

청소년 성격평가질문지(PAI-A) 긍정왜곡 지표의 효율성

박 은 영[†]

대구보호관찰소

홍 상 황

진주교육대학교 교육학과

본 연구는 성인을 대상으로 제작된 긍정왜곡 지표(PIM, DEF, CDF)와 국내 고등학생을 대상으로 개발된 긍정왜곡 지표(KDF)를 청소년 성격평가질문지(PAI-A)에 적용하여 그 조작적 특성과 효율성을 비교하는데 목적이 있다. 고등학생을 대상으로 긍정왜곡 지시문을 사용하여 PAI-A를 실시한 자료와 표준 지시문 자료에 대해 PIM, DEF, CDF, KDF를 산출하였다. 그 결과 긍정왜곡 지시 프로파일은 PIM, RXR, DOM, WRM 및 MAN이 상승하고 대부분의 임상척도가 평균보다 낮았다. 긍정왜곡 지표의 특성을 자세히 비교하기 위해 ROC 분석을 통해 다양한 분할점수의 민감성과 특이성, 각 분할점수의 진단적 효율성의 추정치를 계산했다. 4개 긍정왜곡 지표의 효과크기 및 AUC를 산출한 결과 KDF, PIM, DEF, CDF 순이었다. 끝으로 각 긍정왜곡 지표 분할점수의 진단적 효율성의 추정치를 살펴보면 특정 분할점수에서의 민감성과 특이성의 합이 PIM, DEF, KDF는 큰 반면 CDF는 상대적으로 낮았다. 이러한 연구결과는 KDF가 기존에 성인을 대상으로 제작된 PIM, DEF와 함께 청소년의 긍정왜곡 반응을 탐지하는데 효과적임을 보여주고, 다양한 지표를 동시에 적용할 때 보다 정확하게 반응왜곡을 탐지할 수 있다는 점을 시사한다.

주요어 : 청소년 성격평가질문지(PAI-A), 긍정왜곡, 반응왜곡, ROC, AUC

[†] 교신저자(Corresponding Author) : 박은영, 대구보호관찰소 조사과, 대구시 동구 신암 5동 1503
E-mail : eyp02@korea.kr

심리평가 영역에서 지속되어온 관심거리 중 하나는 평가도구의 정확성이다. 특히 자기보고 검사를 사용할 경우 수검자가 원하면 얼마든지 결과를 왜곡할 수 있어서 검사개발자와 연구자들은 특별히 관심을 두고 있다. 수검자는 검사결과가 당락에 영향을 미칠 수 있다고 가정되는 상황에서 자연스럽게 부정적인 특성을 부인하고 긍정적인 특성을 과장시키고자 할 것이다(Greene, 1991; Lanyon, Dannenbaum, Wolf, & Brown, 1989; Rogers, Sewell, Morey, & Ustad, 1996). 이와 같은 긍정왜곡 또는 긍정적 인상관리 그리고 사회적으로 바람직한 방향으로 반응하려는 태도 등의 반응태세(response set)가 작용했을 경우 검사결과의 신뢰성에 의문을 제기할 수 있다. 실제로 정신적으로 건강한 사람과 심리적 방어 때문에 정신적으로 건강하게 보이는 사람을 구분하기가 어렵다. 또한 검사에서 건강한 것처럼 나타난 사람에게 심리적으로 부정응적인 측면이 지적되는 경우가 종종 있기 때문에 이러한 반응태세를 적절하게 찾아낼 필요가 있다(Shedler, Mayman, & Manis, 1993).

성격평가질문지(Personality Assessment Inventory, PAI; Morey, 1991)는 정신병질을 측정하고 평가하기 위한 자기보고형 도구로, 총 344문항, 4개의 타당성척도, 11개의 임상척도, 5개의 치료고려척도, 2개의 대인관계척도 등 총 22개의 척도로 구성되어 있다. PAI를 구성하는 문항은 측정하고자 하는 증상이나 문제를 구체적이고 직접적으로 기술한다. 또한 모호문항이 없으므로 긍정적이거나 부정적인 인상을 관리하고자 할 경우 검사결과를 쉽게 조작할 수 있다. 즉 PAI 문항은 분명하게 기술되어 있기 때문에 다른 검사보다 수검자가 원하는 방향으로 왜곡될 가능성이 더 높다(Bornstein,

Rossner, Hill & Stepanian, 1994; Peebles & Moore, 1998). 따라서 PAI는 개발단계부터 검사수행이나 결과에 영향을 미치는 반응태세의 영향을 평가하기 위한 4가지 타당성척도를 포함시키고 있다. 뿐만 아니라 긍정적 인상을 관리하려는 노력을 평가하기 위한 지표로는 방어지표(Morey, 1993, 1996)와 Cashel, Rogers, Sewell 및 Martin-Cannici(1995)의 판별함수가 있다. 우리나라에서는 홍상황과 김영환(2001)이 대학생을 대상으로 긍정적 인상관리 반응을 찾아내기 위한 지표를 구성하였고, 홍상황, 한태희 및 장현정(2010)은 고등학생을 대상으로 긍정적 인상관리를 탐지하기 위한 판별지표를 개발하여 입학사정관제 전형에서 효율적으로 사용할 수 있는 방안을 마련하였다.

각각의 척도 또는 지표를 살펴보면 우선 긍정적 인상관리 척도(PIM; Morey, 1991)는 자신에 대해 지나치게 긍정적인 인상을 주고자 하거나 비교적 사소한 결점이나 결함도 부인하려는 태도를 평가하는 문항으로 구성되어 있다. PIM은 긍정적인 인상을 주고자 시도한 것을 반영하는 척도이지만 이외에도 자신의 문제에 대한 인식이 부족하거나 자아상이 고양되어 있을 경우에도 PIM이 상승할 수 있다. Morey(1991)는 긍정왜곡 지시 및 표준지시 하에서 반응한 대학생의 자료를 가장 잘 변별할 수 있는 PIM 분할점수를 연구하였다. 그 결과 PIM 원점수 18(t점수 57)을 적용할 때 방어표본에 대한 민감성과 특이성이 각각 82%, 70%였다. 이와 같은 결과는 Cashel 등(1995)의 연구에서도 동일하게 반복되었다. 이들은 수검자에게 검사결과의 신뢰도를 유지할 수 있는 방법에 대한 정보를 제공한 후 긍정왜곡 지시를 하였고, 분할점수 18을 적용한 결과 민감성과 특이성이 각각 48%, 81%로 다소 떨어졌

다. Fals-Stewart(1996)는 소변검사에서는 양성반응이 나왔음에도 불구하고 법정에서는 물질사용을 부인한 약물남용자들을 대상으로 PIM의 분할점수에 대해 연구했다. 앞선 연구에서와 동일한 분할점수를 적용했을 때 긍정왜곡이 의심되는 프로파일에 대한 민감도와 특이성이 각각 88%, 80%이었다. 뿐만 아니라 낮은 PIM은 방어반응을 배제시킬 수 있는 유용한 준거될 수 있다는 사실도 발견하였다.

방어지표(Defensiveness Index, DEF; Morey, 1996)는 방어적인 반응을 탐지하기 위해 부가적인 지표로 제작되었다. DEF는 일반 정상인 또는 임상 표본에서보다 긍정왜곡 지시를 받은 수검자의 프로파일에서 현저하게 지적되는 PAI 프로파일의 형태적 특성을 나타내는 8개 항목으로 구성되어 있다. 예컨대, PIM ≥ 50 인 경우 가중치 2점, PIM ≥ 45 인 경우 가중치 1점을 추가하는 것이 한 항목이다. Morey (1996)의 연구에서 긍정왜곡 표본의 DEF 평균은 6.23인 반면 정상표본의 평균은 2.81로 상당한 차이가 지적되었다. Morey와 Lanier(1998)는 표준지시 및 왜곡지시를 받은 대학생 표본을 적절히 변별할 수 있는 DEF의 분할점수가 5점인 것으로 지적하였고, 5점을 적용할 경우 민감성과 특이성이 각각 93.3%, 86.7%였다. 그러나 Cashel 등(1995)은 수검자에게 타당성척도에 관한 사전지식을 제공하였고, 그 결과 DEF의 평균은 4.27로 낮아졌다. Baer와 Wetter(1997)는 표준지시와 사전정보를 제공받지 않은 왜곡 표본을 변별하는 DEF의 효과크기가 1.63인 반면 표준지시와 사전정보를 제공받은 왜곡표본을 변별하는 효과크기가 0.20으로 감소된다고 지적하였다. 따라서 방어를 시도하더라도 신뢰성을 확보하려고 하기 때문 사전정보를 제공받을 경우 DEF의 민감도가 떨어

진다고 설명할 수 있을 것이다.

방어적 반응을 탐지하기 위한 다른 지표들은 표준지시와 긍정왜곡 지시 하에서 반응한 프로파일을 변별하기 위해 판별분석을 실시하여 도출된 판별식이다. 첫째, Cashel 판별함수(Cashel Discriminant Function, CDF; Cashel et al., 1995)¹⁾는 교도소 수감자를 대상으로 방어적 반응을 지시하여 새로운 방어지표를 구성하기 위해 단계적 판별분석을 실시하여 BOR, PIM, MAN, RXR, ALC, STR 등 6개 척도로 구성된 판별식을 산출하였다. 이 판별식을 대학생에게 적용한 결과 긍정적 인상관리를 시도한 수검자를 탐지하는 적중률은 87%였으나 정직하게 반응한 대학생을 방어적으로 반응했다고 분류하는 오류부정률도 46.2%로 높았다.

홍상황과 김영환(2001)은 대학생을 대상으로 방어적 반응을 지시하여 이를 찾아내기 위한 지표를 구성하기 위해서 판별분석을 실시하였고, 그 결과 WRM, ANX, NON, INF, RXR, DOM, PIM 등 7개 척도로 구성된 판별식을 산출하였고, 이를 타당화하기 위해 검증표본에 적용한 결과 분류정확률은 92.5%였다. 이 지표와 CDF를 비교해 보면 PIM과 RXR은 공통적으로 포함되어 있으나 그 이외의 긍정왜곡을 탐지하는데 효과적인 것으로 지적된 척도들은 상이하다. 이런 결과는 홍상황과 김영환(2001)은 대학생을 대상으로 판별식을 추출한 반면 Cashel 등 (1995)은 교도소 수감자를 대상으로 판별식을 추출한 점에 기인할 것으로 생각된다.

홍상황 등(2010)은 입학사정관제 전형에서 PAI-A 결과가 당락에 영향을 미치는 점을 고려해서 반응한 프로파일, 즉 긍정왜곡 프로파

1) $CDF = 1.67BOR + .97PIM + .72MAN + .60RXR - .52ALC - .68STR$

일을 찾아내기 위한 지표를 구성하였다. 이들은 고등학교 3학년을 대상으로 표준지시 및 긍정왜곡 지시 하에서 실시한 자료를 이용하여 긍정왜곡을 탐지하는 판별식을 고안하였다. 이를 위해 PAI-A의 22개 척도점수를 예언변인으로 하여 단계적 판별분석을 실시하였고, 그 결과 PIM, ALC, STR로 구성된 판별식을 도출하였다(Korean defensiveness function, KDF)²⁾. 그리고 이 판별식을 검증표본에 적용한 결과 판별식만으로 긍정왜곡 집단 및 표준지시 집단을 분류할 수 있는 분류예언율은 각각 79.78%, 77.66%인 반면 판별식과 표준 지시문을 사용하여 계산한 PIM 척도 원점수의 1.5 표준편차에 해당되는 PIM 원점수(20점)를 같이 적용한 결과 긍정왜곡 집단 및 표준지시 집단에 대한 분류예언율이 각각 92.55%, 85.11%로 상당히 증가한 것을 확인할 수 있었다. 이는 다양한 방어반응 탐지지표를 적용할 경우 왜곡 프로파일을 더욱 정확하게 분류할 수 있음을 시사하는 결과이다.

이상의 연구를 살펴본 결과 PAI 프로파일에 작용된 반응태세를 탐색하기 위한 지표를 고안하여 적용하는 것은 매우 중요하다. 특히 최근 들어 우리나라에서는 학교폭력 가해자 또는 피해자, 학교부적응 등 다양한 이유로 인해 청소년들의 심리, 정서적 상태를 평가하는 비율이 증가하고 있고, 그런 장면에서 청소년 성격평가질문지를 사용하는 사례가 높아지고 있다. 예컨대, 소년범이 경찰조사를 받을 때 범죄심리사 등 전문가들이 비행촉발요인 검사(이수정과 조은경, 2005)와 청소년 성격평가질문지(PAI-A)를 실시하여 재범가능성의 정도를 판단한다. 또한 범죄소년에 대한 적합한 처분을 결정하기 위해 검사가 검사결정전조사

를 의뢰하거나 법원에서 판결전조사를 의뢰할 경우 보호관찰소에서는 다면적 인성검사 및 청소년 성격평가질문지를 포함한 다양한 자기보고형 검사를 사용해서 재범예측 및 적절한 처우에 관한 의견을 포함하는 조사서를 작성하고 있다(박은영, 2008). 비행청소년에게 청소년 성격평가질문지를 실시한 연구(이남희, 이봉건, 2009; 고제원, 이규호, 이민희, 2010)들을 살펴보면 절도범, 폭력범, 사기범 등 비행의 종류나 비행심각성에 따라 타당성척도에서 긍정적 인상관리 또는 부정적 인상관리 척도가 유의미하게 상승하였다. 이와 같은 결과는 비행청소년들이 검사결과가 자신의 처분에 영향을 미칠 것을 염려하여 인상관리를 시도했을 가능성을 시사한다고 볼 수 있다. 이런 연구결과를 볼 때 비행청소년 뿐만 아니라 입학사정관제에 응시하는 청소년들이 자신에 대한 긍정적인 인상을 주고자 긍정왜곡을 할 가능성이 있을 것이다. 더구나 입시전형에서 단순한 성적 중심의 평가에서 탈피하여 성적 이외 학생의 환경, 잠재력 및 소질 등을 종합적으로 판단하여 신입생을 선발하는 입학사정관제도(교육인적자원부, 2007)를 적용한다면 학생들의 적성이나 성격을 보다 정확하게 파악하는 것이 중요한 일일 것이다. 그렇지만 학생들이 긍정왜곡 등의 반응태세를 취한다면 학생에 대한 적절한 평가를 하기 어려울 것이다. 이러한 측면에서 본 연구는 성인을 대상으로 도출된 3가지 지표와 고등학생을 대상으로 도출된 지표를 PAI-A 자료에 적용하여 조작적 특징과 상대적 효율성을 검토하여 청소년 성격평가질문지의 긍정왜곡 반응을 효과적으로 탐지할 수 있는 지표를 마련하는데 목적이 있다.

2) $KDF = -1.649 + .148PIM - .083ALC - .062STR$

방 법

연구대상

경남 소재 2개 고등학교의 고등학교 3학년 98명을 대상으로 긍정적 인상관리 지시문을 사용하여 검사를 실시하였고 수집한 자료 중에서 무성의하게 응답하거나 이중으로 응답하거나 누락된 문항의 수가 17개 이상인 자료, 8부를 제외한 90명의 자료를 사용하였다. 그리고 표준지시에 따른 자료는 청소년용 PAI를 표준화할 때 수집한 표준자료(김영환 등, 2006) 중 고등학교 3학년 90명의 자료를 무선 추출 하여 사용하였다. 긍정적 인상관리 지시문 수검자는 남 34명, 여 56명이었고 평균연령은 17.34($SD=.58$), 표준지시문 수검자는 남 44, 여 46, 평균연령은 16.72($SD=.59$)이었다.

도구

PAI-A. 김영환, 김지혜, 오상우, 이수정, 조은경, 홍상황(2006)이 표준화한 청소년 성격평가질문지(Adolescent Personality Assessment Inventory, PAI-A)를 사용하였다. 이 검사는 수검태도를 평가하기 위한 비일관성(ICN), 저빈도(INF), 부정적 인상(NIM), 긍정적 인상(PIM) 등 4개의 타당성척도, 신체적호소(SOM), 불안(ANX), 불안관련장애(ARD), 우울(DEP), 조증(MAN), 망상(PAR), 정신분열병(SCZ), 경계선적 특징(BOR), 반사회적 특징(ANT), 알코올문제(ALC), 약물문제(DRG) 등과 같은 여러 가지 임상적 문제와 진단집단을 측정하고 평가하기 위한 11개의 임상척도, 공격성(AGG), 자살관념(SUI), 스트레스(STR), 비지지(NON), 치료거부(RXR) 등과 같은 치료에서 중요하게 다루어야

할 사항을 평가하기 위한 5개의 치료고려척도, 지배성(DOM), 온정성(WRM) 등 2개의 대인관계척도, 총 22가지 척도로 구성되어 있다. 이 중 10개의 척도는 복잡한 임상적 구성개념을 측정하고 해석하기 위한 하위척도를 포함하고 있다. 그리고 표준화 연구결과(김영환 등, 2006) 검사의 내적 합치도는 ICN과 INF 척도를 제외한 나머지 척도는 .63~.88, 전체 22개 척도의 검사-재검사 신뢰도는 ICN과 INF 척도를 제외하고 .68~.86 범위로 비교적 양호한 편이었다.

지시문. 반응태세를 유도하기 위해 학생들에게 사용한 지시문은 다음과 같다: (1) 표준지시문: “이 검사지에는 당신이 일상생활에서 경험하는 여러 가지 문제들을 항목들로 모아 두었습니다. 이 항목들을 하나하나 읽어 가면서 자신을 가장 잘 나타내 주는 대답을 하나 골라 주십시오.” (2) 긍정적 인상 지시문: 현재 자신의 감정이나 경험, 태도, 신념에 근거해서 반응하지 마시고 “입학사정관제 선발 전형과 같은 시험에서 본 검사의 결과가 당락에 영향을 미칠 수 있다는 점을 고려하여” 각 문항에 반응하여 주십시오.

자료분석

수집된 자료를 분석하기 위해 SPSS 19.0을 이용하여 각 지시문 유형에 따른 척도의 평균을 산출하였고, 각 집단에서 4개 긍정왜곡 탐지지표의 점수를 산출하여 t 검증을 통해 집단 간 차이를 확인하였다. 또한 지시문에 따른 4가지 지표에 대한 효과의 크기를 추정하기 위해 Cohen의 d 통계치를 계산하였고 지표의 특성을 보다 자세하게 비교하기 위해서 작

동특성 곡선(receiver operating characteristics curve, ROC)을 이용해서 ROC 곡선 아래 영역 추정치(estimated area under the ROC curve, AUC)와 민감성(sensitivity) 및 특이성(specificity)을 산출하였다.

결 과

지시문 유형에 따른 PAI-A 전체척도 점수는 표 1과 같다. 그림 1에서와 같이 두 지시문 유형에 따른 프로파일 형태가 상당히 다르다. 표준지시 집단의 경우 전체 척도들이 평균 T 점수인 50점을 기준으로 비교적 큰 분산 없이 평평하다. 반면 긍정왜곡 지시 집단은 긍정적인 인상을 주려고 노력할 경우 상승할 것으로 예상할 수 있는 척도들 즉 PIM, RXR, DOM 및 WRM이 표준지시 집단보다 상당히 상승한 것을 알 수 있다. 또한 일반적인 정신병리를 측정하는 대부분의 임상척도들이 평균보다 매우 낮은 것을 알 수 있다. 한 가지 특이한 점은 MAN이 상승한 점이다. MAN은 원래 조증과 경조증에서 나타나는 다양한 임상적 특징을 측정하지만 자존감의 고양, 활동수준 증가 및 적극적이고 사교적이며 모험적인 특성을 측정하는 문항이 포함되어 있으므로 긍정적인 인상을 시도할 경우 상승할 수 있을 것이다 (홍상환, 2001). 각 지시문 유형에 따른 4개의 긍정왜곡 탐지지표의 평균과 집단 간의 차이를 비교한 값과 추정된 효과의 크기효과 (Cohen's d)는 표 2와 같다. 4개의 긍정왜곡 탐지지표들 중에서 효과크기가 가장 큰 것은 KDF이고, 그 다음은 PIM 원점수인 반면 효과크기가 가장 낮은 것은 CDF였다. 두 집단에서 KDF 값의 차이를 보면 긍정왜곡 집단은 표준

집단 평균의 1.6표준편차만큼 KDF가 큰 것을 알 수 있다. PIM 원점수 차이를 살펴보면 긍정왜곡 집단은 표준집단 평균의 1.4표준편차만큼 크다. 이러한 결과를 같이 고려해 보면 지시문에 따른 효과가 KDF, PIM 및 CDF 지표에 반영되었다는 것을 시사하는 것이다.

긍정왜곡 탐지지표의 특성을 보다 자세하게 비교하기 위해서 ROC(receiver operating character

표 1. 지시문 유형에 따른 척도의 평균과 표준편차

	지시문 유형	
	긍정왜곡	표준
비일관성(ICN)	40.07(8.40)	49.13(10.31)
저빈도(INF)	49.51(8.84)	49.40(11.08)
부정적 인상관리(NIM)	40.34(6.17)	49.93(9.91)
긍정적 인상관리(PIM)	62.74(9.79)	49.09(9.43)
신체적 호소(SOM)	40.56(6.04)	49.08(10.32)
불안(ANX)	37.71(9.28)	49.58(10.50)
불안관련 장애(ARD)	41.82(6.75)	49.69(10.46)
우울(DEP)	36.87(8.67)	49.11(10.30)
조증(MAN)	48.71(6.58)	50.27(9.73)
망상(PAR)	39.80(9.32)	49.76(8.91)
정신분열병(SCZ)	39.59(7.68)	48.85(9.90)
경계선적 특징(BOR)	36.14(9.95)	49.08(9.8)
반사회적 특징(ANT)	39.40(7.65)	48.74(9.12)
알코올문제(ALC)	43.19(3.70)	50.72(10.78)
약물문제(DRG)	45.99(6.30)	50.72(11.06)
공격성(AGG)	41.32(9.04)	48.74(9.40)
자살관념(SUI)	42.22(6.00)	49.28(9.91)
스트레스(STR)	38.17(9.47)	49.91(9.95)
비지시(NON)	40.39(9.03)	48.40(9.65)
치료거부(RXR)	60.10(9.55)	50.07(10.03)
지배성(DOM)	60.71(11.01)	51.29(10.31)
온정성(WRM)	59.78(11.96)	52.17(9.76)

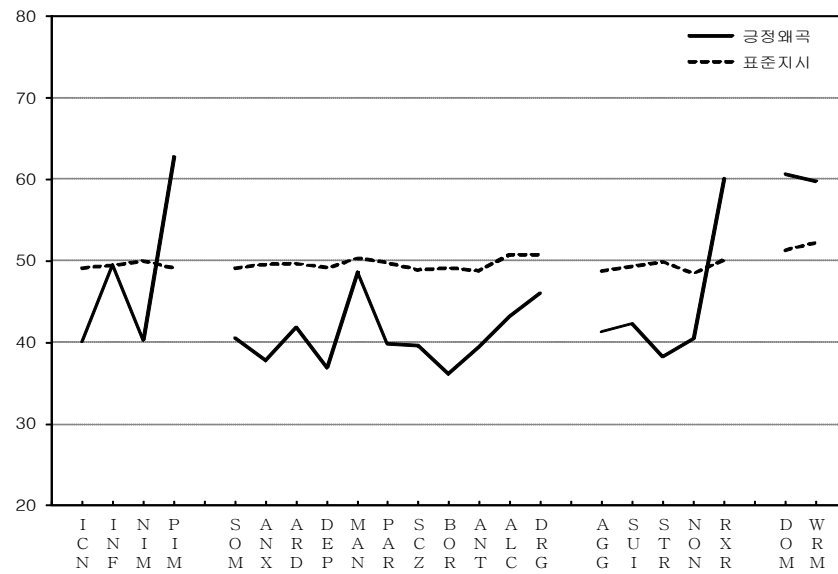


그림 1. 긍정왜곡과 표준 PAI-A 프로파일

표 2. 4가지 긍정왜곡 지표의 평균과 표준편차

	지시문 유형		<i>t</i>	효과크기
	긍정왜곡	표준지시		
PIM 원점수	19.73(4.53)	13.39(4.31)	9.50***	1.44
DEF	5.05(1.76)	3.15(1.69)	4.30***	1.11
CDF	143.76(10.42)	135.63(14.44)	7.22***	.65
KDF	1.68(2.01)	-1.69(1.98)	11.33***	1.70

*** $p < .001$

-istics) 분석을 통해 다양한 분할점수 별로 민감(sensitivity)과 특이성(specificity), 각 분할점수에서 진단적 효율성 또는 탐지효율성의 추정치를 계산하기 위해 Bayes 공리를 적용하여 기저율을 달리하여 계산하였다(표 3). 즉 실제 검사장면 또는 모집단에서 반응왜곡이 일어날 확률이 다르므로 기저율(base rate, BR)을 20%, 30%, 80%로 설정하여 ppp(positive predictive power), npp(negative predictive power)를 산출하였

다. 그리고 이러한 지표들의 ROC 곡선과 AUC(area under curve) 값은 그림 2와 같다. 각 지표의 분할점수를 검증할 때 민감성과 특이성이 높을수록 ROC 곡선이 그래프의 좌측상단에 가까워질 것이고, 그 곡선 아래 영역이 넓어진다. AUC는 척도의 전반적인 효율성을 평가하는데 유용한 지표이다. 완벽한 척도일 경우 AUC는 1.0이고, 유용하지 않은 척도일 경우 .05다. 4개 긍정왜곡 탐지지표의 AUC는 PIM

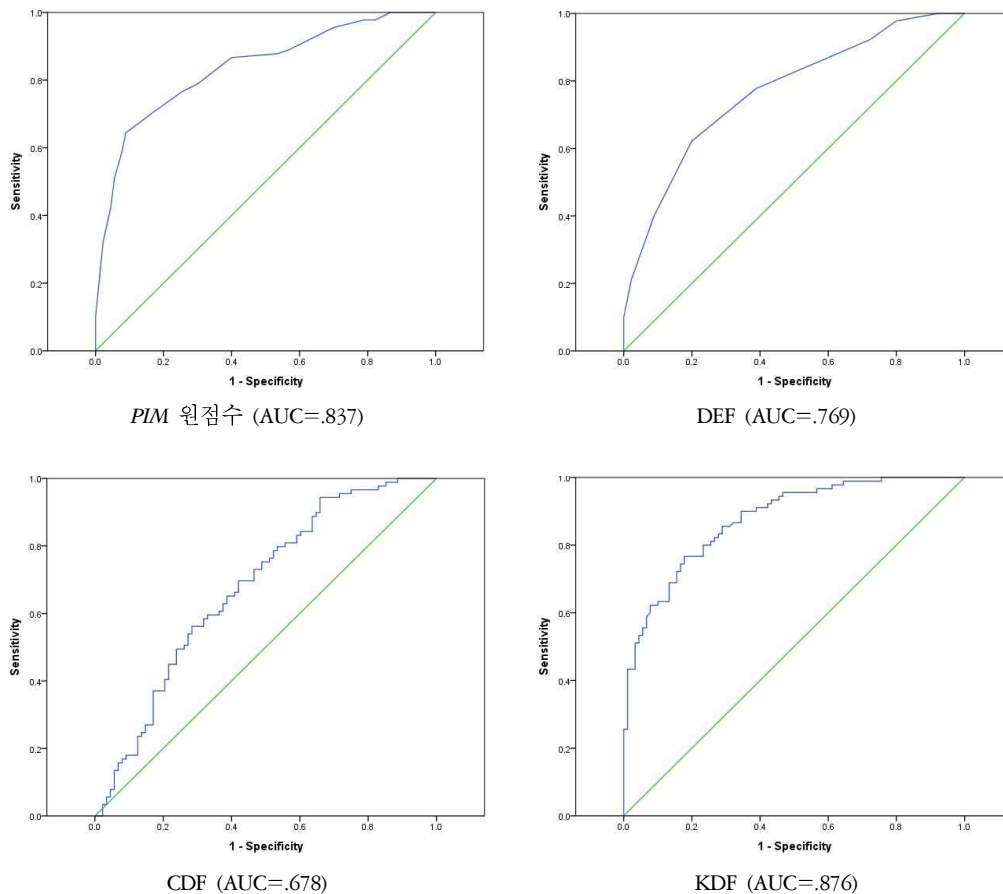


그림 2. 4가지 긍정왜곡 지표의 ROC 곡선

.837, DEF .769, CDF .678, KDF .878이다. 효과 크기에서와 마찬가지로 KDF의 AUC가 가장 높았고, CDF는 낮았다. 그림 2에서도 알 수 있듯이 CDF의 경우 곡선 아래의 영역이 상대적으로 적은 것을 관찰할 수 있다.

청소년 성격평가질문지를 사용하는 실제 장면에서 각 지표를 활용하려면 서로 다른 분할점수의 특성을 이해하는 것이 중요하다. 표 3에는 민감성과 특이성이 가장 높은 분할점수를 진하게 표시하였다. PIM, DEF, KDF는 특정 분할점수에서 민감성과 특이성이 높지만 CDF는 상대적으로 낮다. 표를 보면 기저율이 .2,

.3, .8 일 때 ppp, npp가 매우 다르게 나타나고 있다는 것을 알 수 있다. 즉 최적의 분할점수는 기저율에 따라 반응왜곡 지표의 ppp와 npp를 같이 고려해야 한다는 것을 알 수 있다.

긍정왜곡을 탐지하기 위한 4가지 지표들 간의 관계를 알아보기 위해서 상관을 산출한 결과 표 4에서와 같이 PIM 원점수, DEF 및 KDF 간에는 높은 정적 상관이 지적되었다. 하지만 CDF의 경우 긍정왜곡 지표로 개발되었지만 다른 긍정왜곡 지표와의 상관이 .28~.41로 상대적으로 낮은 것으로 나타났다.

표 3. 4가지 긍정왜곡 지표의 분할점수별 비율추정치

분할점수	민감도	특이성	BR=.2		BR=.3		BR=.8	
			PPP	nPP	PPP	nPP	PPP	nPP
PIM 원점수								
25	.0111	1.000	1.00	.80	1.00	.50	1.00	.20
23	.3222	.9778	.78	.85	.94	.59	.98	.27
22	.4222	.9556	.70	.87	.90	.62	.97	.29
21	.5111	.9444	.70	.89	.90	.66	.97	.33
20	.8111	.9222	.72	.95	.91	.83	.98	.55
18	.7111	.8222	.50	.92	.80	.74	.94	.42
17	.7667	.7444	.43	.93	.75	.76	.92	.44
15	.8667	.6000	.35	.95	.68	.82	.90	.53
13	.8889	.4333	.28	.94	.61	.80	.86	.49
11	.9780	.2110	.24	.97	.55	.91	.83	.71
9	1.0000	.1333	.22	1.00	.54	1.00	.82	1.0
DEF								
8	.1000	1.0000	1.00	.82	1.00	.53	1.00	.22
7	.2111	.9778	.70	.83	.90	.55	.97	.24
6	.4000	.9111	.53	.86	.82	.60	.95	.28
5	.6222	.8000	.44	.89	.76	.68	.93	.35
4	.7778	.6111	.33	.92	.67	.73	.89	.41
3	.9222	.2778	.24	.93	.56	.78	.84	.47
2	.9778	.2222	.24	.98	.56	.91	.83	.71
1	1.0000	.0778	.21	1.00	.52	1.00	.81	1.0
CDF								
173.84	.0000	.9886	.00	.80	.00	.50	.00	.20
155.58	.1348	.9318	.33	.81	.66	.52	.89	.21
150.99	.2247	.8750	.31	.82	.64	.53	.88	.22
144.65	.5056	.7273	.32	.85	.65	.60	.88	.27
140.39	.6517	.6136	.30	.88	.63	.64	.87	.31
138.20	.7079	.5341	.28	.88	.60	.65	.86	.31
135.04	.7978	.4659	.27	.90	.60	.70	.86	.37
134.26	.8090	.4432	.27	.90	.59	.70	.85	.37
125.76	.9551	.2500	.24	.96	.56	.85	.84	.58
117.07	1.0000	.1136	.22	1.00	.53	1.00	.82	1.0
KDF								
2.54	.4000	.9889	.90	.87	.97	.62	.99	.29
2.09	.5111	.9667	.79	.89	.94	.66	.98	.33
1.50	.5889	.9333	.69	.90	.90	.69	.97	.36
.77	.6778	.8667	.56	.91	.84	.73	.95	.40
.16	.7556	.8222	.52	.93	.81	.77	.94	.46
-.38	.8000	.7667	.46	.94	.77	.79	.93	.49
-.68	.8556	.7111	.43	.95	.75	.83	.92	.55
-1.08	.9000	.6556	.40	.96	.72	.87	.91	.62
-1.52	.9333	.5444	.34	.97	.67	.89	.89	.67
-2.55	.9778	.3667	.28	.99	.61	.94	.86	.80

주. BR: base rate, ppp: positive predictive power, npp: negative predictive power

표 4. 긍정왜곡 지표 간의 상관

	PIM	DEF	CDF
PIM 원점수	1.00		
DEF	.74**	1.00	
CDF	.28**	.35**	1.00
KDF	.92**	.74**	.41**

** $p < .01$

논 의

자기보고형 질문지를 사용할 때에는 수검자의 반응태세를 검토하여 검사결과 타당성을 확인한 후 프로파일을 해석하는 것이 중요하다. 최근 학교폭력 가해자 또는 피해자, 학교 부적응 청소년 및 선발전형 등 다양한 장면에서 청소년에게 심리평가를 실시하고 있다. 이러한 맥락이라면 청소년들은 검사결과가 자신의 처분이나 당락에 미칠 영향을 고려하여 긍정적 인상을 주고자 노력하거나 사소한 결점이나 실수를 부인하는 등 긍정왜곡 반응을 할 수 있을 것이다. 그러므로 이러한 긍정왜곡을 효율적으로 탐지하고 판단할 수 있는 지표의 필요성이 제기되었다. 이런 맥락에서 본 연구는 성인을 대상으로 제작된 긍정왜곡 지표와 국내 고등학생을 대상으로 개발된 긍정왜곡 지표를 청소년에게 적용시켜 조작적 특성을 살펴보고 긍정왜곡을 탐지하는 효율성을 비교하고자 하였다. 이를 위해 표준지시 및 긍정왜곡 지시 하에서 반응한 청소년 프로파일과 청소년 성격평가질문지 표준화표본에서 추출한 자료에 PAI의 타당성척도인 PIM, Morey가 제작한 DEF, 판별식으로 구성된 지표 중 성인의 긍정왜곡을 탐지하기 위한 대표적 지표인

CDF, 한국 고등학생을 대상으로 긍정왜곡을 탐지하기 위해 제작된 KDF 등 4개 지표를 적용해서 각 집단에서의 차이와 효과크기 및 조작적 특성과 각 지표의 서로 다른 분할점수의 비율추정치를 산출하여 긍정왜곡 탐지 효율성을 비교하였다.

연구결과를 살펴보면 첫째, 두 지시문 유형에 따른 프로파일을 살펴보면 긍정왜곡 프로파일의 경우 PIM, RXR, DOM, WRM 및 MAN 척도가 상승한 반면 대부분의 임상척도는 표준지시 프로파일보다 상당히 낮은 점이 지적되었다. 이 척도들은 홍상환과 김영환(2001)이 대학생들을 대상으로 긍정왜곡 탐지지표를 구성하기 위해 제작했던 판별식에 모두 포함되어 있다. 이 척도들은 연령에 관계없이 한국인에게는 긍정적으로 지각되는 특성들이 포함되어 있는 것으로 볼 수 있고, 프로파일에서 이 척도들이 상승했다면 긍정왜곡의 가능성을 염두에 두어야 할 것이다.

둘째, 지시유형에 따른 4개 긍정왜곡 탐지 지표의 평균과 집단 간의 차이를 비교하였고 효과크기를 산출하였다. 그 결과 KDF($d=1.70$), PIM($d=1.44$), DEF($d=1.11$), CDF($d=.65$) 순으로 효과가 컸다. 실제로 각 지시유형에서 KDF의 값을 비교해 보면 긍정왜곡지시 집단이 표준지시 집단 평균의 표준편차보다 대략 1.6배 이상 크다. PIM 값을 비교해 보면 긍정왜곡지시 집단이 표준지시 집단 평균의 표준편차보다 1.4배 이상 크다. 이는 수검자들이 긍정왜곡 지시문에 따라 긍정적 인상을 주고자 시도한 점을 두 지표가 효과적으로 반영한다는 것을 시사한다.

셋째, 긍정왜곡 지표의 특성을 보다 자세하게 비교하기 위해서 ROC 분석을 통해 다양한 분할점수 별로 민감성(sensitivity)과 특이성

(specificity), 각 분할점수에서 진단적 효율성 또는 탐지효율성의 추정치를 계산하였다. 실제 검사장면에서 반응왜곡이 일어날 확률이 다르므로 서로 다른 기저율(20%, 30%, 80%)에서의 긍정예언력과 부정예언력을 산출하였고, 각 지표들의 ROC 곡선과 AUC 값을 계산하였다. 그 결과 각 지표의 AUC는 KDF .878, PIM .837, DEF .769, CDF .678이었다. 효과크기와 마찬가지로 KDF의 AUC가 가장 높았고, CDF는 낮았다. 이는 Morey와 Lanier(1998)가 긍정왜곡 지시 하에서 반응한 대학생을 대상으로 PIM, DEF 및 CDF의 효과크기 및 AUC를 산출한 결과와 일치한다. 이들의 연구에서 효과크기는 PIM 1.48, DEF 1.47, CDF 0.71이었다. 또한 AUC는 PIM .948, DEF .941, CDF .878이었다. 따라서 PIM과 DEF는 연령이나 문화권과 관계없이 긍정왜곡을 탐지하는데 효과적이라고 할 수 있을 것이다.

긍정왜곡 지표의 분할점수의 특성을 살펴보면 PIM, DEF, KDF는 특정 분할점수에서의 민감도와 특이성이 높지만 CDF는 상대적으로 낮다. 또한 PIM 원점수와 DEF의 경우 Morey와 Lanier(1998)의 연구에서 최적의 분할점수로 지적된 점수 PIM = 20, DEF = 5와 동일하였다. 하지만 최적의 분할점수라 하더라도 상이한 장면에서는 서로 다른 분할점수가 적용될 수 있다. 즉 입사전형이나 입학사정관제 등의 전형장면에서는 민감도에 치중하겠지만 당락여부를 최종판단하는 단계에서는 특이성이 우선시될 것이다. 또한 긍정왜곡반응이 발생하는 기저율이 다르다면 분할점수의 효율성에도 영향이 있을 것이다. 흔히 긍정적 인상을 주고자 노력할 것으로 예상되는 인사선발 장면이나 재판과정에서 실시된 검사라면 긍정왜곡의 발생기저율이 상당히 높을 것이다. 반면 자발

적으로 심리치료에 참가하고자 내원하는 정신과 외래 장면에서는 왜곡반응의 발생빈도가 낮거나 오히려 부정왜곡의 발생기저율이 높을 수 있을 것이다. 따라서 분할점수를 적용할 때 검사가 실시된 장면을 고려해야 할 것으로 생각된다. 한편 $BR=.2$ 와 $BR=.3$ 간의 ppp 와 npp 차이가 $BR=.3$ 과 $BR=.8$ 간의 ppp 와 npp 차이에 비해 큰 것으로 지적되었다. 본 연구에서는 강제로 긍정왜곡을 지시했던 점으로 인해 실제 기저율에 따른 차이가 반영되지 않았을 가능성이 있을 것으로 추정된다. 따라서 다양한 크기와 유형의 표본을 대상으로 이런 부분을 확인해 볼 필요가 있을 것으로 생각된다.

4개의 긍정왜곡 지표 간의 관련성을 살펴본 결과 KDF와 PIM, DEF 간에는 강한 정적 상관이 지적되었고 특히 KDF를 구성하는 척도 중 PIM이 포함되므로 PIM과 KDF 간의 상관은 .92로 매우 높았다. 반면 CDF는 상관이 .24 ~ .41로 낮았다. 상관패턴 뿐만 아니라 전반적인 결과에서 CDF는 Cashel 등(1995)이 긍정왜곡을 탐지하기 위한 지표로 산출하였으나 한국 고등학생들의 긍정왜곡을 탐지하는 효율성은 저조한 것으로 나타났다. CDF는 PAI 개발 초기에 제작된 긍정왜곡 지표이지만, 앞서 언급한 바와 같이 CDF는 교도소 표본의 자료를 대상으로 판별식을 추출하였기 때문에 일반 고등학생에게 적용하는데 어려움이 있을 것이다. 따라서 한국의 교도소 표본 또는 비행청소년에게 적용시켜 문화권에 따른 차이 또는 표본에 따른 차이를 살펴볼 필요가 있을 것이다.

이상의 연구결과들을 종합해 보면 PIM, DEF 및 KDF는 청소년의 긍정왜곡 반응을 탐지하고 판단하는데 효율적인 것을 알 수 있다. 특히 한국 고등학생들을 대상으로 제작

된 KDF는 다른 지표들 보다 더 효율적인 것으로 지적되었다. 이런 결과를 볼 때 한 가지 지표 보다는 여러 지표를 동시에 적용해서 긍정왜곡 반응을 탐지한다면 오류율을 줄일 수 있을 것으로 생각된다.

본 연구에서는 긍정왜곡 지표의 효율성을 비교하였으나 청소년들의 경우 자신의 정서적, 심리적인 문제를 더욱 부정적으로 지각하거나 자신이 심한 피해를 입은 것으로 보이거나 결과를 부정적으로 왜곡시킬 수도 있으므로 이후 연구에서는 부정왜곡 지표들의 효율성을 비교해 볼 필요도 있을 것으로 생각된다. 또한 본 연구에서는 임의로 긍정왜곡의 기저율을 설정하여 ppp와 npp를 산출하였으나 경찰조사 장면, 입시전형 또는 판결전조사 등 실제로 긍정왜곡이 발생할 수 있는 다양한 장면에서 청소년의 긍정왜곡 기저율을 확인해 봄으로써 청소년과 관련된 평가 장면에서 보다 실질적인 기저율을 적용할 수 있을 것으로 생각된다.

참고문헌

- 고제원, 이규호, 이민희 (2010). 비행청소년의 범죄유형별 PAI 프로파일. 법학연구, 39, 157-176.
- 교육인적자원부 (2007). 대입전형선진화를 위한 입학사정관제 지원계획안.
- 김영환, 김지혜, 오상우, 이수정, 조은경, 홍상황 (2006). 청소년 성격평가 질문지 실시요강. 서울: 학지사.
- 박은영 (2008). 판결전조사에 사용하는 심리검사의 이해. 범죄예방정책연구, 20, 201-249.
- 이남희, 이봉건 (2009). 비행청소년의 범죄유형에 따른 재범 위험성 차이에 관한 연구 - 비행촉발요인 조사서와 PAI 검사결과를 중심으로. 한국심리학회지: 사회 및 성격, 23, 127-140.
- 이수정, 조은경 (2004). 경찰단계에서의 소년범 위험성 평가를 위한 비행촉발요인 조사도구 개발. 한국심리학회지: 사회 및 성격, 19, 27-43.
- 홍상황, 김영환 (2001). PAI 무선반응과 인상관리의 탐지: II. 탐지지표의 구성을 중심으로. 한국심리학회지: 임상, 20, 751-761.
- 홍상황, 한태희, 장현정 (2010, 10). 진주교육대학교 입학사정관제 전형에서 인성검사 (PAI-A)의 활용방안. 2010 입학사정관제 지원사업 연구과제 보고서.
- 홍상황 (2001). PAI 반응태세의 탐지와 진단변별 기능분석. 경북대학교 대학원 박사학위논문.
- Bare, R. A., & Wetter, M. W. (1997). Effects of information about validity scales on underreporting of symptoms on the Personality Assessment Inventory. *Journal of Personality Assessment*, 68, 402-413.
- Bornstein, R. F., Rossner, S. C., Hill, E. L., & Stepanian, M. L. (1994). Face validity and fakability of abjective and projective measures of dependency. *Journal of Personality Assessment*, 63, 363-386.
- Cashel, M. L., Rogers, R., Sewell, K., & Martin-Cannici, C. (1995). The Personality Assessment Inventory and the detection of defensiveness. *Assessment*, 2, 333-342.
- Fals-Stewart, W. (1996). The ability of individuals with psychoactive substance use disorders to escape detection by the Personality Assessment

- Inventory. *Psychological Assessment*, 8, 60-68.
- Greene, R. L. (1991). *The MMPI-2/MMPI-A: An interpretive manual*. Boston: Allyn & Bacon.
- Lanyon, R. I., Dannenbaum, S. E., Wolf, L. L., & Brown, A. R. (1989). Dimensions of deceptive responding in criminal offenders. *Psychological Assessment*, 1, 300-304.
- Morey, L. C. (1991). *Personality Assessment Inventory Manual*. Odessa, FL: Psychological Assessment Resources.
- Morey, L. C. (1993, August). *Defensiveness and malingering indices for the PAI*. Paper presented at the meeting of the American Psychological Association, Toronto, Canada.
- Morey, L. C. (1996). *An interpretive guide to the Personality Assessment Inventory*. Odessa, FL: Psychological Assessment Resources.
- Morey, L. C., & Lanier, V. W. (1998). Operating characteristics for six response distortion indicators for the Personality Assessment Inventory. *Assessment*, 5, 203-214.
- Peebles, J., & Moore, R. J. (1998). Detecting socially desirable responding with the Personality Assessment Inventory: The Positive Impression Management scale and the Defensiveness index. *Journal of Clinical Psychology*, 54, 621-628.
- Rogers, R., Sewell, K. W., Morey, L. C., & Ustad, K. L. (1996). Detection of feigned mental disorders on the Personality Assessment Inventory: A discrimination analysis. *Journal of Personality Assessment*, 67, 629-640.
- Shedler, J., Mayman, M. & Manis, M. (1993). The illusion of mental health. *American Psychologist*, 48, 1117-1131.
- 원 고 접 수 일 : 2012. 06. 14.
수정원고접수일 : 2012. 07. 29.
최종게재결정일 : 2012. 08. 02.

Efficiencies of positive distortion indicators of PAI-A

Eun-Young Park

Deagu Probation Office

Sang-Hwang Hong

Dept. of Education
Chinju National University of Education

This study evaluated the operating characteristics and the efficiencies of PAI-A positive distortion indicators developed on adults(PIM, DEF, CDF) and KDF constructed on Korean high schoolers. High school students administrated PAI-A under two conditions: positive impression instruction and standard instruction. The profile of positive dissimulation had high score on PIM, RXR, DOM, WRM, MAN and almost all clinical scales were lower than average. Through analysis of effect size and ROC curves to evaluate the sensitivity and specificity of the measures using different cutoff scores, the order of the efficiency of positive distortion indicators are KDF, PIM, DEF and CDF. Review of the identification rate estimates for each cutoff score indicates that the sum of sensitivity and specificity on specific cutoff score of KDF, PIM, DEF are high but that of CDF is relatively low. These results suggest that KDF is effective on detecting the positive impression management as much as PIM and DEF developed on adults. Also, adapting more than one indicator are more effective to monitor the response distortion accurately.

Key words : Adolescent Personality Assessment Inventory(PAI-A), positive dissimulation, response distortion, ROC, AUC