

Rasch 모형을 이용한 한국판 사회적 바람직성 척도 단축형(SDS-9)의 타당화*

배 병 훈 이 동 귀* 함 경 애
연세대학교

Stöber(2001)의 사회적 바람직성 척도(Social Desirability Scale-17; SDS-17)는 단일요인의 총 16문항으로 구성되어 있는데 문항수가 상대적으로 적어 설문지 연구에 포함하기에 용이하며, 사회상을 잘 반영하여 문항의 내용과 표현이 적절하고, 18세 이상 성인 대상으로 신뢰도와 타당도가 높은 유용한 척도로 평가되고 있다. 본 연구에서는 SDS-17에 대한 체계적 문항분석 과정을 거쳐 한국판 사회적 바람직성 척도를 타당화 하고자 대학생 302명으로부터 수집된 자료에 대해 Rasch 평정모형을 적용하였다. 그 결과, 문항의 적합도와 곤란도 분석을 바탕으로 9개의 최종 문항을 선정한 후, 추가적으로 수렴 및 변별 타당도 검증을 실시하여 수정된 SDS-9 척도(2요인 9문항)의 타당도를 평가하였다. 끝으로 한국판 사회적 바람직성 척도의 활용방안과 본 연구의 제한점을 논의하였다.

주요어 : 사회적 바람직성 척도(SDS-9), 타당화, Rasch 모형

* 이 논문은 제 1저자에게 수여된 2013년 국가연구장학금(장학증서번호: 2013-0085725669) 지원 보고서의 내용에 바탕을 두고 확장한 것임.

† 교신저자(Corresponding Author): 이동귀 / 연세대학교 심리학과 / (120-749) 서울시 서대문구 연세로 50
Tel: 02-2123-2437 / E-mail: lee82@yonsei.ac.kr

서 론

상담 분야 연구에서, 자기보고식 설문조사는 다양한 변수들 간의 관련성을 이해하기 쉬울 뿐 아니라 사람들의 경험이나 태도와 같은 삶의 다양한 측면들을 반영할 수 있고(Sheperis, Young, & Daniels, 2013), 동시에 많은 사람들에게 실시하기에 용이하다는 뚜렷한 장점이 있어(Ganster, Hennessey, & Luthans, 1983) 널리 사용되고 있다. 반면, 다양한 심리 변수들 간의 관계를 주로 연구 참여자의 주관적 경험 보고에 의존해서 파악하기 때문에 반응편파(response bias)에 의해서 연구결과가 왜곡될 가능성을 배제할 수 없다. 이 때문에 특정한 표본을 대상으로 설문조사를 실시한 연구들의 논의 부분에서는 흔히 연구결과의 일반화 제약, 지나친 인과적 해석의 문제점, 그리고 반응 편파(Moorman & Podsakoff, 1992)를 한계점으로 제시하곤 한다. 반응편파는 응답자 반응의 진실성과 맞닿아 있다. 조사에 사용한 질문지 자체가 응답자의 솔직한 반응을 이끌어내는 방향으로 정교하게 구성되었다 하더라도 만일 응답자가 연구자의 의도를 알아채거나 자신을 솔직하게 드러내지 않는다면 특정 방향으로 반응을 왜곡할 가능성은 여전히 잔존한다. 이러한 왜곡반응은 연구결과의 신뢰도와 타당도에 큰 위협이 되며 왜곡된 연구결과에 바탕을 둔 정책이나 개입이 이루어진다면 혼란과 사회적 손실을 초래할 수도 있다.

응답자 반응편파의 대표적인 예로 ‘사회적 바람직성(socially-desirable responding)’ 혹은 ‘사회적 선호도’를 들 수 있는데, 이는 응답자가 설문을 작성할 때 자신의 실제 의견이나 태도보다는 사회적으로 용인되거나 지지되는 방향으로 응답하려는 경향성을 의미한다. 이는 사회적으로 바람직하다고 판단되는 문항에 동의함으로써 긍정적으로 자신을 묘사하려는 욕구(Crowne & Marlowe, 1960;

Paulhus, 2002)를 반영하며, 응답자가 자신의 실제 생각이나 감정에 기초해서 솔직하게 응답하기보다는 의식적 혹은 무의식적으로 자신에 대한 긍정적 이미지 고양을 위하여 사회적인 또는 규준적인 외부 압력에 부응하는 방향으로 동의하는 경향성을 말한다(Zerbe & Paulhus, 1987). 응답자의 사회적 바람직성이 높을수록 특히 타인의 눈에 부정적으로 보일 수 있는 요소를 포함하는 심리적인 변인들 간의 관련성의 강도나 방향이 실제 응답자의 태도나 행동과는 유리되어 연구결과의 신뢰성과 타당성에 위협이 된다. 요컨대, 사회적 바람직성은 응답자의 반응 상에 체계적 오류를 야기함으로써 측정 결과를 모호하게 하는 대표적인 오염원으로 연구되어 왔다(Cronbach, 1946; Edwards, 1957). 이 때문에 응답자의 반응이 실제보다 과장되거나 축소되는 등 수집된 자료를 왜곡시켜 잘못된 결과를 도출하게 되거나(Fisher, 1993), 변인들 간 관련성을 살펴보는 연구에도 의도치 않은 영향을 미침으로써 연구결과의 타당성을 훼손한다(Ganster, Hennessey, & Luthans, 1983; Rubin & Babbie, 2008). 따라서 자기보고식 질문지를 이용한 태도나 행동을 측정하는 연구가 주를 차지하는 상담 분야에서 응답자의 반응을 왜곡시킬 수 있는 사회적 바람직성에 대한 비평적 고찰이 필요한 시점이다.

문화비교적 측면에서 볼 때, 아시아 국가 및 민족은 서양 국가와 민족에 비해 사회적 바람직성이 경향이 높다(Dudley et al., 2005; Keillor, Owens, & Pettijohn, 2001). 따라서 타인을 많이 의식하는 집단주의 문화권인 우리나라의 설문연구에서는 집단 내의 사회적 가치로 인해 사회적으로 좋은 모습을 보여주는 반응들이 개인의 태도나 행동에 영향을 미칠 가능성이 높다(Johnson, 1998; Lalwani, Shavitt, & Johnson, 2006; Moorman & Podsakoff, 1992). 지금까지 국외에서는 사회적 바

람직성을 측정하는 척도들에 대한 개발이 이루어짐과 동시에 사회적 바람직성 자체를 연구결과의 타당성에 영향을 미치는 주요 변인으로 다룬 연구가 활발히 진행되었지만(Connelly & Chang, 2015; Dodou & Winter, 2014; Krumpal, 2013; Richman et al., 1999), 국내에서는 사회적 바람직성에 대한 연구가 상대적으로 부족하였다. 국내에서는 그간 사회적 바람직성의 효용성에 대해서는 중요하게 생각하면서도 오염원으로서의 사회적 바람직성에 대해서는 별로 고려하지 않았는데, 아마도 그 이유는 첫째, 사회적 바람직성으로 인해 왜곡된 연구결과가 도출될 가능성에 대한 인식 자체가 부족하였기 때문이다. 사회적 바람직성 척도를 개발하거나 타당화 한 연구(김용석, 2010; 박중규, 2002; 이주일, 1988; 정보라, 2004; 최보라, 2007)들은 있어 왔지만, 실제로 사회적 바람직성을 연구의 주요 변인으로 상정하거나 통제 변인으로 사용한 국내 연구들은 그간 부족하였다(손은영, 2006; 어운경, 차정은, 2005). 특히, 상담 분야 연구에서 사회적 바람직성을 통제변인으로 실제로 사용되는 경우는 한국교육학술정보원(RISS) 검색 결과 3편에 불과하였다(구본용, 유제민, 2005; 최혜윤, 이동혁, 2015; 한수경, 정남운, 2014). 둘째, 만일 사회적 바람직성 변인이 반응편파에 미치는 영향을 고려한다고 할지라도, 문항수가 많은 사회적 바람직성 척도를 설문지에 추가하는 것에 대해 일선 연구자들이 부담을 느끼는 것 또한 사회적 바람직성 연구가 부족한 한 원인으로 볼 수 있을 것이다. 문항수가 늘어나게 되면 응답자의 검사 피로도가 증가하고(Anastasi, 1976), 응답의 정확성과 성실성이 떨어지게 되어(김권현 외, 2015) 연구의 타당성이 위협받기 때문에 연구자들은 많은 문항수로 이루어진 사회적 바람직성 척도의 사용을 꺼리게 되기 쉽다. 그러나 연구결과의 타당성을 담보하기 위해서 연구자는 사회적 바람직

성이 다른 변수들의 관계에 미칠 수 있는 오염요소를 통제해야 하며, 따라서 신뢰롭고 타당하면서도 문항수가 적어 사용하기 용이한 한국판 사회적 바람직성 척도가 필요하다.

지금까지 사회적 바람직성을 측정하는데 사용된 대표적인 척도로는 Crowne과 Marlowe(1960)가 개발한 말로-크라운 사회적 바람직성 척도(Marlowe-Crowne Social Desirability Scale; 이하 MCSDS)와 Paulhus(1984)가 개발한 바람직성 반응 균형 척도(Balanced Inventory of Desirable Responding; 이하 BIDR), Stöber(2001)의 사회적 바람직성 척도(Social Desirability Scale-17; 이하 SDS-17)가 있다. 이중 MCSDS는 총 33문항으로 이루어져 있으며, 대학생 집단을 대상으로 일상적인 사회적 선의도를 반영하려는 목적으로 개발되었다. 하지만, 사회적 바람직성의 개념적 특성상 지금 시대를 기준으로 문화나 규범이 반영되어야 함에도 불구하고, 개발된 지 너무 오래 되어 50년 전의 사회적 기준을 반영하고 있기 때문에 현 시대의 기준을 반영하지 못하고 있는 것으로 평가되고 있다(Blake et al., 2006; Stöber, 2001; Zerbe & Paulhus, 1987). 또한 다양한 척도들과 함께 설문지를 구성할 경우에는 문항수가 많아서 검사 피로도가 증가해 응답자들에게 부담이 된다는 지적을 받고 단축형 척도들이 개발되었지만(Reynolds, 1982; Strahan & Gerbasi, 1972), 이전 문항들의 일부를 그대로 사용했기 때문에 전체 문항수만 줄었을 뿐 현 시대의 기준을 반영하지 못하는 문제점은 개선되지 못하였다. 다음으로 BIDR은 자기기만적 고양을 측정하는 20문항과 인상관리를 측정하는 20문항, 총 40문항으로 구성되어 있는데, 자기기만적 고양(self-deceptive enhancement)은 솔직하게 응답하려 하지만 의도하지 못하는 사이에 자신을 과장되게 긍정적으로 묘사하려는 왜곡이며, 인상관리(impression management)는 의식적으로

타인에게 자신의 좋은 인상을 나타내려는 경향성을 의미한다(Paulus, 1984). BIDR은 사회적 바람직성을 자기본위적 편파와 인정에 대한 욕구나 도덕적인 편파, 둘로 구분하였다는데 의미가 있지만, 이 또한 MCSDS에서와 마찬가지로 문항수가 많다는 문제점과 30년 전 개발되었다는 점에서 오늘날의 사회현상을 반영하지 못하는 측면에서 제한점이 있다.

다른 두 척도에 비해 상대적으로 최근인 2000년대 이후에 개발된 SDS-17의 경우 사회적으로 바람직하지만 빈번하게 일어나지 않는 행동을 측정하는 10문항과 사회적으로 바람직하지 않지만 빈번하게 일어나는 행동을 측정하는 6문항을 통해 양쪽을 모두 측정할 수 있다는 장점이 있으며, 더불어 정신병리를 반영할 수 있는 문항을 제외함으로써 이와 관련된 가외변인(extraneous variable)을 사전에 통제하는 방식을 채택했다는 점에서 장점이 있다. 특정한 척도 문항에 심리부적응적인 측면이 혼입되면 해당 척도와 정신건강을 측정하는 변수간의 실질적인 관계가 차폐될 수 있다(Stanton et al., 1994). 또한 ‘나는 때때로 길에 쓰레기를 버린다’, ‘나는 때때로 나의 불쾌한 기분을 남들에게 푼다’, ‘나는 한 번 약속을 하면 무슨 일이 있어도 지킨다’와 같이 문항들을 살펴보면 더라도 실제 현실에서 충분히 있을 수 있는 행동적인 측면을 반영하는 문항들로 구성되어 있기 때문에 응답자들이 문항에 익숙하며, 독특한 문항이 없어 응답자에게 이질감을 주지 않기 때문에 다른 척도들과도 함께 사용하기에 용이하다. SDS-17은 개발 당시에는 총 17문항이었지만, 약물과 관련된 한 문항이 연구윤리적인 측면에서 측정상 애로사항이 있을뿐더러 척도의 본래 목적에도 잘 부합되지 않는다는 지적이 있어 제거됨으로써 실질적으로는 총 16문항의 단일요인으로 사용되어 왔다. SDS-17 척도는 문항수가 비교적 적

어 설문 패키지에 포함하기에 상대적으로 용이하며, 다른 사회적 바람직성 척도들에 비해 상대적으로 최근에 개발되었기 때문에 문항의 내용과 표현이 현대적인 척도라는 평가를 받고 있다(Blake et al., 2006; Stöber, 2001). 또한 이 척도는 원래 개발되었던 독일에서 뿐만 아니라 미국에서도 신뢰도와 타당도가 입증되었으며, 18세 이상에서 89세의 일반인들을 대상으로 척도를 개발하였기 때문에 척도의 적용 대상 범위도 비교적 넓은 유용한 척도로 평가할 수 있다.

SDS-17은 국내에서 김용석, 전수미과 이평순(2008)에 의해 한국어로 이미 번안·타당화 된바 있다. 그러나 이 연구는 타당화 과정에서 다음과 같은 한계점이 있어 연구결과의 일반화측면에서 중요한 한계가 있다. 우선 성인기 전 연령대의 일반인을 대상으로 개발·타당화를 한 원칙도와는 달리, 국내에서 이루어진 이 연구는 가정폭력 가해 남성들만을 대상으로 타당화가 이루어졌다. 그런데 사회적 바람직성에 있어서 성별에 따른 차이가 있기 때문에(Bell & Naugle, 2007), 가정폭력 가해 남성들만으로 타당화가 이루어진 이 척도는 다른 연령이나 여성에 대한 교차타당화를 실시하지 않아 적용 대상의 측면에서 지나치게 제한적이다. 아울러 구성개념 타당도를 확인하기 위하여 최대우도법(maximum likelihood estimation method)을 이용해 요인분석을 실시했는데, 이 연구의 참여자 수가 93명에 불과하여 요인해법을 도출하는 측면에서도 한계점이 있다. 예컨대, 방법론 전문가 Comrey와 Lee(1992)는 요인분석에서 양호한 해법을 도출하기 위해서는 최소 200명 이상이 필요하다고 제안한 바 있다. 요약하면, SDS-17은 상담 분야와 같이 설문조사를 많이 사용하는 경우에 쉽게 반응편파를 확인할 수 있는 유용한 측정도구이지만 현재 타당화된 한국판은 한정된 표본 집단과 적은 표본 수와 같은 타당화 과정의 한계점

들로 인해 다양한 집단들을 대상으로 한 후속연구에서 활용할 때 제한이 있을 가능성이 높다. 즉, SDS-17의 한국판에 대해 보다 정교한 분석방법을 사용한 타당화가 필요한 시점이다.

따라서 이러한 문제의식을 바탕으로 본 연구의 목적은 다음과 같다.

첫째, SDS-17문항에 대해 원저자의 허가를 받아 체계적인 번역 및 역번역 과정을 통해 새롭게 한국어로 번안하고 안정적인 요인해법을 도출할 수 있는 표본(예: 300명 이상의 대학생)을 사용하여 척도의 요인구조를 확인하고자 한다.

둘째, 문항반응이론(Item Response Theory)의 하나인 Rasch 모형(Rasch, 1960)을 적용하여 의미가 지나치게 중첩되는 문항이나 적합도가 떨어지는 문항들을 평가하고 재구성하고자 한다. 이때, Rasch 모형을 적용할 경우 문항의 적합도와 난이도 평가를 통해 척도의 완성도를 높이고 엄격한 기준으로 문항을 제거하거나 수정함으로써 척도의 활용도를 높이면서도 완성도를 함께 확보할 수 있다(이익섭, 홍세희, 신은경, 2007).

셋째, 문항선별 작업을 통해 확정된 최종문항으로 수정된 SDS 척도를 구성하고 이 척도의 수렴·변별 타당도를 확인할 것이다.

본 연구에서 적은 수의 문항으로 한국인의 사회적 바람직성을 타당하게 측정할 수 있는 도구를 확보할 수 있다면, 이는 향후 상담 분야의 다양한 연구 장면에서 연구자가 살펴 보고자하는 변인들 간의 관계에 있어서 사회적으로 용인된 방향으로 반응하는 경향성을 통제변인으로 활용함으로써 보다 신뢰성과 타당성이 있는 연구결과를 도출할 수 있는 중요한 근거를 마련할 수 있을 것이다. 아울러 특정한 구성개념에 대한 새로운 측정도구를 개발할 때 본 연구를 통해 타당화된 SDS 척도가 새로운 도구에 대한 변별타당도의 지표로서도 활용될 수 있을 것이다.

방 법

연구 대상

연구자가 소속된 대학의 학과 연구심의위원회로부터 연구수행과 연구윤리에 관한 사전 승인을 받아 경기도 소재 2개의 4년제 대학 심리학 전공 필수 과목과 심리학 관련 교양 과목을 수강하는 학생 310명을 대상으로 인터넷 온라인 설문을 실시하였다. 설문 시작 전 모든 참여자들은 연구 참여자의 권리에 대한 설명을 읽고 참가 동의를 하였으며, 설문 이후에는 연구의 목적과 예상된 결과에 대한 사후 설명을 읽을 수 있게 하였다. 수집된 310부의 자료 중 불성실한 답변을 한 8부를 제외한 총 302부의 자료를 분석에 이용하였다. 인구통계학적 특성을 살펴보면, 성별 분포는 남자 139명(46.0%), 여자 163명(54.0%)으로 구성되어 있었고, 평균 연령은 21.03세($SD = 2.66$)였다. 학년별 분포는 1학년 87명(28.8%), 2학년 119명(39.4%), 3학년 71명(23.5%), 4학년 25명(8.3%)로 나타났다.

측정도구

한국판 사회적 바람직성 척도(K-SDS-17)

한국 문화에서의 사회적 바람직성을 측정하기 위하여 Stöber(2001)가 개발한 사회적 바람직성 척도(The Social Desirability Scale-17; SDS-17)를 사용하였다. 기존의 SDS-17은 총 16문항의 단일 요인으로 구성되어 있으며, ‘예(1점)’, ‘아니오(0점)’의 이항적 응답범주로 응답하도록 되어 있다. 하지만 본 연구에서는 개발된 사회적 바람직성 척도들의 응답을 점수화하는 방식에 있어서 연속적 점수화 방식인 Likert형 점수화 방식이 더 높은 타당도를 보여준다는 연구(Stöber, Dette, & Musch, 2002)를 바탕으로 본 연구에서는 ‘전혀 그렇지 않다(1점)’

에서 ‘매우 그렇다(5점)’까지 반응할 수 있는 Likert 5점 척도를 사용하였다. 합산한 문항 점수가 높을수록 사회적 바람직성이 높은 것을 의미한다. Stöber(2001)의 연구에서의 문항 간 내적 합치도(Cronbach's α)는 .80으로 나타났으며, Blake 등(2006)의 연구에서는 .64-.92로 나타났다. 본 연구에서 나타난 내적 합치도는 .72였다.

바람직성 반응 균형 척도(BIDR)

본 연구에서 타당화할 SDS-17의 수렴 타당도를 살펴보기 위하여 정보라(2005)가 Paulhus(1998)의 PDS(Paulhus Deception Scale)와 BIDR-7에 기초하여 개발한 사회적 바람직성 척도를 사용하였다. 자기기만적 고양과 인상관리 2개 하위요인으로 이루어져 있으며, 각 10문항씩 총 20개 문항으로 구성되어 있다. ‘전혀 그렇지 않다(1점)’에서 ‘매우 그렇다(5점)’까지 반응할 수 있는 Likert 5점 척도이며, 점수가 높을수록 사회적 바람직성이 높음을 의미한다. Paulhus(1998)의 연구에서 나타난 내적 합치도는 자기기만적 고양이 .70, 인상관리가 .81로 나타났으며, 정보라(2005)의 연구에서는 모두 .69로 나타났다. 본 연구에서 나타난 내적 합치도는 자기기만적 고양 .70, 인상관리는 .71, 전체 .71이었다.

성격 5요인 척도(IPIP)

변별 타당도를 살펴보기 위하여 Goldberg(1999)가 개발한 IPIP(International Personality Item Pool)를 유태용, 이기범, Ashton(2004)이 번안한 척도를 사용하였다. IPIP는 개인의 성격을 신경증(Neuroticism), 외향성(Extroversion), 개방성(Openness), 친화성(Agreeableness), 성실성(Conscientiousness)의 5가지 요인으로 측정한다. 각 요인별 10문항씩 총 50문항으로 구성되어 있으며, ‘전혀 그렇지 않다(1점)’에서 ‘매우 그렇다(5점)’까지 응답할 수 있는

Likert 5점 척도이다. 개인의 특성을 파악하기 위해서 가장 일반적으로 사용되는 검사가 성격검사이며, 사회적 바람직성의 구인타당도를 확인하기 위해 가장 많이 사용되고 있는 개인의 특성 변인이 바로 성격적 특성이다(Griffith, 1997; Ones & Viswesvaran, 1998; Schmit & Ryan, 1993). 따라서 본 연구에서는 Stöber(2001)의 척도개발 연구에서도 변별타당도를 살펴보기 위해 사용한 5요인 성격검사를 활용하였는데, 관련성이 확인된 신경증 10문항과 성실성 10문항만 이용하였다. 유태용, 이기범, Ashton(2004)의 연구에서 나타난 내적 합치도는 신경증 .87, 성실성 .76이었다. 본 연구에서 나타난 내적 합치도는 각각 .88(신경증), .80(성실성)이었다.

연구절차

우선 척도를 타당화 함에 앞서 SDS-17의 원저자에게 이메일을 보내서 척도의 번안·타당화에 대한 사용 허락을 받았다. 더불어 번안 과정상에서 의미가 모호한 문항(예, 2번 문항, 3번 문항)의 경우에는 원저자에게 연락하여 문항의 개발 의도와 정확한 의미를 확인받았다. 이후 연구자들 및 심리학 박사 1명, 박사과정생 1명, 석사과정 2명, 학사 졸업생 5명이 함께 검토하여 SDS-17 문항을 한글로 번역하였다. 다음으로 심리학 석사과정생이며 동시에 미국에서 학부까지 마친 이중 언어(한국어, 영어) 능통자가 영문 SDS-17의 원래 문항을 알지 못하는 상태에서 번안된 한글 문항을 영어로 재번역 하였다. 원래 영어 문항과 영어로 역번역된 문항은 캐나다와 미국에서 학부를 마치고 국내에서 석사과정을 마친 이중 언어(한국어, 영어) 능통자 2명에게 메일을 보내서 내용면에서 두 버전(version) 간 동일성 평정을 실시하였다. 이때, 원래 문항과 역번역된 문항이 얼마나 동일한지를

5점 척도(1: 전혀 동일하지 않음, 5: 완전히 동일함)로 평정하게 한 후, 각자 의미에 차이가 있다고 생각된 문항이 있다면 그 이유에 대해 서술하도록 하였다. 마지막으로 두 명의 평정자 중 한 명이라도 3점 이하로 평정된 문항의 경우에는 문항을 수정 · 재검토 과정을 반복하여 최종적인 번역 문항을 완성하였다.

자료분석

사회적 바람직성 척도를 타당화 하기 위하여 다음과 같은 분석을 실시하였다. 첫째, SPSS 21.0을 이용하여 연구 대상의 인구통계학적 특성을 살펴보기 위하여 빈도분석을 실시하고, 연구에 이용된 척도들의 내적 합치도를 산출하였다. 둘째, Rasch 모형을 적용하기에 앞서 척도 문항들의 일차원성 가정을 만족하는지를 검증하기 위하여 SPSS 21.0을 이용하여 탐색적 요인분석을 실시하였다. 셋째, Facets 3.7(Linacre, 2014)을 이용해서 응답범주에 대한 범주 확률곡선, 문항특성곡선과 95% 신뢰구간, 개별문항의 적합도인 내적합도지수와 외적합도지수, Rasch 모형에 의한 피험자 × 문항 분포도를 살펴보았다. 이때, Rasch 모형 중 다분문항모형(Polytomous Model)인 Andrich(1978)의 평정척도모형(Rating Scale Model: RSM)을 적용하였다. Rasch 모형은 개발 당시 그 목적이 성취도 검사에 대한 타당화이기 때문에 개인의 능력은 심리적 특질(psychological trait)의 정도를, 문항 난이도는 동의하기 어려운 정도를 의미한다(홍세희, 조용래, 2006). 따라서 본 연구에서는 개인의 능력을 사회적 바람직성 정도로, 문항 난이도는 문항에 동의하기 어려운 정도로 사용하였다. 또한 Hong, Kim과 Wolfe(2005)의 절차에 따라 문항의 응답범주 수가 적절한지를 검증한 다음 각 문항의 적합도를 바탕으로 부적합한 문항을 제거하였

고, 문항에 대해 동의하기 어려운 정도를 고려하여 중복되는 난이도를 지닌 문항을 제거하여 최종 문항을 선정하였다. 마지막으로 SPSS 21.0을 이용하여 사회적 바람직성 척도의 수렴 · 변별 타당도를 상관분석을 통해 살펴보았다.

결 과

기술통계

Rasch 모형을 적용해서 사회적 바람직성 척도를 타당화 하기 전에 각 문항들의 기술통계를 살펴보기 위하여 평균과 표준편차를 산출하였고, 그 결과를 표 1에 제시하였다.

분석 결과, 15번 문항인 ‘나는 항상 건강식을 먹는다’(평균 = 1.97, 표준편차 = .73)의 평균이 가장 낮았으며, 전반적으로 역채점을 하는 6개 문항들(1번, 5번, 6번, 10번, 14번, 16번)의 평균이 그렇지 않은 문항들에 비해 높은 것을 확인할 수 있다.

일차원성 검증

문항반응이론을 통해 사회적 바람직성 척도의 문항들의 적절성을 확인하기 위해서는 분석의 가정인 일차원성에 대한 검증이 먼저 이루어져야 한다(Bond & Fox, 2001). 비록 개발 당시와 국내 타당화 연구에서도 모두 단일차원이 지지되었음에도 불구하고, 선행했던 국내 타당화 연구에서의 요인분석 결과가 적은 표본수로 인한 한계가 있기 때문에 본 연구에서 다시 변안한 16문항에 대해 탐색적 요인분석(주축분해법과 프로맥스 회전 방식을 사용)을 실시하였다. 스크리 도표를 살펴 보았을 때 2요인이 지지되는 것으로 확인되었지

표 1. 각 문항의 평균과 표준편차

문항 번호	내용	평균	표준편차
1	<u>나는 때때로 길에 쓰레기를 버린다.</u>	3.71	.93
2	나는 항상 나의 실수를 솔직하게 인정하고 일어날지 모를 부정적인 결과들에 직면한다.	3.02	.69
3	나는 교통수단을 이용할 때 항상 다른 사람들에게 공손하게 대하고 배려한다.	3.12	.73
4	나는 남들이 내 의견에 동의하지 않더라도 항상 그 사람의 의견을 수용한다.	2.87	.76
5	<u>나는 때때로 나의 불쾌한 기분을 남들에게 푼다.</u>	3.48	.88
6	<u>나는 나의 이익을 위해 누군가를 이용한 적이 있다.</u>	3.30	.80
7	나는 대화를 할 때 항상 다른 사람의 말을 항상 주의 깊게 경청하고 끝까지 듣는다.	3.05	.76
8	나는 위급한 상황에 있는 사람을 돕는데 결코 망설이지 않는다.	2.66	.78
9	나는 한 번 약속을 하면 무슨 일이 있어도 지킨다.	2.65	.77
10	<u>나는 때때로 상대방이 없는 자리에서 그 사람에 대해 뒷담화를 한다.</u>	3.33	.78
11	나는 결코 다른 사람의 신세를 지며 살지 않을 것이다.	2.79	.95
12	나는 스트레스를 받을 때라도 항상 다른 사람들에게 친절하고 예의를 지킨다.	2.74	.79
13	논쟁을 할 때 나는 항상 감정에 치우치지 않고 객관적인 태도를 유지한다.	2.57	.79
14	<u>내가 빌린 물건을 되돌려주지 않은 적이 적어도 한 번은 있다.</u>	3.38	.91
15	나는 항상 건강식을 먹는다.	1.97	.73
16	<u>나는 가끔 보답을 기대할 수 있을 때만 도움을 준다.</u>	3.82	.75

주. 밑줄 친 1번, 5번, 6번, 10번, 14번, 16번은 역채점 문항들로 모두 역채점 되었음.

만, 스크리 도표는 요인이 과잉추출되고 연구자의 주관적 판단이 개입될 가능성이 있기 때문에 보다 정확한 통계적 절차로 평행분석을 실시하였다 (Velicer, Eaton, & Fava, 2000). 이 방법은 실제 자료

를 통해서 추출된 고유값(eigenvalue)과 실제 자료와 사례수가 동일한 50개 이상의 무선(random) 추출 표본에서 나온 고유값의 기대치를 비교함으로써 하위요인 수가 우연 발생적인 것 이상으로 나

표 2. 평행분석 결과

요인수	실제 고유값	평균 고유값 기대치	95번째 백분위의 고유값
1	3.077	1.422	1.504
2	1.466	1.327	1.394
3	.642	1.253	1.293
4	.445	1.196	1.238

주. 60개 무선 추출 표본.

타난 것까지를 요인 수로 결정하는 방법이다(김병직 외, 2013).

표 2에서 볼 수 있듯이 요인 수가 2개 이후부터 본 연구의 실제 자료에서 나온 고유값이 무선 추출 표본에서 얻은 고유값보다 낮아지고 있는 것을 통해 확실한 요인 수는 2개로 볼 수 있다. 이후 요인 수를 2개로 고정한 이후 탐색적 요인 분석을 다시 실시한 결과, 요인 1은 총 10개 문항(2번, 3번, 4번, 7번, 8번, 9번, 11번, 12번, 13번, 15번)으로 구성되었으며, ‘나는 남들이 내 의견에 동의하지 않더라도 항상 그 사람의 의견을 수용한다’, ‘나는 한 번 약속을 하면 무슨 일이 있어도 지킨다’와 같이 바람직하지만 빈번하게 일어나지 않는 행동을 측정한다. 각 문항에 동의하는 정도가 높을수록 사회적 바람직성이 높음을 보여준다고 판단되어 ‘긍정형 사회적 바람직성’으로 명명하였다. 이 요인에 대한 문항 간 내적 합치도 계수는 .79로 나타났다. 다음으로 요인 2는 총 6개 문항(1번, 5번, 6번, 10번, 14번, 16번)으로 구성되어 있으며, ‘나는 때때로 길에 쓰레기를 버린다’, ‘나는 나의 이익을 위해 누군가를 이용한 적이 있다’와 같이 바람직하지 않지만 빈번하게 일어나는 행동을 측정한다. 각 문항에 동의하는 정도가 낮을수록 사회적 바람직성이 높음을 보여준다고 판단되어 ‘부정형 사회적 바람직성’으로 명명하였다. 이 요인에 대한 문항 간 내적 합치도 계수는 .61로 나타났다. 이처럼 개발 당시와 타당화 했던 선행연구와는 달리 2개의 요인구조로 가진 것이 확인되었기 때문에 이후의 평정척도모형(Rating Scale Model)을 이용한 분석에서는 각 하위 요인별 문항에 대해 분석을 실시하였다(김혜진, 김옥분, 2009; 이익섭, 홍세희, 신은경, 2007; 홍세희, 조용래, 2006).

응답범주에 대한 범주 확률곡선

사회적 바람직성 척도의 문항 응답범주가 5개인 것이 적절한지를 확인하기 위하여 각 요인에 대하여 범주 확률곡선을 살펴보고, 그 결과를 그림 1과 그림 2에 제시하였다. 1점은 ‘전혀 그렇지 않다’, 2점은 ‘대체로 그렇지 않다’, 3점은 ‘보통이다’, 4점은 ‘대체로 그렇다’, 5점은 ‘매우 그렇다’로 응답한 것을 의미한다. 그림에서 X축은 개인의 사회적 바람직성 태도와 문항에 대해 동의하기 어려운 정도 사이의 로짓 차이(person

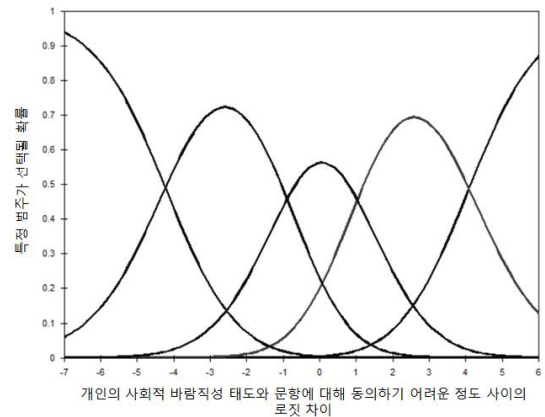


그림 1. 긍정형 사회적 바람직성 요인의 범주확률곡선

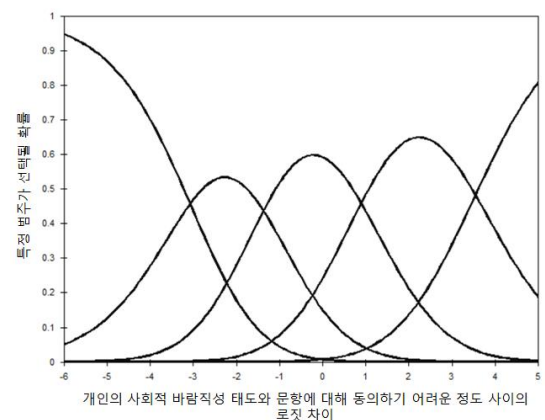


그림 2. 부정형 사회적 바람직성 요인의 범주확률곡선

measure relative to item difficulty)이며, Y축은 특정 범주가 선택될 확률(category probability)로서, 그림에 제시된 각 곡선들은 왼쪽에서부터 1점에서 5점까지의 응답범주를 나타낸다. 각 응답범주가 X축의 특정 영역에서 가장 높게 선택될 확률을 보여야만 그 범주가 유용하다고 할 수 있는데, 그림 1과 그림 2를 통해 알 수 있듯이 각 곡선들이 다른 곡선에 포함되지 않고 범주 간 간격의 차이가 분명히 나타났기 때문에 5점의 응답범주가 적절하였음을 확인할 수 있다.

문항특성곡선에 의한 적합도 검증

다음으로 Rasch 모형 검증을 통해서 나온 각 요인별 이론적 문항특성곡선(가운데 가장 굵은 곡선)과 실제 관찰된 점수의 평균('x'로 표시됨)간의 관계를 나타낸 그래프들을 통해서 각 피험자들의 관찰된 점수의 평균과 그에 따른 95% 신뢰구간(가장 바깥쪽의 가는 두 곡선)을 살펴봄으로써 각 피험자들의 관찰된 점수들이 문항특성곡선을 따라서 얼마나 잘 부합하는지를 시각적으로 살펴보고, 그 결과를 그림 3과 그림 4에 제시하였다.

그 결과 긍정형 사회적 바람직성 요인의 경우 관찰된 점수의 평균(대략 26개의 'x'로 표시됨) 중에서 1개를 제외한 나머지 25개의 'x'가 95% 신뢰구간 안에 위치하였고, 부정형 사회적 바람직성 요인의 경우 모든 관찰된 점수의 평균(x)가 95% 신뢰구간 안에 위치한 것을 알 수 있었다. 이러한 결과를 통해 두 요인 모두 피험자들을 통해서 관찰된 점수들이 문항특성곡선에 적합하다는 것을 확인할 수 있다.

문항적합도 분석

2개 요인 각각에 있어 문항에 응답한 피험자의

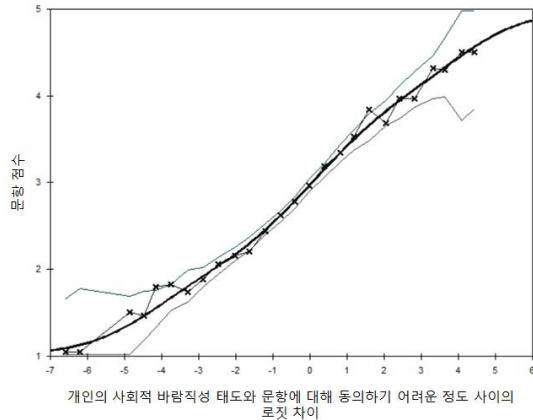


그림 3. 긍정형 사회적 바람직성 요인의 문항특성곡선과 95% 신뢰구간

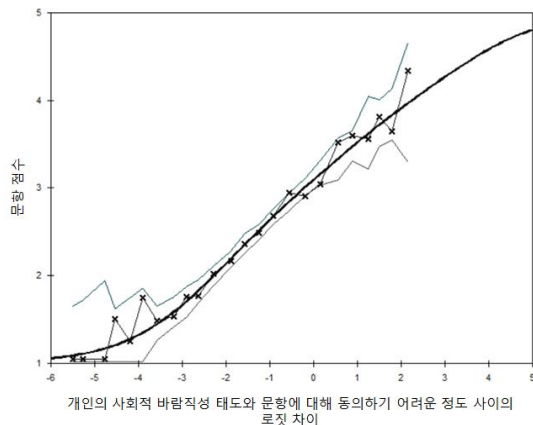


그림 4. 부정형 사회적 바람직성 요인의 문항특성곡선과 95% 신뢰구간

반응이 얼마나 평정척도모형에 적합한지를 모형에 의해 기대되는 확률값과 피험자의 실제 반응값을 비교함으로써 피험자나 문항의 적합도를 검증하는 타당도 지수를 통해 개별 문항의 타당도를 확인하였다.

Rasch 모형에서는 내적합도지수(Infit Mean Square)와 외적합도지수(Outfit Mean Square) 두 가지가 제시된다. 외적합도지수는 피험자의 능력수준에 비

추어서 너무 쉬운 문항을 틀리거나 또는 너무 어려운 문항을 맞힐 경우와 같은 이상 문항반응 형태에 민감한 지수이며, 내적합도지수는 극소수의 이상 문항반응 형태로 말미암아 전체 적합도지수가 부적합으로 판단될 가능성이 높기 때문에 이를 극복하기 위해서 피험자의 능력수준에서 많이 벗어나지 않은 문항의 이상 문항반응 형태에 가중치를 둔 적합도지수이다(Wright & Masters, 1982). 양호한 문항에 대한 두 지수의 기대값은 1이며, 지수가 1보다 클수록 해당 문항은 측정하러

표 3. 문항적합도 분석

요인	문항 번호	내용	동의하기		
			어려운 정도	내적합도	외적합도
긍정형 사회적 바람직성	2	나는 항상 나의 실수를 솔직하게 인정하고 일어날지 모를 부정적인 결과들에 직면한다.	-.70	.79	.81
	3	나는 교통수단을 이용할 때 항상 다른 사람들에게 공손하게 대하고 배려한다.	-.94	.77	.77
	4	나는 남들이 내 의견에 동의하지 않더라도 항상 그 사람의 의견을 수용한다.	-.35	.84	.85
	7	나는 대화를 할 때 항상 다른 사람의 말을 항상 주의 깊게 경청하고 끝까지 듣는다.	-.78	.77	.79
	8	나는 위급한 상황에 있는 사람을 돕는데 결코 망설이지 않는다.	.18	.99	.96
	9	나는 한 번 약속을 하면 무슨 일이 있어도 지킨다.	.20	.91	.90
	11	나는 결코 다른 사람의 신세를 지며 살지 않을 것이다.	-.17	1.67	1.66
	12	나는 스트레스를 받을 때라도 항상 다른 사람들에게 친절하고 예의를 지킨다.	-.03	.85	.83
	13	논쟁을 할 때 나는 항상 감정에 치우치지 않고 객관적인 태도를 유지한다.	.41	1.02	.99
부정형 사회적 바람직성	15	나는 항상 건강식을 먹는다.	2.17	1.48	1.42
	1	나는 때때로 길에 쓰레기를 버린다.	-.44	1.16	1.16
	5	나는 때때로 나의 불쾌한 기분을 남들에게 푼다.	.05	1.04	1.05
	6	나는 나의 이익을 위해 누군가를 이용한 적이 있다.	.44	.79	.80
	10	나는 때때로 상대방이 없는 자리에서 그 사람에 대해 뒷담화를 한다.	.37	.85	.88
	14	내가 빌린 물건을 되돌려주지 않은 적이 적어도 한 번은 있다.	.26	1.17	1.17
	16	나는 가끔 보답을 기대할 수 있을 때만 도움을 준다.	-.67	.97	1.00

주. 음영이 표시된 2번, 3번, 6번, 11번, 15번 문항이 문항적합도 분석을 통해 제거된 문항임.

는 심리적 특질과 관계가 약함을 의미하고, 1보다 작을수록 다른 문항과 내용이 중복되어 있을 가능성이 높음을 의미한다. 두 지수값이 대략 .75 이하이거나 1.30 이상일 경우 적합도가 좋지 않은 것으로 판단된다(McNamara, 1996). 하지만, 이 기준은 절대적인 것이 아니며 연구목적에 따라 기준이 달라질 수도 있고(Karabassos, 1997), 심지어 .50 - 1.50을 적합도 판단 기준으로 하는 연구들도 있다. 본 연구에서는 척도의 활용성을 높이기 위하여 문항의 수를 줄여서 양호한 문항만으로 구성된 척도를 만드는 것이 목적이므로 더 엄격한 기준을 적용하여 내적합도지수와 외적합도지수 모두 .80 이하이거나 1.20 이상인 문항을 제거하였다. 다만 .75 - .80, 그리고 1.20 - 1.30 사이의 문항 적합도 지수를 보이는 문항도 비교적 양호한 문항이라고 볼 수 있으므로 이 문항의 곤란도가 다른 문항이 측정하지 못하는 곤란도일 경우에는 문항을 제거하지 않았다. 사회적 바람직성 척도의 2개 요인에 있어서 문항 각각의 적합도 지수를 제시한 결과는 표 3에 제시하였다.

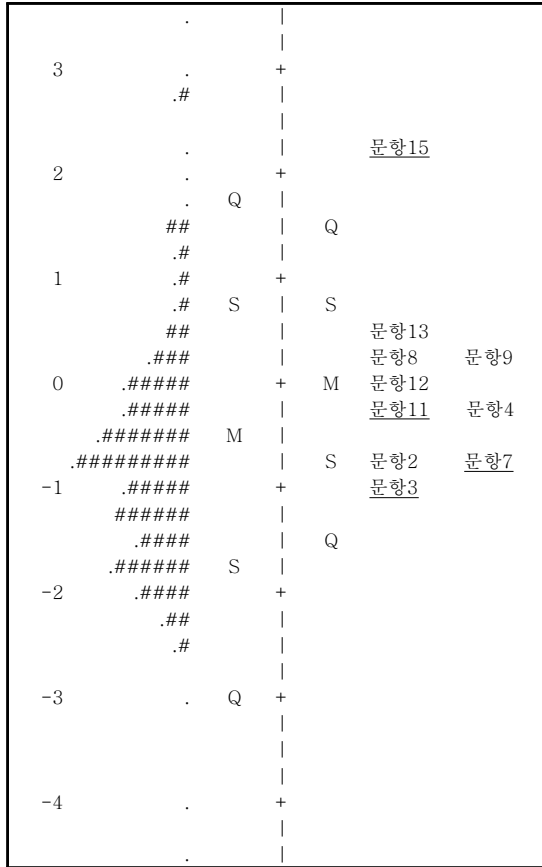
우선 긍정형 사회적 바람직성 요인에 있어서 내적합도지수와 외적합도지수 모두 .80 이하, 그리고 1.20 이상인 문항번호 3번, 7번, 11번, 15번, 총 4문항을 제외한 6개 문항을 1차로 선정하였고, 부정형 사회적 바람직성 요인에 있어서는 6번 문항을 제외한 5개 문항을 1차로 선정하였다. 1차로 선정된 11문항에 대해서는 이후 피험자 × 문항 분포도를 통해 문항에 대해 동의하기 어려운 정도가 중복되는 문항일 경우에 다시 제거함으로써 최종적으로 문항을 선정하였다.

Rasch 모형에 의한 피험자 × 문항 분포도

마지막으로 피험자들의 사회적 바람직성 정도와 문항들의 동의하기 어려운 정도를 한 번에 비

교해서 살펴볼 수 있도록 각 요인에 대해 이 둘을 하나의 그래프에 위치시킨 피험자 × 문항 분포도를 그림 5와 그림 6에 제시하였다. 이 그래프는 개인의 사회적 바람직성 점수와 문항에 대해 동의하기 어려운 정도를 로짓(logit) 척도로 동일하게 변환시켰기 때문에 직접 비교가 가능하다. 직접 비교를 통해 각 문항에 동의하는 정도가 본 연구 집단에 적절한지를 평가할 수 있다. 이때, 그림의 가운데 세로선을 기준으로 왼편에는 개인의 사회적 바람직성점수 분포가, 오른편에는 문항에 대해 동의하기 어려운 정도 분포가 각각 나타나 있다. 두 분포의 범위가 일치할 때, 각 문항에 대해 동의하기 어려운 정도가 개인의 사회적 바람직성 정도의 모든 범위를 측정할 수 있는 정도로 적절하게 분포되었다고 볼 수 있을 것이다(Hong et al., 2005). 그림 5와 그림 6 모두 “#”은 4명의 피험자를 의미하며, “.”은 1명의 피험자를 말한다. “M”은 피험자의 사회적 바람직성 분포와 문항에 동의하기 어려운 정도 분포의 평균 위치를, “S”는 평균으로부터 1표준편차, “T”는 2표준편차 떨어진 위치를 의미한다. 본 연구의 목적이 단축형 검사를 개발하는데 있기 때문에 문항에 대해 동의하기 어려운 정도가 중복될 경우 문항의 유용성이 떨어지므로 문항내용을 검토하여 단축형 검사에서 제외하였다(홍세희, 조용래, 2006).

우선 긍정형 사회적 바람직성 요인 문항들의 경우, 문항 8번과 문항 9번의 동의하기 어려운 정도가 중복되었기 때문에 표 3의 문항적합도 분석 결과를 기준으로 두 문항 중 상대적으로 적합도가 좋지 않은 9번 문항을 제거하였다. 최종적으로 제거된 5개 문항(3번, 7번, 9번, 11번, 15번)을 제외하고 그림 5를 살펴보았을 때, 개인의 사회적 바람직성 점수와 문항에 대해 동의하기 어려운 정도의 분포가 비슷했으나 주로 ±1 표준편차 이내의 평균적인 수준을 측정하는데 문항이 집중되

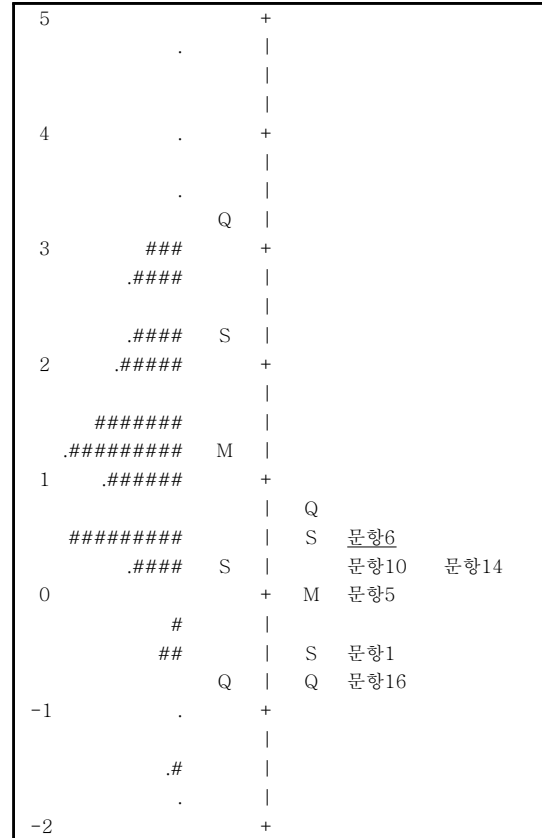


주. 밑줄로 표시한 문항은 문항적합도 분석 결과 적합도가 낮은 문항으로 선정했던 문항들임.

그림 5. 긍정형 사회적 바람직성 요인
피험자 x 문항 분포도

어 있어서 사회적 바람직성 점수가 낮은 응답자에게 적절한 수준의 문항들이 부족한 것으로 나타났다.

다음으로 부정형 사회적 바람직성 요인 문항들을 살펴보면, 문항 10번과 문항 14번의 동의하기 어려운 정도가 중복되었기 때문에 표 3의 문항적합도 분석 결과를 기준으로 상대적으로 적합도가 좋지 않은 14번 문항을 제거하였다. 최종적으로 제거된 2개 문항(6번, 14번)을 제외하고 그림 5를



주. 밑줄로 표시한 문항은 문항적합도 분석 결과 적합도가 낮은 문항으로 선정했던 문항들임.

그림 6. 부정형 사회적 바람직성 요인
피험자 x 문항 분포도

살펴보았을 때, 응답자의 사회적 바람직성의 낮은 점수들에 대해서는 문항들이 적절하게 분포하는 것처럼 보이지만, 사회적 바람직성 점수가 높은 응답자에게 적절한 문항들이 부족한 것으로 나타났다.

종합하면, Rasch 모형을 적용해서 분석한 결과 한국판 사회적 바람직성 척도 단축형(SDS-9)에는 총 9개 문항(긍정형 사회적 바람직성 요인 5개, 부정형 사회적 바람직성 요인 4개)의 문항이 선택되었다. 내적 합치도의 경우 9개 문항의 단축형

경우 .61, 하위요인별로 살펴보면 긍정형 사회적 바람직성 요인(2번, 4번, 8번, 12번, 13번)이 .72, 부정형 사회적 바람직성 요인(1번, 5번, 10번, 16번)이 .47로, 부정형 사회적 바람직성 요인의 내적 합치도가 다소 낮은 것으로 나타났다. 더불어 원래의 16문항 사회적 바람직성 척도와 본 연구에서 타당화한 9문항인 단축형 사회적 바람직성 척도의 상관계수는 매우 높은 정적 상관을 보여주었다($r = .92, p < .001$).

수렴 및 변별 타당도 검증

사회적 바람직성 척도의 타당도를 살펴보기 위하여 사회적 바람직성 척도 단축형의 전체 점수 및 2개 하위요인들에 대해 다른 사회적 바람직성 척도인 BIDR 및 신경증, 성실성을 이용해서 상관 분석 실시하였고, 그 결과를 표 4에 제시하였다. 이때, 신경증과 성실성과는 변별타당도를 BIDR 전체 점수 및 하위요인인 자기기만적 고양, 인상관리와는 수렴타당도를 확인하였다.

우선 수렴타당도를 확인하기 위한 단축형 사회적 바람직성의 전체 점수와 다른 사회적 바람직성 척도인 BIDR과 그 하위요인인 자기기만적 고양, 인상관리의 상관관계를 확인한 결과, 모

든 상관계수들이 통계적으로 유의하였으며, 전체 BIDR($r = .51, p < .001$)과 인상관리($r = .46, p < .001$)와는 중간 크기의 정적 상관관계를 보였으며, 자기기만적 고양($r = .29, p < .001$)과는 작은 크기의 정적 상관관계를 보였다. 참고로 자기기만적 고양과 인상관리의 상관계수의 차이를 분석한 결과(Steiger, 1980), 인상관리의 상관이 자기기만적 고양과의 상관보다 유의하게 큰 것으로 나타났다($Z = 2.09, p < .05$). 하위요인별로 살펴보았을 때, 부정형 사회적 바람직성의 경우 전체 BIDR($r = .40, p < .001$) 및 인상관리($r = .45, p < .001$)와는 중간 크기의 정적 상관관계를 보였으며, 나머지 자기기만적 고양 및 긍정형 사회적 바람직성은 작은 크기의 정적 상관관계를 보였다.

다음으로 변별타당도를 검증하기 위하여 단축형 사회적 바람직성의 전체 점수와 성격 특성 중 신경증, 성실성과의 상관관계를 살펴본 결과, 성격 특성 중 신경증($r = -.25, p < .001$)과는 낮은 크기의 부적 상관관계, 성실성($r = .29, p < .001$)과는 낮은 크기의 정적 상관관계를 보였다. 하위요인별로 살펴보았을 때도 2개 요인 모두 전체 점수와 동일한 방향으로 신경증 및 성실성과 낮은 크기의 상관관계를 보였다.

표 4. 사회적 바람직성 척도 단축형(SDS-9)의 수렴/변별 타당도

($N = 302$)

척도		전체 단축형 사회적 바람직성	긍정형 사회적 바람직성	부정형 사회적 바람직성
수렴타당도	전체 BIDR	.51***	.37***	.40***
	자기기만적 고양	.29***	.29***	.13*
	인상관리	.46***	.26***	.45***
변별타당도	신경증	-.25***	-.18**	-.20**
	성실성	.29***	.26***	.17**

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

논 의

기존의 다른 사회적 바람직성 척도들(예: MCSDS, BIDR)은 연구 및 임상 실제에서 다른 측정도구들과 함께 사용하기에는 문항수가 많고 지나치게 오래 전 사회문화를 반영하고 있다는 문제가 있으며 상대적으로 장점이 많은 SDS-17 척도도 이전 연구자들이 한국어판을 타당화 하는 과정에서 여러 가지 문제가 제기되어 실제로 사용하기에 적절하지 않았다. 이러한 문제를 극복하고 연구에서 응답자가 얼마나 사회적으로 용인되는 방식으로 반응하려는 경향성이 있는가를 신뢰롭고 타당하게 측정하는 도구가 필요하다는 인식에서 본 연구가 진행되었다. 구체적으로, 본 연구는 Stöber(2001)가 개발한 SDS-17에 대해 문항반응 이론의 일종인 Rasch 모형의 적용을 포함한 엄정한 도구 번안·타당화 절차를 적용하여 문항수가 적으면서도 신뢰롭고 타당한, 한국판 사회적 바람직성 척도를 구축하려는 시도였고 그 주요 결과 및 의의는 다음과 같다.

우선, 한국판 SDS-17의 특성에 대해 중요한 결과들을 확인하였는데 그 첫 번째가 바로 SDS-17이 2요인으로 구성되어 있다는 점이다. 이는 선행 연구들(김용석 등, 2008; Stöber, 2001)의 결과와는 차이가 있는데, 요인별 문항들을 살펴보니 빈번하게 일어나지만 바람직하지 않은 행동 문항들과 빈번하게 일어나지는 않지만 바람직한 행동 문항들로 이루어져 있음을 확인할 수 있었다. 이는 Stöber(2001)가 SDS-17 척도 개발 시 문항 반응의 두 가지 측면(빈번하지만 바람직하지 않은 행동, 빈번하지 않지만 바람직한 행동)을 고려한 문항을 개발한 배경과 일치하는 결과이다. 본 연구에서는 빈번하게 일어나지만 바람직하지 않은 행동을 ‘부정형 사회적 바람직성’으로, 빈번하게 일어나지만 바람직한 행동을 ‘긍정형 사회적 바람직성’

으로 명명하였다. 이러한 결과는 한국에서의 사회적 바람직성이라는 구성개념이 단일요인이 아니라는 점을 의미하며, Stöber(2001)의 SDS-17이 설문지를 실시할 때 발생할 수 있는 일시적인 방해요인이나 동기적인 요인으로 보이는 현상인 반응 세트(response set)라기 보다는 시간이나 검사지에 상관없이 어느 정도의 안정성을 보이는 개인의 특성인 반응양식(response style)으로 볼 여지가 있음을 알 수 있다(Paulhus, 2002). 그러나 수렴타당도를 통해서 살펴보았을 때 기존에 사회적 바람직성을 반응양식으로 가정했던 Paulhus(2002)의 BIDR과의 상관크기가 크지 않고 성격구인 중 두 가지 특성과 낮은 크기의 상관이 나타난 것을 통해 살펴보았을 때 SDS-17에서 측정하는 사회적 바람직성의 구성개념을 단순히 성격적인 특성으로 과잉판단하는 것도 자제할 필요가 있다. 다만 본 연구의 표본집단이 대학생으로 한정되어 있는 만큼 사회적 바람직성을 반응편파를 유발하는 가외변인 혹은 혼입효과(Edwards, 1957)으로 볼 수 있는지, 아니면 성격 특성(MacGraw & Costa, 1983; Paulus, 2002)으로 간주할 지를 판단하기 위해서는 추후 연구에서 더 다양한 연령과 인구통계학 특성을 포함한 자료를 수집하여 확인할 필요가 있다. 또한 변인들 간 연구에 사회적 바람직성 척도를 활용하는 측면과 내적 합치도의 측면을 고려할 때, 두 요인을 구분해서 각각 점수화하는 것보다는 단일 점수를 이용하는 것을 권장한다. 본 연구에서는 Rasch 모형의 기본 가정인 단일차원성이 아닌 2요인으로 구성되어 있음이 밝혀졌기 때문에 각 요인에 대해서 개별적으로 Rasch 모형을 적용하였다(홍세희, 조용래, 2006).

다음으로 본 연구결과, 한국판 SDS-17은 두 요인 모두 응답범주의 측면에서 5점 척도로 사용하는 것이 적절하다는 것이 밝혀졌다. 이는 응답범주에 대한 범주 확률곡선을 살펴 본 결과로 확인

되었는데, 각 문항의 적합도와 동의하기 어려운 정도를 함께 고려한 결과이다. 더 나아가 문항적합도 분석, 즉 각각의 문항에 응답한 피험자의 반응이 얼마나 문항반응이론 모형에 적합한지와 개별문항의 동의하기 어려운 정도를 확인하여 문항을 선별한 결과, 한국판 SDS-17의 16개 문항 중 최종 선정된 9개 문항만으로도 사회적 바람직성에 대한 타당한 측정이 이루어질 수 있음이 밝혀졌다. 이는 최종 9문항으로 구성된 한국판 SDS-9 척도의 활용도 측면에서 매우 고무적인 일이다. 수정된 SDS-9는 문항수가 원칙도에 비해 거의 절반 정도밖에 안되면서도 문항의 적절성과 타당성 측면에서 우수한 문항만을 선별한 것이기 때문에 설문조사 패키지에 포함시킬 때 긴 문항수의 압박 문제와 개념 측정의 타당성에 대한 회의를 일소할 수 있다는 뚜렷한 장점이 있다.

동일한 맥락에서 최종 SDS-9에 대해 수렴 및 변별 타당도 증거들에 확보되었다는 점 역시 특기할 만하다. 먼저, 수렴타당도의 측면에서, SDS-9는 기존에 사용되었던 사회적 바람직성 척도의 하나인 BIDR 및 그 하위요인인 자기기만적 고양, 인상관리 모두와 유의한 상관을 보였다. 이는 40개라는 많은 문항으로 사회적 바람직성을 측정해야 했던 BIDR에 비해 단지 9개의 적은 문항으로도 SDS를 사용하여도 유사한 결과가 나올 수 있다는 점을 의미하며, 이런 측면에서 SDS-9가 BIDR에 비해 시간 및 비용 측면에서 보다 효율적인 도구라고 할 수 있다. 동시에, SDS-9가 BIDR의 자기기만적 고양보다는 인상관리 요인과 보다 높은 상관을 보였다는 점도 논의될 필요가 있다. 인상관리 요인은 고의적으로 타인에게 자신을 긍정적으로 나타내려는 시도로써 상황적 요구나 일시적인 동기에 따라 다양하게 나타나므로(Paulhus, 1991), 체면을 중시하는 집단주의 문화권인 한국에서 보다 두드러질 가능성이 있다. 이런 측면에

서 보면 SDS-9에서 측정하고 있는 사회적 바람직성은 개인의 솔직한 의사를 직접적으로 드러내기 보다는 주변을 의식해 행동하거나 대답하는 경향과 보다 관련될 가능성이 있다(Lalwani, Shavitt, & Johnson, 2006).

다음으로 SDS-9의 변별 타당도를 검증하기 위하여 성격 특성 중 Stöber(2001)의 연구에서 관련성이 확인된 신경증, 성실성과의 상관관계를 살펴본 결과, 신경증과는 낮은 크기의 부적 상관관계, 성실성과는 낮은 크기의 정적 상관관계를 보였으나, 이러한 결과는 상관의 크기가 .30 이하로 낮은 상관을 보이는 것은 사회적 바람직성에 대한 성격 5요인 성격검사와의 관련성을 살펴봄으로써 사회적 바람직성이 성격구인과 관련됨을 확인한 Stöber(2001) 및 Barrick과 Mount(1996)의 연구결과와 일치한다. 이는 사회적 바람직성을 성격 구인으로 바라보는 접근들을 뒷받침하는 결과이며, 상담 장면에서 개인의 특질 중 성격과 관련되는 변인들을 연구함에 있어서 사회적 바람직성을 통제변인으로 활용해야 함을 시사한다. 다만 본 연구에서는 5요인 성격 중 신경증과 성실성과의 관련성만을 확인하였기 때문에 추후 연구에서는 5요인 성격구인 모두와의 관련성을 살펴보는 것이 사회적 바람직성의 특성에 대한 보다 구체적인 정보를 제공할 수 있을 것이다.

연구의 제한점 측면에서 살펴보면, 우선 본 연구의 결과는 수도권 지역 대학생 집단에 국한된 것이므로, 결과의 일반화에 제약이 따른다. 따라서 앞으로 지역사회 표본과 다양한 연령층에서 수집된 자료에 대해 Rasch 모형을 적용했을 때에도 이러한 결과가 도출되는지 확인할 필요가 있다. 본 연구에서 활용한 Facet 프로그램은 차별기능문항분석에 대한 결과를 제공해주지 않고 있는데, 추후 연구에서는 성별이나 연령에 따라 각 문항들이 차이가 나는지를 살펴볼 수 있다면 척도

활용에 대한 올바른 기준을 설정할 수 있을 것으로 생각된다. 둘째, 사회적 바람직성에 의한 반응 편파가 다른 변수들 간의 관계를 살펴볼 때 혼입 변수로 작용하는 정도에 대한 보다 분명한 평가를 내리고 연구의 통제변인으로 활용하기 위해서는 사회적 바람직성의 정도를 질적으로 분별할 수 있는 일종의 절단점수(cut-off score)에 대한 기준 작성이 중요하다. 본 연구는 대학생들을 표본으로 타당화 하였는데, 추후 연구에서는 척도의 활용적 측면에서 다양한 연령을 고려해서 표집을 한다면 표준화를 통한 기준 작성이 가능할 것이다. 셋째, 본 연구에서 타당화 한 SDS-9가 비록 엄정한 번역, 문항분석 및 타당화 과정을 거쳐서 확보되었지만 각 문항의 내용이 한국인의 고유한 가치와 문화를 고려한 사회적 바람직성의 개념을 포괄하는가의 여부는 확실치 않다. 왜냐하면 서구에서 개발된 문항 속에 한국인에게 정말 중요한 사회적 바람직성의 요소가 누락되었을 가능성을 완전히 배제할 수는 없기 때문이다. 끝으로, 본 연구에서 타당화한 SDS-9 척도에 포함된 문항들이 사회적 바람직성이라는 속성(attribute) 혹은 특질(trait)의 연속차원 전체를 충분히 포괄하고 있는가의 측면에 대해서는 추후 연구나 후속 논의가 필요하다. 즉, 현재로서는 포함된 문항들의 대부분이 사회적 바람직성 정도(degree)의 측면에서 중간 혹은 낮은 수준으로 구성되어 있어 높은 수준을 포괄하는 문항을 추가하고 이에 따른 후속 타당화 과정을 실시할 필요가 있다. 그럼에도 불구하고 체계적이고 엄정한 타당화 과정을 거쳐 9문항으로 구성된 한국판 사회적 바람직성 척도 단축형(SDS-9)은 앞으로 상담 분야에서 다양한 심리적 변수들 간의 관계를 연구할 때, 사회적으로 용인된 방식으로 반응하는 경향성이 있는 응답자에 의한 반응편파의 가능성이 결과에 유의한 방식으로 영향을 미치는지에 대해 평가하고, 만일 그렇

다면 통제변인으로 활용함으로써 신뢰로운 결과를 도출할 수 있는 효율적인 도구로 자리매김 할 것으로 기대한다.

참고문헌

- 구본용, 유제민 (2005). 성격 및 환경요인과 행복과의 관계. *상담학연구*, 6(1), 1-9.
- 김권현, 유동주, 김형준, 김청택 (2015). 설문지의 길이가 응답의 질에 미치는 영향. *조사연구*, 16(1), 1-48.
- 김병직, 이동귀, 김지근, 김윤상, 황정은 (2013). 자기용서 척도의 개발 및 타당화. *상담학연구*, 14(2), 973-993.
- 김용석 (2010). 사회적 바람직성 척도의 개발. *한국사회복지행정학*, 12(3), 1-39.
- 김용석, 전수미, 이평순 (2003). 한국어판 사회적 바람직성척도(SDS-17)의 신뢰도와 타당도 평가-가정폭력 가해 남성을 대상으로-. *사회복지연구*, 38, 5-32.
- 김혜진, 김옥분 (2009). Rasch 측정모형의 적용을 통한 대학생용 단축형 완벽주의 척도 구성. *교육평가연구*, 22(3), 757-781.
- 박중규 (2009). 청소년 대상 단축형 13문항 Marlowe-Crowne 사회적 바람직성 척도의 요인 구조. *한국심리학회지: 학교*, 6(3), 429-439.
- 손은영, 차정은, 김아영 (2006). 사회적 바람직성 상, 하위 집단 간 성격검사의 구인동등성 검증. *한국심리학회지: 사회 및 성격*, 21(2), 71-87.
- 어윤경 (2007). 심리검사의 운영모드와 설명기입 여부에 따른 사회적 바람직성 변인 탐색. *교육평가연구*, 20(4), 235-258.
- 유태용, 이기범, Ashton, M. (2004). 한국판 HEXACO 성격검사의 구성 타당화 연구. *한국*

- 심리학회지: 사회 및 성격, 18(3), 61-75.
- 이익섭, 홍세희, 신은경 (2007). 장애 정체감 척도의 Rasch 모형 적용. *한국사회복지학*, 59(4), 273-296.
- 이주일 (1988). 이분척도와 Likert형 7점 척도의 신뢰도 타당도 반응양식의 비교: 자기검색척도와 자기의식척도를 대상으로. 석사학위논문, 서울대학교.
- 정보라 (2004). 사회적 바람직성 구인탐색. 석사학위논문, 이화여자대학교.
- 최보라 (2007). 사회적 바람직성 척도 타당화. 석사학위논문, 이화여자대학교.
- 최혜윤, 이동혁 (2015). 한국판 상담자 스트레스 척도의 타당화. *상담학연구*, 14(5), 3217-3236.
- 한수경, 정남운 (2014). 대인간 외상경험이 대인관계문제에 미치는 영향에서 자기체계손상의 매개효과. *인간이해*, 35(1), 149-168.
- 홍세희, 조용래 (2006). 역기능적 신념검사 단축판의 구성: Rasch 평정척도모형의 적용. *한국심리학회지: 임상*, 25(3), 865-880.
- Anastasi, A. (1976). *Psychological testing* (4th ed.). NY: Macmillan.
- Andrich, D. (1978). Rating formulation for ordered response categories. *Psychometrika*, 43, 561-573.
- Barrick, M. B., & Mount, M. K. (1996). Effects of impression management and self-deception on the predictive validity of personality constructs. *Journal of Applied Psychology*, 81(3), 261-272.
- Bell, K. M., & Naugle, A. E. (2007). Effects of social desirability on students self-reporting of partner abuse perpetration and victimization. *Violence and Victims*, 22(2), 243-256.
- Blake, B. F., Valdiserri, J., Neuendorf, K. A., & Nemeth, J. (2006). Validity of the SDS-17 measure of social desirability in the American context. *Personality and Individual Differences*, 40, 1625-1636.
- Bond, T. G., & Fox, C. M. (2001). *Applying the Rasch model: Fundamental measurement in the human sciences*. Mahwah, NJ: Lawrence Erlbaum.
- Christina, H. P., & Alex, C. (2004). The student readiness inventory and the big five: Examining social desirability and college academic performance. *Personality and Individual Differences*, 41, 663-673.
- Comrey, A. L., & Lee, H. B. (1992). *A first course in factor analysis* (2nd ed.). Hillsdale, NJ: Lawrence Erlbaum.
- Connelly, B. S., & Chang, L. (2015). A meta-analytic multitrait multitrait separation of substance and style in social desirability scales. *Journal of Personality. Advance on line publication*. <http://dx.doi.org/10.1111/jopy.12161>.
- Cronbach, J. J. (1946). Response sets and test validity. *Educational and Psychological Bulletin*, 52, 281-302.
- Crowne, D. P., & Marlowe, D. (1960). A new scale of social desirability independent of psychopathology. *Journal of Consulting Psychology*, 24, 349-354.
- Deniz, S. O., Chockalingam, V., & Reiss, A. D. (1998). The effects of social desirability and faking on personality and integrity assessment for personnel selection. *Journal of Applied Psychology*, 81, 660-679.
- Dodou, D., & De Winter, J. C. F. (2014). Social desirability is the same in offline, online, and paper surveys: A meta-analysis. *Computers in Human Behavior*, 36, 487-495.
- Dudley, N. M., McFarland, L. A., Goodman, S. A., Hunt, S. T., & Sydel, E. J. (2005). Racial

- differences in socially desirable responding in selection contexts: Magnitude and consequences. *Journal of Personality Assessment*, 85(1), 50-64.
- Edwards, A. L. (1957). *The social desirability variable in personality assessment and research*. NY: The Dryden Press.
- Fisher, R. (1993). Social desirability bias and the validity of indirect questioning. *Journal of Consumer Research*, 20, 303-315.
- Ganster, D. C., Hennessey, H. W., & Luthans, F. (1983). Social desirability response effects: Three alternative models. *Academy of Management Journal*, 26(2), 321-331.
- Goldberg, L. R. (1999). A broad-bandwidth, public-domain, personality inventory measuring the lower-level facets of several five-factor models. In I. Mervielde, I. Deary, F. De Fruyt & F. Ostendorf (Eds.), *Personality psychology in Europe, Vol. 7* (pp. 7-28). Tilburg University Press: The Netherlands.
- Griffith, R. (1997). *Faking of non-cognitive selection devices: Red herring is hard to swallow*. Unpublished doctoral dissertation, University of Akron, Akron, OH.
- Hong, S., Kim, B., & Wolfe, M. (2005). A psychometric revision of the European American Values Scale for Asian Americans using the Rasch Model. *Measurement and Evaluation in Counseling and Development*, 37, 194-207.
- Johnson, T. P. (1998). *Empirical evidence of an association between individualism/collectivism and the trait social desirability*. Paper presented at the Third ZUMA Symposium on Cross-Cultural Survey Methodology, Leinsweiler, Germany.
- Karabatsos, G. (1997). The sexual experiences survey: Interpretation and validity. *Journal of Outcome Measurement*, 1, 305-328.
- Keillor, B., Owens, D., & Pettijohn, C. (2001). A cross-cultural/cross-national study of influencing factors and socially desirable response biases. *International Journal of Market Research*, 43(1), 63-84.
- Krumpal, I. (2013). Determinants of social desirability bias in sensitive surveys: A literature review. *Quality & Quantity*, 47(4), 2025-2047.
- Lalwani, A. K., Shavitt, S., & Johnson, T. (2006). What is the relation between cultural orientation and socially desirable responding? *Journal of Personality and Social Psychology*, 90(1), 165-178.
- Linacre, J. M. (2014). *Facets Rasch measurement* (Version 3.71.4) [Computer software]. Chicago, IL: Winsteps.com.
- McCrae, R. R., & Costa, P. T. (1983). Social desirability scales: More substance than style. *Journal of Consulting and Clinical Psychology*, 51, 882-888.
- McNamara, T. F. (1996). *Measuring second language performance*. London: Addison Wesley Longman.
- Moorman, R. H., & Podsakoff, P. M. (1992). A meta-analytic review and empirical test of the potential confounding effects of social desirability response sets in organizational behaviour research. *Journal of Occupational and Organizational Psychology*, 65, 131-149.
- Ones, D. S., & Viswesvaran, C. (1998). The effects of social desirability and faking on personality and integrity assessment for personnel selection. *Human Performance*, 11, 245-269.
- Paulhus, D. L. (1984). Two component models of socially desirable responding. *Journal of Personality and Social Psychology*, 46, 589-609.
- Paulhus, D. L. (1991). Measurement and control of

- response bias. In J. P. Robinson, P. R. Shaver, & L. S. Wrightsman (Eds.), *Measures of personality and social psychological attitudes* (pp. 1-59). San Diego, CA: Academic Press.
- Paulhus, D. L. (1998). *Manual for the Balanced Inventory of Desirable Responding (BIDR-7)*. Toronto/Buffalo: Multi-Health systems.
- Paulhus, D. L. (2002). Socially desirable responding: The evolution of a construct. In H. Braun, D. N. Jackson, & D. E. Wiley (Eds.), *The role of constructs in psychological and educational measurement* (pp. 67-88). Hillsdale, NJ: Erlbaum.
- Podsakoff, P. M., MacKenzie, S. B., Lee, J. Y., & Podsakoff, N. P. (2003). Common method biases in behavioral research: A critical review of the literature and recommended remedies. *Journal of Applied Psychology*, 88(5), 879-903.
- Rasch, G. (1960). *Probabilistic models for some intelligence and attainment tests*. Copenhagen: Danish Institute for Educational Research.
- Reynolds, W. M. (1982). Development of reliable and valid short forms of the Marlowe-Crowne Social Desirability Scale. *Journal of Clinical Psychology*, 38, 119-125.
- Richman, W. L., Kiesler, S., Wiesband, S., & Drasgow, F. (1999). A meta-analytic study of social desirability distortion in computer-administered questionnaires, traditional questionnaires, and interviews. *Journal of Applied Psychology*, 84(5), 754-775.
- Rubin, A., & Babbie, E. (2008). *Research methods for social work* (6th ed.). Belmont, CA: Wadsworth.
- Schmit, M. J., & Ryan, A. M. (1993). The Big Five in personnel selection: Factor structure in applicant and nonapplicant population. *Journal of Applied Psychology*, 78, 966-974.
- Sheperis, C. J., Young J. S., & Daniels, M. H. (2013). 상담연구방법론: 양적·질적·혼합적 방법론 [*Counseling research: Quantitative, qualitative, and mixed methods*]. (최한나, 김은하, 김형수 공역). 서울: 학지사(원전은 2010년에 출판).
- Stanton, A. L., Danoff-Burg, S., Cameron, C. L., & Ellis, A. P. (1994). Coping through emotional approach: Problems of conceptualization and confounding. *Journal of Personality and Social Psychology*, 66(2), 350-362.
- Steiger (1980). Tests for comparing elements of a correlation matrix. *Psychological Bulletin*, 87(2), 245-251.
- Stöber, J. (2001). The Social Desirability Scale-17 (SDS-17): Convergent validity, discriminant validity, and relationship with age. *European Journal of Psychological Assessment*, 17(3), 222-232.
- Stöber, J., Dette, D. E., & Musch, J. (2002). Comparing continuous and dichotomous scoring of the balanced inventory of desirable responding. *Journal of Personality Assessment*, 78(2), 370-389.
- Strahan, R., & Gerbasi, K. C. (1972). Short, homogeneous versions of the Marlowe-Crowne Social Desirability Scale. *Journal of Clinical Psychology*, 28, 191-193.
- Wright, B. D., & Masters, G. N. (1982). *Rating scale analysis*. Chicago, IL: Measurement, Evaluation, Statistics, and Assessment Press.
- Zerbe, W. J., & Paulhus, D. L. (1987). Socially desirable responding in organizational behaviour: A reconception. *Academy of Management Review*, 12, 250-264.
- 원고 접수일 : 2015. 10. 15.
수정원고 접수일 : 2015. 12. 08.
게재 결정일 : 2015. 12. 11.

Validation of the Korean Short-Version of Social Desirability Scale (SDS-9) Using the Rasch Model

Bae, Byeong-Hun

Lee, Dong-Gwi

Ham, Kyong-Ae

Yonsei University

The Social Desirability Scale-17 (SDS-17; unidimensional 16 items), developed by Stöber (2001), has distinct strengths including: (a) its convenience in survey administration thanks to its relatively small number of items, (b) appropriateness in its contents that presumably reflect the societal reality, and (c) its sound psychometric properties (i.e., reliability and validity estimates) with adult samples over the age of 18. This study aims to validate the SDS-17 with the sample of 302 Korean college students through systematic and robust processes to assess the appropriateness and difficulty level of each item using the Rasch model. As a result, nine final items with two factors were selected, and then convergent and discriminant validate estimates were calculated to assess the validity of the revised version (SDS-9). This study ends with discussion of the utility of the scale in Korean context as well as its limitations.

Key words : Social Desirability Scale, Validation, Rasch Model