

## 한국판 CPGI와 원본척도(CPGI) 간 측정동등성 및 점수연계 가능성\*

이 순 목\*      윤 수 철      차 정 은      김 종 남      여 성 철  
성균관대학교      리더스인사이트      이화여자대학교      서울여자대학교      건국대학교

본 연구에서는 캐나다 도박행동 변별척도(CPGI)와 이를 번역한 한국판 CPGI(KCPGI) 간 측정 동등성을 확인하고 두 척도 간 점수를 연계시키고자 하였다. 그 이유는 점수 연계를 통하여 두 척도 간 직접적인 점수 비교가 가능하며 국가 간 유행률 비교도 가능하기 때문이다. 측정 동등성 확인을 위해 CPGI에 응답한 캐나다인과 KCPGI에 응답한 한국인 간에 전체, 남자, 여자 집단으로 나누어 3회에 걸쳐 다집단 분석을 실시하였는데, 세부적으로 측정틀 동일성, 측정단위 동일성, 측정원점 동일성, 요인분산 동일성, 및 요인평균 동일성의 검증을 실시하였다. 캐나다 자료 2,089명, 한국 자료 882명의 남녀 통합자료에 대하여 다집단 분석을 실시한 결과, 측정틀 동일성은 성립되는 것으로 판단 가능하였고, 측정단위 동일성은 9개 문항 중 6개 문항에 대해서 성립되었다. 측정원점 동일성은 위의 6개 문항에 대해 성립되었으나, 요인분산 동일성은 기각되었다. 요인평균 비교에서는 한국집단이 캐나다 집단보다 더 높은 것으로 나타났다. 그 다음으로 전체 자료를 남/녀로 구분하여 각 성별로 캐나다와 한국 간 다집단 분석을 하였다. 남자집단 분석을 위해 캐나다 남성 1,019명, 한국 남성 544명의 자료가 사용되었는데, 그 결과는 전체집단의 분석결과와 대부분 유사하게 나타났다. 그 다음으로 여자집단 분석을 위해 캐나다 여성 1,069명, 한국 여성 337명의 자료가 사용되었다. 그 결과 위의 결과와 약간의 차이가 있었는데, 구체적으로 측정틀 동일성 검증에서는 만족할 만한 합치도 지수를 보였으며, 측정단위 동일성은 9개 문항 중 6개 문항에 대해서 성립되었다. 측정원점 동일성은 이들 6개 문항에서 지지되었으며, 요인분산 동일성도 성립되었다. 요인평균 비교에서는 역시 한국이 더 높았다. 마지막으로 KCPGI와 CPGI의 점수 연계를 시도하였고, 그 결과로 두 척도 점수의 전환표를 제시하였다.

주요어 : CPGI, KCPGI, 측정 동등성, 점수연계, 다집단 분석

\* 본 연구는 연구자들이 사행산업통합감독위원회의 용역을 받아 수행한 2011년 「전국민 대상 도박문제 선별척도 및 기준점수 타당화 연구」 보고서의 일부를 토대로 작성된 것임.

† 교신저자(Corresponding Author) : 이순목 / 성균관대학교 심리학과 / 인재개발학과 / (110-745) 서울시 중로 구 명륜동 3가 / Tel : 02-760-0492 / E-mail : smlyhl@chol.com

본 연구에서는 한국어로 번역된 캐나다 도박행동 변별 척도(KCPGI)의 측정적 속성이 원척도(CPGI: Canadian Problem Gambling Index) (Ferris & Wynne, 2001a, 2001b)와 어느 정도 유사한지 측정 동등성(measurement equivalence, measurement invariance, or factorial invariance)을 검증하였다. 모든 변안척도는 원척도를 변안하는 과정이 타당한지 그리고 실제 자료에 적용시켜 얻은 측정구조가 원척도와 동등한지를 제시해야 한다. 변안과정에 대해서는 번역-역번역 과정의 타당성을 제시하게 되며(김아영, 임은영, 2003; Brislin, 1970; Eremenco, Cella, & Arnold, 2005), 측정구조에 대해서는 변안척도와 원척도 간의 측정동등성 또는 요인동등성을 제시해야 한다.

변안과정의 타당성을 보이기 위해서는 변안과정이라는 질적인 측면의 상세한 서술이 필요하며, 측정동등성을 보이기 위해서는 먼저 원척도의 자료(원자료 또는 요약자료)를 확보해야 하며 언어가 다른 두 국가 간 다집단 분석을 실시하여 문항 자료에 기초한 심리측정적 속성의 동등성을 입증해야 한다. 이러한 동등성이 성립하면 두 집단 간에 해당 구성개념에 대한 측정체계가 동등하므로 점수의 의미가 동등하며, 두 집단 간 측정치 비교에 기초하여 이론변수(요인) 수준에서의 집단차이를 추론할 수가 있고, 나아가 두 집단 간 점수연계(linking)를 시도할 수가 있다. 점수 간 연계가 되지 않으면 변안척도에서의 점수가 원척도에서 몇 점인지를 이야기할 수 없고, 임상장면 척도의 경우 국가 간에 유행률 비교가 거의 불가능하다. 현재 국내의 임상장면에서 많은 변안척도들이 이러한 문제를 안고 있는 가운데, 본 연구에서는 국민일반을 대상으로 도박행동 수준을 변별하는 목적으로 개발된

최초의 척도인 원본 CPGI와 한국판 CPGI(김아영, 차정은, 권선중, 이순목, 2011)에 대해서 측정동등성 및 점수의 연계를 시도하였다.

## CPGI의 개요 및 원본과 한국판 간 측정 동등성 검증의 필요성

### CPGI 개요

CPGI는 캐나다에서 일반인을 대상으로 개발된 도박 문제행동 변별척도로서 그 의도는 기존의 척도들을 보완하고자 한 것이다. 기존의 도박행동 척도들이 주로 임상장면에서 고수준의 도박행동(중독)을 변별하는 것에 목적을 둔 것이 사회공중건강의 관점에서 지향하는 목적에는 제한적이라고 보고, 고수준뿐만 아니라 저수준(문제가 없는 도박행동)까지도 변별하고자 한 것이다. 이렇게 CPGI가 도박행동의 극단만을 대상으로 하지 않고 보다 폭넓게 측정하게 됨에 따라 기존 척도들인 SOGS(The South Oaks Gambling Screen; Lesieur & Blume, 1987), MAGS(Massachusetts Gambling Screen; Shaffer, LaBrie, Scanlan, & Cummings, 1994), NODS(The National Opinion Research Center DSM Screen for Gambling Problems; National Gambling Impact Study Commission, 1999) 등의 목적보다는 다목적적인 척도가 되었다. 즉, 문제의 식별을 넘어 문제의 예방을 위한 정책적 시사를 제공하기 위한 목적이 포함된 척도가 되었다고 할 수 있다. CPGI는 현재 여러 나라(예: 캐나다, 미국, 영국, 호주, 노르웨이, 아이슬란드)에서 사용하고 있으며 그 측정구조는 그림 1과 같다.

국내에서도 이미 CPGI의 세 가지 번역본이 사용되고 있는 중이다. 문화관광연구원(2008)

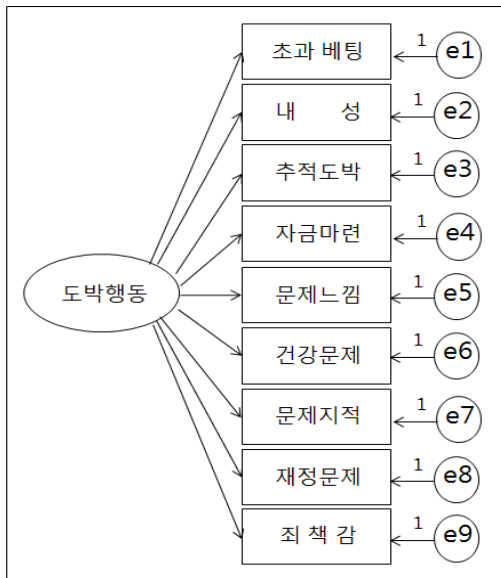


그림 1. CPGI 1요인구조 모형(9문항)

에서 번역한 것에 대한 타당화 예비단계가 학술지에 보고되었고(이경희, 2009), 2010년의 번역본(김교현, 조성겸, 권선중, 이동형, 2010)이 도박실태조사에 사용된 적이 있다. 끝으로 2011년에 김아영, 권선중, 차정은, 이순목(2011)이 역번역 절차 및 국내의 하위집단들 간에 다집단 분석을 통하여 타당화를 한 것이 가장 최근의 번역본이다. 2008년의 번역본은 충분히 타당화되었다고 볼 수 없고(이홍표, 2008), 2010년의 번역본은 아직 타당화 연구와 관련하여 학술지에 보고되지 않았으며, 가장 최근의 번역본(김아영 등, 2011)이 비교적 충실한 타당화를 보고하고 있다.

원본과 번역본간 측정동등성 검증의 필요성

심리검사의 사용에서 측정기능의 동등성에 관심을 가지는 경우는 다음의 경우이다. 즉, 상이한 집단 간에 하나의 검사가 사용될 때와

유사한(대안적인) 두 개의 검사가 하나의 자료에 사용되었을 때이다. 어떤 경우든, 상이한 자료집단 간에 측정구조(요인구조)가 동등한지의 여부에 관심을 가진다. 측정동등성에 대한 연구가 서구에서는 1990년대부터, 국내에서는 2000년대에 들어와 많이 실시되고 있다. 이렇게 측정기능의 동등성을 검증하는 이유는 하나의 측정도구가 집단에 걸쳐 일반화하여 사용될 수 있는지 또는 유사한 측정도구 간에 어느 정도까지 동등성이 성립하고 어느 정도까지 차별성이 있는지를 파악하기 위한 것이다.

집단간 척도의 일반화 가능성에 대한 연구는 문화가 다른 하위집단 간에 척도나 모형이 비교문화적으로 일반화될 수 있는지를 검토하는 방향으로 확대되고 있다. 집단 간에 문화가 다르다면 통계적으로는 검사점수의 분포가 달라지며, 검사제작에서는 규준이나 기준점수가 달라진다. 과거에는 규준이 다를 수 있는 가능성을 확인하기 위하여 하위집단 간 분산(variance)을 비교하고 분산이 크게 다르지 않을 때 평균비교를 실시하였다. 이러한 전통적인 방법론은 구조방정식 모형의 발전과 더불어, 집단간 측정모형의 비교(예: 김정호 등, 2005; 손은영, 차정은, 김아영, 2007; 방희정 등, 2007; 하정, 홍세희, 유성경, 2005; 조용래, 김정호, 2002; 한소영, 신희천, 2007; 최대정, 조현주, 박동건, 2005; Sohn, 2001)로 대체되고 있다. 측정모형의 비교를 통상적으로는 측정동등성 검증이라고 한다. 측정모형의 동등성 검증은 이론모형의 동등성 검증까지 포함하는 다집단 분석의 한 부분이다(이순목, 금은희, 이찬순, 2010; 이순목, 김인혜, 2009; 이순목, 김한조, 2011).

현재 국내 연구에서 하위집단 간 비교에 측

정동등성 검증을 이용하고 있으나, 이러한 접근은 국가 간에도 가능하다. 예로서 Cheung, Murrmann, Murrmann, 및 Becker(2004)는 미국과 홍콩 간 비교문화적 연구에서 측정동등성 확보가 없이 집단 간 점수를 비교하는 것이 위험함을 지적하였다. 즉, 측정의 동등성이 확보되지 않고는 미국 점수와 홍콩에서의 점수는 비교할 수 없다는 것이다.

이러한 측정동등성의 문제는 외국 척도를 번안하여 사용하는 경우 반드시 검토해야 하는 부분이다. 외국척도를 번안하여 국내 임상 장면에서 적절한 기준이나 기준점수를 만들어 사용하는 것이 현재의 관행이다. 그러나 연구 결과의 국제화 또는 국제 유병률 비교라는 과제를 다루기 위해서는 국내판의 타당화만으로 부족하다. 예로서 CPGI를 캐나다에서 사용해서 얻은 점수와 국내 번역본에서의 점수를 같은 의미로 해석가능할 수 있어야 연구 결과가 쉽게 국제화될 수 있고, 국제 유병률 비교도 가능하다. 이것은 번역본의 타당화를 넘어서, 원본척도와 측정동등성의 확보를 필요로 한다.

도박연구 분야에서는 그 동안 임상장면에서 NODS(National Gambling Impact Study Commission, 1999)나 SOGS(Lesieur & Blume, 1987)도 국제적으로 번안되어 사용되었으나, 일반인을 대상으로 한 CPGI가 보다 많은 나라에서 사용되고 있고 그에 따라 국가마다 유병률도 각기 다른 것으로 보고되고 있다. 통상적으로 유병률에서 언급되는 도박행동의 수준을 3수준과 그 아래 2수준이라고 할 때, 일반인에게 CPGI를 사용해서 조사한 3수준과 2수준의 유병률 결과는 다음과 같다. 캐나다의 경우 Ferris와 Wynne(2001b)에서 0.9%와 2.4%로, 한국의 경우 문화관광연구원(2008)에서는 2.3%와 7.2%로,

김교현 등(2010)에서는 1.7%와 4.4%로 보고하고 있다. 이렇듯 유병률 조사결과가 국가 간에도 다르고 번역본 간에도 차이를 알 수 있다.

각국에서 캐나다 원본을 번안해서 또는 영어권의 경우 그대로 사용하고 있으나 국가 간 또는 번역본 간 측정동등성의 검증이 되지 않은 상태이므로 유병률이 다른 것이 과연 척도의 기능이 달라서인지, 척도의 기능은 같은데 실제로 집단 간 즉, 국가 간 또는 연구 표본 간 도박행동 수준에 차이가 있어서 그런 것인지를 알 수가 없다. 만일에 척도의 기능이 다르다면 집단 간 점수의 의미가 다르므로 해석을 위해서 연계할 수 있는 가능성이 없고, 척도의 기능이 같다면 점수의 의미가 같으므로 집단 간 동등한 의미를 가지는 점수 짝을 찾아 연계하는 것이 가능하다. 연계가 가능하면 집단은 달라도 하나의 측정체계를 가지고 점수가 산출되므로 상이한 집단 간 직접적 비교가 가능하다. 따라서 본 연구에서는 국내 번역본과 캐나다 원본척도 간의 측정동등성을 검토하여 어느 정도 CPGI의 측정기능이 일반화될 수 있는지, 그리고 국내에서 얻은 점수를 캐나다에서의 점수로 전환하여 해석하는 것이 가능한지를 검토하는 것이 목적이다.

## 본 연구에 사용된 다집단 분석의 틀

다집단 분석은 구조방정식모형의 일반화 또는 차별화를 위한 방법론으로서, 상이한 모집단에 동일한 측정치를 사용하여 수집된 자료에 모형을 합치시켜 일반화 또는 차별화를 시도하는 것이다(이순목, 1993; Jöreskog, 1971; Sorbom, 1974; Meredith, 1993). 일반화 연구의

예로는 척도나 모형에 대하여 집단 간 일반화의 시도, 비교문화적 검토, 또는 기존의 검사 양식과 동일한 구성개념을 측정하는 대안적 검사 양식의 개발에서 두 양식이 평행한지의 검토 등을 들 수 있다. 또한 차별화된 연구의 예는, 서로 다른 시점에서 모형의 변화를 파악하거나 시점에 따라 진행되는 조직의 변화를 연구하는 경우이다(Vandenberg & Lance, 2000).

다집단 분석은 일련의 순차적인 동일성 검증 단계를 포함하는데, 첫 번째 단계는 측정틀 동일성 검증이다. 집단 간 측정틀의 동일성이 성립되고 나면, 그 다음으로 측정단위의 동일성을 검증한다. 전반적/부분적 측정단위 동일성의 성립은 다음 단계인 측정원점 동일성 검증으로 진행할 수 있게 해준다. 전반적/부분적 측정원점의 동일성이 성립하게 되면 다집단 분석의 궁극적인 목적이라고 할 수 있는 집단 간 요인평균 비교가 가능해진다. 경우에 따라서 측정원점 동일성 성립 여부와 상관없이 집단 간 요인평균 비교를 하는 것이 집단간 정확한 평균비교가 가능하게 됨을 주장하는 연구도 있으나(이순목, 금은희, 이찬순, 2010), 일반적인 경우 요인평균의 선행단계로서 측정원점 동일성을 검증한다. 그 외에도 요인분산/공분산 동일성 검증 및 측정오차분산 동일성 검증 등이 가능하다. 이러한 순차적인 검증단계를 통해, 선행 단계의 검증이 성립된 모형은 다음 단계의 모형과 비교시에 기저모형이 된다.

본 연구에서는 다집단 분석을 이용하여 CPGI 번역본(KCPGI)과 원본 척도(CPGI)간 측정의 체계가 동등한지를 검토하였다. 우선 다집단 분석의 용어에 대한 국내 문헌에서의 번역이 다양하므로(이순목, 김한조, 2011, 표 2

참조), 본 연구에서는 심리측정 이론의 오랜 전통인 수량산출론(scaling theory)(이순목, 2002; Michell, 1986)에 기초하여 번역된 이순목과 동료들(이순목, 금은희, 이찬순, 2010; 이순목, 김인혜, 2009; 이순목, 김한조, 2011)의 번역을 따랐다. 자세한 절차는 방법 부분에서 제시되었다.

## 방 법

### 설계

CPGI와 KCPGI 간에 측정동등성을 검증하는 것은 하나의 척도를 두 집단에 실시하거나 두 개의 척도를 하나의 집단에 실시하여 측정동등성을 검증하는 경우보다 훨씬 어렵다. 두 개의 척도를 두 개의 집단에 실시하는 경우이기 때문이다. 비록 CPGI의 번역과정(김아영 등, 2011)이 타당화되었다고 해도 CPGI와 KCPGI가 완벽하게 동일한 것은 아니므로, 두 개의 검사양식(alternate forms)을 서로 상이한

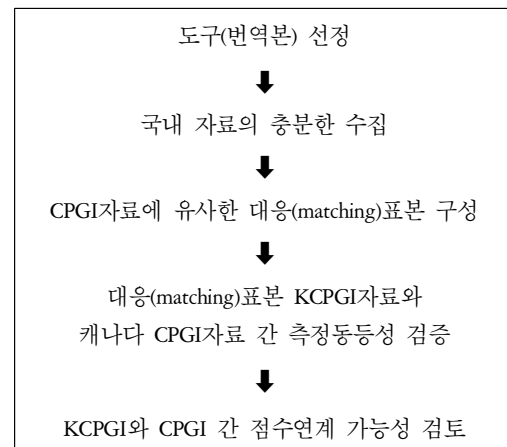


그림 2. 연구설계

집단에 실시한 것으로 볼 수 있다. 본 연구의 설계는 그림 2와 같다.

우선 CPGI의 한국어 번역본 가운데 가장 적절한 것으로 도구를 선정하고, 그 도구를 사용하여 국내자료를 충분한 규모로 수집한 후, CPGI 자료(Ferris & Wynne, 2001b)에서 제시된 수준별 인구비율에 대등하게 표집을 하여 대응표본을 구성한다. 대응표본이 필요한 이유는 실험할 때 집단간 동등성을 확보한 다음 처치를 하는 원리와 유사하다. 즉, CPGI 자료에서 도박행동 수준별 인구비율을 참조하여 국내 자료도 그와 대등한 비율로 구성했을 때, 측정체계의 동등성 검토 과정에서 불필요한 오염을 최대한 배제할 수 있다. 그런 다음 CPGI와 KCPGI 간 측정동등성을 검증하고 그 결과에 따라 점수간 연계 가능성을 검토할 수 있다.

## 도구

도구로서는 이미 국내에 세 가지 번역본(김아영 등, 2011; 김교현 등, 2010; 이경희, 2009)이 있으나 가장 최근에 정교한 역번역 및 상이한 응답자 집단간 문항의 편향여부에 대한 검토 등을 거쳐 보다 충실하게 타당화된 김아영 등(2011)의 것을 선택하였다. 그 번역본은 KCPGI-2011이라고 하는데 본 연구에서는 혼동의 우려가 없을 경우 KCPGI로 기술하였다.

## 국내 자료 수집 및 CPGI자료에 대응하는 표본구성

먼저 국내에서 KCPGI를 실시하여 충분한 수의 자료를 수집하였다. 이때, KCPGI와 함께 제 3의 척도를 실시하였다. 그 이유는 CPGI의

측정체계에 비추어 KCPGI의 측정체계가 동일한지 검증하기 위해 사용될 대응표본 구성은 KCPGI와 독립적으로 이루어져야 하기 때문이다. CPGI에서 사용된 수준별 구분의 개념 및 논리(표 1 참조)를 이 제 3의 척도에 적용하여, CPGI에서의 수준별 인구 구성비율(Ferris & Wynne, 2001b, Table 14)에 준하는 한국 표본(대응표본)의 구성이 필요한 것이다. 대응표본의 응답자들이 KCPGI에 응답한 자료를 가지고 캐나다 CPGI 자료와 비교하여야 측정동등성 검증이 가능하다. 따라서 본 연구에서는 남녀 통합자료에서 CPGI에서의 개념화에 따라 구분된 수준별 인구 구성비율이 CPGI와 대응표본 간에 유사하도록 하였다.

대응표본 구성에 사용될 제 3의 척도로는, CPGI의 개발목적에 참조하여 두 가지를 선택하였다. 김교현, 권선중, 김세진, 이순목(2011)에서 저수준 도박행동자 중심으로 개발된 L척도와 이순목, 김교현 등(2011)에서 고수준 도박행동자 중심으로 개발된 H척도를 KCPGI와 함께 실시하였다. CPGI의 개발목적은, 기존의 고수준 도박행동은 물론 일반 인구에서의 도박행동까지 측정하는 것이다(Ferris & Wynne, 2001b, pp.1-2 참조). 기존의 의학모형에서 병적 도박 중심의 측정을 목표로 했다면, 공중건강 모형에서는 도박행동의 연속선에서 보다 넓은 범위를 포괄하여 측정하는 것을 목표로 한다(예: Neal, Delfabbro, & O'Neil, 2005). 그런데 고수준의 도박행동은 저수준의 도박행동과 현저하게 다르다. 그 차이가 양적으로 제시된 것은 물론(예: 김교현 등, 2011, 표 1 참조), 질적으로도 다른 것을 알 수 있다. 즉, 김교현 등(2011)에서 저수준 도박행동을 잘 나타내는 문항들만으로 척도를 구성하여 자료를 수집하고 요인분석을 한 결과(564명), 도박 무관심이나

표 1. 본 연구와 타 연구자들의 도박행동 수준별 개념에 대한 비교

| 본 연구에서의 용어 —>                | 1수준                                       | 2수준  | 3수준  |
|------------------------------|---|--|--|
| Shaffer와 동료들<br>(1996, 1997) | level 1 gambling,<br>non-problem gambling | level 2 gambling,<br>sub-clinical gambling       | level 3 gambling,<br>pathological gambling |
| NODS(NGISC, 1999)            | low-risk                                  | at-risk, problem                                 | pathological gambler                       |
| CPGI(Ferris & Wynne, 2001a)  | 처음 논의<br>최종 분류                            | non-problem gambling<br>non-problem G+low risk G | at-risk gambling<br>moderate gambling      |
|                              |   |  | problem gambling<br>problem gambling       |

주. NODS에서 low-risk=0점, at-risk=1~2점, problem=3~4점, pathological=5점 이상.

CPGI에서 non-problem=0점, low risk=1~2점, moderate=3~7점, problem=8점 이상.

사교도박과 같이 기존의 의학모형 관점의 개념화(예: DSM-IV의 10가지 준거)에는 전혀 포함되지 않은 요인들이 발견되었다. 도박행동의 고수준 집단과 저수준 집단은 구성개념이 크게 다른 것이다. 따라서 본 연구에서는 KCPGI와 함께 사용된 제 3의 척도로서 L척도(김교현 등, 2011)와 H척도(이순목, 김교현 등, 2011)를 사용하였다.

일반인 대상으로 KCPGI와 L척도를 사용하여 자료수집을 하였고, 고수준에 해당되는 이용객 대상으로는 KCPGI와 H척도를 사용하여 자료수집을 하였다.

#### 캐나다 자료

원본 CPGI에 대한 전체 응답자는 총 3,120명이었으나, 결측치가 모두 제거되고(listwise) 분석에 사용된 자료는 2,089명(남자=1,069, 여자=1,019, 성별 미상=1)이었다(CPGI 제2저자인 Wynne으로부터 요약 자료 받음). 남녀 통합자료에서, 도박행동 수준별(표 1 참조) 구성 비율은 3수준에 0.9%, 2수준에 2.4%, 1수준에 82.7%(low risk: 6.8%; non-problem: 75.9%), 0수준(비도박) 14%였다(Ferris & Wynne, 2001b).

#### CPGI자료에 유사한 대응표본 구성

CPGI에서 사용하는 도박행동 수준별 인구비율에 대응하도록 국내 인구에서의 대응표본을 구성하는 작업은 3단계 과정이 필요하였다. 1단계는 도박행동 수준 분류의 체계를 분명히 하는 의사결정이 필요하고, 2단계는 CPGI 수준분류의 고수준에 해당되는 대응표본 구성이 필요하며, 끝으로 CPGI 수준 분류의 저수준에 해당되는 대응표본 구성이 필요하다.

1단계로 현재 기존척도들에서 사용되는 수준 분류와 CPGI에서의 수준 분류가 상이한 부분이 있으므로(표 1 참조) 국내 전문가들의 워크숍을 통해서 합의를 도출하고자 하였다. 2008년 11월 1일 14인의 전문가들이 모여서 1차 워크숍을 하여 각 수준에 대한 개념화를 하였고(이순목, 현명호, 최삼욱, 김종남, 김수진, 2008), 2009년 12월 5일 일부가 교체된 14인의 전문가들이 다시 모여 최종적으로 표 2와 같은 수준 정의에 합의하였다(이순목, 김교현 등, 2009).

수준 정의 과정에서 2수준과 3수준은 Shaffer와 동료들(1996, 1997)의 체계와 일관성 있게 논의되었고, CPGI와도 대체로 일관성이 있었

표 2. 본 연구에서 사용되는 각 수준의 정의

| 수 준 | 정 의  |
|-----|--|
| 0수준 | 도박행동의 경험이 없거나, 경험이 있더라도 자발적 동기와 흥미가 없고 장차 도박행동에 참여할 의사가 없는 상태                                    |
| 1수준 | 재미나 사교의 목적으로 도박행동을 하며 도박에 사용하는 시간과 금액의 조절이 가능하고 일상생활과 역할기능에 지장이 없는 상태                            |
| 2수준 | 도박에 사용하는 시간과 금액이 증가하고, 도박행동 및 결과를 숨기며, 개인의 조절능력(통제력)을 일부 상실하여 일상생활과 역할기능에 피해가 초래될 정도로 도박에 몰입된 상태 |
| 3수준 | 도박행동으로 인해서 일상생활과 역할 기능이 심각하게 손상되었거나 통제력 상실이 심화된 상태   |

출처: 이순목, 김교헌 등, 2009, p.61, 표 75.

다. CPGI에서의 수준 분류는 ‘Shaffer 등(1997)의 체계에 대응하도록’(Ferris & Wynne, 2001a, p.18) 하였기 때문이다.

그러나, 본 연구에서의 0수준과 1수준은 Shaffer 등(1996, 1997) 및 CPGI와 다르게 정의되었다. 0수준은 Shaffer 등 및 CPGI, 모두 비도박(도박을 전혀 한 적이 없음)을 의미하지만, 그것은 도박을 물리적 대상으로 보는 관점이다. 전문가들의 워크숍에서 ‘도박’ 아닌 ‘도박행동’으로 행동과학적인 개념화를 하면서 다양한 행동은 구성개념에 대한 지표(indicator)로 보게 되었다. 즉, 자료수집 시 목표로 하는 측정대상은 자료에 잠재된 이론개념인 ‘도박행동(복수의 요인으로 분화 가능)’이라는 논의가 힘을 얻어 ‘0’수준 역시 물리적 관점의 0이 아닌 심리적 관점의 ‘0’수준으로 정의하게 되었다. 전혀 도박을 안하는 것은 물론 동기가 없는 도박행동 역시 ‘0’수준으로 보게 된 것이다. 한편 1수준은 동기는 있으나 건전한 동기에 머물러 있는 상태로 한정하게 되었다.

이러한 개념 정의에 따라 2009년의 워크숍에서는, 수준 간 경계점수(기준점수)를 정할

때 사용할 문항별 점수들을 산출하였다. 즉, 수정된 Angoff(1971) 방식을 3라운드에 걸쳐 실시하였는데, L척도 문항들에 대하여 0수준과 1수준 간의 경계에 있는 사람이 보일 수 있는 점수를 산출하였고, H척도 문항들에 대하여 1-2수준 경계, 2-3수준 경계에 있는 사람들이 보일 수 있는 점수를 산출하였다. 자료수집 후 다양한 방법으로, H척도 및 L척도 사용 시 수준을 구분하는 기준점수를 설정하였다(이순목, 김아영 등, 2011, 7장 참조). 도박행동의 고수준과 저수준 간에 질적인 차이가 있으므로, 2, 3수준에 해당되는 사람들에게는 H척도를 실시하고 0, 1수준에 해당되는 사람들에게는 L척도를 사용하여 자료수집을 하는 것으로 논의가 모아졌다.

다음은 2단계로 고수준에 해당되는 도박행동자들의 대응표본을 구성하는 것이다. 우선 H척도를 강원랜드 카지노, 경마 본장 및 장외 매장, 경륜 본장 및 장외 매장에서 이용객 1,047명에게 실시하였고 1,030명의 유효자료를 확보하였다. 이들 자료에 2009년 워크숍에서 도출된 H척도 문항들의 기준점수를 적용하여

2수준과 3수준의 도박자를 식별한 다음 각 수준 해당 응답자 중에서 무선적으로 일부를 뽑아 그들의 KCPGI 점수를 대응표본에 포함하였다. CPGI 자료에서의 비율을 참조하여, 대응표본에서 2수준이 3%, 1수준이 1%가 되도록 하였다.

마지막 3단계는, CPGI 체계에서 저수준에 해당하는 응답자들을 확보해서 그들의 KCPGI 점수를 대응표본에 포함하는 작업이다. 그런데 CPGI의 저수준에 대응하는 표본을 구성하는 데에는 어려움이 있었다. CPGI의 저수준은 표 1에서와 같이 0(비도박)수준, 1-1수준(non-problem), 1-2수준(low risk)으로 되어 있으나, 본 연구에서는 0수준과 1수준으로 구분되며 L장면 인구를 분류할 때 표 2와 같이, CPGI와는 다르게 개념화가 되어 있기 때문이다.

본 연구에서 대응표본을 구성하기 위해서 거칠긴 하지만 CPGI의 0, 1-1, 1-2수준을 본 연구의 0수준 및 1수준에 맞게 해석하고 그에 준해서 본 연구에서의 0수준과 1수준의 비율을 산정하여 그 비율만큼을 대응표본에 포함하였다. CPGI에서 0수준(14%)은 전혀 도박을 해 본 적이 없는 사람, 1-1수준(75.9%)은 도박을 하되 문제가 없는 도박자이다. 또한 1-2수준(6.8%)은 비문제 도박을 저위험(low-risk)으로 보는 NODS에서의 명칭을 빌려와서 1수준의 상단부분을 범주화한 것이다. 즉, CPGI의 1-2수준도 종래의 비문제 도박에 포함되긴 하되 척도(27점 만점)에서 1~2점인 사람을 대상으로 하는 범주이다. 이제 과제는 CPGI의 1-1수준(0점) 75.9% 중 어느 정도가 심리적 0수준에 해당할 것인가를 추정하는 것이다. 1-1수준에서 심리적 0수준에 해당되는 인구가 제외되면 그 나머지와 1-2수준이 함께 본 연구의 1수준에 포함된다. 양적인 정보가 없는 가운데, 오

류의 가능성을 최소화하기 위하여 중앙값(약 38%)을 추정치로 사용한다면, 대응표본에서 0수준은 52(=14+38)%를, 1수준은 44.8(=38+6.8)%를 차지하게 된다. 이것은 1-1 수준에서 심리적 0수준에 해당되는 사람들의 비율이 알려지지 않은 가운데 중앙값을 사용한 결과이므로, 본 연구에서 0수준과 1수준의 비율은 대략 유사한 정도로 보기로 하였다. 그렇다면 대응표본에서 2수준(3%)과 3수준(1%)이 차지하는 비율을 뺀 나머지인 96%의 반인 48%을 각각 0수준 및 1수준으로 볼 수가 있다. 그에 따라, L척도 응답자 중 표 2의 0수준 및 1수준 해당자는 각각 대응표본 크기에서 48%씩을 차지하도록 구성되었고, H척도 응답자 중 2수준 및 3수준 해당자는 대응표본에서 각각 3%와 1%가 되도록 하였다. 그 결과 0수준 423명, 1수준 423명, 2수준 27명, 3수준 9명이 표집되었다(882명, 남자=554, 여자=337, 성별 미상=1). 전체 응답자로부터의 표집은 엑셀에서 난수를 부여해서 수동으로 표집하였다. 대응표본에 포함될 때는 저수준 집단은 L척도 점수를, 고수준 집단은 H척도 점수를 참조해서 표집되었으나 일단 대응표본에 포함된 후에는 그들의 KCPGI 점수만을 CPGI와의 측정동등성 분석에 사용하였다. CPGI 및 KCPGI의 남녀통합 요약자료는 부록에 제시하였다.

#### CPGI와 KCPGI간 측정동등성 검증 단계

##### 측정틀 동일성(configural invariance) 검증

본 연구에서는 CPGI 9개 문항에 대하여 1요인(‘도박행동’)이 설정된 모형이 측정의 틀이다. 측정틀 동일성이 성립되면, 집단 간에 개념 이해의 틀이 동일하다고 해석할 수 있다.

본 연구에서는 CPGI에 응답한 캐나다인과

KCPGI에 응답한 한국인 간에 전체, 남자, 그리고 여자 자료에 대해 세 번에 걸친 다집단 분석을 실시하였다.

#### 측정단위 동일성(metric invariance) 검증

측정단위 동일성은 구조방정식모형에서 요인계수를 집단 간 동일하게 제약함으로써 검증할 수 있다. 여기서 요인계수는 형태계수(pattern loading)이므로 측정단위 동일성은 요인 형태계수 동일성인데(Alwin & Jackson, 1981), Yoon과 Millsap(2007)은 형태 동일성(pattern invariance)으로 부르기도 한다. 본 연구에서는 CPGI에 대한 캐나다와 한국 자료 간에 도박행동에 대한 개념구조가 동일하게 되면(측정틀 동일성), 각 집단 내 도박행동(이론변수)이 CPGI 문항들(측정변수)로 구체화되는 방식이 동일한지를 측정단위 동일성 검증을 이용하여 확인하였다.

#### 측정원점 동일성(scalar invariance) 검증

측정원점은 요인의 척도 상에서 '영점'에 해당되는 지점(Bollen, 1989)으로, 측정원점 동일성 검증은 요인점수 '0'에 해당하는 측정치(측정변수 방정식에서의 절편)가 집단 간에 동일한지를 검증하는 것이다. 본 연구에서 캐나다와 한국자료가 도박행동에 대하여 동일한 원점을 갖는지 검증하였으며, 도박행동에 대하여, 측정틀, 측정단위 그리고 측정원점이 동일하게 되면, 도박행동(이론변수)에서 동등한 수준에 있는 응답자들은 측정변수인 척도점수에서도 동일한 수준의 값을 가진다고 볼 수 있다.

#### 요인분산 동일성(factor variance) 검증

요인분산은 진점수 사용의 폭, 즉 구성개념

의 척도 상에서의 변산 정도를 나타낸다. 분산이 클수록 집단 내 구성개념 척도에서 1단위의 간격이 넓다는 것을 의미한다. 본 연구에서 캐나다와 한국 간 요인분산 동일성이 성립하게 되면, 두 집단에서의 개념구조에서 도박행동의 변산이 유사함을 의미한다.

#### 요인평균(factor mean) 비교

집단 간 측정체계 즉, 측정틀, 측정단위 그리고 측정원점이 동일하면 요인평균을 비교하는 것이 의미가 있게 된다. 집단들 중 한 집단을 참조집단(요인평균=0)으로 고정하고 측정원점 동일화 제약을 하게 되면 나머지 집단들의 요인평균들이 참조집단에 대비되는 값으로 산출되며 이들의 크기를 비교하기 위해서 Cohen (1988)의 효과크기( $d$ )를 계산하여 그 크기 정도를 해석한다. 본 연구에서 도박행동에 대한 캐나다와 한국 집단 간 측정체계가 동일하고, 요인평균의 차이가 크지 않다면, 이는 두 집단 모두 도박에의 접근 및 행동적으로 관여하는 정도가 유사함을 의미한다.

#### 모형 합치도 지수의 선택적 사용

모형 합치도 지수는 자료와 모형의 합치도 정도를 수치로 나타낸 값으로 모형을 평가하는 데 사용된다. 단일 집단에서 주로 사용하는 합치도 지수들로는 표본 크기에 민감한 검증용 합치도인 카이제곱치, 그리고 이를 보완할 수 있는 판단적 합치도로서 표본 크기에 상대적으로 덜 민감한 TLI(터커-루이스 계수, 비표준화합치도 NNFI, Bentler & Bonett, 1980), RMSEA(개략화오차평균, Steiger & Lind, 1980), CFI(비교합치도, Bentler, 1990), SRMR(표준화 원소간 평균차이, Bentler, 1995) 등이 있다. 이 합치도 지수들은 다집단 분석에서 기저모형이

없는 측정틀 동일성 검증 시 모형의 합치도 판단에서도 사용된다. 이들 합치도 지수들에 대하여 최근의 보다 엄격한 기준제시(Hu & Bentler, 1999)가 있긴 하나, 전통적인 기준은 CFI와 TLI는 보통 .90 이상일 때 적합하다고 보며, RMSEA는 .08 이하, SRMR은 .10 이하일 경우 적절하다고 본다(Vandenberg & Lance, 2000). 본 연구에서는 이 전통적인 기준값을 참고하였다.

측정단위 동일성 검증부터는 기저모형과 비교 검증되는 제약모형 간의 합치도 차이( $\Delta$ 합치도)를 참조하는데, 이들은 단일집단 분석에서의 합치도들과 독립적이면서 서로 중복이 되지 않아야 한다. Cheung과 Rensvold(2002)는 그러한 기준을 감안하여 부적 합치도에 기초한  $\Delta\chi^2$ , 정적 합치도에 기초한  $\Delta$ CFI,  $\Delta$ mc, 및  $\Delta$ gamma hat을 추천하였다. 단, gamma hat에 대해서는 Cheung과 Rensvold(2002)에서 집단 수효가 반영되지 않은 공식을 사용한 결과를 기초로 기준값을 추천했다는 문제가 있어, 본 연구에서  $\Delta$ gamma hat을 제시는 했어도 해석에서 크게 참고하지 않았다. RMSEA의 경우 다집단 분석에 대해 제시된 기준치는 없으나 참고로 함께 제시하였다.

#### 집단 간 점수 연계(linking)

본 연구에서 비교되는 자료집단은 원척도에 대한 캐나다 자료와 번역척도에 대한 한국자료로서 검사이론의 관점에서 보면 두 가지 관점에서 비교할 수가 있다. 첫째는 동일한 검사에 대하여 상이한 모집단(예: 남/녀, 한국/캐나다)에서 자료를 구한 것으로 본다면 문항이나 검사에서 집단별 편향(bias)이나 차별적 기능(differential item functioning)이 있는지를 검토

해야 한다(Holland & Thayer, 1988). 결과로 차별적 기능의 문항이 없거나, 약간은 있어도 문항 아닌 척도 수준에서의 동등성이 성립하면 두 집단 간 척도점수는 같은 의미로 해석이 가능하다. 또 하나의 관점은, 하나의 개념을 측정하는 검사에 대한 2개의 상이한 양식(예: A형/B형, 한국판/캐나다판)으로 보는 것이다. 이때는 양식 간에 상이한 점수분포를 가져올 수 있으므로 적절한 방식으로 자료수집을 하여 두 점수분포 간에 연계화(linking)를 해야 할 것이다. 즉, 동일한 의미를 갖는 점수 짝들을 찾는 과정인 동등화가 필요하다(Kolen & Brennan, 2004).

그런데 어느 관점도 본 연구에 정확히 맞지는 않는다. 첫째 관점을 취하자면 CPGI와 KCPGI가 동일한 검사라는 확인이 있어야 한다. 비록 번안과정이 타당하고 측정구조에서 측정 동등성이 밝혀지고(차별기능 없음) 요인 평균이 동일하다고 해도 척도가 사용되는 문화적 환경이 다를 수 있으므로 완벽한 동등이란 어려운 일이다. 두 번째 관점은 ‘도박행동’이라는 개념을 측정하는 2개의 양식이라고 할 때, 동등화를 위한 자료수집이 선행되어야 동등화가 가능하다. 그런데 본 연구에서와 같이 두 개 양식 간에 공통문항이 전혀 없는 가운데, 상이한 집단에서 자료가 수집되었다면 연계화를 할 수가 없다. 그렇다면 본 연구에 가까운 것은 첫번째 관점인데, 본 연구에서 측정 동등성이 지지되는 강도에 따라 연계가 허용되는 정도가 달라질 것이다. 즉, CPGI와 KCPGI에 의한 측정이 동등한 정도가 곧, 두 척도 간 점수의 호환성을 좌우할 것이다. CPGI의 한국어 번안과정이 타당화된 가운데(김아영 등, 2011), 본 연구에서 캐나다와 한국 자료 간 측정 동등성이 충분히 지지되면 두

척도 간 점수의 호환 내지 연계 가능성을 시도할 수 있다.

### 결 과

3차례에 걸쳐 다집단 분석을 하였다. 전체 자료, 남자집단, 여자집단을 사용한 분석이었다. 각 분석에서 세부적으로는 측정틀 동일성 검증, 측정단위 동일성 검증, 측정원점 동일성 검증, 요인분산 동일성 검증, 요인평균 비교를 하였다. 소프트웨어로는 Mplus Version 6을 사용하였다.

#### 남녀통합 자료에 대한 분석

남녀 구분 없이 캐나다 자료 2,089명, 한국 자료 882명에 대하여 다집단 분석을 하였다.

#### 측정틀 동일성 검증

측정틀 동일성을 검증하기 전에 먼저 요인에 척도제공을 위한 참조지표 결정은 Vandenberg (2002)가 제안한 방법, 즉 집단별로 탐색적(사각회전) 요인분석을 하여 집단 간에 요인계수가 크고 유사한 문항을 참조지표로 결정하는 방법을 선택하였다. 그 결과 8번 문항을 측정틀 동일성 검증을 위한 참조지표로 결정하였다. 각 집단에 1요인 모형을 합치시킨 결과를 표 3에 제시하였다.

표 3을 보면, 두 집단 모두 TLI와 CFI가 .90 이상, RMSEA와 SRMR이 .08 이하로 만족할 만한 합치도 지수를 보였다. 따라서 캐나다, 한국 두 국가의 전체집단에 대한 측정틀 동일성 검증을 실시하였다. 전체집단의 모형에 대한 전반적 합치도들을 볼 때 두 집단 간 측정틀 동일성이 성립한다고 판단하여 다음 단계의 검증을 실시하였다. 결과는 표 4, 표 5, 및 표 6에 제시되었다.

표 3. 모형 합치도 및 측정틀 동일성 검증결과(남녀통합)

| 모형   | $\chi^2$ | df | p    | RMSEA | 90% CI      | TLI  | CFI  | SRMR |
|------|----------|----|------|-------|-------------|------|------|------|
| 캐나다  | 419.985  | 27 | .000 | .083  | (.077~.091) | .909 | .932 | .037 |
| 한국   | 148.971  | 27 | .000 | .072  | (.061~.083) | .909 | .932 | .038 |
| 전체집단 | 568.956  | 54 | .000 | .080  | (.074~.086) | .909 | .932 | -    |

주. 캐나다(n=2,089), 한국(n=882)

표 4. 캐나다와 한국 간 측정단위 비교(남녀통합)

| 문항  |      | 1    | 2    | 3    | 4    | 5    | 6    | 7    | 8    | 9    |
|-----|------|------|------|------|------|------|------|------|------|------|
| 캐나다 | 요인계수 | .224 | .142 | .171 | .069 | .165 | .143 | .199 | .132 | .352 |
|     | 표준오차 | .007 | .005 | .006 | .003 | .004 | .004 | .008 | .003 | .013 |
| 한국  | 요인계수 | .117 | .142 | .171 | .163 | .165 | .143 | .199 | .132 | .173 |
|     | 표준오차 | .008 | .005 | .006 | .009 | .004 | .004 | .008 | .003 | .013 |

주. 측정단위가 다른 경우 음영을 하였음.

표 5. 동일화 제약 결과(남녀통합)

| 모형                              | $\chi^2$ | df | p    | RMSEA | CFI  | mc    | Gamma hat |
|---------------------------------|----------|----|------|-------|------|-------|-----------|
| 모형1: 측정틀 동일화제약                  | 568.956  | 54 | .000 | .080  | .932 | .9169 | .9284     |
| 모형2: 전반 측정단위 동일화제약              | 916.631  | 62 | .000 | .096  | .887 | .8660 | .8866     |
| 모형3: 부분 측정단위 동일화제약 <sup>a</sup> | 766.989  | 61 | .000 | .088  | .907 | .8879 | .9044     |
| 모형4: 부분 측정단위 동일화제약 <sup>b</sup> | 697.076  | 60 | .000 | .085  | .916 | .8983 | .9129     |
| 모형5: 부분 측정단위 동일화제약 <sup>c</sup> | 602.211  | 59 | .000 | .079  | .928 | .9126 | .9248     |
| 모형6: 측정원점 동일화제약 <sup>d</sup>    | 657.705  | 64 | .000 | .079  | .922 | .9049 | .9184     |
| 모형7: 요인분산 동일화제약                 | 783.960  | 65 | .000 | .086  | .905 | .8860 | .9028     |

주. 캐나다( $n=2,089$ ), 한국( $n=882$ ), 1종오류 수준=.05

- a. 4번 문항을 자유모수로 해제
- b. 4번 문항과 9번 문항을 자유모수로 해제
- c. 4번, 9번, 그리고 1번 문항을 자유모수로 해제
- d. 이미 모형5에서 3개 문항이 단위차별성이 있어서 원점 동일성 검증에서도 이들 3개는 자유모수로 하고 나머지 6개 문항에서만 원점 동일성 제약을 가하였음.

표 6. 측정 동일성 검증 결과(남녀통합)

| 검증대상                | 비교        | $\Delta\chi^2$ | $\Delta df$ | p    | $\Delta CFI^a$ | $\Delta mc$ | $\Delta Gamma\ hat$ | 해석 |
|---------------------|-----------|----------------|-------------|------|----------------|-------------|---------------------|----|
| 전반 단위 동일성           | 모형1 대 모형2 | 347.675        | 8           | .000 | -.045          | -.051       | -.0418              | 기각 |
| 부분 단위 동일성1          | 모형1 대 모형3 | 198.033        | 7           | .000 | -.025          | -.029       | -.0240              | 기각 |
| 부분 단위 동일성2          | 모형1 대 모형4 | 128.120        | 6           | .000 | -.016          | -.019       | -.0155              | 기각 |
| 부분 단위 동일성3          | 모형1 대 모형5 | 33.255         | 5           | .000 | -.004          | -.004       | -.0036              | 유지 |
| 원점 동일성 <sup>b</sup> | 모형5 대 모형6 | 55.494         | 5           | .000 | -.006          | -.008       | -.0064              | 유지 |
| 요인분산 동일성            | 모형6 대 모형7 | 126.255        | 1           | .000 | -.017          | -.019       | -.0156              | 기각 |

주. 캐나다( $n=2089$ ), 한국( $n=882$ )

- a. Cheung과 Rensvold(2002)가 제시하는 기준치:  $\Delta CFI=-.01$ ,  $\Delta mc=-.02$ ,  $\Delta gamma\ hat=-.001$ . 이 값들 보다 크면 (0에 가깝거나, 0보다 큰 값) 기저모형에 비해 제약모형의 합치도가 큰 차이없거나, 나은 것으로 본다.
- b. 단위 동일성 있는 6개 문항에서의 원점 동일성임.

**측정단위 동일성 검증**

검증방식은 전반적 제약방식(HC: Holistic Constraints)을 사용하였고(김한조, 2010; Lee & Kim, 2011; Stark, Chernyshenko, & Drasgow, 2006), 이론변수에 대한 척도제공은 Yoon과

Millsap(2007)의 방식을 사용하였다. 즉, 참조집단인 캐나다집단의 요인분산을 1.0으로 고정하고, 한국집단의 요인분산을 자유모수로 하였으며, 두 집단 간에 요인계수들에 동일화제약을 한 모형이다. 전반적 측정단위 동일성을

검증하기 위한 동일화 제약을 하고(표 5, 모형 2) 검증해 본 결과(표 6) 측정단위 동일성이 기각되었다. 수정지수(MI)를 감안하여 4번 문항을 자유모수로 해제하여 부분적인 단위 동일화제약1(모형3)을 실시하였다.

그러나 모형3 또한 기각이 되었기 때문에 Cheung과 Rensvold (2002)가 제시하는 합치도 기준치에 적합한 모형을 찾기까지 MI값이 큰 차례대로 하나씩 자유모수로 하여 검증하였다. 우선 9번을 자유모수로 하여 검증하였고, 거기서 제약된 요인계수의 MI값을 검토하여 가장 큰 1번 문항을 자유모수로 하여 부분단위 동일성 검증을 실시하였다(표 6의 부분단위 동일성 1, 2, 3). 결과로 6개의 문항에서만 두 집단 간 단위동일성이 성립되었다. 측정단위 차별성이 있는 세 문항은 표 4에서 음영으로 표시되었다.

**측정원점 동일성 검증**

측정단위 동일성 검증에서 사용된 Yoon과 Millsap의 방식을 계속 유지하여 측정원점 동일성 검증까지 실시하였다. 1번, 4번, 9번 문항을 제외한 6개 문항에 대하여 전반적 측정원점동일성 검증을 실시한 결과, 두 집단 간 측정원점 동일성 검증이 성립하였다(표 6의 원점 동일성).

따라서 6개 문항에서 도박행동의 심리적 수준이 0(무관심)일 때 응답자들이 문항에 반응하는 기본 위치는 캐나다와 한국 집단 간에 동일하다고 볼 수 있다.

**요인분산 동일성 검증**

앞서의 단위 동일성과 원점 동일성 검증에서는 Yoon과 Millsap의 방식을 사용하였기에 참조지표를 설정하지 않고 참조집단(캐나다)의 요인분산을 1.0으로 고정하였다. 그러나 이제 두 집단의 요인분산을 모두 자유모수로 하여 각기 계산해야 하므로 두 집단에서 참조지표에 해당하는 측정변수의 요인계수를 1.0으로 고정할 필요가 있다. 그에 따라 요인분산 동일성 검증을 실시하기 전에, 측정틀 동일성 검증에서 사용한 8번 문항을 참조지표로서 사용하였다. 두 집단 간 요인분산(캐나다=.017, 한국=.041)을 비교한 결과 요인분산 동일성은 기각되었다. 즉 요인분산이 한국자료에서 캐나다에 비해 2배를 조금 넘었는데, 이는 한국 집단에서의 도박행동 측정에서 변산도가 상대적으로 큰 것을 의미한다.

**요인평균 비교**

참조집단(캐나다 집단)의 요인평균을 0으로 고정하고 원점 동일화 제약이 주어지면 요인평균은 참조집단에 대비되는 값으로 산출된다. 앞에서 요인분산 동일성이 기각되었기 때문에 원점 동일성 모형(모형 6)으로 돌아가서 두 집단의 요인평균을 살펴보았다. 모형 6에서 캐나다 집단 요인평균이 0이고, 한국집단의 평균은 .100( $p=.000$ )이었다. 이 차이의 효과크기를 파악하기 위해서 Cohen의  $d$ (Cohen, 1988)를 계산하였다. 그 결과,  $d=.644$ 로 두 집단의 요인평균이 중간크기 정도의 차이를 보인다는

표 7. 집단 간 문항 평균점수(남녀통합)

| 문항    | 1    | 2    | 3    | 4    | 5    | 6    | 7    | 8    | 9    | 척도평균  | 표준편차  |
|-------|------|------|------|------|------|------|------|------|------|-------|-------|
| 캐나다집단 | .090 | .046 | .077 | .015 | .028 | .026 | .060 | .021 | .129 | .491  | 1.458 |
| 한국집단  | .098 | .121 | .293 | .127 | .196 | .102 | .213 | .107 | .238 | 1.495 | 1.937 |

것을 알 수 있었다. 참고로 문항 및 척도평균을 표 7에 제시하였다.

요인평균과 척도 총점 평균에서 한국 집단이 캐나다 집단보다 더 높았다. 측정의 틀, 단위, 그리고 원점에서 두 나라 간에 상당한 정도의 동일성이 지지되므로 측정치에 의한 추론이 요인점수에 의한 추론과 일관성 있는 것이다. 전체 응답자들의 척도 평균이 캐나다가 .491인데 비해 한국이 1.495로서 캐나다 집단보다 1점 정도 높다.

#### 남/녀 각 집단에 대한 분석

그 다음으로 남자집단과 여자집단에 대하여 각각 CPGI와 KCPGI 간 다집단 분석을 하였다. 앞서의 남녀통합 자료에서와 동일한 방식으로 실시하였고, 그 결과 역시 유사하다. 특히 남자집단 결과가 더욱 유사하였다. 남/녀 집단 각각에 대한 요약자료는 제 1저자에게 요청하면 제공이 가능하다.

#### 남자집단 간 분석

CPGI에 대한 캐나다 남성 응답자는 총 1,419명이었으나, 분석에 사용된 자료는 1,019명이었다. 이에 반하여 한국 남성자료는 544명이었다. 측정틀 동일성 검증을 하기 전에 집단별 탐색적 요인분석을 하여 두 집단에 동일한 참조지표를 사용하기로 하였다. 그 결과 8번 문항을 선택하였다. 캐나다와 한국 남자집단에 1요인 모형을 적용시킨 결과, 전반적 합치도들이 대체로 만족할 만한 합치도 지수를 보였다. 전반적 측정단위 동일성은 기각되었고, 1번, 4번, 9번을 제외한 6개 문항에서만 두 집단 간 단위동일성이 성립됨을 파악할 수 있었다.

측정단위 동일성이 성립하는 6개 문항에 대하여 전반적 측정원점동일성 검증을 실시한 결과, 동일성이 지지되었다. 따라서 측정단위 동일성이 성립하는 6개 문항에서, 도박행동의 심리적 수준이 0(무관심)일 때 응답자들이 문항에 반응하는 기본 위치는 두 집단 간에 동일함을 의미한다.

두 집단 간 요인분산(캐나다=.027, 한국=.056) 동일성은 기각되었고, 요인분산은 한국 자료에서 캐나다에 비해 약 2배를 넘었다. 이는 도박행동 개념의 변산도가 한국집단에서 상대적으로 큰 것을 의미한다. 요인점수 수준에서 캐나다 평균이 0이고 한국의 평균이 .117 ( $p=.000$ )로서 Cohen의  $d$ 로 계산한 결과, 효과크기가 .608로 두 집단의 요인평균이 중간크기 정도의 차이를 보인다는 것을 알 수 있었다. 척도평균은 캐나다가 자료에서 .7인데 비해 한국 자료에서 1.8 정도로서 한국이 약 1점 더 높다.

#### 여자집단 간 분석

CPGI에 대한 캐나다 여성 응답자는 총 1,700명이었으나 분석에 사용된 자료는 1,069명이었다. 그에 반해 한국 여성자료는 337명의 것이었다. 측정틀 동일성 검증을 하기 전에, 집단별 탐색적 요인분석을 하였고 결과로 8번 문항을 참조지표로 선정하였다. 캐나다와 한국 여자집단에 1요인 모형을 합치시킨 결과, 전반적 합치도들은 두 집단 간 측정틀 동일성이 성립한다고 보기에 완벽한 정도는 아니지만 어느 정도 만족할 만한 수준이었으며, 두 국가 간 자료의 비교목적상 다음 단계의 검증을 실시하였다. 전반적 측정단위 동일성은 기각되었고 1번, 4번, 및 8번을 제외한 6개의 문항에서만 두 집단 간 단위동일성이 성립되

었다.

측정단위 동일성이 성립하는 6개 문항에 대하여 전반적 측정원점 동일성 검증을 실시한 결과 동일성이 지지되었다. 이는 측정단위 동일성이 성립하는 6개 문항에서 도박행동의 심리적 수준이 0(무관심)일 때 응답자들이 문항에 반응하는 기본 위치는 두 집단 간에 동일함을 의미한다.

그 다음 두 집단 간 요인분산(분산=.051) 동일성이 성립하였고, 요인점수 수준에서 캐나다 평균이 0일 때 한국평균은 .12( $p=.000$ )로서 Cohen의  $d$ 는 .531이었다. 이는 중간 정도의 효과크기로, 두 집단의 요인평균에는 차이가 있는 것으로 판단되었다. 척도평균은 캐나다 자료에서 .334였고, 한국 자료에서 1.037로서 한국이 약 0.7점 더 높다.

#### KCPGI 점수와 CPGI 점수의 연계(linking)

KCPGI와 CPGI 자료가 완벽하게 동등성을 보이는 것은 아니지만 측정 동등성이 충분히 지지되고 있다. 이를 바탕으로 남녀 통합자료에서 KCPGI와 CPGI 점수의 연계를 시도하였다. 도박행동이라는 이론변수가 측정변수로 변하는 과정을 나타내는 전환식은 요인분석이라는 1차식의 모형이다. 즉, CPGI 점수와 KCPGI 점수가 모두 도박행동(F)의 1차식으로 표현되고, 그에 따라 CPGI와 KCPGI 점수 간에도 1차식의 관계가 성립함을 의미한다. 도박행동이라고 하는 요인의 점수(F)에서 하나의 값은 CPGI 점수(Y)로도 표현되고, KCPGI 점수(X)로도 표현된다. 만일에 어떤 도박행동자의 요인점수 F가 표준화 점수이고(실제로 요인은 표준화 변수임), 측정의 오차를 감안하여 관찰된 표준점수가 Z라 하자. 그 Z가 캐나다에서

는 CPGI로, 한국에서는 KCPGI로 측정되면서 각각 원점수로는 Y와 X가 얻어진다고 하자. 그러면 Z와 X 그리고 Z와 Y 간에 다음 등식이 성립한다.

$$\frac{Y - \bar{Y}}{S_Y} = Z_Y \dots \dots (1)$$

$$\frac{X - \bar{X}}{S_X} = Z_X \dots \dots (2)$$

그런데 평균적으로 CPGI의 요인값(F)이 0일 때 KCPGI의 요인값(F)은 .10이었다(남녀통합자료의 요인평균비교 참조). 그렇다면 관찰된 표준점수에서도,  $Z_Y = 0$ 일 때  $Z_X = .10$ 이 기대되므로 둘 간의 관계는 아래와 같다.

$$Z_Y + .10 = Z_X \dots \dots (3)$$

식 (1), (2), 및 (3)에 표 7의 값을 대입하면  $(Y-.491)/1.458+.10=(X-1.495)/1.937$  즉,

$$Y = .7527X - .7761 \dots (4)$$

로서 Y와 X간의 1차식이 된다. 이것이 전환식으로, KCPGI 점수를 CPGI 점수로 바꾸려면 X의 값에 KCPGI 점수(0점~27점까지)를 대입하면 된다. 이것은 검사이론의 연계화(linking) 이론에서 추정(calibration)(Kolen & Brennan, 2004)의 개념을 적용한 것이다. 전환식에 X값의 전체범위인 0에서 27점까지를 대입하면 Y값을 얻을 수 있고 CPGI와 KCPGI 간 전환표를 표 8과 같이 만들 수 있다.

또한 대응표본에서 KCPGI의 점수분포를 표 9에 제시하였다. 표 8을 보면 CPGI로 2점(KCPGI로는 3점)이 될 때까지가 두 집단 간에

1점 차이가 관찰되는 경우이고 이때까지 대응 표본에서의 누적빈도는 86.2%이다. 두 집단 간에 점수 차이가 없는 0점 응답자들이 48%이고 캐나다 보다 한국에서 4점 이상의 응답자(전체의 13.8%)의 점수가 좀 높아도 전체적으로는 평균이 1점 차이밖에 나지 않는 것이

다(표 7 참조). 그렇다면 표 8의 전환표는, KCPGI에서 비교적 빈도가 높은(CPGI 보고서에서는 수준별 빈도만 제시되었고 점수별 빈도는 제시되지 않았음) 0, 1, 2, 3, 4점을 보이는 응답자들에 대한 전환이 가장 정확할 것이고, 식(4) 역시 이들의 점수를 대폭 반영하는 전환식이라 할 수 있다. 따라서 표 8에서 KCPGI로 5점부터의 범위에 대한 전환은 그 이전 자료에 준하여 통계적 연장(projection)을 한 것에 불과하므로 주의해서 참조해야 할 것이다.

그런데 현실적으로 표 8에서 중요한 부분은 CPGI 점수로 3점(2수준이 되는 기준점수)과 8점(3수준이 되는 기준점수)이다. 고수준에 대한 유병률은 대체로 2수준과 3수준을 기준으로 보고하기 때문이다. 그 두 점수에 대응하는 KCPGI 점수는 표 8에서 각각 5점과 11점이다. 이것이 현실적으로 타당화되는지에 대한 검토를 해보았다. 즉, 한국 자료에서 2수준과 3수준은 H척도의 실시결과를 참조해서 대응표본의 4%(882명 가운데 35~36명)가 되도록 표집된 것이다. 이들이 표집될 때 KCPGI 점수를 참조한 것은 아니었으나, 표 9의 결과를 보면 KCPGI 5점에서 25점까지가 36명으로서, KCPGI 5점을 2수준(CPGI로 3점)의 시작으로 볼 수 있다. 그리고 3수준은 대응표본의 1%(882명 가운데 9명)가 되도록 표집된 것인데, 표 9의 결과를 보면 KCPGI 점수로 11점

표 8. KCPGI 점수와 CPGI 점수의 연계

| KCPGI | CPGI | KCPGI | CPGI |
|-------|------|-------|------|
| 0     | 0    | 14    | 10   |
| 1     | 0    | 15    | 11   |
| 2     | 1    | 16    | 11   |
| 3     | 2    | 17    | 12   |
| 4     | 2    | 18    | 13   |
| 5     | 3    | 19    | 14   |
| 6     | 4    | 20    | 14   |
| 7     | 5    | 21    | 15   |
| 8     | 5    | 22    | 16   |
| 9     | 6    | 23    | 17   |
| 10    | 7    | 24    | 17   |
| 11    | 8    | 25    | 18   |
| 12    | 8    | 26    | 19   |
| 13    | 9    | 27    | 20   |

주. CPGI에서 0점: non-problem gambling,  
 1-2점: low-risk G, 3-7점: moderate risk G,  
 8점 이상: problem G.

표 9. 대응표본에서 KCPGI 점수분포

| 점수  | 0    | 1    | 2    | 3    | 4    | 5    | 6    | 7    | 8    | 9    | 10   | 11   | 15   | 16   | 21   | 25    | Total |
|-----|------|------|------|------|------|------|------|------|------|------|------|------|------|------|------|-------|-------|
| 빈도  | 423  | 151  | 99   | 87   | 86   | 2    | 5    | 4    | 7    | 6    | 3    | 2    | 3    | 1    | 2    | 1     | 882명  |
| %   | 48.0 | 17.1 | 11.2 | 9.9  | 9.8  | .2   | .6   | .5   | .8   | .7   | .3   | .2   | .3   | .1   | .2   | .1    | 100.0 |
| 누적% | 48.0 | 65.1 | 76.3 | 86.2 | 96.0 | 96.2 | 96.8 | 97.3 | 98.1 | 98.8 | 99.1 | 99.3 | 99.6 | 99.7 | 99.9 | 100.0 |       |

이상이 9명이 된다. 즉, 3수준의 시작(CPGI로 8점)은 KCPGI로 11점이 된다. 이러한 결과는 표 8에서처럼 KCPGI 점수 0, 1, 2, 3, 4, 5, 그리고 11점을, 각각 CPGI의 0, 0, 1, 2, 2, 3, 그리고 8점에 대응하여 사용하는 것에 무리가 없음을 시사한다.

## 논 의

본 연구는 캐나다 도박행동 변별척도 CPGI와 이를 번역한 KCPGI 간 측정 동등성을 확인하고 두 척도 간 점수를 연계시키고자 실시되었다. 도박 수준별 인구비율을 유사하게 조정하여 캐나다인 2,089명의 자료와 한국인 882명의 자료가 분석에 사용되었고, 분석방법으로 구조방정식 모형에서의 다집단 분석을 실시하였다.

분석 결과, 본 연구에서 KCPGI와 CPGI의 측정의 틀이 동일하고, 측정단위와 측정 원점에서 완벽하진 않으나 상당한 동등성이 확인되었다. 즉, 개념구조가 같고 측정의 단위와 원점이 문항들의  $\frac{2}{3}$ 에서 동등한 것으로 나타났다. 또한 요인 분산의 차이가 있긴 하나(한국과 캐나다 간에 2배 정도) 실질적 차이로 간주되지 않는 4배 이내(Johnson & Wichern, 2001)이므로, 한국과 캐나다 간에 '도박행동'이란 개념에 대하여 응답자들이 가지고 있는 측정의 체계는 동등하다고 볼 수 있다. 즉, 척도가 집단에 따라 차별적으로 높거나 낮은 점수를 산출하지는 않는다고 할 수 있다. 한편, 한국 표본에서 요인 평균이 캐나다 표본에 비해서 높게 나왔는데, 이는 두 나라 간 국민일반의 도박행동 수준에서 한국이 더 높음을 의미한다. 그에 따라 개발된 전환식에 의해 KCPGI

와 CPGI 간 점수연계를 시도하였고, KCPGI 점수로 0, 1, 2, 3, 4, 5, 그리고 11점은 CPGI 점수로 각각 0, 0, 1, 2, 2, 3, 및 8점인 것으로 나타났다. 통상 도박 유병률이 2수준(CPGI로 3점)과 3수준(CPGI로 8점)에 대해서 보고되는 것을 감안하면 아주 유용한 전환 결과라 할 수 있다. 또한 저수준에 대해서도 CPGI에서 사용하는 0, 1, 2점에 해당하는 KCPGI 점수를 알 수가 있게 되었다. 이러한 점수 변환을 통해 국내에서 얻은 점수를 원본척도의 점수로 전환할 수 있었다.

또한 전체 집단에 대한 분석 외에도 성별에 따른 차이를 밝히고자 남녀 집단을 구분하여 다집단 분석을 실시하였다. 그 결과, 남녀 집단 간에 약간의 차이가 나타났다. 남자 집단의 경우, 전체집단에서 나타난 결과와 대부분 일치하는 결과를 보였으나, 여자 집단의 경우에는 요인분산 동일성이 성립되었고 9번 문항에서의 반응이 조금 달랐다. 남성에 비해 여성은 한국과 캐나다 공통으로 평균이 낮고 점수의 변산이 작기 때문에, 두 나라 여성 간 비교는 남성의 경우보다 더 유사한 결과를 보인 것이다.

이상의 결과에서 다음과 같은 본 연구의 의의를 찾을 수 있다. 첫째, 본 연구의 주요 내용은 두 국가 간 척도의 측정 동등성이고 점수 연계 가능성은 부수적 내용이지만, 실제로 기여하는 비중으로 보면 후자가 더 중요할 수 있다. 측정의 동등성에 대한 검증은 서구에서 이미 1990년대에 많은 소개가 되었고 실제 적용에 있어 꽃을 피웠으므로 특별히 새로운 방법이 아니다. 국내에서도 일찍이 소개되었고(예: 이순목, 1993), 2000년대에 와서는 심리학, 교육학 분야에서 각기 10개가 넘는 경험 논문들이 다집단 분석(손원숙, 2002), 잠재평균 분

석(홍세희, 황매향, 이은설, 2005), 구인동등성(손은영, 차정은, 김아영, 2007), 요인구조 동등성(최대정, 조현주, 박동건, 2005) 등의 제목으로 발표되었다. 그런데 이러한 연구들에서의 목적은 대체로 측정구조의 일반화에 있으나, 본 연구는 측정구조의 일반화에 기반하여 두 측정체계 간 점수연계를 시도했는데 새로운 의의를 찾을 수 있다. 심리학의 각 분야에서 유사한 목적으로 다수의 척도가 개발되지만 연구자가 다르면 척도들 간의 관계성이 잘 파악되지 않는다. 이들에 대해 척도간 측정 동등성을 검토하고 점수간 연계를 한다면 사용된 척도에 관계없이 연구결과의 통합이 용이해지고, 응답자의 도박행동 수준의 판정이나 유병률이 단일한 척도에서 이루어지는 결과를 가져올 것이다. 점수연계 덕분에, 하나의 척도를 사용한 경우와 동일한 상황이 될 수 있는 것이다. 이러한 점수연계의 방법론과 실제 효과는 동일한 구성개념의 측정을 목적으로 개발된 유사한 척도들, 국내/국외의 척도들, 원본 척도와 단축형 척도들 간에 파악될 수 있는 여지가 매우 크다.

둘째, 국내의 임상 장면에서 많은 외국척도가 번안되어 사용되고 있으나, 원본 척도와의 측정 동등성을 확인하고 척도 간 점수를 연계한 것은 본 연구가 최초라 할 수 있다. 본 연구의 방법을 통해 번안척도와 원본척도 간 점수의 전환이 가능해졌으며 발병률이나 유병률 연구에서 국가 간 비교도 손쉽게 이루어질 수 있는 토대를 마련했다고 할 수 있다. 예로서 CPGI에 대한 번역본만 해도 국내에 3개가 있다(김교현 등, 2010; 김아영 등, 2011; 이경희, 2009). 이들이 동일한 원본인 캐나다 척도(CPGI)를 번안한 것임을 감안한다면, 사용된 연구에 따라서 산출된 결과를 서로 충분히 연

계할 필요가 있다. 또한 유사한 목적으로 개발된 상이한 외국 척도들의 번안물 간에도 측정 동등성 및 점수 연계를 해볼 수 있는 방법적 기반이 마련되었다고 할 수 있다. 예로서, 도박 연구 분야에서 CPGI가 NODS 및 SOGS와 개발 목적은 상이하나 제작된 결과물은 유사한 기능을 가지고 있음이 인식되고 있으므로(Ladouceur et al., 2005), 국내에서도 KCPGI와 KNODS 및 KSOGS간 측정 동등성 연구는 척도간 일반화 가능성의 범위와 한계를 밝혀줄 것이다.

셋째, 측정의 동등성 확인을 위해 다양한(전체집단, 남성집단, 여성집단) 다집단 분석 방법을 사용한 점이다. 척도를 개발하고 어떤 집단에 타당화되었다고 할 때에, 집단내 하위 집단에 걸쳐 그 척도가 차별적으로 기능하지 않음을 암묵적으로 가정하는 경우가 대부분인데 본 연구에서는 그 가정을 명시적으로 검증해 본 것이다. 앞으로 이러한 암묵적 가정이 검증없이 수용되는 관행에서 벗어나는데, 본 연구가 모범 사례가 될 수 있다.

본 연구의 제한점과 향후 연구 및 적용에 대한 제언은 다음과 같다. 첫째, 본 연구는 CPGI와 KCPGI의 점수 연계를 시도한 최초의 연구로 충분한 표본을 확보하여 다양하고 보다 엄밀한 연구방법을 분석에 사용하였으나, 본 연구 결과를 확증하기 위해서는 반복연구가 필요하다. 특히 두 국가 간에 점수연계를 명시적 목표로 내걸고 정교한 표집을 한다면 이상적이다. 그럴 경우, 두 국가 간 질적으로 동등한 표본을 확보하기 용이할 것이다. 본 연구에서 저수준의 분류 체계가 CPGI와 달라 캐나다 집단의 비율에 정확히 일치하는 국내 표본을 구성하는데 한계가 있었기 때문에, 이를 극복한다면 본 연구에서 관찰된 두 국가

간 점수 차이에 대해서도 보다 세부적 검토가 가능할 것이다. 둘째, 캐나다 표본에 비해 한국 표본에서 척도 평균과 요인 평균이 더 높게 나타났고, 요인분산에서도 두 집단 간 차이가 있는 것으로 나타났는데, 이에 대한 보다 정확한 해석을 위해서는 추가적인 비교연구가 필요하다. 우선 변산도가 높은 것은 한국 표본에서 극단적으로 큰 값들이 많음을 의미하는데 캐나다의 원자료에 접근하지 못하므로 비교확인은 어렵다. 대표치의 값이 한국판에서 높은 것은 본 연구의 원자료는 모두 자기보고 설문 결과이나 캐나다 자료는 모두 전화면접에 의한 것으로 이러한 자료수집 방식의 차이가 평균점수의 차이를 초래했을 가능성이 있다. 앞으로 두 국가간 합의된 방식으로 동시에 자료수집이 된다면 보다 정확한 해석이 가능할 것이다. 셋째, 캐나다 자료는 2,000명이 넘는 표본을 대상으로 한 것이나 한국 자료는 900명에 조금 못 미쳤다. 이렇게 표본 크기에서 2배 이상의 차이가 남으로 인해 결과에서 다소 간의 편향이 발생했을 가능성을 배제하기 어렵다. 이후 연구에서는 한국 자료에서 보다 큰 표본을 확보하는 것이 바람직할 것이다. 넷째, 앞으로 KCPGI에 대한 연구를 기반으로 하여 다른 번역본 도박행동 검사들(예: KSOGS, KNODS)과의 점수 연계화가 시도될 수 있다. 연계화가 달성되면 어떤 척도를 사용하든 다른 척도의 점수로 바꾸어 볼 수 있어 연구 결과의 일반화가 용이해지고 유병률 통계가 단일화될 수 있을 것이다. 끝으로 척도가 만능은 아니므로, 항상 임상분야 전문가의 전문성과 경험이 척도 사용에 함께 할 때 보다 적절히 활용될 수 있음을 부언한다.

## 참고문헌

- 김교현, 권선중, 김세진, 이순목 (2011). 저수준 도박행동 연구를 위한 개념화 및 척도개발. 한국심리학회지: 일반, 30(2), 599-628.
- 김교현, 조성겸, 권선중, 이동형 (2010). 사행산업 이용 실태 조사. 서울: 사행산업통합감독위원회.
- 김아영, 임은영 (2003). 타문화권 척도 변안과정에서 적용하는 절차들 간의 효과비교. 한국심리학회지: 일반, 22(1), 89-113.
- 김아영, 차정은, 권선중, 이순목 (2011). CPGI의 한국판 제작 및 타당화. 한국심리학회지: 일반, 30(4), 1011-1038.
- 김정호, 오수성, 오상우, 김상훈, 김학렬, 박상학 (2005). KEDI-WISC와 K-WISC-III의 구조 및 측정 동일성 검증: 임상표본을 대상으로. 한국심리학회지: 임상, 24(2), 413-426.
- 김한조 (2010). 다집단 분석에서의 부분 동일성 전략들: 적절한 전략을 찾기 위한 시뮬레이션 연구. 성균관대학교 석사학위 청구논문.
- 문화관광연구원 (2008). 사행산업 이용실태 조사 분석 및 총량 조정 연구. 서울: 사행산업통합감독위원회.
- 방희정, 윤진영, 김아영, 조혜자, 조숙자, 김현정 (2007). 한국 성인의 관계적 자기 구성요인 탐색 및 척도개발. 한국심리학회지: 사회문제, 13(3), 23-63.
- 손원숙 (2002). 성별에 따른 “수학에 대한 태도”의 차원성 연구: 다집단 확인적 요인 분석의 적용. 교육심리연구, 16(3), 235-251.
- 손은영, 차정은, 김아영 (2007). 사회적 바람직성 상, 하위 집단 간 성격검사의 구인동

- 등성 검증. 한국심리학회지: 사회 및 성격, 21(2), 71-87.
- 이경희 (2009). 한국판 캐나다 문제도박 척도 (CPGI)의 타당화를 위한 예비연구. 한국심리학회지: 건강, 14, 667-675.
- 이순목 (1993). 중급 LISREL. 한국심리학회 동계 연수회 자료집.
- 이순목 (2002). 사회과학을 위한 측정의 원리. 서울: 학지사.
- 이순목, 금은희, 이찬순 (2010). 다집단분석의 문제: 평균구조분석에서의 측정원점 동일성 검증필요 여부. 교육평가연구, 23(2), 391-416.
- 이순목, 김교현, 최용성, 이홍표, 김종남, 김수진 (2009). 한국형 도박중독 변별척도 개발 연구(2기 보고서). 서울: 사행산업통합감독위원회.
- 이순목, 김교현, 최용성, 이홍표, 김종남, 김수진, 권선중, 김세진, 김인혜 (2011). 한국형 도박행동 H척도의 개발: 고수준 도박자 대상의 연구용 척도. 한국심리학회지: 일반, 30(4), 1053-1082.
- 이순목, 김아영, 권선중, 김종남, 차정은, 김인혜 (2011). 전국민 대상 도박문제 선별척도 및 기준점수 타당화 연구. 서울: 사행산업통합감독위원회.
- 이순목, 김인혜 (2009). 구조방정식모형에서 다집단 분석의 문제 및 대안으로서의 다특질 다상황 다방법 모형. 교육평가연구, 22(1), 219-242.
- 이순목, 김한조 (2011). 구조방정식 모형의 일반화 또는 집단차 연구를 위한 다집단 분석의 관행과 문제점. 사회과학(성균관대), 43(1), 63-112.
- 이순목, 현명호, 최삼욱, 김종남, 김수진 (2008). 한국형 도박중독 변별척도 개발 연구(1기 보고서). 서울: 사행산업통합감독위원회.
- 이홍표 (2008). 도박중독 진단에 대한 논의: 토론. 한국사회의 도박중독 문제, 어떻게 풀어나가나? pp.45-52. 2008년도 제2차 중독심리공동교육 교재. 한국심리학회 중독심리전문가 위원회.
- 조용래, 김정호 (2002). 한국판 Beck Depression Inventory의 확인적 요인분석: 대학생과 임상표본 간 구조 및 측정동일성 검증. 한국심리학회지: 임상, 21(4), 843-857.
- 최대정, 조현주, 박동건 (2005). 조직 구성원들의 정서 경험: 요인구조의 동등성 및 타당도 분석. 한국심리학회지: 산업 및 조직, 18(2), 385-411.
- 하정, 홍세희, 유성경 (2005). 아동-청소년용 한국형 남녀평등의식검사(KGES-AC) 타당화. 한국심리학회지: 일반, 24(2), 105-127.
- 한소영, 신희천 (2007). 커플 기본 심리적 욕구 중요성 척도의 다집단 요인분석과 잠재평균분석. 한국심리학회지: 상담 및 심리치료, 19(2), 447-465.
- 홍세희, 황매향, 이은설 (2005). 청소년용 여성진로장벽 척도의 잠재평균분석. 교육심리연구, 19(4), 1159-1177.
- Alwin, D. F., & Jackson, D. J. (1981). *Applications of simultaneous factor analysis to issues of factorial invariance*. In D. D. Jackson & E. P. Borgana (Eds.), *Factor analysis and measurement in sociological research: A multidimensional perspective*. Beverly Hills, CA: Sage.
- Angoff, W. H. (1971). Scales, norms, and equivalent scores. In R. L. Thorndike (Ed.), *Educational measurement* (2nd ed., pp.508-600). Washington, D. C.: American Council on

- Education.
- Bentler, P. M. (1990). Comparative fit indexes in structural models. *Psychological Bulletin*, 107(2), 238-246.
- Bentler, P. M. (1995). *EQS Structural equations program manual*. Encino, CA: Multivariate Software.
- Bentler, P. M., & Bonett, D. G. (1980). Significance tests and goodness of fit in the analysis of covariance structure. *Psychological Bulletin*, 88, 588-606.
- Bollen, K. A. (1989). *Structural equations with latent variables*. Wiley & Sons.
- Brislin, R. W. (1970). Back-translation for cross-cultural research. *Journal of Cross-Cultural Psychology*, 1(3), 185-216.
- Cheung, G. W., Murrmann, K. F., Murrmann, S. K., & Becker, C. (2004). Cross-cultural comparisons of restaurant service expectations using non-invariant measurement items. *Journal of Hospitality and Tourism Research*, 28(4), 375-390.
- 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*(2nd ed.). Hillsdale, NJ: LEA.
- Eremenco, S. L., Cella, D., & Arnold, B. J. (2005). A comprehensive method for the translation and cross-cultural validation of health status questionnaire. *Evaluation and The Health Professions*, 28(2), 212-232.
- Ferris, J., & Wynne, H. (2001a). *The Canadian Problem Gambling Index: User manual*. Canada: Canadian center on substance abuse.
- Ferris, J., & Wynne, H. (2001b). *The Canadian Problem Gambling Index: Final report*. Report to the Canadian inter-provincial task force on problem gambling.
- Holland, P. W., & Thayer, D. T. (1988). Differential item performance and the Mantel-Haenszel procedure. In *Test validity*. Edited by Howard Wainer and Henry I. Braun, 129-145.
- Hu, L., & Bentler, P. M. (1999). Cutoff criteria for fit indexes in covariance structure analysis: Conventional criteria versus new alternatives. *Structural Equation Modeling*, 6, 1-55.
- Johnson, R. A., & Wichern D. W. (2001). *Applied Multivariate Statistical Analysis, 5th ed.* Prentice-Hall.
- Jöreskog, K. G. (1971). Simultaneous factor analysis in several populations. *Psychometrika*, 36(4), 409-426.
- Kolen, M. J., & Brennan, R. L. (2004). *Test equating, scaling, and linking: Methods and practices*. 2nd Ed. New York: Springer-Verlag.
- Ladouceur, R., Jacques, C., Chevalier, S., Sevigny, S., & Hamel, D. (2005). Prevalence of pathological gambling in Quebec in 2002. *Canadian Journal of Psychiatry*, 50(8), 451-456.
- Lee, S., & Kim, H. (2011). *Interaction between test strategies and invariance/noninvariance conditions in testing for partial (metric) invariance in structural equation modeling*. Paper presented at the 17th International meeting of the psychometric society, Hong Kong, July 21, 2011.
- Lesieur, H. R., & Blume, S. B. (1987). The South Oaks Gambling Screen(SOGS): A new

- instrument for the identification of pathological gamblers. *American Journal of Psychiatry*, 144(9), 1184-1188.
- Meredith, W. (1993). Measurement invariance, factor analysis and factorial invariance. *Psychometrika*, 58(4), 525-543.
- Michell, J. (1986). Measurement scales and statistics: A clash of paradigms. *Psychological Bulletin*, 100, 398-407.
- National Gambling Impact Study Commission (NGISC, 1999). *Gambling impact and behavior study*. Washington, D. C.: Author.
- Neal, P., Delfabbro, P., & O'Neil, M. (2005). *Problem gambling and harm: Towards a national definition*. Gambling Research Australia. Australia.
- Shaffer, H. J. & Hall, M. N. (1996). Estimating the prevalence of adolescent gambling disorders: A quantitative synthesis and guide toward standard gambling nomenclature. *Journal of Gambling Studies*, 12(2), 193-214.
- Shaffer, H. J., Hall, M. N., & Vander Bilt, J. (1997). *Estimating the prevalence of disordered gambling behavior in the United States and Canada: A meta-analysis*. Harvard medical school, division of addition, December 10, 1997.
- Shaffer, H. J, LaBrie, R., Scanlan, K. M., & Cummings, T. N. (1994). Pathological Gambling among adolescents: Massachusetts Gambling Screen(MAGS). *Journal of Gambling Studies*, 10(4), 339-362.
- Sohn, W. (2001). Equivalence of constructs measured by two different language versions of 16PF. *Korean Journal of Psychology: General*, 21(1), 91-116.
- Sorbom, D. (1974). A general method for studying differences in factor means and factor structures between groups. *British Journal of Mathematical and Statistical Psychology*, 27, 229-239.
- 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(6), 1292-1306.
- Steiger, J. H., & Lind, J. C. (1980). *Statistically-based tests for the number of common factors*. Paper presented at the annual spring meeting of the psychometric society in Iowa city. May 30, 1980.
- Vandenberg, R. J. (2002). Toward a further understanding of and improvement in measurement invariance methods and procedures. *Organizational Research Methods*, 5(2), 139-158.
- Vandenberg, R. J., & Lance, C. E. (2000). A review and synthesis of the measurement invariance literature: Suggestions, practices, and recommendations for organizational research. *Organizational Research Methods*, 3(1), 4-70.
- Yoon, M., & Millsap, R. E. (2007). Detecting violations of factorial invariance using data-based specification searches: A Monte Carlo study. *Structural Equation Modeling*, 14(3), 435-463.
- 원고접수일 : 2011. 11. 7.  
1차 수정 원고접수일 : 2011. 12. 30.  
게재결정일 : 2012. 1. 28.

## Measurement Equivalence and Linking Between Canadian Problem Gambling Index and Its Korean Version

**Soonmook, Lee**

Sungkyunkwan U.

**Soocheol, Yun**

Leaders Insight

**Jungeun, Cha**

Ewha Womans U.

**Jongnam, Kim**

Seoul Women's U.

**Sung Chil, Yeo**

Konkuk U.

The purpose of this study is to investigate measurement equivalence between CPGI(Canadian Problem Gambling Index) and KCPGI, and to link scores of the two scales. To do so, we conducted multiple group analysis with CPGI data(Canadian data) and KCPGI data(Korean data) for total group, male group, and female group. Specifically, we tested configural invariance, metric invariance, scalar invariance, invariance of factor variances, and factor mean differences. We used data from 2,089 Canadians and 882 Koreans. The results are as follows. At first, configural invariance was supported. Metric invariance and scalar invariance were partially supported. However, factor variances were statistically different, although the difference was not practically large. Factor mean comparison showed that Koreans were higher than Canadians. The results of male and female groups were similar to those of the total group, although there was a little difference in the female group. Finally, we attempted linking scores between CPGI and KCPGI. As a result we could provide a conversion table between the two scale scores. Lastly, we discussed the implications and the limitations of this study.

*Key words : CPGI, KCPGI, measurement equivalence, score linking, multiple group analysis*

부 록

다집단 분석에 사용된 자료

- 캐나다 남녀 통합자료 공분산행렬

|       | CPGI1   | CPGI2   | CPGI3   | CPGI4   | CPGI5   | CPGI6   | CPGI7   | CPGI8   | CPGI9   |
|-------|---------|---------|---------|---------|---------|---------|---------|---------|---------|
| CPGI1 | 0.127   |         |         |         |         |         |         |         |         |
| CPGI2 | 0.036   | 0.071   |         |         |         |         |         |         |         |
| CPGI3 | 0.043   | 0.028   | 0.100   |         |         |         |         |         |         |
| CPGI4 | 0.016   | 0.015   | 0.010   | 0.018   |         |         |         |         |         |
| CPGI5 | 0.036   | 0.020   | 0.028   | 0.010   | 0.047   |         |         |         |         |
| CPGI6 | 0.028   | 0.017   | 0.019   | 0.009   | 0.027   | 0.041   |         |         |         |
| CPGI7 | 0.048   | 0.032   | 0.036   | 0.019   | 0.040   | 0.029   | 0.177   |         |         |
| CPGI8 | 0.029   | 0.022   | 0.021   | 0.009   | 0.021   | 0.019   | 0.022   | 0.030   |         |
| CPGI9 | 0.086   | 0.053   | 0.062   | 0.024   | 0.058   | 0.050   | 0.079   | 0.042   | 0.371   |
| Mean  | 0.09000 | 0.04600 | 0.07660 | 0.01480 | 0.02780 | 0.02580 | 0.06030 | 0.02060 | 0.12930 |
| N     | 2089    | 2089    | 2089    | 2089    | 2089    | 2089    | 2089    | 2089    | 2089    |

- 한국 남녀 통합자료 공분산행렬

|       | CPGI1   | CPGI2   | CPGI3   | CPGI4   | CPGI5   | CPGI6   | CPGI7   | CPGI8   | CPGI9   |
|-------|---------|---------|---------|---------|---------|---------|---------|---------|---------|
| CPGI1 | 0.122   |         |         |         |         |         |         |         |         |
| CPGI2 | 0.054   | 0.131   |         |         |         |         |         |         |         |
| CPGI3 | 0.051   | 0.078   | 0.281   |         |         |         |         |         |         |
| CPGI4 | 0.040   | 0.045   | 0.079   | 0.145   |         |         |         |         |         |
| CPGI5 | 0.043   | 0.045   | 0.070   | 0.052   | 0.200   |         |         |         |         |
| CPGI6 | 0.037   | 0.043   | 0.052   | 0.065   | 0.052   | 0.134   |         |         |         |
| CPGI7 | 0.037   | 0.045   | 0.073   | 0.065   | 0.070   | 0.070   | 0.258   |         |         |
| CPGI8 | 0.046   | 0.037   | 0.058   | 0.065   | 0.055   | 0.057   | 0.069   | 0.126   |         |
| CPGI9 | 0.045   | 0.049   | 0.087   | 0.062   | 0.073   | 0.053   | 0.080   | 0.065   | 0.289   |
| Mean  | 0.09841 | 0.12143 | 0.29274 | 0.12732 | 0.19580 | 0.10249 | 0.21281 | 0.10658 | 0.23776 |
| N     | 882     | 882     | 882     | 882     | 882     | 882     | 882     | 882     | 882     |