

부정문항이 포함된 척도의 요인구조 및 방법효과 검증과 남녀 간의 차이 비교: Rosenberg 자기존중감 척도를 중심으로

최 수 미
부산대학교

조 영 일[†]
성신여자대학교

하나의 척도에 긍정 및 부정문항을 동시에 포함하고 있는 검사들의 경우 신뢰도가 과소 추정되거나 단일요인이 아닌 서로 다른 요인으로 밝혀지는 등 문제점들이 제기되었다. 본 연구는 긍정 및 부정문항을 동시에 포함하는 Rosenberg 자기존중감 척도의 요인구조와 관련해서 제기되는 문제점들을 해결하기 위해 본 척도를 가장 잘 설명하는 요인모형을 찾고자 하였다. 또한, 자기존중감 척도의 성차간 요인구조의 차이검증과 방법효과를 고려한 후 성차간 자기존중감의 요인평 균을 비교하였다. 본 연구는 한국 아동·청소년패널조사의 1차년도 조사자료 중에서 중학교 1학년 2,351명(남학생 1,177명, 여학생 1,174명)의 응답을 사용하였다. 연구결과, 자기존중감 척도의 문항 간의 공분산을 설명하는 여섯 개의 요인모형 중, 일반요인과 방법요인을 고려한 2요인 모형이 가장 적합한 것으로 나타났다. 둘째, 성차간 측정모형의 동일성 검증을 실시한 결과, 측정모형의 동일성 가정이 지지되었다. 셋째, 자기존중감척도의 부정문항에 대한 방법효과를 고려한 2요인구조모형의 남녀 간에 부정문항에 대한 반응 수준에 차이가 있음을 확인하였다. 끝으로, 본 연구의 제한점과 의의를 제시하였다.

주요어 : Rosenberg 자기존중감척도, 방법효과, 요인구조모형, 측정모형의 동일성, 성차

[†] 교신저자: 조영일, 성신여자대학교 심리학과, (136-742) 서울시 성북구 보문로 34다길 2
Tel: 02-920-7593, Fax: 02-920-2040, E-mail: yicho@sungshin.ac.kr

사회과학 분야의 대부분의 조사연구들은 연구변인을 측정하기 위해 자기보고식 검사를 제작하여 사용해왔다. 하지만, 자기보고식 검사들에서는 문항에 응답하는 사람들의 경향성 혹은 방법 효과들로 인해 체계적이고(systematic) 관련성이 없는(irrelevant) 편파들(bias)이 발생된다(American Educational Research Association, 1999; Quilty, Oakman, & Risko, 2006).

현재까지 많은 연구자들은 심리검사 문항의 내용을 평가하는 것과 더불어서 반응과정에서의 정확한 측정과 내적 타당도를 저해하는 요인들을 적절하게 처리하기 위한 노력들에 관심을 가져왔다(Bagozzi, 1993). 예를 들어, 응답자들의 반응경향성에 의한 왜곡을 막기 위한 시도로써 연구자들은 긍정문항과 부정문항을 넣어 질문지를 구성한다. 이에 덧붙여서, 실제 몇몇 연구자들은 논리적으로 반대되는 문항을 함께 넣어 개발하는 것을 필수적인 것으로 주장한다(e.g., Paulhus, 1991; Winkler, Kanouse, & Ware, 1982). 다른 형식으로 진술된 문항이라 할지라도 같은 개념을 측정한다는 가정을 바탕으로(홍세희, 노연경, 정송, 2011), 이러한 방법들은 응답자의 묵종(acquiescence), 동의편파(agreement bias) 등을 감소시킬 수 있다(Quilty et al., 2006). 하지만 긍정적이거나 부정적인 진술형태의 문항들이 함께 포함된 심리적 평정 척도의 사용은 몇 가지 문제점들을 불러일으킨다. 예를 들어, 척도의 신뢰도가 종종 감소되고(Schrisheim & Hill, 1981), 척도의 요인 구조가 긍정문항과 부정문항 각각을 반영하는 서로 다른 요인으로 밝혀지는 등(Watson & Clark, 1991, 1992) 체계적인 편파(systematic bias)가 유발될 수 있다는 주장이 제기되었다(Horan, DiStefano, & Motl, 2003).

긍정 혹은 부정문항의 사용으로 인하여 유발되는 방법효과들(method effects)을 단순히 오차(noise)를 반영하는 것으로 간주될 수 있지만, 몇몇 연구자들은 요인구조를 해석함에 있어서 방법효과들을 고려해야 한다고 주장하였다(Motl & DiStefano, 2002; Tomas & Oliver, 1999). 보다 구체적으로, 긍정적 혹은 부정적으로 기술된 문항들로 인한 방법효과들은 응답자 반응을 왜곡할 수도 있고, 더 나아가 잘못된 해석을 유발할 수 있을 수 있다고 한다. 따라서 이러한 형태의 자기보고식 검사도구의 경우, 긍정과 부정의 두 가지 진술형태를 가진 문항들이 측정하고자 하는 구성개념을 측정하는데 실질적으로 차이가 있는지 혹은 단순히 진술양식만을 구분하는 방법효과를 구성하는 것인지에 관한 보다 정확한 진단이 필요하다.

이처럼, 동일 개념을 측정하기 위해 긍정문항과 부정문항의 2가지 진술형태를 가진 자기보고식 검사도구들 중에서 교육학이나 심리학 등 사회과학 분야에서 가장 널리 사용되고 있는 도구중 하나로 Rosenberg가 개발하고 타당화한 자기존중감 척도를 들 수 있다(Marsh, 1996). Rosenberg의 자기존중감 척도는 5개의 긍정문항과 5개의 부정문항으로 총 10개 문항으로 구성되었고, 단일 차원의 전반적인 자기존중감(general self esteem)을 측정하는 척도로 개발되었다(Tomas & Oliver, 1999). 하지만, Rosenberg의 자기존중감 척도의 요인 구조에 대해서는 일관되지 않은 연구 결과들이 보고되고 있다. 보다 구체적으로, 몇몇 연구자들은 Rosenberg의 자기존중감 척도가 타당하고 신뢰할만한 단일차원의 척도라고 보고하고 있지만(이자영, 남숙경, 이미경, 이지희, 이상민, 2009; Crandall, 1973; Silbert & Trippett, 1965), 다른

연구들은 Rosenberg의 척도가 긍정적인 자존감과 부정적인 자존감인 2개의 요인으로 구성된 요인구조를 가진다고 보고했다(이미리, 2005; Marsh, 1986).

심지어, Rosenberg(1979)조차도 자기존중감은 단일차원으로 구성된 개념이라고 했던 종전의 입장(Rosenberg, 1965)을 바꿨다. 척도가 두 개의 요인으로 구성된다고 하였다. 2개의 긍정적 및 부정적 자기평가가 동일한 성격차원을 측정하는 것이라고 한다할지라도(Rosenberg, 1979), 이것들은 개인의 전반적인 자기존중감을 측정하기 위해 개발된 단일 차원의 자기존중감 척도의 개발의도와는 일치되지 않는다. 게다가, 많은 연구들이 Rosenberg 척도 문항에 대한 반응들이 2요인 구조를 나타낸다는 분석 결과들을 보고해오고는 있지만, 2개의 요인들이 실질적으로 의미 있는 구성개념을 나타내는지 혹은 실질적으로 관계없는 방법효과로 인하여 나타내는지에 관한 논쟁들이 지속되어 오고 있다(Marsh, 1996).

이처럼, 국외에서는 Rosenberg 자기존중감 척도의 요인구조에 관한 검증 노력들이 활발하게 진행되어왔다. 이에 반해서, 국내에서는 Rosenberg 자기존중감 척도가 충분한 타당화 작업과 요인구조의 검증 없이 사용되고 있다(최보가, 전귀연, 1993). 하지만, 최근 국내에서도 이미리(2005)의 연구를 통해 Rosenberg 자기존중감 척도의 요인구조에 대한 의문이 제기되었다. 저자는 인간의 부정적인 정서와 긍정적인 정서가 서로 독립적이라는 선행연구들(Larsen & Ketelaar, 1991; Watson, 1988; 이미리, 2005 재인용)을 언급하며, 자기존중감 척도를 구성하는 두 요인인 긍정적 자아평가와 부정적 자아평가를 서로 독립적 차원이라 주장하였다. 이러한 근거를 바탕으로 저자는 탐색적

요인분석 결과를 통해 나타난 2개의 요인, 즉 긍정적 자아평가와 부정적 자아평가로 자기존중감을 구성하여 연구를 수행하였다.

또한, 탐색적 요인분석을 사용한 정병삼(2010)의 연구에서도 척도가 긍정적 자아평가와 부정적 자아평가라는 2개의 하위요인으로 구분되었다. 하지만, 이들의 연구들은 계속해서 제기되고 있는 자기존중감 척도의 부정문항으로 인한 방법효과에 대한 고려없이 척도를 2요인 구조로 하여 연구를 진행하였으며, 이는 연구의 제한점이라 할 수 있다. 또한, 탐색적 요인분석이 완전한 요인구조들을 구별할 수 없다는 점을 고려해봤을 때(Quilty, Oakman, & Risko, 2006), 탐색적 요인분석을 실시한 결과를 통해 단일차원의 자아존중감 척도를 2요인 구조로 하여 연구를 진행한 것 역시 연구의 제한점으로 볼 수 있다.

이미리(2005)의 연구와 정병삼(2010)의 연구에 의해 제기된 자기존중감 척도의 요인구조에 대한 검증을 위한 연구가 아직 초기단계이지만 국내 몇몇 연구자들에 의해서도 이를 밝히기 위한 노력이 시도되고 있다. 최근 홍세희, 노연경, 정송(2011)은 확인적 요인분석을 통한 긍정문항과 부정문항으로 구성된 자기존중감 척도의 다양한 요인구조들을 평가하고 비교 검증하는 연구를 수행하였다. 그 결과 2개의 별도의 요인들을 설정한 요인모형보다는 부정적인 자아평가 문항들의 방법효과를 통제 한 상태에서의 단일요인구조모형이 적합한 것으로 보고되었다.

이상에서 살펴본 바와 같이, 국외처럼 활발하지는 않다할지라도 국내에서도 Rosenberg 자기존중감 척도의 요인구조에 관한 검증 노력들이 점차적으로 이루어지기 시작하였다. 그러나 Rosenberg 자기존중감 척도의 요인구조에

관한 국내 연구들(이미리, 2006; 정병삼, 2010; 홍세희, 노연경, 정송, 2011)은 모두 한국청소년패널조사 데이터의 6개 문항으로 구성된 축약된 척도를 사용하였다는 제한점을 지닌다. 이처럼 축약된 버전은 요인에 대한 지표변수들의 비율 뿐 아니라 종종 연구에 사용되었던 원척도(full scale)의 구성타당도에 비하여 타당도가 감소할 수 있다(Quilty, Oakman, & Risko, 2006). 따라서 본 연구에서는 Rosenberg 자기존중감의 원척도를 사용하여 다양한 요인구조를 가정한 모형들을 확인적 요인분석으로 검증하고 비교하여 원척도의 요인구조를 가장 잘 설명하는 모형을 찾고자 한다.

한편, Rosenberg 자기존중감 척도는 자기존중감에 관한 성차연구에 많이 사용되어왔다. 예를 들어, Kling과 그의 동료들은 1987년과 1995년 사이에 발간된 218편의 자기존중감 논문들을 통해 자기존중감의 성차에 관한 메타분석을 실시하였다. 그 결과, 통계적으로 유의한 자기존중감의 성차를 확인하였다($d=.22$; Kling et al., 1999). 하지만, DiStefano와 Motl (2009)은 남녀 간의 자기존중감의 차이는 부정문항들과 관련된 불균형적인 반응들에 의해 유발될 수 있다고 주장하였다. 다시 말해서, 어느 한쪽의 성별이 부정문항과 관련하여 더 강한 혹은 약한 반응을 함에 따라서 자기존중감의 평균에서 차이를 일으킬 수 있다고 설명하면서 부정문항들의 방법효과로 인한 집단간 차이에 대한 검증도 이루어져야한다고 하였다. 이에 본 연구의 두 번째 연구목적은 자기존중감 척도의 성차간 요인구조의 차이 검증과 방법효과를 고려한 성차간 자기존중감의 요인평균을 비교하는 것이다.

이상의 논의를 바탕으로 한 본 연구의 연구문제를 정리하면 다음과 같다. 첫째, Rosenberg

자기존중감 척도에 적합한 요인구조는 무엇인지에 관해 검증해보고자 한다. 먼저, 기존의 연구들에서 논의된 것과 같이 Rosenberg의 자기존중감 검사가 단일차원 척도인지 아니면 2개의 요인으로 구성된 척도인지를 검증하였다. 이들은 그림 1의 모형 1과 모형 2로 표시되었다. 둘째, 자기존중감 척도의 요인구조와 관련하여 비일관된 결과들이 보고되는 원인을 검증하고자 하였다. 이러한 목표를 달성하기 위하여, 긍정적인 혹은 부정적인 형식을 가진 문항의 사용으로 인하여 발생하는 방법요인을 모형화 하였다. 보다 구체적으로 그림 1의 모형 3과 5는 긍정적인 형식을 가진 문항들의 고유분산 간의 공통성을 모형화 하였고 모형 4와 6은 부정적인 형식을 가진 문항들의 고유분산 간의 공분산을 모형화 하였다. 마지막으로, 척도 내에서 공통된 문항 형태를 사용함으로써 인하여 발생한 방법효과를 모형화 하는 방법을 검증하기 위하여 문항형태를 공유하는 문항들 간의 공분산을 허용하는 모형(모형 3과 4)과 문항들 간의 공분산을 방법요인으로 설명하는 모형(모형 5와 6)의 합치도를 비교하였다. 이와 같은 연구문제를 검증하기 위해 본 연구는 Rosenberg 자기존중감 척도의 문항간 공분산을 설명하는 6개의 요인구조 모형을 설정하였다(그림 1 참조).

둘째, 그림 1에서 제시한 6개의 연구모형 중 가장 좋은 적합도를 보이는 모형이 남녀 집단 간에 동일한 구성타당도를 보이는지에 관해 검증해보고자 한다.

셋째, 방법효과를 고려한 후에 자기존중감이 요인 수준에서 남녀 집단 간 차이가 있는지를 검증해보고자 한다.

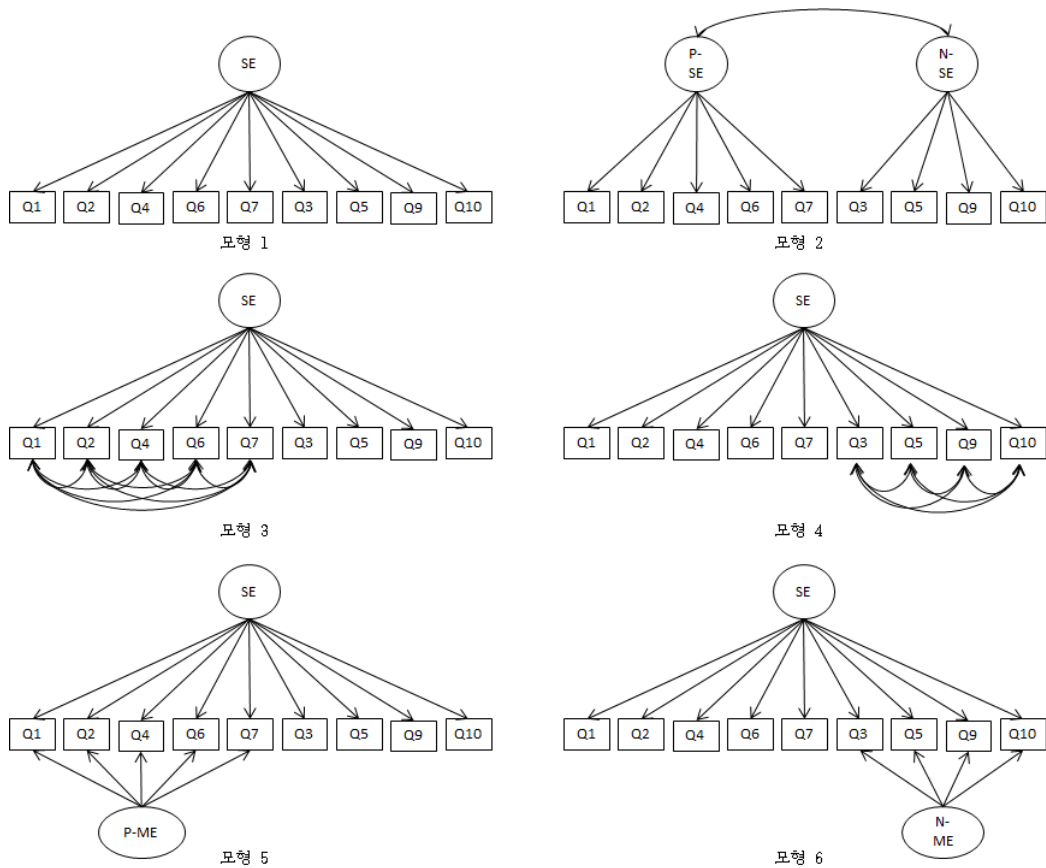


그림 1. 연구모형

연구방법

연구대상

본 연구는 한국청소년정책연구원이 2010년부터 아동과 청소년을 대상으로 실시하고 있는 한국 아동·청소년 패널조사의 1차년도 중 1 패널 자료를 사용하였다. 연구의 참가자들은 2009년도에 중학교 1학년이었던 학생이다. 한국 아동·청소년 패널조사는 층화다단계집락표집(stratified multi-stage cluster sampling)방법을 활용하였다. 전국을 16개의 특별시, 광역시,

및 도 별로 나누고 이를 다시 지역적 근접성을 기준으로 총 27개의 집락으로 나누었다. 각 집락의 전체 중학교 1학년 학생 수에 비례하여 전체 목표 표본수(2,200명)를 집락별로 표본수를 할당하였다. 이 후에, 각 집락의 학교 수(전체 78개 학교가 자료 수집에 동의함)와 학교를 결정하고, 선택된 중학교 1학년의 학급 중 하나를 무선적으로 선택하였다. 무선 선택 후, 조사원이 해당 학급을 방문하여 학생들을 대상으로 집단면접방법으로 조사하였다. 본 연구에는 총 2,350명(남학생 1,176명, 여학생 1,174명)의 중학교 1학년 학생이 조사에

포함되었다.

측정도구

자기존중감 척도(Self-Esteem Scale)

패널조사에서 사용된 자기존중감 척도는 국내에서 번역되어 사용 중인 Rosenberg의 자아 존중감 척도(Rosenberg, 1965, 1989)를 사용하였다. 번역된 Rosenberg의 자아 존중감 척도는 다른 연구들에서 자기존중감을 측정하기 위해 사용되었다(이미리, 2006; 정병삼, 2010). 패널 조사에서 사용된 자기존중감 척도는 자기 자신에 대한 태도를 측정하는 자기보고식 검사로서 긍정적인 모습 5개 문항과 부정적인 모습 5개 문항으로 총 10문항으로 구성되었다. 하지만, 본 연구에서는 번역상의 오류가 있는 것으로 주장된 8번 문항을 제외한 총 9문항을 분석에 포함하였다. 최근 이자영 등(2009)은 자기존중감 척도의 8번 문항(“나는 나를 좀 더 존중할 수 있었으면 좋겠다.”)이 미국을 비롯한 서양과 동양의 문화 차이를 반영하지 못하고 있다고 주장하였고, 특히 이들 연구에서 8번 문항이 다른 9개 문항들과 부적인 상관을 보임으로써 번역된 척도의 신뢰도를 낮춤과 동시에 자기존중감을 적절하게 측정하지 못하고 있다고 보고하였다. 이후, 이자영 등(2009)은 사회문화적인 측면을 고려한 문항수정을 제안하였고, 세 가지 형태의 수정문항들을 비교하였다. 하지만, 아동·청소년 패널조사에서는 수정된 문항을 사용하지 않고, 원래의 문항을 그대로 사용하였기에 본 연구에서는 8번 문항을 제외한 총 9개 문항(긍정 5개 문항, 부정 4개 문항)을 사용하였다.

각 문항은 4점 Likert 척도로 ‘매우 그렇다’

1점에서부터 ‘전혀 그렇지 않다’ 4점으로 응답하도록 하였다. 긍정적인 자아 평가 문항을 역코딩하였고, 점수가 높을수록 긍정적인 자기관을 나타낸다고 보았다. 본 연구에서 사용된 자기존중감 척도의 신뢰도는 .873이었고, 성별로는 남학생 집단 .848, 여학생 집단 .883으로 Cronbach's α 값이 대체로 안정적이고 양호한 것으로 나타났다.

분석방법

본 연구에서는 Rosenberg 자기존중감 척도의 요인 구조를 파악하기 위해서 M-plus 6.0 (Muthén & Muthén, 2010)을 활용하여 확인적 요인분석(confirmatory factor analysis)이 수행되었다. 이 분석에서는 완전정보최대우도법(Full Information Maximum Likelihood Estimation Method; FIML)이 모수를 추정하기 위하여 사용되었다. 확인적 요인분석에서는 다양한 모형합치도 지수(fit index)를 활용하여 모형을 평가한다. χ^2 검증 방법은 모형과 자료의 완벽한 일치도를 평가하기 때문에 너무 엄격하고 표본크기에 민감하게 반응하기 때문에(Browne & Cudeck, 1993), 여러 대안적인 합치도 지수들이 개발되었고, 또한 제안되어 왔다.

본 연구에서는, 사례 수에 덜 민감한 것으로 알려진 RMSEA(Root Mean Square Error of Approximation)와 TLI(Tucker-Lewis Index)를 사용하였다. 이와 별도로, CFI(Comparative Fit Index)도 모형의 합치도를 평가하기 위해서 사용되었다. RMSEA는 그 값이 적을수록 모형이 자료를 잘 설명하고 있음을 의미한다. RMSEA의 값을 0.5이하로 가지면 매우 좋은 모형이고, 0.8이하의 값을 가지면 괜찮은 모형이고, 0.10

이상의 값을 가지면 모형 합치도가 좋지 않다고 이야기 한다(Browne & Cudeck, 1993). 이에 반해서, TLI와 CFI는 그 값이 클수록 모형이 자료를 잘 설명하고 있음을 가리킨다. Hu와 Bentler(1999)는 시뮬레이션 분석을 활용하여 TLI와 CFI를 0.95이상의 값을 가진 모형은 좋은 합치도를 보이는 것으로 해석될 수 있다고 하였다.

전체 중학교 1학년 집단을 대상으로 Rosenberg 자기존중감 척도의 9개의 문항들 간의 상관을 가장 잘 설명하는 요인구조를 확인한 후, 해당하는 모형이 남녀학생 간에 차이가 있는지 검증하였다. 두 집단 간의 요인모형의 동일성을 검증하는 구인동등성 검증을 실시하기 위해서, 다집단 확인적 요인분석을 실시하였다. 보다 구체적으로, 형태 동일성, 측정단위 동일성, 측정원점 동일성의 가정을 순서대로 검증하였다. 이러한 가정들을 포함한 모형들은 내재된 관계이기 때문에, χ^2 차이검증을 실시하였다. 하지만 χ^2 차이검증 역시 표본의 크기에 민감하기 때문에(Anderson & Gerbing, 1988; Sreenkamp & Baumgartner, 1998), 다른 합치도지수(예, TLI, CFI, RMSEA, McDonald mc, 그리고 Gamma hat)의 크기 및 변화 또한 고려하였다(Cheung & Rensvold, 2002).

다집단 확인적 요인분석으로 남녀 중학생 집단 간의 구인동등성을 검증한 후에, 자기존중감이 남녀중학생 집단 간의 평균차이가 있는지를 검증하였다. 측정오차를 통제하면서 남녀 중학생 집단 간의 평균차이를 검증하기 위해서 잠재평균분석을 수행하였다.

연구결과

경쟁모형의 비교

본 연구에서는 Rosenberg의 자기존중감 척도의 문항간 공분산을 설명하는 6개의 모형을 그림 1과 같이 설정하였다. 모형 1은 전반적인 자기존중감 요인을 가정하는 1요인 모형이다. 모형 2는 긍정적인 자기존중감과 부정적인 자기존중감을 요인으로 가정하는 2요인 모형이다. 모형 3과 4에서는 전반적인 자기존중감의 1요인 모형을 가정하는 동시에 문항 형식의 유사성을 설명하기 위해서 긍정적 혹은 부정적 자기를 묻는 문항들의 오차 간의 공분산을 허용한 모형이다. 모형 5와 6은 전반적인 자기존중감의 1요인과 긍정적 혹은 부정적 자기를 묻는 문항들의 공분산을 설명하는 2차적 요인으로서 방법효과를 가정했다. 그림 1에서 전반적인 자기존중감은 SE로, 긍정적 자기존중감은 P-SE로, 부정적 자기존중감은 N-SE로, 긍정적 자기를 묻는 문항들의 방법효과는 P-ME로, 부정적 자기를 묻는 문항들의 방법효과는 N-ME로 표시하였다. 그림의 간명성을 위해서 오차항은 표시하지 않았다.

6개의 모형들에 대한 모형 합치도 지수를 표 1에 보고하였다. 6개의 모형들이 서로 내재된(nested) 관계가 아니기 때문에 6개의 모형들 간의 비교에서는 χ^2 차이검증을 실시할 수 없다. 하지만, 모형 1과 2는 서로 내재된 관계가 있기 때문에, 1요인 모형과 2요인 모형을 χ^2 차이검증을 활용하여 모형비교를 수행할 수 있다. 비슷한 논리로, 모형 1, 3과 5은 서로 내재된 관계이기 때문에 모형들 간의 비교에서 χ^2 차이검증을 실시하였다. 마찬가지로, 모형 1, 4와 6은 서로 내재된 관계이기 때문에 모형들 간의 비교에서 χ^2 차이검증을 실시하였다. 덧붙여서, 각 모형들의 합치도 지수의

표 1. 모형의 합치도 지수

| 모형 | $\chi^2(df)$ | TLI | CFI | RMSEA [90% CI] |
|------|-----------------|------|------|-------------------|
| 모형 1 | 1613.78*** (27) | .761 | .821 | .158 [.152, .165] |
| 모형 2 | 299.08*** (26) | .957 | .969 | .067 [.060, .074] |
| 모형 3 | 167.58*** (17) | .964 | .983 | .061 [.053, .070] |
| 모형 4 | 181.71*** (21) | .969 | .982 | .058 [.050, .065] |
| 모형 5 | 272.46*** (22) | .954 | .972 | .070 [.062, .077] |
| 모형 6 | 182.16*** (23) | .972 | .982 | .054 [.047, .062] |

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

크기를 비교함으로써 Rosenberg 자기존중감 척도의 문항들 간의 공분산을 가장 잘 설명하는 모형을 찾았다.

전반적인 자기존중감을 가정하는 1요인 모형(모형 1)이 가장 나쁜 모형 합치도를 보였다 ($\chi^2(27)=1,613.78$ $p < .001$, TLI=.761, CFI=.821, RMSEA=.158, 90%신뢰구간=[.152, .165]). 긍정적인 자기평가와 부정적인 자기평가라는 요인들을 가정하는 2요인 모형(모형 2)은 1요인 모형(모형 1)에 비하여 상대적으로 좋은 모형 합치도를 보였다, $\Delta\chi^2=1,314.70$, $\Delta df=1$, $p < .001$.

둘째, 긍정적인 문항 형태의 공통성을 설명하는 모형 비교가 이루어졌다. 모형 5는 모형 1에 비하여 상대적으로 좋은 모형 합치도를 보였다, $\Delta\chi^2=1,341.72$, $\Delta df=5$, $p < .001$. 이와 더불어, 상대적 합치도 지수들의 값들도 모형 1에 비하여 모형 5를 선호하였다(표 1 참조). 모형 3은 모형 5에 비하여 상대적으로 좋은 모형 합치도를 보였다, $\Delta\chi^2=104.88$, $\Delta df=5$, $p < .001$. 이와 더불어, 상대적 합치도 지수들의 값들도 모형 5에 비하여 모형 3을 선호하였다(표 1 참조). 따라서 모형 3이 긍정적인 문항형태를 가진 문항들의 고유분산 간의 공분산을 상대적으로 잘 설명하였다.

셋째, 부정적인 문항 형태의 공통성을 설명하는 모형 비교가 이루어졌다. 모형 6은 모형 1에 비하여 상대적으로 좋은 모형 합치도를 보였다, $\Delta\chi^2=1,431.62$, $\Delta df=4$, $p < .001$. 이와 더불어, 상대적 합치도 지수들의 값들도 모형 1에 비하여 모형 6을 선호하였다(표 1 참조). 모형 6은 모형 4에 비하여 상대적으로 좋은 모형 합치도를 보였다, $\Delta\chi^2=0.45$, $\Delta df=2$, $p=.80$. 이와 더불어, 상대적 합치도 지수들의 값들도 모형 4에 비하여 모형 6을 선호하였다(표 1 참조). 따라서 모형 6이 부정적인 문항형태를 가진 문항들의 고유분산 간의 공분산을 상대적으로 잘 설명하였다.

마지막으로, 위의 모형 비교에서 상대적으로 좋은 합치도를 보인 세 개의 모형들(모형 2, 3, 6)의 모형의 간명도 및 요인계수의 크기 등을 고려한 적합도 수준을 비교하였다. 세 개의 모형들 중에서 모형 6이 상대적으로 높은 수준의 합치도 지수를 보였다, $\chi^2(23)=182.16$, $p < .001$, TLI=.972, CFI=.982, RMSEA=.054, 90%신뢰구간=[.047, .062]. 특히, 모형 3과 6은 비슷한 수준의 합치도 지수를 보이지만, 모형 6이 모형 3에 비하여 간명한 모형이었다. 따라서, 모형 6을 최종모형으로

선택하였다. 보다 구체적으로, 전반적인 자기 존중감 요인과 함께 부적인 자기에 대해서 묻는 문항들의 오차항 간의 공분산을 설명하는 부정적 문항의 방법효과를 가정한 모형 6이 Rosenberg의 자기존중감 척도의 문항간의 상관을 가장 잘 설명했다. 이에 덧붙여서, 모형 6에서는 2번 문항과 4번 문항의 수정지수의 값이 54.83으로 모형을 수정할 것이 권고되었다. 하지만, 모형 6의 적합도 지수의 값들이 공분산을 고려하지 않은 모형의 적합도 수준이 높음을 나타내고, 두 문항들의 고유요인들이 관련될 수 있는 이론적 근거가 명확하지 않기 때문에 2번 문항과 4번 문항 간의 공분산을 추정하지 않았다. 모형 내의 다른 수정지수의 값들은 상대적으로 작았고($MIs < 33$), 이론적인 근거도 미약하여 모형의 수정이 이루어지지

않았다.

표 2에서 모형 6의 확인적 요인분석 결과가 보고되었다. Rosenberg의 자기존중감 척도의 문항들이 전반적인 자기존중감과 부정적 문항으로 인한 방법 효과 요인들에 의해서 잘 설명됨을 알 수 있다 (문항들의 표준화 요인계수 $> .46$). 이와 덧붙여서 전반적인 자기존중감의 분산($\sigma^2=0.304$, $se=0.015$, $p<.001$)과 부정적 문항으로 인한 방법효과 요인들의 분산($\sigma^2=0.228$, $se=0.016$, $p<.001$)은 통계적으로 유의하였다.

모형의 측정단위 · 원점 동일성 검증

전반적인 자기존중감과 함께 부정적 문항의 방법효과 요인을 가정한 모형이 남녀 중학생

표 2. 전반적인 자기존중감과 부정적 문항에 대한 방법효과를 가정한 요인구조(모형 6)의 요인계수

| 요인 | 문항 | 비표준화 요인계수 | 표준오차 | 표준화 요인계수 |
|------------|------------------|-----------|----------|----------|
| 전반적인 자기존중감 | Q1 | 1.000*** | | .747 |
| | Q2 | 1.080*** | 0.032 | .751 |
| | Q3 | 0.707*** | 0.035 | .451 |
| | Q4 | 0.866*** | 0.028 | .674 |
| | Q5 | 0.904*** | 0.034 | .588 |
| | Q6 | 0.992*** | 0.031 | .692 |
| | Q7 | 1.075*** | 0.036 | .663 |
| | Q9 | 0.886*** | 0.038 | .522 |
| | Q10 | 0.786*** | 0.036 | .489 |
| | 부정적 문항으로 인한 방법효과 | Q3 | 1.000*** | |
| Q5 | | 0.801*** | 0.045 | .416 |
| Q9 | | 1.393*** | 0.062 | .656 |
| Q10 | | 1.316*** | 0.060 | .654 |

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

집단 모두에서 적합한 모형인지를 다집단 확인적 요인분석을 활용하여 검증하였다. 특히, 다집단 분석에서 참조지표(reference indicator)의 선정이 측정모형의 동일성 검증에 영향을 미치기 때문에, 이순목, 이찬순, 이현정, 여성칠(2012)의 연구에서 제안한 방식에 따라서 참조지표를 설정하였다. 측정모형의 동일성 검증의 결과는 표 3에서 제시되었다. 첫째로, 남녀 중학생 집단 간의 형태 동일성을 검증하였다. 즉, 전반적인 자기존중감과 부정적 문항의 방법효과를 잠재변수로 가정한 모형들이 남녀 중학생 집단 각각에 적합한 모형인지를 검증하였다. 남녀 중학생 집단 모두에서 전반적인 자기존중감과 부정적 문항의 방법효과를 잠재변수로 가지는 모형이 적합한 것으로 나타났다($\chi^2(46)=217.95, p<.001, TLI=.970, CFI=.981, RMSEA=.056, 90\%신뢰구간=[.049, .064]$).

남녀 중학생 집단 간의 형태 동일성이 존재함을 확인한 후에, 형태 동일성을 가정하는 모형을 기저모형(baseline model)으로 사용하여 이후의 측정단위 동일성 검증을 실시하였다. 형태 동일성을 검증한 후에, 측정단위 동일성의 가정을 검증하였다. 보다 구체적으로, 남녀 중학생 집단 간의 동일한 형태의 요인구조를

가정(형태 동일성 가정)함에 더붙어서 전반적인 자기존중감과 부정적 문항의 방법효과와 잠재 변수들에 대한 요인계수의 값이 두 집단 간에 동일하다는 제약을 추가하였다. 이 요인계수의 동일화 제약을 가한 측정단위 동일성 모형의 χ^2 값과 자유도를 기저모형의 그것들과 비교하였다. 특히, 기저모형과 측정단위 동일성 모형들은 내재된 모형이기 때문에 χ^2 차이 검증을 실시하였다. 표 3에서 나타나듯이, 기저모형(모형 1)과 측정단위 동일성 모형(모형 2)간의 χ^2 값의 차이는 $\alpha=.05$ 수준에서 통계적으로 유의하다, $\Delta\chi^2(\Delta df)=28.98(11), p<.001$. 이것은 측정단위 동일성 가정이 기각됨을 의미한다. 하지만, χ^2 차이 검증 역시 표본크기에 민감하기 때문에 다른 합치도 지수의 모형간의 변화 역시 살펴보았다. 예를 들어서, 비록 모형 2의 CFI, mc값이 모형 1의 그것들에 비해서 상대적으로 나빠졌지만, 그 변화의 크기가 Cheung과 Rensvold가 제안한 기준값($\Delta CFI=.02, \Delta mc=.01$)보다 작았다(본 연구에서는 ΔCFI 이 .002이고, Δmc 는 .001이었다.). 따라서, 측정단위 동일성을 가정한 모형 2는 기저모형이 모형 1보다 상대적 합치도 지수가 유의하게 나빠지지 않았기 때문에 측정단위 동일성

표 3. 측정모형의 동일성 평가

| 모형 | $\chi^2 (df)$ | $\Delta\chi^2 (\Delta df)$ | TLI | CFI | mc | RMSEA [90%CI] |
|------------------------------|---------------|----------------------------|------|------|------|-------------------|
| 모형 1: 형태 동일성 | 218.94 (46) | | .970 | .981 | .980 | .057 [.049, .064] |
| 모형 2: 측정단위 동일성 | 247.92 (57) | 28.98*** (11) | .973 | .979 | .979 | .060 [.053, .066] |
| 모형 3: 측정단위 및 측정 원점 동일성 | 292.71 (64) | 44.79*** (7) | .971 | .974 | .974 | .055 [.049, .062] |
| 모형 4: 측정단위, 측정 원점 및 요인분산 동일성 | 296.89 (66) | 4.18 (2) | .972 | .974 | .974 | .055 [.048, .061] |

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

가정이 성립되는 것으로 보였다. 이러한 결과는 Rosenberg의 자기존중감 척도가 남녀 중학생 집단에서 잠재요인들과 그에 해당하는 척도들 간의 관계가 동일함을 의미하여 두 집단에서 하나의 척도로 활용될 수 있음을 의미한다(홍세희, 황매향, 이은설, 2005).

남녀 중학교 1학년 학생들 간에 측정단위 동일성의 가정이 지지됨을 확인하였다. 다음 단계로서, 두 집단 간에 문항들의 절편이 동일한지(측정원점 동일성)를 검증하였다. 이를 위해서, 측정단위 동일성을 가정한 모형 2가 기저모형으로 사용되었고, 측정원점 동일성을 가정한 모형 3이 비교 모형으로 설정되었다. 보다 구체적으로, 모형 3에서는 측정 변수들의 요인 계수가 집단 간에 동일함을 가정함(모형 2)과 더불어서 측정 변수들의 절편에도 집단 간의 동일화 제약을 추가되었다. 표 3에 나타나듯이, χ^2 차이 검증은 모형 3에 비하여 모형 2를 선호하였다, $\Delta\chi^2(\Delta df)=44.77(7)$, $p<.001$. 그러나, 비록 모형 2의 CFI, mc값이 모형 1의 그것들에 비해서 상대적으로 나빠졌지만, 그 변화의 크기가 Cheung과 Rensvold가 제안한 기준값($\Delta CFI=.02$, $\Delta mc=.01$)보다 작았다(본 연구에서는 ΔCFI 이 .005이고, Δmc 는 .005이었다). 따라서 측정원점 동일성 가정이 지지되는 것으로 해석했다. 측정의 행태·단위·원점 동일성 가정의 성립은 남녀 중학교 1학년 학생들에서 나타나는 집단 간의 Rosenberg의 자기존중감 척도의 평균 차이는 전반적인 자기존중감 요인 혹은 부정적 문항들에서 나타나는 방법요인을 실제 차이를 반영할 수 있다고 볼 수 있다(홍세희, 황매향, 이은설, 2005).

남녀 중학교 1학년 학생 집단 간의 자기존중감 및 방법 요인의 잠재평균 비교

남녀 중학교 1학년 학생 집단 간의 측정단위 및 측정원점 동일성 가정이 충족되었기 때문에, 두 집단 간의 자기존중감 및 부정적 문항들의 방법효과 요인들에 대한 잠재평균 비교분석을 실시하였다. 잠재평균 비교 시 두 집단 이상에서 요인의 평균을 동시에 추정하는 것이 불가능하다. 그리하여 한 집단(참조 집단)의 요인 평균을 0으로 고정한 다음에 다른 집단의 요인 평균을 추정하는 방법이 사용된다(홍세희, 황매향, 이은설, 2005; Hancock, 1997). 이 과정에서, 다른 집단의 요인 평균이 통계적으로 0과 다른지를 검증할 수 있는 장점이 있다.

본 연구에서는 남자 중학교 1학년 집단을 참조 집단으로 설정하고, 여자 중학교 1학년 집단의 전반적인 자기존중감과 부정적 문항에 대한 부적 방법효과의 잠재 요인의 평균을 추정하였다. Cohen's d 의 효과 크기를 계산하는데 있어서, 잠재 요인들의 표준편차를 사용한다. 이를 위해서 두 집단 간의 요인 분산의 동일성 가정을 검증하였다. 표 3에서 집단 간의 요인 분산의 동일성을 가정한 모형 4가 측정원점 동일성을 가정한 모형 3에 비해서 합치도 지수가 나빠지지 않았기 때문에 남녀 중학생 집단 간의 잠재 요인들의 분산이 동일함을 가정하는 것이 지지되었다. 따라서, 모형 4에서 추정된 잠재 요인의 표준편차를 활용하여 Cohen's d 를 추정하였다.

표 4는 여자 중학생 집단의 전반적인 자기존중감과 부적 방법효과의 잠재 요인들의 평균 및 효과크기(Cohen's d)를 보고하였다. Cohen(1988)은 d 값이 0.2, 0.5, 그리고 0.8을 각각 작

표 4. 전반적인 자기존중감 및 부정적 문항에 대한 방법요인에 대한 남녀 중학교 1학년 집단 간의 잠재평균 차이 분석

| 요인 | 남학생 | 여학생 | 효과크기 (<i>d</i>) |
|------------|------|-----------|-------------------|
| 전반적인 자기존중감 | 0.00 | -0.053 * | .10 |
| 부정적 방법효과 | 0.00 | -0.058 ** | .14 |

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

은 차이, 중간 차이, 그리고 큰 차이로 해석할 것을 제안했다. 부정적인 문항들에 대한 방법효과를 통제한 후에, 여자 중학교 1학년생들이 남자 중학교 1학년생들보다 전반적인 자기존중감에서 통계적으로 낮은 수준을 보였으나, 그 집단 간의 차이는 작았다.

논 의

연구 참여자가 심리검사에서 일정한 방향으로 편향되게 반응하는 것을 방지하기 위해서 종종 사용되어지는 역문항의 사용은 심리검사의 요인구조에 영향을 주어, 초기에 의도된 요인구조와는 다른 형태를 가지게 되는 경우가 자주 발생하게 된다. 이러한 문제로 인하여, 역문항이 포함된 심리검사의 요인구조에 관해서는 일치된 견해가 존재하지 않는 경우가 많으며, 연구자들이 역문항이 포함된 심리검사의 요인구조의 해석에 어려움을 경험하고 있다.

이에 본 연구에서는 긍정문항과 부정문항을 동시에 사용하고 있는 Rosenberg 자기존중감 원척도의 요인구조를 탐색하고, 남녀 간의 요인구조의 동질성과 전반적인 자기존중감과 부정적 방법효과와의 수준 차이를 검증하였다. 본 연구에서의 연구문제들에 대한 결과는 아래와

같다.

첫 번째, Rosenberg 자기존중감 척도의 문항 간의 공분산을 설명하기 위해서 설정된 6개의 요인모형 중에서 일반적인 자기존중감 요인 뿐만 아니라 부정적인 문항들에 대한 방법요인을 설정한 모형(모형 6: 그림 1 참조)이 Rosenberg 자기존중감 척도에 가장 적합하였다. 이는 자기존중감 척도가 긍정적 자기평가와 부정적 자기평가라는 2개의 요인으로 구성된다는 기존의 국내 연구결과들(이미리, 2006; 정병삼, 2010)과는 일치하지 않는다. 이는 앞서도 밝혔듯이, 기존 연구들이 사용한 연구방법의 한계로 인하여 발생한다고 볼 수 있다. 보다 구체적으로, 본 연구에서는 역문항으로 인하여 발생하는 방법효과를 적절하게 통제하기 위해서 확인적 요인분석을 실시하였다. 이에 반하여, Rosenberg 자기존중감 척도의 2요인구조를 주장한 기존의 연구들은 방법효과를 적절하게 고려하지 못한 탐색적 요인분석을 사용하였다(홍세희 등, 2011). 이로 인하여 기존의 국내 연구들에서는, Rosenberg 자기존중감 척도의 문항 간의 공분산을 보다 잘 설명할 수 있는 요인구조는 검증의 대상이 되지 못하였다. 하지만, 본연구의 결과는 확인적 요인분석을 사용하여 방법효과를 효과적으로 고려한 국외 연구(Quilty et al., 2006)와 국내 기존 연구(홍세희 등, 2011)의 결과와 일치한다. 첫 번

재 연구결과가 의미하는 것은 Rosenberg 자기존중감 척도를 사용하는 경우에 부정적인 문항들에 대한 방법 요인을 적절하게 통제하지 않은 경우에는 자기존중감의 수준을 정확하게 측정할 수 없다는 것이다. 보다 구체적으로, 역문항들의 고유요인들 간의 공분산을 적절하게 설명하는 방법요인을 고려하지 않으면, 측정하고자 하는 심리적 구성개념(예를 들어, 자기존중감, 우울)의 수준을 과대추정하거나 혹은 과소추정하게 된다. 따라서 역문항을 사용한 척도에서 역문항의 사용으로 인하여 발생하는 방법요인을 요인구조에 포함시켜야 할 것이다.

덧붙여서, 본 연구에서는 부정문항과 긍정문항을 각각의 문항 형태를 설명하는 방법효과를 고려한 2개의 요인모형으로 비교하였다. 두 요인 구조 모형의 비교에서 부정적 문항을 설명하는 방법효과를 추가한 요인모형이 긍정적 문항을 설명하는 방법효과를 추가한 요인모형에 비하여 Rosenberg 자기존중감 척도의 문항 간의 공분산을 설명하는데 보다 적합하였다. 이러한 결과는 기존의 국내외 연구들(홍세희 등, 2011; Horan et al., 2003; Quilty et al., 2006)을 지지하는 것이다. 즉, 일반적으로 부정문항은 수검자들에게 평시에 접하지 않는 익숙지 않은 문항형태이기 때문에 이들이 긍정문항에 비하여 부정문항에 대하여 고유한 반응형태를 보인다고 볼 수 있다(홍세희 등, 2011).

두 번째로, 첫 번째 연구문제를 통해서 확인된 요인구조가 남녀 초기청소년들에게 동시에 적용가능한지에 대하여 조사하였다. 이를 위해서 남녀 초기청소년들 간의 측정단위 동일성 검증을 실시하고 두 집단 간의 측정모형이 동일함을 확인하였다. 이는 남녀 초기청소

년 집단 모두에서 전반적인 자기존중감 요인 뿐만 아니라 역문항으로 인하여 발생하는 방법효과 요인이 관찰됨을 의미한다. 동시에, 본 연구에서 사용된 9개의 문항들과 전반적인 자기존중감 요인과 방법효과 요인과의 관계가 남녀 초기청소년 집단에서 동일함을 의미한다.

세 번째로, 남녀 초기청소년들의 측정모형의 동일성을 확인한 후에 일반적인 자기존중감 요인과 부정적인 문항들에 대한 방법요인들에 대한 남녀 간의 수준 차이가 있음을 확인하였다. 보다 구체적으로, 여자 초기청소년들은 일반적인 자기존중감 요인 뿐만 아니라 부정적인 문항들의 형태를 설명하는 방법요인에서도 남자 청소년들에 비하여 비록 통계적으로 유의하지만 Cohen's d 가 작은 수준으로 점수가 낮음을 확인하였다. 이를 통해서 유추될 수 있는 것은 연구자들이 부정적인 문항들에 대한 방법 요인을 적절하게 통제하지 않는다면 Rosenberg 자기존중감 척도를 사용하여 남녀 청소년 간의 자기존중감의 수준을 비교하게 되면 그 차이가 과대 추정될 수 있음을 의미한다. 따라서, 남녀 청소년 간의 자기존중감 수준을 정확하게 비교하기 위해서는 부정적인 문항들에 대한 방법 요인의 남녀 청소년의 차이를 적절하게 통제해야 할 것이다.

이에 덧붙여서, 본 연구에서 발견된 남녀 청소년 간의 자기존중감 차이(Cohen's $d=.10$)를 Kling과 동료들(1999)의 메타 연구의 값(Cohen's $d=.22$)과 비교하면 흥미로운 결과가 도출된다. 즉, 방법효과를 고려한 이후에 자기존중감 수준은 그렇지 않은 경우에 비해서 남녀 간의 차이가 줄어든다는 것이다. 방법요인에 의해서 나타나는 성차를 고려하지 않음으로 인하여 자기존중감의 성차가 과대 추정될 수 있음을 보여주고 있다. 보다 흥미로운 사실은 본

연구에서 발견된 자기존중감(Cohen's $d=.10$)과 방법효과(Cohen's $d=.14$)의 남녀 차이를 합하면 메타 연구의 값(Cohen's $d=.22$)과 유사해 진다는 것이다. 하지만 이와 같은 연구결과를 일반화시키기에는 국내에서의 메타연구가 진행되어야 할 것이다.

본 연구는 몇 가지 제한점을 가지고 있다. 첫 번째, 비록 층화다단계집락표집을 사용하여 결과의 일반화 가능성을 높였지만, 연구는 중학교 1학년만을 대상으로 이루어졌기 때문에 다른 청소년 집단에서도 동일한 연구결과가 반복되는지에 대한 확인이 필요하다. 두 번째, Rosenberg 자기존중감 척도의 8번 문항을 분석에서 제외함으로 인하여 연구결과의 일반화 가능성이 떨어졌다. 따라서 추후의 연구에서는 8번 문항을 제외하는 대신에 이자영, 남숙경, 이미경, 이지희, 이상민(2009)의 연구에서 제안된 문항을 사용하여 총 10개의 문항으로 이루어진 Rosenberg 자기존중감 척도를 사용하여야 할 것이다. 보다 구체적으로, 이자영 등(2009)의 연구에서 추가된 문항을 포함한 Rosenberg 자기존중감 척도의 요인구조를 이순목, 이찬순, 이현정, 여성철(2012)에서 이루어진 분석절차를 활용하여 경험적으로 검증함으로써 번역타당도 및 2요인 모형의 일반화 가능성을 평가하여야 할 것이다. 세 번째, 다집단 분석(multiple-group analysis)에서 요인구조의 동일성 검증의 순서의 변화를 고려해 볼 수 있다. 본 연구에서와 달리, 명시적으로 구분되는 집단 별로 측정모형을 타당화한 이후에 집단 간 측정모형의 일반화가능성을 검증할 수 있다.

이러한 제한점에도 불구하고, 본 연구는 Rosenberg 자기존중감 척도를 사용하여 자기존중감의 수준을 측정하고 남녀 간의 그 수준을

비교할 때, 부정적으로 진술된 문항들에 방법효과 요인을 고려해야 함을 증명했다는 데 의의가 있다. 그러므로 Rosenberg 자기존중감 척도를 사용하여 자기존중감 수준의 변화 형태를 연구하는 잠재성장모형(latent growth curve models) 혹은 자기존중감을 구조방정식모형(structural equation modeling) 속의 하나의 요인으로 사용하는 연구모형을 검증하는 경우에는 역문항들에 대한 방법효과 요인을 고려해야 할 것이다. 덧붙여서, 연구자들이 Rosenberg 자기존중감 척도를 활용하여 남녀 청소년 간의 자기존중감의 수준차이를 비교하는 경우에는 부적 문항으로 인한 방법효과를 효과적으로 통제할 경우(즉, 방법효과를 방법요인 혹은 측정오차 간의 공분산을 요인구조에 포함하고 일반요인의 요인점수(factor score)를 추정하여 사용할 수 있음)에만 집단간의 자기존중감의 차이를 정확하게 측정할 수 있을 것이다.

참고문헌

- 이미리 (2005) 청소년기 자아존중감과 가족, 친구, 학업, 여가, 직업 변인들의 관계: 긍정적 자아평가와 부정적 자아평가를 중심으로. *한국청소년연구*, 16(2), 263-293.
- 이순목, 이찬순, 이현정, 여성철 (2012). 캐나다 도박행동 척도에서 개념적 구조 및 심리 측정적 특성의 일반화 가능성: 남녀간 다집단 분석. *한국심리학회지: 일반*, 31(1), 1-26
- 이자영, 남숙경, 이미경, 이지희, 이상민 (2009). Rosenberg의 자아존중감 척도: 문항수준 타당도분석. *한국심리학회지: 상담 및 심리치료*. 21(1), 173-189.

- 정병삼 (2010) 부모-자녀관계에 착과 부모지도 감독이 청소년의 자아존중감의 변화에 미치는 종단적 영향. *한국청소년연구*, 21(4), 5-30.
- 최보가, 전귀연 (1993). 자아존중감 척도 개발에 관한 연구 I. *대한가정학회지*, 31(2), 1-14.
- 홍세희, 노연경, 정 송 (2011). 부정문항이 포함된 검사의 요인구조: 자아존중감 검사의 예. *교육평가연구*, 24(3), 713-732.
- 홍세희, 황매향, 이은설 (2005). 청소년용 여성 진로장벽 척도의 잠재평균비교. *교육심리연구*, 19(4), 1159-1177.
- American Educational Research Association. (1999). *Standards for educational and psychological testing*. Washington, DC: Author.
- Anderson, J. C. & Gerbing, D. W. (1988). Structural equation modeling in practice: A review and recommended two-step approach, *Psychological Bulletin*, 103(3), 411-423.
- Bagozzi, R. P. (1993). Assessing construct validity in personality research: Applications to measures of self-esteem. *Journal of Research in Personality*, 27, 49-87.
- Browne, W., & Cudeck, R. (1993). Alternative ways of assessing model fit. In K. A. Bollen & J. S. Long(Eds.), *Testing structural equation models* (pp.136-162). Newbury Park, CA: Sage.
- Carmines, E. G., & Zeller, R. A. (1979). *Reliability and validity assessment*. Beverly Hills, CA: Sage.
- Cheung, G. W., & Rensvold, R. B. (2002). Evaluating goodness-of-fit indexes for testing measurement invariance. *Structural Equation Modeling*, 9(2), 233-255.
- Cohen, J. (1988). *Statistical power analysis for the behavioral sciences*. Hillsdale, NJ: Lawrence Erlbaum Associates, Inc.
- Crandall, R. (1973). The measurement of self-esteem and related constructs. In J. P. Robinson & P. Shaver(Eds.), *Measurements of social psychological attitudes* (pp.45-167). Ann Arbor, MI: Institute for Social Research.
- Silbert, E., & Trippett, J. (1965). Self-esteem: Clinical assessment and measurement validation. *Psychological Reports*, 16, 1017-1071.
- DiStefano, C., & Motl, R. W. (2009). Self-esteem and method effects associated with negatively worded items: Investigating factorial invariance by sex. *Structural Equation Modeling: A Multidisciplinary Journal*, 22, 134-146.
- Hancock, G. R. (1997). Structural equation modeling methods of hypothesis testing of latent variable means, *Measurement Evaluation in Counseling and Development*, 30(2), 91-105.
- Horan, P. M., DiStefano, C., & Motl, R. W. (2003). Wording effects in self esteem scales: Methodological artifact or response style? *Structural Equation Modeling*, 10, 444-455.
- Hu, L. & Bentler, P. M. (1999). Cutoff criteria for fit indexes in covariance structure analysis: Conventional criteria versus new alternatives. *Structural Equation Modeling*, 6(1), 1-55.
- Kling, K. C., Hyde, J. S., Showers, C. J., & Buswell, B. N. (1999). Gender differences in self-esteem: A meta-analysis. *Psychological Bulletin*, 125, 470-500.
- Larsen, R. J., & Ketelaar, T. (1991). Personality and susceptibility to positive and negative emotional states. *Journal of Personality and Social Psychology*, 61(1), 132-140.

- Marsh, H. W. (1986). The bias of negatively worded items in rating scales for young children: A cognitive-developmental phenomenon. *Developmental Psychology, 22*, 37-49.
- Marsh, H. W. (1996). Positive and negative global self-esteem: A substantively meaningful distinction or artifacts?. *Journal of Personality and Social Psychology, 70*(4), 810-819.
- Motl, R. W., & DiStefano, C. (2002). Longitudinal invariance of self-esteem and method effects associated with negatively worded items. *Structural Equation Modeling, 9*, 562-578.
- Muthén, L. K., & Muthén, B. O. (2010). *Mplus 6.0 {Computer Software}* Los Angeles CA: Muthén & Muthén.
- 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. 17-59). San Diego, CA: Academic.
- Quilty, L. C., Oakman, J. M., & Risko, E. (2006). Correlates of the Rosenberg self-esteem scale method effects. *Structural Equation Modeling, 13*(1), 99-117.
- Rosenberg, M. (1979). *Conceiving the self*. New York: Basic Books.
- Rosenberg, M. (1965). *Society and the adolescent child*. Princeton, NJ: Princeton University Press.
- Schrisheim, C. A., & Hill, K. D. (1981). Controlling acquiescence bias by item reversals: The effect on questionnaire validity. *Education and Psychological Measurement, 41*, 1101-1114.
- Steenkamp, E. M., & Baumgartner, H. (1998). Assessing measurement invariance in cross-national consumer research. *Journal of Consumer Research, 25*(1), 78-107.
- Tomas, J. M., & Oliver, A. (1999). Rosenberg's self-esteem scale: Two factors or method effects. *Structural Equation Modeling, 6*, 84-98.
- Watson, D., & Clark, L. (1991). Self-versus peer ratings of specific emotional traits: Evidence of convergent and discriminant validity. *Journal of Personality and Social Psychology, 62*, 489-505.
- Watson, D., & Clark, L. (1992). Affects separable and inseparable: On the hierarchical arrangement of negative affects. *Journal of Personality and Social Psychology, 62*, 489-505.
- Watson, D. (1988). Intraindividual and interindividual analyses of positive and negative affects: Their relation to health complaints, perceived stress and daily activities. *Journal of Personality and Social Psychology, 54*(6), 1020-1030.
- Winkler, J. D., Kanouse, D. E., & Ware, J. E. (1982). Controlling acquiescence response set in scale development. *Journal of Applied Psychology, 67*, 555-561.
- 1차원고접수 : 2013. 03. 26.
수정원고접수 : 2013. 05. 31.
최종게재결정 : 2013. 08. 27.

Measurement Invariance of Self-Esteem between Male and Female Adolescents and Method Effects Associated with Negatively Worded Items

Sumi Choi

Pusan National University

Young Il Cho

Sungshin Women's University

In many scales in social science, negatively worded items were often employed to prevent bias of response tendency. However, using the negatively worded items in the scale might reduce its reliability and distort its factor structure. In order to clarify those problems, researchers have suggested to employ a factor structure including a method factor accounting for negatively worded items. In this study, Rosenberg's Self-Esteem scale which employs two types of items(i.e., positively and negatively worded items) was investigated to uncover its factor structure. Additionally, measurement invariance and latent mean difference between males and females were studied. In conclusion, the factor model considering the method factor of negatively worded items was preferred over others. Measurement invariance between males and females were supported and females showed the lower self-esteem level than males after controlling for the effect of the method effect. Finally, limitation and implication of the current study are discussed.

Key words : Rosenberg's self-esteem scale, method effect, factor structural model, measurement equivalence, gender difference

부 록

부록 1. 전체 집단의 상관 행렬 (N=2,350)

| | Q1 | Q2 | Q3 | Q4 | Q5 | Q6 | Q7 | Q9 | Q10 |
|------|--------|--------|--------|--------|--------|--------|--------|--------|------|
| Q1 | 1 | | | | | | | | |
| Q2 | .54*** | 1 | | | | | | | |
| Q3 | .31*** | .32*** | 1 | | | | | | |
| Q4 | .53*** | .56*** | .29*** | 1 | | | | | |
| Q5 | .44*** | .46*** | .47*** | .38*** | 1 | | | | |
| Q6 | .55*** | .49*** | .35*** | .43*** | .38*** | 1 | | | |
| Q7 | .48*** | .51*** | .35*** | .39*** | .40*** | .50*** | 1 | | |
| Q9 | .40*** | .37*** | .57*** | .32*** | .58*** | .37*** | .38*** | 1 | |
| Q10 | .38*** | .36*** | .55*** | .28*** | .56*** | .35*** | .36*** | .68*** | 1 |
| 평균 | 3.03 | 2.64 | 2.86 | 2.99 | 2.78 | 2.96 | 2.95 | 2.98 | 2.81 |
| 표준편차 | 0.71 | 0.77 | 0.84 | 0.69 | 0.82 | 0.76 | 0.84 | 0.90 | 0.86 |

부록 2. 남학생 집단의 상관 행렬 (N=1,176)

| | Q1 | Q2 | Q3 | Q4 | Q5 | Q6 | Q7 | Q9 | Q10 |
|------|--------|--------|--------|--------|--------|--------|--------|--------|------|
| Q1 | 1 | | | | | | | | |
| Q2 | .54*** | 1 | | | | | | | |
| Q3 | .26*** | .28*** | 1 | | | | | | |
| Q4 | .55*** | .57*** | .25*** | 1 | | | | | |
| Q5 | .42*** | .44*** | .43*** | .36*** | 1 | | | | |
| Q6 | .56*** | .49*** | .35*** | .44*** | .35*** | 1 | | | |
| Q7 | .43*** | .48*** | .29*** | .37*** | .37*** | .50*** | 1 | | |
| Q9 | .35*** | .34*** | .55*** | .30*** | .56*** | .36*** | .32*** | 1 | |
| Q10 | .32*** | .31*** | .53*** | .29*** | .55*** | .34*** | .29*** | .66*** | 1 |
| 평균 | 3.03 | 2.66 | 2.90 | 2.98 | 2.81 | 3.01 | 3.04 | 3.05 | 2.87 |
| 표준편차 | 0.74 | 0.80 | 0.86 | 0.71 | 0.84 | 0.78 | 0.86 | 0.90 | 0.86 |

부록 3. 여학생 집단의 상관 행렬 (N=1,174)

| | Q1 | Q2 | Q3 | Q4 | Q5 | Q6 | Q7 | Q9 | Q10 |
|------|--------|--------|--------|--------|--------|--------|--------|--------|------|
| Q1 | 1 | | | | | | | | |
| Q2 | .55*** | 1 | | | | | | | |
| Q3 | .37*** | .35*** | 1 | | | | | | |
| Q4 | .51*** | .55*** | .33*** | 1 | | | | | |
| Q5 | .45*** | .50*** | .52*** | .41*** | 1 | | | | |
| Q6 | .53*** | .50*** | .35*** | .42*** | .41*** | 1 | | | |
| Q7 | .53*** | .54*** | .39*** | .42*** | .44*** | .49*** | 1 | | |
| Q9 | .46*** | .40*** | .59*** | .35*** | .59*** | .39*** | .44*** | 1 | |
| Q10 | .44*** | .41*** | .57*** | .29*** | .58*** | .36*** | .41*** | .70*** | 1 |
| 평균 | 3.04 | 2.62 | 2.82 | 3.00 | 2.74 | 2.91 | 2.86 | 2.91 | 2.76 |
| 표준편차 | 0.69 | 0.73 | 0.81 | 0.66 | 0.80 | 0.75 | 0.86 | 0.90 | 0.85 |