	
				왜곡문항		왜곡문항		왜곡문항		왜곡문항		왜곡문항		왜곡문항		왜곡문항		왜곡문항		왜곡문항	
				20%	50%	20%	50%	20%	50%	20%	50%	20%	50%	20%	50%	20%	50%	20%	50%		
100	Small	0.279	0.253	0.242	0.194	0.197	0.135	0.282	0.250	0.230	0.180	0.171	0.114	0.251	0.220	0.180	0.140	0.098	0.065		
	Large	0.233	0.174	0.175	0.108	0.114	0.058	0.235	0.173	0.165	0.091	0.103	0.043	0.208	0.157	0.142	0.086	0.073	0.036		
300	Small	0.361	0.315	0.321	0.259	0.273	0.198	0.356	0.317	0.314	0.254	0.263	0.188	0.334	0.300	0.273	0.224	0.199	0.143		
	Large	0.300	0.210	0.228	0.129	0.154	0.071	0.296	0.209	0.222	0.127	0.146	0.064	0.269	0.190	0.189	0.104	0.107	0.044		
500	Small	0.379	0.335	0.342	0.279	0.298	0.215	0.379	0.336	0.338	0.275	0.287	0.206	0.365	0.324	0.314	0.255	0.244	0.175		
	Large	0.320	0.221	0.245	0.138	0.168	0.075	0.317	0.221	0.241	0.136	0.162	0.072	0.299	0.208	0.219	0.119	0.132	0.057		

주. 수렴율이 80%미만인 조건은 색칠하여 표시하였음



<그림 8> 응답 왜곡 유형에 따른 1종 오류 비율의 변화

2. 자료의 특성에 따른 검정력 비율

6개의 조건(검사 길이, 왜곡하기 쉬운 문항의 비율, 피검자의 수, 응답 왜곡 피검자의 비율, 응답 왜곡 유형, 응답 왜곡 크기) 하에서 검정력이 변화하는 양상을 살펴보았다. 각 독립변수들의 개별적인 효과를 살펴보기 위해 분석의 주가 되는 독립변수 외의 조건들은 동일하게 하여 통제하였다.

1) 검사 길이에 따른 검정력

검사 길이에 따른 검정력을 <표 14>와 <표 15>에 제시하였다. 그 결과, 모든 조건에서 검사 길이가 증가할수록 검정력이 증가하는 것으로 나타났다. 수렴율이 80%이상이었다는 조건들 중 변화량이 가장 큰 조건은 500명의 긍정 왜곡 조건에서 faker의 비율이 4:1이고 왜곡 문항의 비율이 20%이며, 왜곡의 크기가 작았을 때였으며, 그 수치는 0.028이었다. 변화량이 가장 작았던 조건은 전체 피검자 수가 500명이고 faker의 비율이 1:1이며 왜곡 문항의 비율이 50%, 왜곡의 크기가 작은 부정 왜곡 조건으로, 그 수치는 0.102였다. <그림 9>는 위의 두 조건을 기술한 것으로, Biggest는 변화량이 가장 컸던 경우를, Smallest는 가장 작았던 경우를 의미한다. 변화량에 차이가 있으나 검사 길이가 길어질수록 검정력이 높아지는 양상을 동일하게 보임을 알 수 있다.

추가적으로 수렴율을 살펴보았을 때, 전체 피검자 수가 100명인 조건에서는 수렴율이 모두 80% 미만이었으며, 긍정 왜곡 조건의 경우 검사 길이가 길어질수록 수렴율 또한 80%미만으로 낮아졌다. 특히 긍정 왜곡 조건에서 왜곡 크기가 큰 경우, 검사 길이가 중간 길이 이상일 때 일관되게 낮은 수렴율을 보였다.

<표 14> 긍정왜곡 조건에서의 검사 길이에 따른 검정력

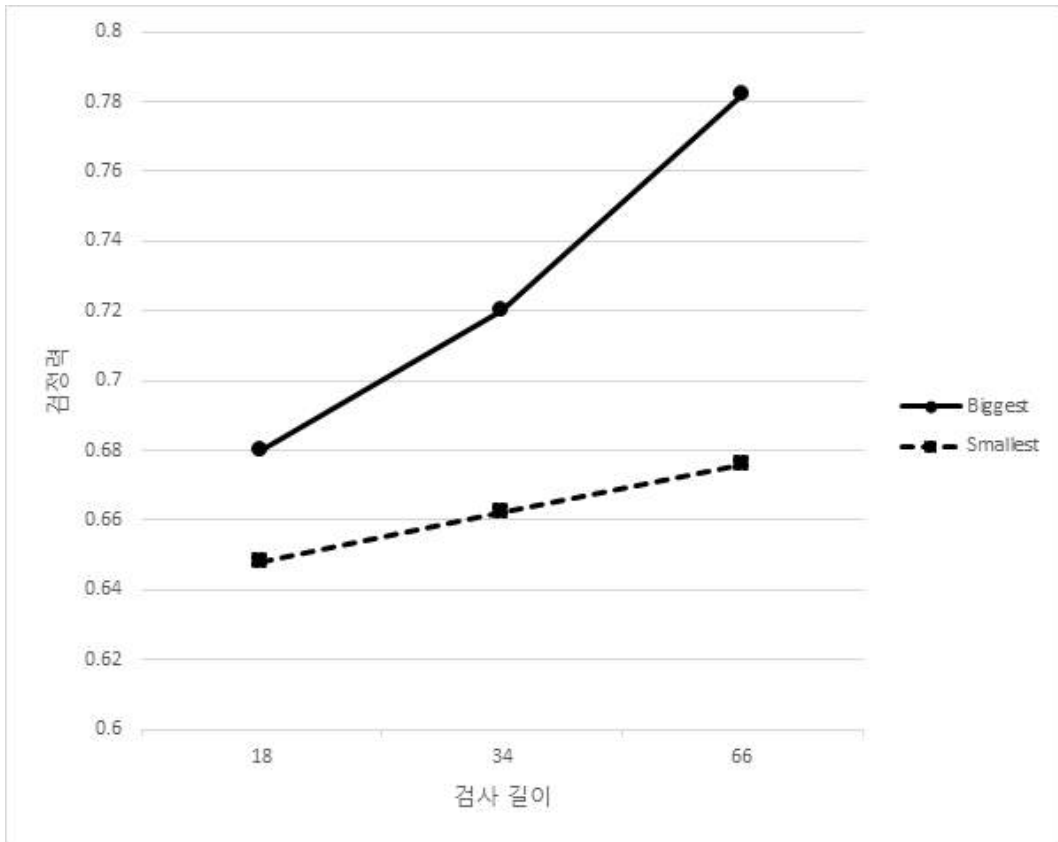
전체 표본 크기	검사 길이	집단 간 표본 크기의 비율											
		1:1				3:2				4:1			
		왜곡문항 20%		왜곡문항 50%		왜곡문항 20%		왜곡문항 50%		왜곡문항 20%		왜곡문항 50%	
		응답왜곡 크기		응답왜곡 크기		응답왜곡 크기		응답왜곡 크기		응답왜곡 크기		응답왜곡 크기	
		Small	Large	Small	Large	Small	Large	Small	Large	Small	Large	Small	Large
100	18	0.714	0.735	0.742	0.780	0.730	0.744	0.759	0.782	0.759	0.759	0.788	0.805
	34	0.749	0.776	0.795	0.833	0.754	0.775	0.791	0.833	0.777	0.782	0.803	0.830
	66	0.791	0.825	0.840	0.900	0.795	0.811	0.827	0.915	0.781	0.791	0.850	0.884
300	18	0.647	0.679	0.678	0.731	0.645	0.677	0.685	0.737	0.672	0.700	0.716	0.754
	34	0.674	0.715	0.722	0.791	0.670	0.710	0.723	0.790	0.686	0.719	0.731	0.785
	66	0.714	0.763	0.777	0.832	0.706	0.754	0.773	0.849	0.703	0.740	0.764	0.838
500	18	0.620	0.654	0.659	0.722	0.626	0.656	0.661	0.723	0.648	0.675	0.686	0.734
	34	0.646	0.696	0.705	0.784	0.647	0.693	0.703	0.782	0.662	0.699	0.711	0.774
	66	0.687	0.747	0.757	0.840	0.680	0.740	0.755	0.841	0.676	0.733	0.749	0.839

주. 수렴율이 80%미만인 조건은 색칠하여 표시하였음

<표 15> 부정왜곡 조건에서의 검사 길이에 따른 검정력

전체 표본 크기	검사 길이	집단 간 표본 크기의 비율											
		1:1				3:2				4:1			
		왜곡문항 20%		왜곡문항 50%		왜곡문항 20%		왜곡문항 50%		왜곡문항 20%		왜곡문항 50%	
		응답왜곡 크기		응답왜곡 크기		응답왜곡 크기		응답왜곡 크기		응답왜곡 크기		응답왜곡 크기	
		Small	Large	Small	Large	Small	Large	Small	Large	Small	Large	Small	Large
100	18	0.720	0.801	0.760	0.882	0.748	0.822	0.791	0.884	0.782	0.845	0.828	0.899
	34	0.763	0.843	0.803	0.917	0.770	0.845	0.820	0.929	0.816	0.888	0.864	0.943
	66	0.798	0.897	0.870	0.967	0.822	0.908	0.876	0.973	0.841	0.945	0.911	0.975
300	18	0.648	0.781	0.701	0.881	0.658	0.783	0.711	0.889	0.698	0.810	0.761	0.900
	34	0.678	0.819	0.745	0.932	0.677	0.820	0.752	0.939	0.718	0.848	0.792	0.944
	66	0.721	0.872	0.801	0.969	0.720	0.872	0.807	0.970	0.743	0.889	0.838	0.972
500	18	0.624	0.776	0.680	0.872	0.636	0.784	0.687	0.879	0.669	0.800	0.729	0.896
	34	0.652	0.811	0.720	0.929	0.651	0.813	0.724	0.932	0.681	0.831	0.760	0.943
	66	0.696	0.866	0.782	0.970	0.694	0.868	0.785	0.972	0.708	0.876	0.809	0.975

주. 수렴율이 80%미만인 조건은 색칠하여 표시하였음



<그림 9> 검사 길이에 따른 검정력의 변화

2) 왜곡문항 비율에 따른 검정력

왜곡문항 비율에 따른 검정력은 <표 16>, <표 17>과 같다. 분석 결과, 한 조건을 제외한 모든 조건에서 왜곡문항의 비율이 높아질수록 검정력이 높아지는 경향을 보였다. 단, 피검자 수 500명의 부정 왜곡 조건에서 faker의 비율이 4:1이고 전체 문항 수가 66문항이며 응답 왜곡의 크기가 클 경우, 왜곡문항 비율이 높아질수록 검정력이 낮아지는 경향이 나타났다. 수렴율이 80%이상으로 관찰된 조건들 중 가장 큰 변화량을 보인 조건은 전체 피검자 500명의 부정왜곡 조건에서 응답 왜곡 피검자의 비율이 4:1일 때, 검사 길이가 66문항이며 응답 왜곡의 크기가 큰 경우였다. 해당 조건에서의 검정력 변화량은 0.119였다. 반면 가장 작은 변화량을 보인 조건은 전체 피검자 300명의 긍정왜곡 조건에서 응답 왜곡 피검자의 비율이 1:1일 때, 검사 길이가 18문항이며 응답 왜곡의 크기가 작은 경우로, 해당 조건에서의 변화량은 0.031이었다. 위 조건들을 각각 Biggest와 Smallest로 명명하여 <그림 10>에서 제시하였다. 증가폭에 차이가 존재하지만 두 조건 모두 왜곡문항 비율이 20%일 때보다 50%일 때 검정력이 더 큰 것을 알 수 있다.

그러나 전체 피검자 수가 100명인 조건에서는 수렴율이 모두 80% 미만인 것으로 나타났다. 긍정 왜곡 조건의 경우 검사 길이가 길어질수록, 왜곡문항의 비율이 커질수록 수렴율 또한 80%미만으로 낮아졌다. 또한 긍정 왜곡 조건에서는 응답왜곡의 크기가 클 때 왜곡 문항의 비율이 50%인 경우 일관되게 낮은 수렴율을 보였다.

<표 16> 긍정왜곡 조건에서의 왜곡 문항 비율에 따른 검정력

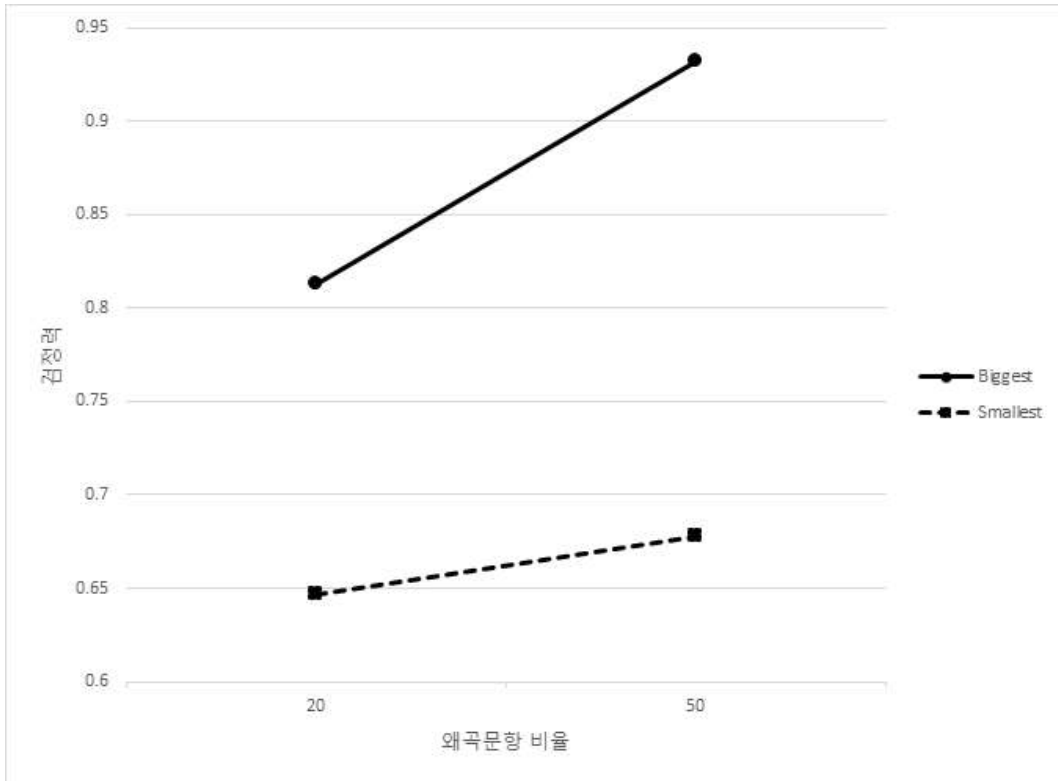
전체 표본 크기		왜곡 문항 비율		집단 간 표본 크기의 비율																	
				1:1				3:2				4:1									
				18문항		34문항		66문항		18문항		34문항		66문항		18문항		34문항		66문항	
				응답왜곡 크기		응답왜곡 크기		응답왜곡 크기		응답왜곡 크기		응답왜곡 크기		응답왜곡 크기		응답왜곡 크기		응답왜곡 크기		응답왜곡 크기	
		Small	Large	Small	Large	Small	Large	Small	Large	Small	Large	Small	Large	Small	Large	Small	Large	Small	Large		
100	20%	0.714	0.735	0.749	0.776	0.791	0.825	0.730	0.744	0.754	0.775	0.795	0.811	0.759	0.759	0.777	0.782	0.781	0.791		
	50%	0.742	0.780	0.795	0.833	0.840	0.900	0.759	0.782	0.791	0.833	0.827	0.915	0.788	0.805	0.803	0.830	0.850	0.884		
300	20%	0.647	0.679	0.674	0.715	0.714	0.763	0.645	0.677	0.670	0.710	0.706	0.754	0.672	0.700	0.686	0.719	0.703	0.740		
	50%	0.678	0.731	0.722	0.791	0.777	0.832	0.685	0.737	0.723	0.790	0.773	0.849	0.716	0.754	0.731	0.785	0.764	0.838		
500	20%	0.620	0.654	0.646	0.696	0.687	0.747	0.626	0.656	0.647	0.693	0.680	0.740	0.648	0.675	0.662	0.699	0.676	0.733		
	50%	0.659	0.722	0.705	0.784	0.757	0.840	0.661	0.723	0.703	0.782	0.755	0.841	0.686	0.734	0.711	0.774	0.749	0.839		

주. 수렴율이 80%미만인 조건은 색칠하여 표시하였음

<표 17> 부정왜곡 조건에서의 왜곡 문항 비율에 따른 검정력

전체 표본 크기		왜곡 문항 비율		집단 간 표본 크기의 비율																	
				1:1				3:2				4:1									
				18문항		34문항		66문항		18문항		34문항		66문항		18문항		34문항		66문항	
				응답왜곡 크기		응답왜곡 크기		응답왜곡 크기		응답왜곡 크기		응답왜곡 크기		응답왜곡 크기		응답왜곡 크기		응답왜곡 크기		응답왜곡 크기	
		Small	Large	Small	Large	Small	Large	Small	Large	Small	Large	Small	Large	Small	Large	Small	Large	Small	Large		
100	20%	0.720	0.801	0.763	0.843	0.798	0.897	0.748	0.822	0.770	0.845	0.822	0.908	0.782	0.845	0.816	0.888	0.841	0.945		
	50%	0.760	0.882	0.803	0.917	0.870	0.967	0.791	0.884	0.820	0.929	0.876	0.973	0.828	0.899	0.864	0.943	0.911	0.975		
300	20%	0.648	0.781	0.678	0.819	0.721	0.872	0.658	0.783	0.677	0.820	0.720	0.872	0.698	0.810	0.718	0.848	0.743	0.889		
	50%	0.701	0.881	0.745	0.932	0.801	0.969	0.711	0.889	0.752	0.939	0.807	0.970	0.761	0.900	0.792	0.944	0.838	0.972		
500	20%	0.624	0.776	0.652	0.811	0.696	0.866	0.636	0.784	0.651	0.813	0.694	0.868	0.669	0.800	0.681	0.831	0.708	0.93		
	50%	0.680	0.872	0.720	0.929	0.782	0.970	0.687	0.879	0.724	0.932	0.785	0.972	0.729	0.896	0.760	0.943	0.809	0.86		

주. 수렴율이 80%미만인 조건은 색칠하여 표시하였음



<그림 10> 왜곡문학 비율에 따른 검정력의 변화

3) 응답 왜곡 피검자들의 비율에 따른 검정력

응답 왜곡 피검자들의 비율(이하 faker 비율)에 따른 1종 오류의 비율을 살펴본 결과는 <표 18>, <표 19>와 같다. 분석 결과, faker 비율이 변화할 때마다 검정력 또한 변화하였으나 그 양상은 다소 비일관적이었다. 먼저, faker 비율 조건을 제외한 나머지 72개 조건 중 과반수를 차지하는 47개의 조건에서 faker의 비율이 작아질수록 검정력이 증가하는 경향이 나타났다. 47개 조건들에는 부정 왜곡 조건에서 34문항 중 왜곡 문항 20%조건과 66문항 중 왜곡 문항 20%조건을 제외한 전부가 포함되었다. 또한 긍정 왜곡 조건에서 34문항 이하 조건 일부 역시 포함되었다. 또 다른 11개의 조건들에서는 faker 비율이 3:2일 때 검정력이 가장 낮았다. 위 조건들은 이전의 47개 조건들과 비교했을 때, faker 비율이 4:1인 경우 검정력이 가장 높았다는 공통점이 있다. 반면 나머지 14개의 조건에서는 faker의 비율이 증가할수록 검정력이 감소하는 양상을 보였다. 이러한 경향은 긍정 왜곡 조건에서의 66문항 조건들 전체와 34문항 조건 일부에서 관찰되었으며, 이 중 대다수는 80% 미만의 수렴율을 보인 조건들이었다.

추가적으로, 전체 피검자 수가 100명인 조건에서 수렴율은 모두 80%를 넘지 못했으며, 긍정 왜곡의 경우 왜곡 문항이 50%일 때 응답 왜곡의 크기가 크다면 일관되게 낮은 수렴율을 보였다. 특히 검사 길이가 66문항으로 클 경우 대다수의 조건들에서 faker의 비율에 관계없이 수렴율이 낮았다.

<표 18> 긍정왜곡 조건에서의 응답 왜곡 피검자 비율에 따른 검정력

전체 표본 크기	집단 간 표본크기 비율	전체 문항 수											
		18문항				34문항				66문항			
		왜곡문항 20%		왜곡문항 50%		왜곡문항 20%		왜곡문항 50%		왜곡문항 20%		왜곡문항 50%	
		응답왜곡 크기		응답왜곡 크기		응답왜곡 크기		응답왜곡 크기		응답왜곡 크기		응답왜곡 크기	
		Small	Large	Small	Large	Small	Large	Small	Large	Small	Large	Small	Large
100	1:1	0.714	0.735	0.742	0.780	0.749	0.776	0.795	0.833	0.791	0.825	0.840	0.900
	3:2	0.730	0.744	0.759	0.782	0.754	0.775	0.791	0.833	0.795	0.811	0.827	0.915
	4:1	0.759	0.759	0.788	0.805	0.777	0.782	0.803	0.830	0.781	0.791	0.850	0.884
300	1:1	0.647	0.679	0.678	0.731	0.674	0.715	0.722	0.791	0.714	0.763	0.777	0.832
	3:2	0.645	0.677	0.685	0.737	0.670	0.710	0.723	0.790	0.706	0.754	0.773	0.849
	4:1	0.672	0.700	0.716	0.754	0.686	0.719	0.731	0.785	0.703	0.740	0.764	0.838
500	1:1	0.620	0.654	0.659	0.722	0.646	0.696	0.705	0.784	0.687	0.747	0.757	0.840
	3:2	0.626	0.656	0.661	0.723	0.647	0.693	0.703	0.782	0.680	0.740	0.755	0.841
	4:1	0.648	0.675	0.686	0.734	0.662	0.699	0.711	0.774	0.676	0.733	0.749	0.839

주. 수렴율이 80%미만인 조건은 색칠하여 표시하였음

<표 19> 부정왜곡 조건에서의 응답 왜곡 피검자 비율에 따른 검정력

전체 표본 크기	집단 간 표본크기 비율	전체 문항 수											
		18문항				34문항				66문항			
		왜곡문항 20%		왜곡문항 50%		왜곡문항 20%		왜곡문항 50%		왜곡문항 20%		왜곡문항 50%	
		응답왜곡 크기		응답왜곡 크기		응답왜곡 크기		응답왜곡 크기		응답왜곡 크기		응답왜곡 크기	
		Small	Large	Small	Large	Small	Large	Small	Large	Small	Large	Small	Large
100	1:1	0.720	0.801	0.760	0.882	0.763	0.843	0.803	0.917	0.798	0.897	0.870	0.967
	3:2	0.748	0.822	0.791	0.884	0.770	0.845	0.820	0.929	0.822	0.908	0.876	0.973
	4:1	0.782	0.845	0.828	0.899	0.816	0.888	0.864	0.943	0.841	0.945	0.911	0.975
300	1:1	0.648	0.781	0.701	0.881	0.678	0.819	0.745	0.932	0.721	0.872	0.801	0.969
	3:2	0.658	0.783	0.711	0.889	0.677	0.820	0.752	0.939	0.720	0.872	0.807	0.970
	4:1	0.698	0.810	0.761	0.900	0.718	0.848	0.792	0.944	0.743	0.889	0.838	0.972
500	1:1	0.624	0.776	0.680	0.872	0.652	0.811	0.720	0.929	0.696	0.866	0.782	0.970
	3:2	0.636	0.784	0.687	0.879	0.651	0.813	0.724	0.932	0.694	0.868	0.785	0.972
	4:1	0.669	0.800	0.729	0.896	0.681	0.831	0.760	0.943	0.708	0.876	0.809	0.975

주. 수렴율이 80%미만인 조건은 색칠하여 표시하였음

4) 피검자 수에 따른 검정력

피검자 수에 따른 검정력을 <표 20>과 <표 21>에 제시하였다. 그 결과, 일부 조건들을 제외한 대다수의 조건들에서 피검자 수가 증가할수록 검정력이 감소하는 것으로 나타났다. 비일관적인 양상을 보이는 일부 조건들은 부정 왜곡에서의 왜곡 문항 비율 50% 조건이었다. 그러나 전체 피검자 수가 100명인 조건에서는 수렴율이 모두 80% 미만이었고, 긍정 왜곡 조건의 경우 검사 길이가 길어질수록 수렴율 또한 80%미만으로 낮아졌다. 특히 긍정 왜곡 조건에서 왜곡 크기가 큰 경우, 검사 길이가 중간 길이 이상일 때 일관되게 낮은 수렴율을 보였다.

<표 20> 긍정왜곡 조건에서의 전체 피검자 수에 따른 검정력

집단 간 표본크기 비율	전체 표본 크기	전체 문항 수											
		18문항				34문항				66문항			
		왜곡문항 20%		왜곡문항 50%		왜곡문항 20%		왜곡문항 50%		왜곡문항 20%		왜곡문항 50%	
		응답왜곡 크기	응답왜곡 크기	응답왜곡 크기	응답왜곡 크기	응답왜곡 크기	응답왜곡 크기	응답왜곡 크기	응답왜곡 크기	응답왜곡 크기	응답왜곡 크기	응답왜곡 크기	응답왜곡 크기
		Small	Large	Small	Large	Small	Large	Small	Large	Small	Large	Small	Large
1:1	100	0.714	0.735	0.742	0.780	0.795	0.833	0.795	0.833	0.791	0.825	0.840	0.900
	300	0.647	0.679	0.678	0.731	0.722	0.791	0.722	0.791	0.714	0.763	0.777	0.832
	500	0.620	0.654	0.659	0.722	0.705	0.784	0.705	0.784	0.687	0.747	0.757	0.840
3:2	100	0.730	0.744	0.759	0.782	0.754	0.775	0.791	0.833	0.795	0.811	0.827	0.915
	300	0.645	0.677	0.685	0.737	0.670	0.710	0.723	0.790	0.706	0.754	0.773	0.849
	500	0.626	0.656	0.661	0.723	0.647	0.693	0.703	0.782	0.680	0.740	0.755	0.841
4:1	100	0.759	0.759	0.788	0.805	0.777	0.782	0.803	0.830	0.781	0.791	0.850	0.884
	300	0.672	0.700	0.716	0.754	0.686	0.719	0.731	0.785	0.703	0.740	0.764	0.838
	500	0.648	0.675	0.686	0.734	0.662	0.699	0.711	0.774	0.676	0.733	0.749	0.839

주. 수렴율이 80%미만인 조건은 색칠하여 표시하였음

<표 21> 부정왜곡 조건에서의 전체 피검자 수에 따른 검정력

집단 간 표본크기 비율	전체 표본 크기	전체 문항 수											
		18문항				34문항				66문항			
		왜곡문항 20%		왜곡문항 50%		왜곡문항 20%		왜곡문항 50%		왜곡문항 20%		왜곡문항 50%	
		응답왜곡 크기	응답왜곡 크기	응답왜곡 크기	응답왜곡 크기	응답왜곡 크기	응답왜곡 크기	응답왜곡 크기	응답왜곡 크기	응답왜곡 크기	응답왜곡 크기	응답왜곡 크기	
		Small	Large	Small	Large	Small	Large	Small	Large	Small	Large	Small	Large
1:1	100	0.720	0.801	0.760	0.882	0.763	0.843	0.803	0.917	0.798	0.897	0.870	0.967
	300	0.648	0.781	0.701	0.881	0.678	0.819	0.745	0.932	0.721	0.872	0.801	0.969
	500	0.624	0.776	0.680	0.872	0.652	0.811	0.720	0.929	0.696	0.866	0.782	0.970
3:2	100	0.748	0.822	0.791	0.884	0.770	0.845	0.820	0.929	0.822	0.908	0.876	0.973
	300	0.658	0.783	0.711	0.889	0.677	0.820	0.752	0.939	0.720	0.872	0.807	0.970
	500	0.636	0.784	0.687	0.879	0.651	0.813	0.724	0.932	0.694	0.868	0.785	0.972
4:1	100	0.782	0.845	0.828	0.899	0.816	0.888	0.864	0.943	0.841	0.945	0.911	0.975
	300	0.698	0.810	0.761	0.900	0.718	0.848	0.792	0.944	0.743	0.889	0.838	0.972
	500	0.669	0.800	0.729	0.896	0.681	0.831	0.760	0.943	0.708	0.876	0.809	0.975

주. 수렴율이 80%미만인 조건은 색칠하여 표시하였음

5) 응답 왜곡 유형에 따른 검정력

응답 왜곡 유형에 따른 검정력은 <표 22>와 <표 23>과 같다. 분석 결과, 모든 조건에서 부정 왜곡의 검정력이 더 높은 것으로 관찰되었다. 수렴율이 80% 이상으로 관찰된 조건들 중 가장 큰 변화량을 보인 조건은 피검자 수 500명의 큰 크기 왜곡 조건에서 faker 비율이 4:1일 때 34문항 중 왜곡 문항이 50%였던 경우였다. 이 경우 변화량이 0.169로 가장 높았다. 반면 가장 작은 변화량을 보인 조건은 피검자 수 300명의 작은 크기의 왜곡 조건에서 faker 비율이 1:1일 때 18문항 중 왜곡 문항이 20%였던 경우로, 그 수치는 0.001이었다. 위 조건들 모두 감소폭에 차이가 존재하지만 부정 왜곡의 경우가 긍정 왜곡의 경우보다 더 높은 검정력을 보이고 있음을 알 수 있다.

그러나 전체 피검자 수가 100명인 조건에서는 수렴율이 모두 80% 미만인 것으로 나타났다. 긍정 왜곡의 경우 왜곡의 크기가 작을 때, 문항 수가 많을수록, 왜곡 문항의 비율이 높을수록 수렴율이 낮았다. 또한 긍정 왜곡은 왜곡의 크기가 클 경우 faker 비율이 크다면 피검자 수에 관계없이 낮은 수렴율을 보였다. 또한 긍정 왜곡은 검사 길이가 중간 길이 이상의 조건 하에서도 일관되게 낮은 수렴율을 보였다. 반면 부정 왜곡의 경우 왜곡 크기가 크고 faker의 비율이 작을 때, 문항의 길이가 긴 조건에서 수렴율이 낮은 것으로 나타났다.

<표 22> 작은 크기의 왜곡 조건에서의 왜곡 유형에 따른 검정력

전체 표본 크기	왜곡 유형	집단 간 표본 크기 비율																			
		1:1						3:2						4:1							
		18문항		34문항		66문항		18문항		34문항		66문항		18문항		34문항		66문항			
		왜곡문항		왜곡문항		왜곡문항		왜곡문항		왜곡문항		왜곡문항		왜곡문항		왜곡문항		왜곡문항			
		20%	50%	20%	50%	20%	50%	20%	50%	20%	50%	20%	50%	20%	50%	20%	50%	20%	50%		
100	긍정	0.714	0.742	0.749	0.795	0.791	0.840	0.730	0.759	0.754	0.791	0.795	0.827	0.759	0.788	0.777	0.803	0.781	0.850		
	부정	0.720	0.760	0.763	0.803	0.798	0.870	0.748	0.791	0.770	0.820	0.822	0.876	0.782	0.828	0.816	0.864	0.841	0.911		
300	긍정	0.647	0.678	0.674	0.722	0.714	0.777	0.645	0.685	0.670	0.723	0.706	0.773	0.672	0.716	0.686	0.731	0.703	0.764		
	부정	0.648	0.701	0.678	0.745	0.721	0.801	0.658	0.711	0.677	0.752	0.720	0.807	0.698	0.761	0.718	0.792	0.743	0.838		
500	긍정	0.620	0.659	0.646	0.705	0.687	0.757	0.626	0.661	0.647	0.703	0.680	0.755	0.648	0.686	0.662	0.711	0.676	0.749		
	부정	0.624	0.680	0.652	0.720	0.696	0.782	0.636	0.687	0.651	0.724	0.694	0.785	0.669	0.729	0.681	0.760	0.708	0.809		

주. 수렴율이 80%미만인 조건은 색칠하여 표시하였음

<표 23> 큰 크기의 왜곡 조건에서의 왜곡 유형에 따른 검정력

전체 표본 크기	왜곡 유형	집단 간 표본 크기 비율																			
		1:1						3:2						4:1							
		18문항		34문항		66문항		18문항		34문항		66문항		18문항		34문항		66문항			
		왜곡문항		왜곡문항		왜곡문항		왜곡문항		왜곡문항		왜곡문항		왜곡문항		왜곡문항		왜곡문항			
		20%	50%	20%	50%	20%	50%	20%	50%	20%	50%	20%	50%	20%	50%	20%	50%	20%	50%		
100	긍정	0.735	0.198	0.776	0.833	0.825	0.900	0.744	0.782	0.775	0.833	0.811	0.915	0.759	0.805	0.782	0.830	0.791	0.884		
	부정	0.801	0.174	0.843	0.917	0.897	0.967	0.822	0.884	0.845	0.929	0.908	0.973	0.845	0.899	0.888	0.943	0.945	0.975		
300	긍정	0.679	0.259	0.715	0.791	0.763	0.832	0.677	0.737	0.710	0.790	0.754	0.849	0.700	0.754	0.719	0.785	0.740	0.838		
	부정	0.781	0.210	0.819	0.932	0.872	0.969	0.783	0.889	0.820	0.939	0.872	0.970	0.810	0.900	0.848	0.944	0.889	0.972		
500	긍정	0.654	0.269	0.696	0.784	0.747	0.840	0.656	0.723	0.693	0.782	0.740	0.841	0.675	0.734	0.699	0.774	0.733	0.839		
	부정	0.776	0.221	0.811	0.929	0.866	0.970	0.784	0.879	0.813	0.932	0.868	0.972	0.800	0.896	0.831	0.943	0.876	0.975		

주. 수렴율이 80%미만인 조건은 색칠하여 표시하였음

6) 응답 왜곡 크기에 따른 검정력

응답 왜곡 크기에 따른 검정력을 <표 24>와 <표 25>에 제시하였다. 그 결과, 한 조건을 제외하고 모든 조건에서 응답 왜곡의 크기가 커질수록 검정력이 증가하는 것으로 나타났다. 예외가 된 조건에서는 응답 왜곡 크기에 따른 증감이 관찰되지 않았으나, 해당 조건은 수렴율이 80%미만이였다. 수렴율이 80%이상인 조건들 중 변화량이 가장 큰 조건은 피검자 수 500명의 부정 왜곡 조건에서 faker의 비율이 1:1이고 전체 34문항 중 왜곡 문항의 비율이 50%인 조건으로, 그 정도는 0.209였다. 반대로 변화량이 가장 작았던 조건은 피검자 수 500명의 긍정 왜곡 조건에서 faker의 비율이 4:1이고 전체 18문항 중 왜곡 문항의 비율이 20%였던 조건으로, 그 수치는 0.027 이었다. <그림 11>은 위의 두 조건을 기술한 것으로, Biggest는 변화량이 가장 컸던 경우를, Smallest는 가장 작았던 경우를 의미한다. 증가량에 차이가 존재하지만 모두 응답 왜곡의 크기가 큰 경우 검정력이 상승한 것을 확인할 수 있다.

추가적으로 수렴율을 살펴보았을 때, 전체 피검자 수가 100명인 조건에서는 수렴율이 모두 80% 미만인 것으로 나타났다. 특히 긍정 왜곡 조건의 경우 66문항 중 왜곡 문항 비율이 50%일 때 일관되게 낮은 수렴율을 보고했다. 또한 긍정 왜곡 조건에서 응답 왜곡의 크기가 클 경우 18문항 중 20%가 왜곡 문항이었던 조건을 제외하고 대부분의 조건에서 수렴율이 낮은 것으로 나타났다.

<표 24> 긍정 왜곡 조건에서의 응답 왜곡 크기에 따른 1종 오류 비율

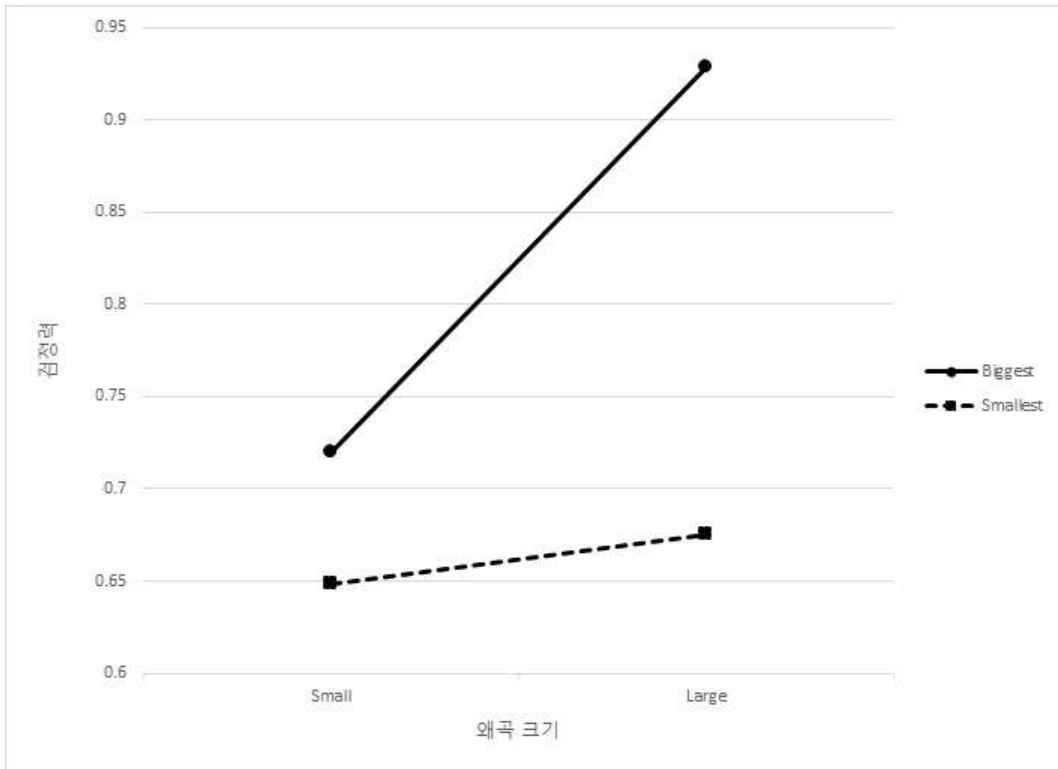
전체 표본 크기		응답 왜곡 크기		집단 간 표본 크기 비율																	
				1:1				3:2				4:1									
				18문항		34문항		66문항		18문항		34문항		66문항		18문항		34문항		66문항	
				왜곡문항		왜곡문항		왜곡문항		왜곡문항		왜곡문항		왜곡문항		왜곡문항		왜곡문항		왜곡문항	
				20%	50%	20%	50%	20%	50%	20%	50%	20%	50%	20%	50%	20%	50%	20%	50%		
100	Small	0.714	0.742	0.749	0.795	0.791	0.840	0.730	0.759	0.754	0.791	0.795	0.827	0.759	0.788	0.777	0.803	0.781	0.850		
	Large	0.735	0.780	0.776	0.833	0.825	0.900	0.744	0.782	0.775	0.833	0.811	0.915	0.759	0.805	0.782	0.830	0.791	0.884		
300	Small	0.647	0.678	0.674	0.722	0.714	0.777	0.645	0.685	0.670	0.723	0.706	0.773	0.672	0.716	0.686	0.731	0.703	0.764		
	Large	0.679	0.731	0.715	0.791	0.763	0.832	0.677	0.737	0.710	0.790	0.754	0.849	0.700	0.754	0.719	0.785	0.740	0.838		
500	Small	0.620	0.659	0.646	0.705	0.687	0.757	0.626	0.661	0.647	0.703	0.680	0.755	0.648	0.686	0.662	0.711	0.676	0.749		
	Large	0.654	0.722	0.696	0.784	0.747	0.840	0.656	0.723	0.693	0.782	0.740	0.841	0.675	0.734	0.699	0.774	0.733	0.839		

주. 수렴율이 80%미만인 조건은 색칠하여 표시하였음

<표 25> 부정 왜곡 조건에서의 응답 왜곡 크기에 따른 1종 오류 비율

전체 표본 크기		응답 왜곡 크기		집단 간 표본 크기 비율																	
				1:1				3:2				4:1									
				18문항		34문항		66문항		18문항		34문항		66문항		18문항		34문항		66문항	
				왜곡문항		왜곡문항		왜곡문항		왜곡문항		왜곡문항		왜곡문항		왜곡문항		왜곡문항		왜곡문항	
				20%	50%	20%	50%	20%	50%	20%	50%	20%	50%	20%	50%	20%	50%	20%	50%		
100	Small	0.720	0.760	0.763	0.803	0.798	0.870	0.748	0.791	0.770	0.820	0.822	0.876	0.782	0.828	0.816	0.864	0.841	0.911		
	Large	0.801	0.882	0.843	0.917	0.897	0.967	0.822	0.884	0.845	0.929	0.908	0.973	0.845	0.899	0.888	0.943	0.945	0.975		
300	Small	0.648	0.701	0.678	0.745	0.721	0.801	0.658	0.711	0.677	0.752	0.720	0.807	0.698	0.761	0.718	0.792	0.743	0.838		
	Large	0.781	0.881	0.819	0.932	0.872	0.969	0.783	0.889	0.820	0.939	0.872	0.970	0.810	0.900	0.848	0.944	0.889	0.972		
500	Small	0.624	0.680	0.652	0.720	0.696	0.782	0.636	0.687	0.651	0.724	0.694	0.785	0.669	0.729	0.681	0.760	0.708	0.809		
	Large	0.776	0.872	0.811	0.929	0.866	0.970	0.784	0.879	0.813	0.932	0.868	0.972	0.800	0.896	0.831	0.943	0.876	0.975		

주. 수렴율이 80%미만인 조건은 색칠하여 표시하였음



<그림 11> 응답 왜곡 크기에 따른 검정력의 변화

VI. 논의

1. 연구 결과 정리 및 논의

Ferrando와 Anguiano-Carrasco(2013)는 응답 왜곡으로 인한 능력(θ) 변화를 문항 수준에서 모델링 하는 방법을 제안하였다. 이는 IRT 우도비 검정 과정(Likelihood-Ratio test, LR test)에 문항 모수를 변화시킨 다집단 분석을 추가한 것으로, Ferrando와 Anguiano-Carrasco(2013)는 구조방정식을 이용한 조정 과정이 응답 왜곡 식별의 정확성을 향상시키리라 기대한 바 있다. 실제로 선행 연구에서 해당 지수를 사용하여 응답 왜곡을 식별했을 때 66%의 적중률(hit rate)을 보이는 것으로 나타났다. 그러나 이는 경험 자료를 기반으로 한 결과이므로 피검자 수와 문항 수 등 여러 조건들이 제한되어있기 때문에 외적 타당도가 확보되기 어렵다는 한계를 지닌다. 따라서 시뮬레이션 연구를 통해 Ferrando와 Anguiano-Carrasco의 지수가 다양한 조건 하에서 어떤 결과를 도출하는지를 확인하여 일반화 가능성을 높이고자 하였다. 본 연구에서 조작되는 조건은 검사 길이, 왜곡 문항 비율, 왜곡 응답자 비율, 피검자 수, 왜곡 유형, 왜곡 크기로 총 6개였으며, 연구문제1에서는 각 조건 하에서의 1종 오류 비율을, 연구문제2에서는 동일 조건에서의 검정력을 살펴보았다. 본 연구에서의 결과를 정리하면 다음과 같다.

먼저 연구문제1에서 1종 오류 비율의 변화 정도가 수용 가능한 수준인지를 평가하기 위해 Bradley(1978)가 제안한 1종 오류의 강건성 기준(criteria of robustness)을 차용하였다. Bradley(1978)는 구조방정식에서의 1종 오류 수준을 평가하기 위해 다음과 같은 기준을 제안한 바 있다. 그

에 따르면 유의수준 α (nominal alpha level)를 기준으로, $\alpha_{nominal} \pm 0.1\alpha_{nominal}$ 는 보수적인 기준 (conservative criterion), $\alpha_{nominal} \pm 0.5\alpha_{nominal}$ 는 관대한 기준 (liberal criterion)이라 할 수 있다. 일반적으로 사용하는 유의수준이 .05임을 반영하여 (Shaughnessy, Zechmeister & Zechmeister, 2000), 본 연구에서도 유의수준을 .05라 가정했을 때 1종 오류의 변화가 Bradley의 기준을 초과하는지를 관찰하였다. 이 때 1종 오류의 변화율이 해당 기준 내에서 움직일 때에는 조건의 변화에 영향을 받지 않은 것으로, 기준을 초과했을 때에는 영향을 받은 것으로 해석할 수 있다.

연구 결과, 검사 길이에 따른 1종 오류의 변화량이 모두 관대한 기준과 엄격한 기준 이상인 것을 확인하였다. 즉, 1종 오류가 검사 길이의 변화에 따라 달라지고 있었으며, 이는 검사 길이에 따라 1종 오류 수준이 달라질 것이라는 가설 1-1을 지지하는 결과라 할 수 있다. 구체적으로, 검사 길이가 길어질수록 1종 오류 수준이 낮아지고 있음이 확인되었다. 따라서 Ferrando와 Anguiano-Carrasco의 지수는 문항의 수가 많은 검사에서 오류의 가능성이 보다 적은 것으로 판단된다.

다음으로 왜곡 문항 비율 조건에서 1종 오류의 변화율은 엄격한 기준 이상이었다. 따라서 왜곡 문항 비율에 따라 1종 오류가 달라질 것이라는 가설 1-2는 지지되었다. 그러나 왜곡 문항이 전체의 20%일 때보다 전체의 50%일 때 더 적은 1종 오류를 보이는 것은 기존의 선행 연구와 상반된 결과이다. 모수적 검정 중 하나인 LR test의 경우, DIF 문항의 비율이 .5 이상일 때 해당 문항을 엄밀히 DIF 집단의 문항이라 지칭하기 어렵다는 문제가 있었다 (박찬호, 2009). 때문에 LR test와 다른 검정법을 비교한 박찬호 (2009)의 연구에서 DIF 문항의 비율이 50%일 때 LR test는 낮은 검정력을 보고하였다. 본 연구에서도 왜곡 피검자의 비율에 따른 조건에서 동일한

이유로 선행 연구와 유사한 양상을 보이는 것으로 확인 되었으나, 해당 문제는 문항 수준에서는 나타나지 않은 것으로 확인된다.

응답 왜곡 피검자의 비율에 따른 1종 오류 비율 또한 엄격한 기준 이상으로 변화하였다. 그러므로 응답 왜곡 피검자의 비율에 따라 1종 오류의 비율이 변화할 것이라는 가설 1-3은 지지되었다. 변화 양상은 다소 비 일관적이었으나, 대다수의 조건들에서 정직 응답자와 왜곡 응답자의 비율 (honest:faker)이 1:1일 때 1종 오류의 비율이 가장 높았다. 그 외의 조건들에서는 honest:faker=3:2일 때 1종 오류가 가장 높거나, honest:faker=1:1일 때와 3:2일 때의 값이 동일하게 가장 높았다. 위 조건들은 모두 honest:faker=4:1인 경우 1종 오류가 가장 낮았다는 공통점이 있다. 이와 같이 faker의 비율이 높아질수록 높은 1종 오류 비율을 보이는 것은 박찬호(2009)가 지적했던 바와 같은 맥락으로 해석 가능하다. 즉, 이러한 결과는 LR test에서 왜곡 응답자의 수가 50%이상일 때 어떤 집단을 참조 집단이라 명명할 수 있는지의 문제에서 비롯되었다고 할 수 있다. 다만 1:1 조건과 3:2 조건의 결과가 일정부분 유사했으므로, Ferrando와 Anguiano-Carrasco의 지수는 faker의 비율이 40%이상일 때 비교적 높은 1종 오류 비율을 보임을 알 수 있다.

피검자 수에 따른 1종 오류 비율의 변화율 또한 엄격한 기준 이상인 것으로 관찰되어, 가설 1-4 역시 지지되었다. 구체적으로, 피검자 수가 커질수록 1종 오류의 비율이 증가하였다. 선행 연구에서 LR test의 수행을 비교한 결과 표본 크기가 커질수록 1종 오류가 증가했음을 고려했을 때(강태훈, 2015), 본 연구에서 사용한 지수는 LR test의 일부를 결합시킨 것이므로 이와 같은 결과는 해당 검정법의 특성을 따랐기 때문이라 해석할 수 있다.

왜곡 유형에 따른 1종 오류 비율의 변화량은 총 108개의 조건 중 96개 조건이 엄격한 기준 이상을 만족하는 것으로 나타나, 가설 1-5는 부분적으

로 지지되었다. 구체적으로, 기준을 충족한 조건들에서 부정적 응답 왜곡이 긍정적 응답 왜곡보다 1종 오류가 일관되게 더 낮았다. 따라서 본 지수는 부정적 응답 왜곡 상황에 보다 더 효과적으로 기능함을 알 수 있다. 이는 LR test가 부정 왜곡 조건에서 더 높은 1종 오류를 보고했다는 선행 연구와 대치되는 결과로(Liu, Douglas, & Henson, 2009), 해당 지수만의 특징인 것으로 보인다.

마지막으로 왜곡 크기에 따른 1종 오류의 변화를 살펴본 결과, 변화량이 엄격한 기준 이상인 것으로 나타나 가설 1-6이 지지되었다. 특히 왜곡의 크기가 클 경우 1종 오류 수준이 더 낮았으므로 일반 문항과 왜곡 문항의 차이가 분명 할 때 오류의 가능성이 더 낮아질 것으로 사료된다.

연구문제2에서는 각 조건의 변화가 검정력에 영향을 미치는지를 확인하였다. 그 결과, 6개의 조건들 아래에서 세부 조건의 변화에 따라 검정력이 달라지는 것이 관찰되었으며, 그 양상은 연구문제1의 결과와 유사했다. 먼저 검사 길이가 증가할수록 검정력이 상승하는 것으로 나타나, 가설 2-1은 지지되었다. 특히 긴 길이의 검사(66문항) 조건에서는 평균적으로 약 .8 수준의 검정력을 보였으며, 이는 긴 문항에서 우수한 수행을 보이는 것으로 알려진 H^T , l , $D(\theta)$ 와 유사한 수준이었다(Karabatsos, 2003). 따라서 Ferrando와 Anguiano-Carrasco의 지수는 문항의 수가 많은 검사에서 보다 효과적으로 기능한다는 결론을 내릴 수 있다.

왜곡 문항의 비율에 따른 검정력의 변화를 살펴본 결과, 왜곡 문항 비율이 증가함에 따라 검정력이 증가하는 것이 관찰되었다. 따라서 왜곡 문항의 비율에 따라 검정력이 달라질 것이라는 가설 2-2는 지지되었다. 선행 연구에서 DIF 문항이 50%이상 일 때 LR test에서 문제가 발생할 수 있음을 우려하였으나(박찬호, 2009), LR test와 다집단 분석이 결합되었을 때 해당 문제는 발견되지 않는 것으로 나타났다. 특히 박찬호(2009)의 연구에서

피검자가 300명이고 DIF 문항이 50%일 때 검정력이 0.3대에 그쳤던 것에 반해, 본 연구에서는 동일 조건 하에서 최소 0.68, 최대 0.97의 검정력을 보고하면서 검정력이 2배 이상 상승했음을 알 수 있다. 따라서 Ferrando와 Anguiano-Carrasco의 지수는 LR test의 주요한 단점을 보완하고 있는 것으로 여겨진다.

왜곡 응답자의 비율에 따른 검정력의 변화를 살펴보았을 때, 조건에 따라 검정력이 변화하고 있으므로 가설 2-3 역시 지지되었다. 변화 양상은 다소 혼재된 경향이 있었으나, 대다수의 경우 4:1 조건에서 가장 높은 검정력을 보이는 것으로 나타났다. 이는 다집단 분석의 동일성 연구에서 집단 간 표본 크기의 차이가 커질수록 좋지 않은 1종 오류와 검정력을 보고한 것과는 대조되는 결과이다(González-Romá et al., 2006; Hancock, 2000). 따라서 LR test와 다집단 분석을 결합했을 경우 다집단 분석에서 보고되었던 문제가 다소 완화되는 것으로 보인다.

피검자 수에 따른 검정력을 살펴본 결과, 조건의 변화에 따라 검정력이 달라지는 것이 확인되었다. 이는 가설 2-4를 지지하는 결과였으며, 구체적으로 피검자 수가 증가할수록 검정력은 감소하는 추세였다. LR test의 경우 model misfit에 민감하기 때문에 표본의 크기가 커질수록 수행이 나빠진다는 선행 연구와 관련하여(박찬호, 2009), 이와 같은 결과는 LR test의 특성이 반영되었기 때문이라 해석 가능하다.

응답 왜곡 유형에 따른 검정력 역시 차이가 있음이 나타나면서, 가설 2-5는 지지되었다. 1종 오류와 마찬가지로, 부정적 응답 왜곡의 경우에 더 높은 검정력을 보임으로써 해당 지수는 부정적 응답 왜곡에 더 효과적으로 기능하는 것으로 나타났다. 가장 대중적으로 사용되는 개인 합치도 지수인 I_2 와 ζ_2, UB, H^T 모두 부정적 응답 왜곡 조건에서 가장 낮은 검정력을 보였으며, 특히 그 값이 0.1 전후였음을 감안하면(Sinharay, 2017), 평균적으로

로 약 0.76의 검정력을 보인 것은 괄목할 만한 결과라 할 수 있다.

마지막으로, 응답 왜곡의 크기에 따라 검정력이 변화하는 것으로 나타났다. 따라서 가설 2-6은 지지되었다. 특히 응답 왜곡의 크기가 커질수록 검정력이 증가하는 양상을 보였으므로 Ferrando와 Anguiano-Carrasco의 지수는 왜곡의 정도가 심할 때 더 잘 기능하는 것으로 판단된다.

본 연구에서는 전체 반복 횟수 500번 중 모형이 수렴되어 모형 합치도 지수를 산출하는 것이 가능했던 사례의 비율을 계산함으로써 수렴율을 추가적으로 살펴보았다. 그 결과, 피검자 수가 100명일 경우 80%보다 낮은 수렴율을 보였으며, 특히 긍정 왜곡 조건의 경우 검사 길이가 길어질수록 수렴율 역시 낮아지는 경향을 보였다. 이는 지수 산출 과정에서 구조 방정식을 활용하기 때문으로 분석된다. 여러 연구자들은 구조 방정식에서 충분히 큰 표본을 확보하지 못할 경우 모수 추정에 문제가 있을 수 있음을 지적해 왔다(Anderson & Gerbing, 1988; Chou & Bentler, 1995; Hoyle & Kenny, 1999; Holbert & Stephenson, 2002). 그들이 제안한 N:q 법칙³⁾에 따르면, 추정해야 할 모수 1개 당 사례 20명이 필요하며(Tanaka, 1987; Jackson, 2003; Kline, 2011), 그를 충족시키지 못할 경우에도 최소 5:1의 비율이 유지되어야 한다(Bentler & Chou, 1987). 따라서 본 연구에서 보고된 낮은 수렴율을 통해 Ferrando와 Anguiano-Carrasco의 지수는 100명 이하의 소표본에서는 제대로 기능하기 어려우며, 문항 수가 많은 검사일 경우 그에 비례하는 충분한 사례수를 확보하여야 함을 알 수 있다.

요약하면, Ferrando와 Anguiano-Carrasco의 지수는 검사 길이가 길수록, 왜곡 문항의 비율이 높을수록, 왜곡 응답자의 비율이 낮을수록 더 잘 기능하는 것으로 나타났다. 또한 피검자 수가 작을수록, 왜곡의 크기가 클수록

3) N= 표본크기, q=추정하고자 하는 모수의 개수

더 효과적으로 기능하였으며, 특히 부정적 응답 왜곡을 더 정확히 식별하였다. 이와 같은 양상은 기존의 LR test의 효과성을 연구한 결과들과 일부 유사하다 할 수 있다. 그러나 LR test를 포함하여 다수의 개인 합치도 지수를 분석한 연구들과 비교했을 때, 본 연구에서 보다 높은 검정력을 보임을 알 수 있다. 이는 선행 연구들이 모두 changing-persons paradigm을 사용한 것에 반해 본 연구에서는 changing-items paradigm을 사용했기 때문으로 해석 가능하다. 따라서 본 연구에서의 결과는 changing-items paradigm을 사용하였을 때 응답 왜곡 식별의 정확성을 향상시킬 수 있다는 Ferrando와 Anguiano-Carrasco(2013)의 주장과 일맥상통하다 할 수 있다.

2. 종합 논의 및 연구의 제한점

본 연구는 Ferrando와 Anguiano-Carrasco(2013)가 제안한 새로운 방식을 시뮬레이션을 이용하여 검증하는 것을 목표로 하였다. 연구 결과, 해당 방식을 사용하여 개인 합치도 지수를 산출하였을 때 LR test 및 다집단 분석과 일부 유사한 경향을 보이는 것으로 확인되었다. 그러나 changing-items paradigm을 이용할 경우 changing-persons paradigm을 이용한 기본적 방식보다 검증력이 더 향상되는 것으로 나타났다. 이는 Ferrando와 Anguiano-Carrasco(2013)가 주장한 바와 같이, LR test 과정에 구조 방정식을 도입하여 모형 합치도 지수를 산출함으로써 정교한 모델링이 가능했기 때문이었다. 또한 이는 실제 검사와 피검자들을 이용한 Ferrando와 Anguiano-Carrasco(2013)의 결과와 유사한 결과였다. 그들의 연구에서 k 값을 1로 고정했을 때 75%의 적중률과 17%의 1종 오류가 보고되었다. 본 연구에서의 전체 검정력과 1종 오류의 평균은 각각 0.78과 0.23으로 선행 연구와 유사했으므로 Ferrando와 Anguiano-Carrasco의 지수를 실제 상황에 적용하는 것은 충분히 가능하리라 판단된다. 구체적으로, 그들은 개인 합치도 지수를 계산하기 위해 honest 집단과 fake 집단을 임의로 처치하는 실험 연구를 실시하였다. 이러한 방법으로 개인 합치도 지수를 산출하는 것은 시뮬레이션 연구의 결과와 큰 차이가 없으므로, 실제 장면에서 역시 동일한 조건으로 모의 검사를 진행하여 개인 합치도 지수를 계산한 뒤 사용할 수 있을 것이다. 특히 개인 합치도 지수는 추가적인 타당도 문항을 필요로 하지 않기 때문에 한 번의 모의 검사를 실시한다면 기존의 검사들에서도 타당도에 관한 정보를 획득할 수 있을 것으로 보인다. 또한 본 연구에서 추가적으로 밝혀낸 바와 같이 Ferrando와 Anguiano-Carrasco의 지수는 부정적 응답 왜곡 식별에 효과적이므로 임

상, 혹은 입대 상황과 같은 특수한 장면에서 유용할 것으로 보인다. 추가적으로, 긍정적 응답 왜곡 식별에는 탁월하나 부정적 응답 왜곡 식별에는 취약한 것으로 알려진 I_z 등의 개인 합치도 지수와 병행하여 사용한다면 (Sinharay, 2017), 검사 결과의 타당성을 보다 정확히 판별할 수 있을 것이다.

본 연구는 Ferrando와 Anguiano-Carrasco의 지수가 효과적인 상황을 밝혀내고 경험 연구의 신뢰도를 향상시켰다는 의의가 있으나, 절단값 K를 1로 한정시켰다는 제한점을 지닌다. 본 연구에서는 다른 외부 조건들의 영향력을 검증하기 위해 절단값 K를 1로 고정하였다. 그 결과, 1종 오류가 명목 유의 수준인 .05를 초과하는 것으로 나타났다. Ferrando와 Anguiano-Carrasco는 K값의 조정을 통해 1종 오류 비율을 명목 유의 수준으로 통제할 수 있음을 밝혔으므로 후속 연구를 통해 이를 시뮬레이션으로 검증하는 것이 필요하다. 구체적으로, 조작된 조건들 하에서 ROC (Receiver operating characteristic) Curve를 확인하여 적절한 K값을 탐색해 볼 수 있을 것이다. 또한 이러한 K값의 조정에는 외부 조건들을 고려하는 것이 필요해 보인다. 예로, 피검자 수의 경우 문항 수와 비례하여 많은 수를 확보해야하지만 지수의 특성상 피검자 수가 커지면 1종 오류가 증가하고 검정력이 감소하는 경향을 보였다. 따라서 후속 연구에서는 조건들에 따른 K의 가변성을 검증함과 함께 조건과 K값 사이의 함수를 밝혀야 할 필요성이 있다. 후속 연구를 통해 상황 별 적절한 K값이 밝혀진다면 Ferrando와 Anguiano-Carrasco의 지수를 실제 장면의 다양한 상황에 적용할 시 정확성이 보다 향상될 것으로 기대된다.

또한 본 연구는 검사를 이분 문항으로 제한하였다는 한계를 지닌다. 이는 경험 자료를 이용한 선행 연구와의 비교를 위해 조건을 선행 연구와 일치시켰기 때문이었으나 심리학에서는 일반적으로 Likert 척도가 이용되므로(성

태제, 1997), 연구 결과를 모든 검사에 적용하는 것에 어려움이 있다. 따라서 후속 연구에서는 등급반응모형(Grade response model) 등을 이용하여 응답 범주가 3개 이상인 상황에서도 Ferrando와 Anguiano-Carrasco의 지수가 효과적으로 기능하는지 탐색하여야 할 필요가 있다.

참 고 문 헌

- 강유진. (2012). 구조방정식모형을 이용한 비균일적 차별문항기능의 탐지. 성균관대학교 일반대학원.
- 강태훈 (2015). 자료크기와 집단 간 피험자 수의 차이가 차별기능문항 추출의 제1종 오류에 미치는 영향. *교육평가연구*, 28, 577-600.
- 김명소, 이현주 (2006). 성격검사의 형식이 응답왜곡(faking)에 미치는 효과. **한국심리학회지: 산업 및 조직**, 19, 371-393.
- 박찬호 (2009). 차별적 기능 문항 추출을 위한 모수적 방법과 비모수적 방법 비교. *교육평가연구*, 22, 1161-1181.
- 성태제. (1998). 다분문항반응이론(등급반응모형)에 의한 학구적 실패내성척도의 문항분석과 피험자 특성추정. *교육심리연구*, 12, 203-218.
- 성태제. (2001). 문항반응이론의 이해와 적용. 파주:교육과학사.
- 성태제. (2014). 현대 기초통계학 : 이해와 적용(제6판). 서울: 학지사.
- 윤미리. (2015). Evaluation of Various Person-fit Indices in Detecting Social Desirability Bias. 고려대학교 대학원.
- 이순목, 이찬순, 이현정, 여성칠. (2012). 캐나다 도박행동 척도에서 개념적 구조 및 심리측정적 특성의 일반화 가능성. **한국심리학회지: 일반**, 31, 1-26.
- 이종구, 한영석. (2008). 텔레마케터 선발용 역량검사의 구성형식별 신뢰도와 타당도. **한국심리학회지: 산업 및 조직**, 21, 411-428.
- 지여운. (2014). 선발 장면에서 경고 메시지(Warnings)가 응답 왜곡(Faking)에 미치는 영향. 호서대학교 일반대학원.
- Anderson, J. C., & Gerbing, D. W. (1988). Structural equation modeling in

- practice: A review and recommended two-step approach. *Psychological bulletin*, *103*, 411.
- Bentler, P. M., & Chou, C.-P. (1987). Practical issues in structural modeling. *Sociological Methods & Research*, *16*, 78-117.
- Bradley, J. V. (1978). Robustness?. *British Journal of Mathematical and Statistical Psychology*, *31*, 144-152.
- Browne, M. W., Cudeck, R., Bollen, K. A., & Long, J. S. (1993). *Testing structural equation models*.
- Byrne, B. M. & Stewart, S. M. (2006). The MACS approach to testing for multigroup invariance of a second-order structure: A walk through the process. *Structural Equation Modeling*, *13*, 287-321.
- Camilli, G., & Shepard, L. A. (1994). Methods for identifying biased test items. Thousand Oaks, CA: Sage.
- Chou, C. P., & Bentler, P. M. (1995). Estimates and tests in structural equation modeling. Sage Publications, Inc.
- Christiansen, N. D., Goffin, R. D., Johnston, N. G., & Rothstein, M. G. (1994). Correcting the 16PF for faking: Effects on criterion related validity and individual hiring decisions. *Personnel Psychology*, *47*, 847-860.
- Dicken, C. F. (1959). Simulated patterns on the Edwards Personal Preference Schedule. *Journal of Applied Psychology*, *43*, 372.
- Dimitrov, D. M., & Smith, R. M. (2006). Adjusted Rasch person-fit statistics. *Journal of Applied measurement*, *7*, 170.
- Douglas, E. F., McDaniel, M. A., & Snell, A. F. (1996, August). THE VALIDITY OF NON-COGNITIVE MEASURES DECAYS WHEN

- APPLICANTS FAKE. In *Academy of Management Proceedings* (Vol. 1996, No. 1, pp. 127-131). Briarcliff Manor, NY 10510: Academy of Management.
- Dunnette, M. D., McCartney, J., CARLSON, H. C., & KIRCHNER, W. K. (1962). A study of faking behavior on a forced choice self description checklist. *Personnel Psychology, 15*, 13-24.
- Dwight, S. A., & Alliger, G. M. (1997). Reactions to overt integrity test items. *Educational and Psychological Measurement, 57*, 937-948.
- Dwight, S. A., & Donovan, J. J. (2003). Do warnings not to fake reduce faking?. *Human Performance, 16*, 1-23.
- Ellingson, J. E., Sackett, P. R., & Hough, L. M. (1999). Social desirability corrections in personality measurement: Issues of applicant comparison and construct validity. *Journal of Applied psychology, 84*, 155.
- Eysenck, S. B., Eysenck, H. J., & Barrett, P. (1985). A revised version of the psychoticism scale. *Personality and individual differences, 6*, 21-29.
- Ferrando, P. J., & Anguiano-Carrasco, C. (2013). A structural model - based optimal person-fit procedure for identifying faking. *Educational and Psychological Measurement, 73*, 173-190.
- Ferrando, P. J., & Chico, E. (2001). Detecting dissimulation in personality test scores: A comparison between person-fit indices and detection scales. *Educational and Psychological Measurement, 61*, 997-1012.
- Frei, R. L., Griffith, R. L., Snell, A. F., McDaniel, M. A., & Douglas, E. F. (1997). Faking of non-cognitive measures: Factor invariance using

- multiple groups LISREL. In G. Alliger (Chair), Faking matters. Symposium conducted at the meeting of the Society of Industrial and Organizational Psychology, St. Louis, MO.
- Goffin, R. D., & Christiansen, N. D. (2003). Correcting personality tests for faking: A review of popular personality tests and an initial survey of researchers. *International Journal of Selection and assessment, 11*, 340-344.
- González-Romá, V., Hernández, A., & Gómez-Benito, J. (2006). Power and type I error of the mean and covariance structure analysis model for detecting differential item functioning in graded response items. *Multivariate Behavioral Research, 41*, 29-53.
- Hancock, G. R., Lawrence, F. R., & Nevitt, J. (2000). Type I error and power of latent mean methods and MANOVA in factorially invariant and noninvariant latent variable systems. *Structural Equation Modeling, 7*, 534-556.
- Harley, D. C. (1952). Estimation of the dosage mortality relationship when the dose is subject to error(Technical Report No.15). Stanford,CA:Stanford University, Applied Mathematics and Statistical Laboratory.
- Hicks, L. E. (1970). Some properties of ipsative, normative, and forced-choice normative measures. *Psychological bulletin, 74(3)*, 167.
- Holbert, R. L., & Stephenson, M. T. (2002). Structural equation modeling in the communication sciences, 1995 - 2000. *Human Communication Research, 28*, 531-551.
- Hough, L. M. (1998). Effects of intentional distortion in personality

- measurement and evaluation of suggested palliatives. *Human Performance*, 11, 209-244.
- Hough, L. M., Eaton, N. K., Dunnette, M. D., Kamp, J. D., & McCloy, R. A. (1990). Criterion-related validities of personality constructs and the effect of response distortion on those validities. *Journal of applied psychology*, 75, 581.
- Hoyle, R. H., & Kenny, D. A. (1999). Sample size, reliability, and tests of statistical mediation. *Statistical strategies for small sample research*, 1, 195-222.
- Hu, L. T., & Bentler, P. M. (1999). Cutoff criteria for fit indexes in covariance structure analysis: Conventional criteria versus new alternatives. *Structural equation modeling: a multidisciplinary journal*, 6, 1-55.
- Ironson, G. H., & Davis, G. A. (1979). Faking high or low creativity scores on the Adjective Check List. *The Journal of Creative Behavior*.
- Jackson, D. L. (2003). Revisiting sample size and number of parameter estimates: Some support for the N: q hypothesis. *Structural equation modeling*, 10, 128-141.
- Johnson, C. E., Wood, R., & Blinkhorn, S. F. (1988). Spuriouser and spuriouser: The use of ipsative personality tests. *Journal of Occupational Psychology*, 61, 153-162.
- Kaplan, D. & George, R. (1995). A study of the power associated with testing factor mean differences under violations of factorial invariance. *Structural Equation Modeling*, 2, 101-118.
- Karabatsos, G. (2003). Comparing the aberrant response detection

- performance of thirty-six person-fit statistics. *Applied Measurement in Education*, *16*, 277-298.
- Kline, R. B. (2011). Principles and practice of structural equation modeling (3rd ed.). New York, NY: The Guilford Press.
- Kroger, R. O., & Turnbull, W. (1975). Invalidity of validity scales: The case of the MMPI. *Journal of Consulting and Clinical Psychology*, *43*, 48.
- Kuncel, N. R., & Borneman, M. J. (2007). Toward a new method of detecting deliberately faked personality tests: The use of idiosyncratic item responses. *International Journal of Selection and Assessment*, *15*, 220-231.
- Kuncel, N. R., & Borneman, M. J. (2007). Toward a new method of detecting deliberately faked personality tests: The use of idiosyncratic item responses. *International Journal of Selection and Assessment*, *15*, 220-231.
- Li, M. N. F., & Olejnik, S. (1997). The power of Rasch person-fit statistics in detecting unusual response patterns. *Applied Psychological Measurement*, *21*, 215-231.
- Liu, Y., Douglas, J. A., & Henson, R. A. (2009). Testing person fit in cognitive diagnosis. *Applied psychological measurement*, *33*, 579-598.
- McDonald, R. P. (1999). Test theory: A unified treatment. Mahwah, NJ: Lawrence Erlbaum.
- McNicol, D. (1972). A primer of signal detection theory. London, England: George Allen & Unwin.
- McNicol, D. (2005). A primer of signal detection theory. Psychology Press.

- Meijer, R. R., & Sijtsma, K. (1995). Detection of aberrant item score patterns: A review of recent developments. *Applied Measurement in Education, 8*, 261-272.
- Meijer, R. R., & Sijtsma, K. (2001). Methodology review: Evaluating person fit. *Applied psychological measurement, 25*, 107-135.
- Morgeson, F. P., Campion, M. A., Dipboye, R. L., Hollenbeck, J. R., Murphy, K., & Schmitt, N. (2007). Reconsidering the use of personality tests in personnel selection contexts. *Personnel psychology, 60*, 683-729.
- Muthén, B. (1984). A general structural equation model with dichotomous, ordered categorical, and continuous latent variable indicators. *Psychometrika, 49*, 115-132.
- Ones, D. S., Viswesvaran, C., & Reiss, A. D. (1996). Role of social desirability in personality testing for personnel selection: The red herring. *Journal of Applied Psychology, 81*, 660.
- Paulhus, D. L. (1984). Two-component models of socially desirable responding. *Journal of personality and social psychology, 46*, 598.
- Raju, N. S., Fortmann-Johnson, K. A., Kim, W., Morris, S. B., Nering, M. L., & Oshima, T. C. (2009). The item parameter replication method for detecting differential functioning in the polytomous DFIT framework. *Applied Psychological Measurement, 33*, 133-147.
- Reise, S. P., Widaman, K. F., & Pugh, R. H. (1993). Confirmatory factor analysis and item response theory: two approaches for exploring measurement invariance. *Psychological bulletin, 114*, 552.
- Rogers, H., & Swaminathan, H. (1993). A comparison of logistic regression

- and Mantel-Haenszel procedures for detecting differential item functioning. *Applied Psychological Measurement*, *17*, 105-116.
- Ruch, F. L. (1942). A technique for detecting attempts to fake performance on a self-inventory type of personality test. *Studies in personality*, 229-234.
- Schmit, M. J., & Ryan, A. M. (1992). Test-taking dispositions: A missing link?. *Journal of Applied Psychology*, *77*, 629-637.
- Schmit, M. J., & Ryan, A. M. (1993). The Big Five in personnel selection: Factor structure in applicant and nonapplicant populations. *Journal of Applied Psychology*, *78*, 966-974.
- Schmit, M. J., Ryan, A. M., Stierwalt, S. L., & Powell, A. B. (1995). Frame-of-reference effects on personality scale scores and criterion-related validity. *Journal of Applied Psychology*, *80*, 607-620.
- Shaughnessy, J. J., Zechmeister, E. B., & Zechmeister, J. S. (2000). *Research methods in psychology*. McGraw-Hill.
- Sinharay, S. (2017). How to compare parametric and nonparametric person fit statistics using real data. *Journal of Educational Measurement*, *54*, 420-439.
- Spencer, D. (1938). The frankness of subjects on personality measures. *Journal of Educational Psychology*, *29*, 26-35.
- Stark, S., Chernyshenko, O. S., & Drasgow, F. (2006). Detecting differential item functioning with confirmatory factor analysis and item response theory: Toward a unified strategy. *Journal of Applied Psychology*, *91*, 1292-1306.
- Stephenson, M. T., & Holbert, R. L. (2003). A Monte Carlo simulation of

- observable versus latent variable structural equation modeling techniques. *Communication Research*, *30*, 332–354.
- Tanaka, J. S. (1987). "How big is big enough?": Sample size and goodness of fit in structural equation models with latent variables. *Child development*, 134–146.
- Tendeiro, J. N., & Meijer, R. R. (2014). Detection of invalid test scores: The usefulness of simple nonparametric statistics. *Journal of Educational Measurement*, *51*, 239–259.
- Waller, N. G., Tellegen, A., McDonald, R. P., & Lykken, D. T. (1996). Exploring nonlinear models in personality assessment: Development and preliminary validation of a negative emotionality scale. *Journal of Personality*, *64*, 545–576.
- Wang, W. C., Shih, C. L. (2010). MIMIC Methods for Assessing Differential Item Functioning in Polytomous Items. *Applied Psychological Measurement*, *34*, 166–180.
- Worthington, D. L., & Schlottmann, R. S. (1986). The predictive validity of subtle and obvious empirically derived psychological test items under faking conditions. *Journal of Personality Assessment*, *50*, 171–181.
- Zickar, M. J., & Drasgow, F. (1996). Detecting faking on a personality instrument using appropriateness measurement. *Applied psychological measurement*, *20*, 71–87.
- Zickar, M. J., & Robie, C. (1999). Modeling faking good on personality items: An item-level analysis. *Journal of Applied Psychology*, *84*, 551–563.

ABSTRACT

Detection of Faking using Person Fit

Yuna Jang

Department of Psychology

The Graduate School of

Sungshin University

The purpose of this study is to verify the effectiveness of a new person fit proposed by Ferrando and Anguiano–Carrasco(2013). For that, Monte Carlo Simulation is used to check the ratio of the type 1 error and the power under various conditions. The condition this study framed is as follows; 1) Length of the study(short, medium, long), 2) Ratio of fakeable item(20%, 50%), 3) Sample size, 4) Ratio of fakeable item (honest:faker=1:1, 3:2, 4:1), 5) Type of faking(faking good, faking bad), 6) Amount of faking(Small, Large) The result under each condition is, the longer the test is the more type 1 error decreased and the power increased, and it appears identically under the condition of a fakeable item. Also the bigger the difference of the ratio of a faker, the more it has a tendency for type 1 error to decreased and the power increased. As the sample size decreased the power increased, though the drawback was this model was not converged in case of the amount of sample

is under 100. This index is more effective to faking bad, and the larger the faking the better fakers are detected. This study has significant in the sense that it verified the person fit of Ferrando and Anguiano–Carrasco(2013) by simulation, though its limitation is cut–off value was limited as 1.

Key Words: Person fit, Faking, Monte Carlo simulation

부 록

<부록 1> 모형 적합도 및 수렴율

<부록 1> 모형 적합도 및 수렴율

검사 길이	왜곡 문항 비율	전체 피검자 수	왜곡 응답자 비율	왜곡 유형	왜곡 크기	χ^2			CFI		TLI		RMSEA		수렴율
						<i>M</i>	<i>SD</i>	<i>p</i>	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>	
18	20%	100	1:1	긍정	LARGE	306.802	14.555	0.205	0.914	0.061	0.908	0.068	0.037	0.015	0.616
18	20%	100	1:1	긍정	SMALL	306.583	14.563	0.207	0.914	0.061	0.908	0.067	0.037	0.015	0.708
18	20%	100	1:1	부정	LARGE	307.825	15.197	0.198	0.904	0.068	0.898	0.075	0.038	0.015	0.714
18	20%	100	1:1	부정	SMALL	306.765	15.316	0.208	0.908	0.068	0.903	0.075	0.037	0.016	0.786
18	20%	100	3:2	긍정	LARGE	305.365	13.892	0.217	0.917	0.064	0.911	0.070	0.036	0.014	0.588
18	20%	100	3:2	긍정	SMALL	305.184	14.136	0.220	0.915	0.067	0.909	0.075	0.036	0.014	0.67
18	20%	100	3:2	부정	LARGE	305.750	13.907	0.213	0.912	0.064	0.906	0.070	0.037	0.014	0.65
18	20%	100	3:2	부정	SMALL	305.051	14.191	0.222	0.912	0.069	0.906	0.077	0.036	0.014	0.738
18	20%	100	4:1	긍정	LARGE	302.575	12.730	0.243	0.918	0.060	0.912	0.066	0.034	0.014	0.556
18	20%	100	4:1	긍정	SMALL	302.810	12.809	0.241	0.918	0.062	0.913	0.068	0.034	0.014	0.576
18	20%	100	4:1	부정	LARGE	301.305	14.527	0.266	0.924	0.058	0.920	0.065	0.032	0.016	0.404

18	20%	100	4:1	부정	SMALL	300.713	13.336	0.268	0.925	0.060	0.920	0.067	0.031	0.015	0.61
18	20%	300	1:1	긍정	LARGE	296.221	17.598	0.335	0.981	0.019	0.983	0.025	0.015	0.011	0.856
18	20%	300	1:1	긍정	SMALL	295.024	17.779	0.350	0.982	0.019	0.984	0.026	0.014	0.011	0.95
18	20%	300	1:1	부정	LARGE	297.595	17.191	0.318	0.976	0.022	0.978	0.029	0.016	0.011	0.986
18	20%	300	1:1	부정	SMALL	295.630	17.840	0.345	0.980	0.021	0.982	0.028	0.015	0.011	0.994
18	20%	300	3:2	긍정	LARGE	295.106	17.768	0.353	0.982	0.021	0.984	0.027	0.014	0.011	0.894
18	20%	300	3:2	긍정	SMALL	294.109	16.919	0.363	0.983	0.019	0.986	0.026	0.014	0.011	0.976
18	20%	300	3:2	부정	LARGE	298.608	18.607	0.314	0.975	0.026	0.977	0.032	0.016	0.011	0.968
18	20%	300	3:2	부정	SMALL	294.915	16.777	0.352	0.981	0.020	0.984	0.027	0.014	0.011	0.99
18	20%	300	4:1	긍정	LARGE	300.010	15.158	0.283	0.975	0.022	0.975	0.026	0.017	0.010	0.922
18	20%	300	4:1	긍정	SMALL	299.111	15.058	0.293	0.976	0.022	0.976	0.026	0.017	0.010	0.934
18	20%	300	4:1	부정	LARGE	302.511	16.377	0.258	0.969	0.025	0.969	0.029	0.019	0.010	0.904
18	20%	300	4:1	부정	SMALL	299.745	15.057	0.286	0.974	0.023	0.974	0.027	0.017	0.010	0.956
18	20%	500	1:1	긍정	LARGE	292.828	19.049	0.384	0.990	0.011	0.993	0.016	0.010	0.009	0.902
18	20%	500	1:1	긍정	SMALL	291.080	19.171	0.406	0.991	0.011	0.994	0.016	0.009	0.009	0.986

18	20%	500	1:1	부정	LARGE	293.460	18.918	0.374	0.989	0.013	0.991	0.019	0.010	0.009	0.994
18	20%	500	1:1	부정	SMALL	291.724	19.340	0.398	0.990	0.012	0.993	0.017	0.010	0.009	0.996
18	20%	500	3:2	긍정	LARGE	290.339	18.646	0.413	0.991	0.011	0.994	0.016	0.009	0.008	0.906
18	20%	500	3:2	긍정	SMALL	289.655	18.291	0.420	0.992	0.011	0.995	0.016	0.009	0.008	0.982
18	20%	500	3:2	부정	LARGE	292.514	18.041	0.384	0.989	0.013	0.991	0.018	0.010	0.008	0.996
18	20%	500	3:2	부정	SMALL	290.550	18.150	0.410	0.991	0.011	0.994	0.017	0.009	0.008	1
18	20%	500	4:1	긍정	LARGE	296.838	18.640	0.334	0.987	0.013	0.988	0.017	0.012	0.009	0.954
18	20%	500	4:1	긍정	SMALL	295.583	18.452	0.348	0.988	0.013	0.989	0.017	0.011	0.009	0.992
18	20%	500	4:1	부정	LARGE	300.083	19.153	0.299	0.984	0.015	0.984	0.019	0.013	0.009	0.966
18	20%	500	4:1	부정	SMALL	295.963	18.534	0.345	0.987	0.014	0.989	0.018	0.011	0.009	0.994
18	50%	100	1:1	긍정	LARGE	302.951	15.573	0.197	0.910	0.066	0.902	0.075	0.039	0.015	0.426
18	50%	100	1:1	긍정	SMALL	301.700	14.903	0.206	0.914	0.063	0.907	0.070	0.038	0.015	0.584
18	50%	100	1:1	부정	LARGE	305.210	16.598	0.180	0.896	0.069	0.887	0.077	0.041	0.015	0.604
18	50%	100	1:1	부정	SMALL	302.333	15.193	0.201	0.903	0.070	0.895	0.078	0.038	0.015	0.704
18	50%	100	3:2	긍정	LARGE	302.147	15.095	0.202	0.912	0.066	0.905	0.075	0.038	0.015	0.494

18	50%	100	3:2	긍정	SMALL	301.355	14.835	0.210	0.912	0.069	0.905	0.077	0.037	0.015	0.624
18	50%	100	3:2	부정	LARGE	303.138	14.660	0.191	0.906	0.066	0.898	0.073	0.039	0.014	0.562
18	50%	100	3:2	부정	SMALL	300.393	14.124	0.216	0.909	0.073	0.901	0.082	0.037	0.014	0.704
18	50%	100	4:1	긍정	LARGE	299.550	13.411	0.222	0.913	0.062	0.906	0.070	0.036	0.014	0.446
18	50%	100	4:1	긍정	SMALL	297.792	12.772	0.239	0.919	0.060	0.912	0.067	0.034	0.014	0.518
18	50%	100	4:1	부정	LARGE	293.729	13.715	0.293	0.940	0.053	0.936	0.061	0.029	0.017	0.236
18	50%	100	4:1	부정	SMALL	292.881	12.525	0.298	0.935	0.056	0.931	0.066	0.028	0.015	0.564
18	50%	300	1:1	긍정	LARGE	290.637	16.886	0.340	0.982	0.018	0.983	0.025	0.015	0.011	0.728
18	50%	300	1:1	긍정	SMALL	289.506	16.372	0.353	0.983	0.018	0.985	0.024	0.014	0.011	0.924
18	50%	300	1:1	부정	LARGE	294.927	18.008	0.293	0.969	0.028	0.969	0.036	0.017	0.011	0.946
18	50%	300	1:1	부정	SMALL	290.586	16.926	0.342	0.979	0.022	0.981	0.029	0.015	0.011	0.974
18	50%	300	3:2	긍정	LARGE	290.933	17.355	0.336	0.981	0.019	0.983	0.026	0.015	0.011	0.774
18	50%	300	3:2	긍정	SMALL	289.645	17.050	0.354	0.982	0.019	0.985	0.026	0.014	0.011	0.94
18	50%	300	3:2	부정	LARGE	296.019	18.890	0.285	0.968	0.030	0.968	0.037	0.018	0.011	0.952
18	50%	300	3:2	부정	SMALL	289.344	15.976	0.354	0.981	0.021	0.983	0.028	0.014	0.010	0.986

18	50%	300	4:1	긍정	LARGE	295.220	16.456	0.284	0.975	0.023	0.974	0.028	0.018	0.010	0.834
18	50%	300	4:1	긍정	SMALL	293.551	16.642	0.304	0.976	0.024	0.976	0.031	0.017	0.010	0.95
18	50%	300	4:1	부정	LARGE	300.940	19.760	0.231	0.961	0.034	0.959	0.040	0.021	0.010	0.834
18	50%	300	4:1	부정	SMALL	293.873	15.195	0.295	0.973	0.024	0.973	0.030	0.017	0.010	0.948
18	50%	500	1:1	긍정	LARGE	287.347	17.538	0.386	0.991	0.010	0.993	0.015	0.010	0.009	0.78
18	50%	500	1:1	긍정	SMALL	284.968	17.690	0.420	0.992	0.011	0.995	0.016	0.009	0.009	0.97
18	50%	500	1:1	부정	LARGE	288.123	17.290	0.374	0.987	0.014	0.990	0.020	0.010	0.008	0.992
18	50%	500	1:1	부정	SMALL	285.371	17.913	0.412	0.991	0.011	0.994	0.017	0.009	0.009	0.996
18	50%	500	3:2	긍정	LARGE	289.582	17.823	0.357	0.990	0.011	0.992	0.016	0.011	0.009	0.786
18	50%	500	3:2	긍정	SMALL	286.631	18.132	0.397	0.991	0.011	0.994	0.016	0.010	0.009	0.966
18	50%	500	3:2	부정	LARGE	290.493	19.373	0.353	0.986	0.016	0.988	0.022	0.011	0.009	0.984
18	50%	500	3:2	부정	SMALL	286.476	17.921	0.399	0.990	0.012	0.993	0.018	0.010	0.009	0.996
18	50%	500	4:1	긍정	LARGE	291.593	17.938	0.333	0.988	0.012	0.989	0.017	0.012	0.009	0.844
18	50%	500	4:1	긍정	SMALL	289.668	17.315	0.356	0.989	0.012	0.990	0.016	0.011	0.008	0.976
18	50%	500	4:1	부정	LARGE	296.713	19.497	0.280	0.981	0.017	0.981	0.021	0.014	0.009	0.936

18	50%	500	4:1	부정	SMALL	290.199	19.068	0.351	0.988	0.016	0.989	0.020	0.011	0.008	0.994
34	20%	100	1:1	긍정	LARGE	1134.178	16.618	0.141	0.908	0.040	0.904	0.042	0.031	0.005	0.36
34	20%	100	1:1	긍정	SMALL	1134.572	16.728	0.139	0.906	0.043	0.902	0.045	0.031	0.005	0.488
34	20%	100	1:1	부정	LARGE	1141.634	24.323	0.121	0.891	0.048	0.887	0.050	0.033	0.007	0.522
34	20%	100	1:1	부정	SMALL	1136.402	16.430	0.131	0.898	0.047	0.894	0.049	0.032	0.005	0.59
34	20%	100	3:2	긍정	LARGE	1137.514	16.644	0.127	0.903	0.045	0.899	0.047	0.032	0.005	0.372
34	20%	100	3:2	긍정	SMALL	1137.870	16.122	0.125	0.901	0.046	0.897	0.048	0.032	0.005	0.516
34	20%	100	3:2	부정	LARGE	1141.864	21.387	0.117	0.891	0.049	0.887	0.051	0.033	0.006	0.462
34	20%	100	3:2	부정	SMALL	1137.246	17.011	0.129	0.895	0.049	0.891	0.051	0.032	0.005	0.534
34	20%	100	4:1	긍정	LARGE	1125.809	18.574	0.184	0.910	0.050	0.907	0.053	0.028	0.006	0.412
34	20%	100	4:1	긍정	SMALL	1125.338	17.879	0.186	0.912	0.051	0.909	0.053	0.028	0.006	0.436
34	20%	100	4:1	부정	LARGE	1124.455	22.209	0.199	0.915	0.047	0.912	0.049	0.027	0.008	0.238
34	20%	100	4:1	부정	SMALL	1120.463	17.563	0.212	0.920	0.043	0.917	0.045	0.026	0.006	0.48
34	20%	300	1:1	긍정	LARGE	1109.287	21.264	0.287	0.984	0.012	0.983	0.013	0.012	0.006	0.762
34	20%	300	1:1	긍정	SMALL	1108.624	21.207	0.292	0.984	0.012	0.984	0.013	0.012	0.006	0.946

34	20%	300	1:1	부정	LARGE	1117.507	23.129	0.240	0.976	0.016	0.976	0.017	0.014	0.005	0.97
34	20%	300	1:1	부정	SMALL	1110.049	21.557	0.283	0.982	0.013	0.982	0.014	0.012	0.006	0.978
34	20%	300	3:2	긍정	LARGE	1110.956	21.315	0.277	0.983	0.012	0.982	0.013	0.013	0.005	0.768
34	20%	300	3:2	긍정	SMALL	1111.020	21.277	0.277	0.983	0.012	0.982	0.013	0.013	0.006	0.938
34	20%	300	3:2	부정	LARGE	1118.255	22.613	0.234	0.976	0.015	0.976	0.016	0.014	0.005	0.966
34	20%	300	3:2	부정	SMALL	1111.215	21.780	0.276	0.982	0.013	0.981	0.014	0.013	0.006	0.976
34	20%	300	4:1	긍정	LARGE	1121.199	20.021	0.212	0.974	0.014	0.973	0.015	0.015	0.004	0.864
34	20%	300	4:1	긍정	SMALL	1120.714	19.473	0.214	0.974	0.014	0.973	0.015	0.015	0.004	0.908
34	20%	300	4:1	부정	LARGE	1128.450	23.691	0.180	0.967	0.018	0.966	0.019	0.017	0.005	0.832
34	20%	300	4:1	부정	SMALL	1121.283	19.041	0.210	0.973	0.015	0.972	0.015	0.015	0.004	0.9
34	20%	500	1:1	긍정	LARGE	1102.096	25.806	0.344	0.993	0.007	0.993	0.008	0.008	0.005	0.808
34	20%	500	1:1	긍정	SMALL	1100.381	24.922	0.354	0.993	0.007	0.994	0.008	0.007	0.005	0.978
34	20%	500	1:1	부정	LARGE	1104.338	23.887	0.326	0.991	0.008	0.991	0.009	0.008	0.005	0.998
34	20%	500	1:1	부정	SMALL	1100.440	24.111	0.353	0.993	0.007	0.993	0.008	0.007	0.005	0.998
34	20%	500	3:2	긍정	LARGE	1103.254	24.390	0.334	0.992	0.007	0.993	0.008	0.008	0.005	0.846

34	20%	500	3:2	긍정	SMALL	1102.273	24.536	0.340	0.992	0.007	0.993	0.008	0.008	0.005	0.976
34	20%	500	3:2	부정	LARGE	1108.060	26.182	0.303	0.990	0.009	0.990	0.010	0.009	0.005	0.99
34	20%	500	3:2	부정	SMALL	1102.965	24.203	0.335	0.992	0.007	0.992	0.009	0.008	0.005	0.994
34	20%	500	4:1	긍정	LARGE	1109.987	22.855	0.286	0.989	0.008	0.989	0.009	0.009	0.005	0.89
34	20%	500	4:1	긍정	SMALL	1108.834	22.584	0.293	0.990	0.008	0.990	0.009	0.009	0.005	0.964
34	20%	500	4:1	부정	LARGE	1119.458	23.231	0.229	0.985	0.009	0.985	0.010	0.011	0.004	0.95
34	20%	500	4:1	부정	SMALL	1110.370	21.663	0.282	0.989	0.008	0.989	0.009	0.009	0.005	0.98
34	50%	100	1:1	긍정	LARGE	1132.409	18.445	0.110	0.897	0.043	0.892	0.045	0.033	0.005	0.13
34	50%	100	1:1	긍정	SMALL	1131.985	17.086	0.109	0.896	0.042	0.891	0.044	0.033	0.005	0.282
34	50%	100	1:1	부정	LARGE	1135.281	19.722	0.101	0.882	0.047	0.876	0.049	0.034	0.005	0.406
34	50%	100	1:1	부정	SMALL	1130.018	17.600	0.117	0.890	0.047	0.885	0.049	0.033	0.005	0.536
34	50%	100	3:2	긍정	LARGE	1127.776	17.813	0.126	0.904	0.042	0.900	0.044	0.032	0.005	0.18
34	50%	100	3:2	긍정	SMALL	1128.419	16.556	0.122	0.900	0.043	0.895	0.045	0.032	0.005	0.332
34	50%	100	3:2	부정	LARGE	1131.701	21.050	0.116	0.894	0.048	0.889	0.050	0.033	0.006	0.372
34	50%	100	3:2	부정	SMALL	1126.999	17.672	0.129	0.897	0.048	0.892	0.050	0.032	0.005	0.558

34	50%	100	4:1	긍정	LARGE	1120.186	16.401	0.158	0.909	0.047	0.904	0.049	0.030	0.005	0.254
34	50%	100	4:1	긍정	SMALL	1119.759	18.410	0.164	0.908	0.047	0.904	0.049	0.030	0.006	0.326
34	50%	100	4:1	부정	LARGE	1105.800	18.204	0.241	0.936	0.033	0.933	0.034	0.025	0.007	0.092
34	50%	100	4:1	부정	SMALL	1105.588	15.528	0.238	0.930	0.038	0.927	0.040	0.025	0.006	0.4
34	50%	300	1:1	긍정	LARGE	1100.517	21.259	0.279	0.983	0.011	0.983	0.013	0.013	0.006	0.348
34	50%	300	1:1	긍정	SMALL	1097.416	21.627	0.300	0.985	0.011	0.985	0.013	0.012	0.006	0.808
34	50%	300	1:1	부정	LARGE	1111.380	24.353	0.219	0.968	0.020	0.967	0.021	0.015	0.005	0.966
34	50%	300	1:1	부정	SMALL	1099.295	21.743	0.288	0.981	0.013	0.981	0.015	0.012	0.006	0.976
34	50%	300	3:2	긍정	LARGE	1102.077	20.896	0.268	0.982	0.012	0.982	0.013	0.013	0.005	0.45
34	50%	300	3:2	긍정	SMALL	1100.180	20.842	0.281	0.983	0.012	0.983	0.013	0.012	0.006	0.83
34	50%	300	3:2	부정	LARGE	1113.631	24.150	0.206	0.968	0.018	0.967	0.020	0.016	0.005	0.912
34	50%	300	3:2	부정	SMALL	1100.829	20.368	0.275	0.981	0.013	0.980	0.015	0.013	0.005	0.968
34	50%	300	4:1	긍정	LARGE	1110.397	20.541	0.217	0.975	0.014	0.973	0.015	0.015	0.005	0.504
34	50%	300	4:1	긍정	SMALL	1109.062	20.827	0.225	0.975	0.015	0.974	0.016	0.015	0.005	0.762
34	50%	300	4:1	부정	LARGE	1127.832	25.781	0.141	0.957	0.020	0.955	0.021	0.018	0.004	0.762

34	50%	300	4:1	부정	SMALL	1113.623	21.548	0.200	0.970	0.017	0.968	0.018	0.016	0.004	0.918
34	50%	500	1:1	긍정	LARGE	1093.053	25.277	0.335	0.992	0.007	0.993	0.008	0.008	0.005	0.392
34	50%	500	1:1	긍정	SMALL	1090.507	24.766	0.353	0.993	0.007	0.994	0.008	0.008	0.005	0.884
34	50%	500	1:1	부정	LARGE	1099.223	23.694	0.291	0.988	0.010	0.987	0.011	0.009	0.005	0.986
34	50%	500	1:1	부정	SMALL	1091.582	24.015	0.344	0.992	0.007	0.993	0.009	0.008	0.005	0.988
34	50%	500	3:2	긍정	LARGE	1093.624	24.835	0.331	0.992	0.007	0.993	0.008	0.008	0.005	0.444
34	50%	500	3:2	긍정	SMALL	1091.580	24.184	0.344	0.993	0.007	0.993	0.008	0.008	0.005	0.906
34	50%	500	3:2	부정	LARGE	1104.718	26.300	0.261	0.986	0.011	0.985	0.012	0.010	0.005	0.98
34	50%	500	3:2	부정	SMALL	1093.424	23.018	0.329	0.991	0.007	0.992	0.009	0.008	0.005	0.99
34	50%	500	4:1	긍정	LARGE	1101.734	22.364	0.273	0.989	0.007	0.989	0.008	0.010	0.004	0.512
34	50%	500	4:1	긍정	SMALL	1098.730	21.943	0.292	0.990	0.007	0.990	0.008	0.009	0.005	0.886
34	50%	500	4:1	부정	LARGE	1118.116	27.591	0.190	0.980	0.012	0.979	0.013	0.013	0.004	0.906
34	50%	500	4:1	부정	SMALL	1101.886	21.097	0.270	0.988	0.008	0.988	0.009	0.010	0.004	0.982
66	20%	100	1:1	긍정	LARGE	4334.941	23.437	0.096	0.892	0.041	0.890	0.042	0.024	0.002	0.072
66	20%	100	1:1	긍정	SMALL	4335.680	22.627	0.094	0.895	0.037	0.893	0.038	0.024	0.002	0.152

66	20%	100	1:1	부정	LARGE	4348.872	28.845	0.076	0.883	0.041	0.881	0.042	0.025	0.003	0.202
66	20%	100	1:1	부정	SMALL	4336.081	19.990	0.092	0.893	0.039	0.891	0.040	0.024	0.002	0.258
66	20%	100	3:2	긍정	LARGE	4333.614	20.174	0.097	0.905	0.035	0.903	0.035	0.024	0.002	0.118
66	20%	100	3:2	긍정	SMALL	4335.260	21.603	0.094	0.897	0.041	0.895	0.041	0.024	0.002	0.158
66	20%	100	3:2	부정	LARGE	4338.824	29.144	0.092	0.896	0.044	0.894	0.044	0.024	0.003	0.212
66	20%	100	3:2	부정	SMALL	4334.995	20.589	0.094	0.896	0.042	0.894	0.042	0.024	0.002	0.302
66	20%	100	4:1	긍정	LARGE	4312.025	30.409	0.148	0.902	0.042	0.901	0.042	0.022	0.003	0.086
66	20%	100	4:1	긍정	SMALL	4304.504	25.457	0.163	0.917	0.033	0.915	0.034	0.021	0.003	0.124
66	20%	100	4:1	부정	LARGE	4306.113	35.683	0.166	0.913	0.035	0.912	0.036	0.021	0.004	0.058
66	20%	100	4:1	부정	SMALL	4297.842	25.721	0.181	0.919	0.034	0.917	0.034	0.020	0.003	0.172
66	20%	300	1:1	긍정	LARGE	4273.978	21.792	0.251	0.983	0.006	0.983	0.007	0.010	0.002	0.422
66	20%	300	1:1	긍정	SMALL	4273.535	22.071	0.253	0.983	0.007	0.983	0.007	0.010	0.002	0.84
66	20%	300	1:1	부정	LARGE	4288.811	27.636	0.208	0.977	0.010	0.976	0.010	0.011	0.002	0.952
66	20%	300	1:1	부정	SMALL	4274.758	23.009	0.249	0.982	0.008	0.982	0.008	0.010	0.002	0.966
66	20%	300	3:2	긍정	LARGE	4271.482	23.407	0.260	0.984	0.007	0.984	0.007	0.010	0.002	0.468

66	20%	300	3:2	긍정	SMALL	4272.762	24.500	0.257	0.984	0.007	0.983	0.007	0.010	0.002	0.814
66	20%	300	3:2	부정	LARGE	4292.922	28.493	0.196	0.976	0.010	0.975	0.010	0.011	0.002	0.94
66	20%	300	3:2	부정	SMALL	4275.557	22.799	0.246	0.982	0.007	0.982	0.007	0.010	0.002	0.966
66	20%	300	4:1	긍정	LARGE	4297.188	21.916	0.180	0.974	0.009	0.973	0.009	0.012	0.002	0.55
66	20%	300	4:1	긍정	SMALL	4295.096	20.453	0.185	0.974	0.008	0.974	0.008	0.011	0.001	0.668
66	20%	300	4:1	부정	LARGE	4314.235	29.130	0.142	0.967	0.011	0.966	0.011	0.013	0.002	0.702
66	20%	300	4:1	부정	SMALL	4298.547	21.782	0.176	0.972	0.009	0.972	0.009	0.012	0.001	0.824
66	20%	500	1:1	긍정	LARGE	4257.442	32.001	0.316	0.994	0.004	0.994	0.005	0.006	0.003	0.532
66	20%	500	1:1	긍정	SMALL	4256.279	32.226	0.320	0.994	0.004	0.994	0.005	0.006	0.003	0.912
66	20%	500	1:1	부정	LARGE	4266.847	31.151	0.281	0.991	0.005	0.991	0.005	0.007	0.002	0.988
66	20%	500	1:1	부정	SMALL	4257.666	31.184	0.315	0.993	0.005	0.993	0.005	0.006	0.003	0.992
66	20%	500	3:2	긍정	LARGE	4261.326	28.825	0.299	0.993	0.004	0.993	0.004	0.007	0.002	0.554
66	20%	500	3:2	긍정	SMALL	4259.858	28.828	0.305	0.993	0.004	0.993	0.004	0.006	0.002	0.912
66	20%	500	3:2	부정	LARGE	4271.873	29.984	0.263	0.990	0.005	0.990	0.005	0.007	0.002	0.988
66	20%	500	3:2	부정	SMALL	4260.029	28.154	0.304	0.993	0.004	0.993	0.004	0.007	0.002	0.99

66	20%	500	4:1	긍정	LARGE	4271.498	24.120	0.261	0.990	0.004	0.990	0.004	0.007	0.002	0.618
66	20%	500	4:1	긍정	SMALL	4271.447	24.499	0.261	0.990	0.004	0.990	0.004	0.007	0.002	0.866
66	20%	500	4:1	부정	LARGE	4294.994	32.279	0.193	0.985	0.006	0.985	0.006	0.009	0.002	0.928
66	20%	500	4:1	부정	SMALL	4273.885	24.078	0.253	0.989	0.005	0.989	0.005	0.008	0.002	0.956
66	50%	100	1:1	긍정	LARGE	4317.504	11.371	0.086	0.890	0.033	0.887	0.034	0.025	0.001	0.008
66	50%	100	1:1	긍정	SMALL	4316.689	21.772	0.091	0.897	0.035	0.894	0.036	0.024	0.002	0.04
66	50%	100	1:1	부정	LARGE	4334.265	31.880	0.070	0.879	0.037	0.877	0.037	0.026	0.003	0.182
66	50%	100	1:1	부정	SMALL	4317.438	18.658	0.089	0.890	0.036	0.888	0.037	0.025	0.002	0.244
66	50%	100	3:2	긍정	LARGE	4316.156	15.182	0.089	0.910	0.037	0.908	0.038	0.024	0.002	0.01
66	50%	100	3:2	긍정	SMALL	4316.296	23.606	0.093	0.890	0.044	0.888	0.045	0.024	0.002	0.05
66	50%	100	3:2	부정	LARGE	4331.144	28.187	0.072	0.884	0.029	0.882	0.030	0.026	0.003	0.16
66	50%	100	3:2	부정	SMALL	4313.176	17.707	0.096	0.896	0.034	0.894	0.035	0.024	0.002	0.276
66	50%	100	4:1	긍정	LARGE	4299.060	30.217	0.131	0.897	0.050	0.895	0.051	0.023	0.003	0.05
66	50%	100	4:1	긍정	SMALL	4289.642	25.675	0.150	0.911	0.030	0.909	0.031	0.022	0.003	0.062
66	50%	100	4:1	부정	LARGE	4286.133	35.409	0.164	0.920	0.032	0.918	0.032	0.021	0.005	0.016

66	50%	100	4:1	부정	SMALL	4271.657	18.963	0.194	0.920	0.028	0.918	0.029	0.019	0.002	0.152
66	50%	300	1:1	긍정	LARGE	4253.354	21.313	0.252	0.984	0.006	0.984	0.007	0.010	0.002	0.038
66	50%	300	1:1	긍정	SMALL	4254.843	22.503	0.248	0.983	0.007	0.983	0.007	0.010	0.002	0.552
66	50%	300	1:1	부정	LARGE	4281.127	27.974	0.173	0.968	0.012	0.967	0.012	0.012	0.002	0.89
66	50%	300	1:1	부정	SMALL	4258.432	22.549	0.236	0.980	0.008	0.980	0.008	0.010	0.002	0.936
66	50%	300	3:2	긍정	LARGE	4244.335	29.161	0.288	0.986	0.008	0.986	0.008	0.009	0.003	0.048
66	50%	300	3:2	긍정	SMALL	4251.370	24.887	0.261	0.984	0.007	0.984	0.007	0.010	0.002	0.524
66	50%	300	3:2	부정	LARGE	4291.882	33.456	0.150	0.965	0.013	0.964	0.014	0.012	0.002	0.904
66	50%	300	3:2	부정	SMALL	4258.787	21.967	0.235	0.980	0.008	0.980	0.008	0.010	0.002	0.938
66	50%	300	4:1	긍정	LARGE	4283.200	21.955	0.164	0.973	0.009	0.972	0.009	0.012	0.001	0.112
66	50%	300	4:1	긍정	SMALL	4276.230	22.564	0.183	0.974	0.009	0.974	0.009	0.012	0.002	0.378
66	50%	300	4:1	부정	LARGE	4316.044	39.420	0.104	0.956	0.014	0.955	0.015	0.014	0.002	0.634
66	50%	300	4:1	부정	SMALL	4281.110	22.146	0.169	0.970	0.010	0.970	0.010	0.012	0.002	0.83
66	50%	500	1:1	긍정	LARGE	4243.584	23.959	0.288	0.993	0.004	0.993	0.004	0.007	0.002	0.066
66	50%	500	1:1	긍정	SMALL	4235.220	30.942	0.323	0.994	0.004	0.994	0.004	0.006	0.003	0.738

66	50%	500	1:1	부정	LARGE	4256.967	29.723	0.246	0.987	0.006	0.987	0.006	0.008	0.002	0.972
66	50%	500	1:1	부정	SMALL	4240.301	29.838	0.303	0.992	0.005	0.992	0.005	0.006	0.002	0.976
66	50%	500	3:2	긍정	LARGE	4249.478	29.465	0.271	0.992	0.004	0.992	0.004	0.007	0.002	0.08
66	50%	500	3:2	긍정	SMALL	4239.467	28.309	0.305	0.993	0.004	0.993	0.004	0.006	0.002	0.726
66	50%	500	3:2	부정	LARGE	4266.132	28.298	0.216	0.986	0.006	0.986	0.006	0.008	0.002	0.976
66	50%	500	3:2	부정	SMALL	4241.655	28.075	0.297	0.992	0.004	0.992	0.005	0.007	0.002	0.982
66	50%	500	4:1	긍정	LARGE	4252.092	25.781	0.259	0.990	0.005	0.990	0.005	0.007	0.002	0.124
66	50%	500	4:1	긍정	SMALL	4249.081	24.149	0.269	0.991	0.004	0.991	0.004	0.007	0.002	0.64
66	50%	500	4:1	부정	LARGE	4298.350	41.705	0.142	0.979	0.009	0.978	0.009	0.010	0.002	0.86
66	50%	500	4:1	부정	SMALL	4258.091	23.576	0.238	0.988	0.005	0.988	0.005	0.008	0.002	0.968