

사회심리적 건강 측정 도구의 신뢰도 및 타당도 검정

김 정 희*

I. 서 론

1. 연구의 필요성 및 목적

지난 20여년 동안 스트레스는 근로자와 경영주, 보건의료전문인 모두에게 큰 관심이 되어왔다. 산업화가 가속화됨에 따라 오늘날, 주요 상병 발생 요인은 환경과 생활양식의 변화에서 기인된 다요인적인 것으로 간주되고 있으며, 특히, 첨단 산업사회에서 많은 종류의 직업과 작업환경 속에서 일하는 근로자들은 많은 심리적, 신체적 부담을 안고 살아가고 있는 것으로 보고되고 있다.

1985년 미국의 National Health Interview Survey 결과에 따르면, 근로자의 30%가 자신들의 건강을 가장 많이 위협하는 근로요인으로 직업성 스트레스를 들고 있다(Shilling & Brackbill, 1987). 또한 스트레스 관련 비용으로 1970년대에는 연간 750~900억 달러를 소비하였고, 1980년대 말에는 그 비용이 연간 1,500억 달러로 늘어났으며, 스트레스로 인한 경제적 총 손실이 GNP의 10%를 상회하였다(이중목, 1989, 1997). 1981~1982년 스트레스와 관련된 근로자 보상청구율이 전체 직업관련 질환의 11%에 이르고 있어(Murphy, 1996), 스트레스 관리의 중요성은 일상적인 관심사로 강조되고 있다.

스트레스는 누적되면 심리적으로 신경이 예민해져 걱정과 불안, 초조, 긴장 등의 현상으로 신경쇠약과 우울,

의욕상실 등이 나타나고, 극도의 경우에는 자살에 이르게도 하며, 만성적 스트레스는 생리적으로 자율신경계와 내분비계에 영향을 끼쳐 정신적, 신체적 균형을 파괴시켜 각종 질병을 유발시킨다. 또한 행동적인 측면에서도 소식(小食), 불면, 음주·흡연량의 증가, 약물남용 등의 결과가 나타나기도 하며, 조직의 능률과 인화를 저해하고 지각, 결근, 이직, 보상요구 등을 증가시키기도 한다(김문석, 1990 : 이우천, 1997).

이렇듯 현대에서 스트레스는 심리적 건강 측면과 관련된 중요한 개념이 되었으나, 스트레스의 정도를 어떻게 측정하고, 어떻게 관리할 것인지의 문제는 아직 명확하지 않은 실정이다(이채용과 이종영, 1996).

다른 영역의 건강 측정과는 달리 심리적 건강 수준을 측정하는데는 많은 논란의 여지가 있다. 이는 건강이 복합적 성격을 띠면서도 또 다른 한편으로는 추상적 성격을 띠고 있으며(장세진, 차봉석, 박종구, 이은경, 1994), 주로 주관적인 건강을 측정하고, 또한 심리적 건강의 측정 범위와 내용에 대한 기준 설정이 어렵기 때문이다(MacDowell & Newell, 1996).

그동안 국내에서 이루어진 스트레스에 대한 연구들 대부분은 외국에서 개발된 도구를 번역해 사용해 왔는데, 주로 조사목록의 성격을 갖는 SCL-90(symptom checklist-90)과 MMPI, Zung의 SAS(Self-rating Anxiety Scale)와 SDS(Self-rating Depression Scale),

* 서울대학교 간호대학 시간강사 및 박사과정생

생활사건척도 등을 많이 사용하여 왔다(권남희, 1995; 김성현, 윤진희, 박혜순, 1992; 백기청, 1991; 장세진 등, 1994). 이 도구들은 많은 연구들을 통하여 그 타당도와 신뢰도가 국내에서도 인정된 도구이기는 하나, 행동적이며 신체적인 증상 등 정신건강의 부정적 측면을 다루고 있고 건강한 사람과 그렇지 못한 사람들간의 구분을 하기 위한 측정도구의 성격이 강하며, 개인용으로 제작된 도구들이어서 집단을 대상으로 사용하는데 많은 제약이 있음이 지적되고 있다(장세진, 1993).

외국의 경우, 집단을 대상으로 지역사회 인구집단에 사용된 대표적인 도구로는 Langner의 22-item scale, Dupuy의 GWB(General well-being Schedule)와 이를 수정, 보완한 Rand의 MHI(Mental Health Inventory), 그리고 Goldberg의 GHQ(General Health Questionnaire) 등이다(Derose, 1997). 이중 GHQ는 지역사회 조사연구에서 심리적 안정상태를 측정할 수 있는 도구로 인정을 받고 있다(Derose, 1997; MacDowell과 Newell, 1996).

장세진(1993, 1994)은 이 GHQ-60을 기초로 우리나라의 현실에 맞게 수정하여 탐색 연구를 통해 45개 항목으로 축소된 집단용 스트레스 측정도구인 사회심리적 건강측정도구(Psychosocial Well-being Index, 이하 PWI)를 제작하였으며, 장세진 등(1994)은 PWI의 Cronbach's $\alpha=0.94$, GHQ-60, GHQ-30, GHQ-28, GHQ-20, GHQ-12와의 상관관계 0.91-0.99로 높은 신뢰도와 타당도를 확보하고 있음을 보고하였다.

그러나, 몇몇 연구에서 GHQ의 요인분석이 각 연구마다 다르고(Hobbs, P., Ballinger, C. B., & Smith, A. H. W., 1983; Sheck, 1993), 그 결과 GHQ-60을 근거로 개발된 PWI의 요인분석 또한 각 연구마다 다르게 나타나(장세진 등, 1994, 이채용과 이종영, 1996), 사회심리적 건강 속성이 일관적으로 나타나지 않았다. 이채용과 이종영(1996)은 의과 대학생 133명을 대상으로 PWI의 타당도 및 신뢰도를 연구한 바 있는데, 그 결과 PWI는 신뢰성있는 도구라고 판단할 수 있으나, 4가지 하부개념이 장세진 등(1994)의 제시와는 달리 3번째, 4번째 개념이 일치하지 않았다. PWI의 모태가 되는 GHQ의 경우, 신체적 증상, 불안과 불면증, 사회적 역기능, 심각한 우울 등 4가지 하부 개념으로 이루어져 있으며, GHQ가 정신 질환에서 최소 공배수들과 관련되어 있어서 논리적으로는 단면적인 척도가 더 타당하다(Sheck, 1993). 또한 이채용과 이종영(1996)은 2차원적인 도구의 가능성을 제시하기도 하였던 바, PWI의

단면적 도구로서의 가능성에 대한 접근과 PWI의 신뢰도 및 타당도 확보와 관련된 연구가 필요하다.

뿐만 아니라 이 도구를 이용한 연구에서는 GHQ-60 전문을 이용하거나(오수성, 이광섭, 손석준, 최진수와 이정애, 1995), 도구의 타당도 및 신뢰도를 제시하지 않고 사용하는(고맹님, 1996; 김성아 등, 1997; 이명희, 박정일, 구정완, 임현우, 1997; 차봉석과 고상백, 장세진, 박창식, 1996) 등 연구방법상의 문제가 제시되어 도구의 신뢰도 및 타당도에 대한 반복연구가 필요할 것으로 생각된다.

이에 본 연구는 육체적 불구 보다는 긍정적인 측면의 건강과, 본질적으로 광범위하고 추상적이며 주관적인 내용을 포함하고 있는 스트레스를 측정하는 도구(장세진 등, 1994)로써 PWI의 타당도 및 신뢰도를 검증하고자 시도하였다.

본 연구의 구체적 목적은 다음과 같다.

1. 사회심리적 건강측정도구(PWI)의 신뢰도를 검증한다.
2. PWI의 타당도를 검증한다.
3. PWI의 하부개념을 확인한다.

II. 문헌 고찰

1. 총체적 스트레스 측정과 사회심리적 건강측정도구(PWI)

조사연구에서 흔히 사용되는 심리적 건강에 대한 측정은 주로 주관적이며, 이는 측정하려고 하는 변수의 측정가능성에 대한 문제제기를 함축하고 있다. 주관적 측정 도구는 실험이나 물리적 측정을 통해서 추론이 거의 불가능한 통증이나 우울 등을 측정할 때 여러 가지 정보를 제공해줄 수 있으나, 주관적 판단을 수량화하여 이를 평가하는데는 개인간의 오차가 발생할 수 있으며 또한 동일한 수준의 느낌을 갖고 있다고 하여 동일하게 표현될 수 없기 때문이다(장세진 등, 1994).

그러나, 개인이 인지한 주관적인 건강상태는 신체적 기능 뿐 아니라 사망율, 만성질환 유병률 등을 예측하는데 중요하고 민감한 자료로 사용 가능하며(Idler, 1991), 점차로 주관적 판단에 근거하는 개인의 건강 지표가 수용되고 있고(장세진 등, 1994), 또한 불가능하게 생각되어온 스트레스 개념의 정량화가 Holmes 등이 생활변화(social readjustment rating scale)라는 개념을 도입하여 사회적응 평가 척도를 제작함으로써 스트레

스의 양을 정량화하려는 시도가 되었다(정도연, 1986).

주관적 정신건강을 측정하는데 활용되어온 시각은 여러 가지가 있는데, 심리적 건강에 대한 초기의 측정은 심리적 건강을 생활에 대한 만족도에 초점을 두었으며, 두 번째 시각으로는 경험에 대한 정서적 반응 즉, 일상적 경험에 의해 변화되는 감정의 상태를 심리적 건강으로 간주하였고, 세 번째의 심리적 건강을 측정하는 접근은 심리적 긴장(psychological distress)을 검사하는 것이었다. 이러한 접근은 척도가 진단적 분류보다는 일반적인 심리적 긴장의 증상을 검사하는데 주안점을 두게 되었고, 그런 의미에서 심리적 건강이란 바로 긴장의 증상이 없는 상태로 설명되었다(MacDowell과 Newell, 1996).

건강의 결과를 평가(사정)하는데는 흔히 상황적(condition-specific)측정과 총체적(generic)측정이 사용된다. 상황적 측정은 특정 의학적 상태를 반영하는 중후나 증상에 초점을 둔다. 또한 인간의 삶에 대한 질병의 직접적인 결과를 평가할 수 있으며, 그 결과 건강의 미묘한 변화에 민감하게 측정할 수 있다. 총체적 측정은 건강상태에 대한 전체적인 영향을 평가하는 포괄적인 측정이다. 이는 다면적이거나 단면적일 수 있으며, 다른 유형의 질병이나 치료, 환자 등에 적용될 수 있으며, 건강의 신체적, 심리적, 사회적 측면을 포착하기 위하여 고안되었다. 이는 또한 안녕에서부터 불능까지의 전 건강범위에 따라 건강을 측정하기 위하여 사용되는데(Maciejewski, 1997), 현재 스트레스를 측정하는 방법은 크게 3가지로 분류한다. 첫째, 생화학적(내분비적) 측정법이며, 둘째, 임상적 면담과 상담을 통한 개인의 성격, 행동양식을 측정하는 것이다. 셋째, 설문지를 통한 정신사회적 스트레스 측정법이다(차철환, 1993).

심리적 건강상태에 대한 측정은 부정적 스트레스(distress)에 대하여 행동적이며 신체적 증상을 내포하는 증상 체크 목록 즉, 건강한 사람과 그렇지 못한 사람들 간의 구분을 하기 위한 측정도구의 성격이 강하고 정신건강의 부정적 측면만을 취급한 측정과 안녕에 대한 긍정적, 부정적 느낌을 직접 질문하는 측정 방법이 있다. 따라서, 스트레스 수준을 파악하는 요인으로 단지 신체적, 행동적 증상들을 묻는 것도 의미가 있을 수 있으나 그와 함께 긍정적 측면에서의 정신건강이 반영되어야 하고, 이는 건강에 대한 WHO의 개념과 맥을 같이 하는 것이며 이를 고려한 측정도구의 개발이 요청되고 있다(MacDowell과 Newell, 1996).

이러한 맥락에서 정상적으로 일상생활을 영위해 나가

고 있는 사람들을 대상으로 스트레스 수준을 평가하고 스트레스 수준과 상관성을 갖는 변수들을 발견하거나 조정기전을 탐색할 수 있는 지표로서 개발된 PWI는 기존 스트레스 측정 도구들과는 다르다고 할 수 있다(이채용과 이종영, 1996).

PWI는 스트레스 수준을 측정함에 있어 다루어져야 할 내용들이 포괄적으로 들어 있고, 다양한 연구 대상 집단을 통한 신뢰도와 타당도 검사 결과 측정도구로서 높이 평가받아 왔던 GHQ를 기초로 우리나라의 현실을 고려하여, 장세진 등(1994)이 근로자 297명(기술직 및 사무직, 관리직 포함)을 대상으로 한 탐색적 연구를 통해 2단계 수정을 거친후 45개 문항으로 재구성한 설문지이다. 이는 스트레스 수준을 반영시킬 수 있는 부정적 측면과 긍정적 측면을 동시에 포괄하는 도구로, 최근 몇 주간의 육체적, 심리적 상태를 묻는 항목으로 구성되어 있다.

GHQ의 각 문항은 의문문으로 되어있으며, 우울, 불안, 사회적 손상, 염려(신체적 증상) 등 4가지 하부 개념을 갖고 있으나 PWI는 재구성되면서 서술문의 형식을 가지게 되었다. 각 문항에 대해 '전혀 그렇지 않다', '이따금 그렇다', '자주 그렇다', '항상 그렇다'의 Likert 4점 척도(0-1-2-3)로 구성되어 0점에서 135점의 점수를 가질 수 있다.

PWI는 정상인의 스트레스 수준을 측정하기 위해 고안되었으며 정신과적 증상이나 행동양식들을 정상인들과 구별하거나 진단하기 위해 만들어진 것은 아니다(장세진 등, 1994). 따라서 스트레스로 인한 생리적, 심리적 반응 증상 뿐 아니라 일상적 생활을 수행해 나가는데 필요한 사회적 수행능력이나 자기신뢰 그리고 일반건강의 측면을 도구에 포함시켰다. 즉 사회적 역할 수행 및 자기신뢰, 우울, 수면장애 및 불안, 일반 건강 및 생명력 등 4가지 개념으로 이루어져 있다.

개발당시 도구의 Cronbach's $\alpha=0.94$, 반분법에 의한 Guttman의 신뢰도 계수(rc)는 0.88이었으며, GHQ-60, GHQ-30, GHQ-28, GHQ-20, GHQ-12와의 상관관계 0.91-0.99로 높은 신뢰도와 타당도를 보이고 있다.

그러나, 몇몇 연구에서 GHQ의 요인분석이 다르고(Hobbs, P., Ballinger, C. B., & Smith, A.H.W., 1983; Sheck, 1993), 그 결과 GHQ-60을 근거로 개발된 PWI의 요인분석 또한 각 연구마다 다르게 나타나(장세진 등, 1994, 이채용과 이종영, 1996), 사회심리적 건강 속성이 일관적으로 나타나지 않았다. 이채용과 이

종영(1996)은 의과 대학생 133명을 대상으로 PWI의 타당도 및 신뢰도를 연구한 바 있는데, 그 결과 PWI는 신뢰성있는 도구라고 판단할 수 있으나, 4가지 하부개념이 장세진 등(1994)의 제시와는 달리 3번째, 4번째 개념이 일치하지 않았다. 따라서, 이 도구의 사용에 있어서 P/MI가 측정하고자 하는 것이 건강의 정신사회적 측면인지, PWI가 사회심리적 건강을 신뢰성 있게 측정하는 도구인지의 문제가 먼저 해결되어야 할 것이다. 또한 P/MI의 모태가 되는 GHQ의 경우, 신체적 증상, 불안과 불면증, 사회적 역기능, 심각한 우울 등 4가지 하부개념으로 이루어져 있다. 또한 그 동안의 연구결과에서 단면적인 척도인지, 다면적인 척도인지 불명확하게 나타나고 있지만, Goldberg는 GHQ가 정신질환에서 최소 공배수들과 관련되었다고 지적하였으므로 논리적으로는 단면적인 척도가 더 타당하다(Shek, 1993). 또한 이채용과 이종영(1996)의 연구에서는 2차원적인 도구의 가능성을 제시하기도 하였지만, PWI의 하부개념에 대해서도 확인이 필요하다.

뿐만 아니라 이 도구를 이용한 연구에서는 GHQ-60 중 문항 이용하거나(오수성, 이광섭, 손석준, 최진수와 오정애, 1995), 도구의 타당도 및 신뢰도를 제시하지 않고 사용하는(고맹님, 1996; 김성아 등, 1997; 차봉석과 오상배, 장세진, 박창식, 1996) 등 연구방법상의 문제가 제시되어 도구의 신뢰도 및 타당도에 대한 반복연구가 필요할 것으로 생각된다.

2. 일반건강측정표

GHQ(The General Health Questionnaire)는 현재의 진단가능한 정신이상(nonpsychotic)을 발견하기 위해 개발된 자가보고형 도구로, 임상 뿐 아니라 지역사회 단위로 한 조사 등 일차의료 상황에 광범위하게 사용되고 있다(MacDowell과 Newell, 1996).

GHQ는 첫째, 개인의 건강한 기능 수행의 어려움을 발견하고, 둘째, 스트레스의 새로운 현상이 나타난 경우를 발견하기 위해 개발되었다(Goldberg, 1979). 이는 심신과적 우울이나 정신분열증과 같은 심각한 불건강 상태를 발견하기 위하여 고안된 것이 아니며, 스트레스와 관련된 성격장애 혹은 적응 장애 등을 다룬다. 즉, GHQ가 스트레스의 다양한 유형들을 포함하고는 있지만, 이는 정신과적 이상의 정도를 판별하거나 진단하기 위해 개발된 것은 아니다(MacDowell과 Newell, 1996).

GHQ는 스트레스의 4가지 개념 즉, 우울, 불안, 사회적 손상, 염려(신체적 증상)을 다루고 있다.

GHQ의 원본은 60개의 항목으로 구성된 GHQ-60으로 높은 타당도를 갖고 있는 것으로 보고되고 있으나, Goldberg(1979)는 이를 사용하기 어려울 경우 축소판인 GHQ-30, GHQ-20, GHQ-12, GHQ-28를 사용하기를 제안하였다. GHQ-28은 신체적 증상, 불안과 불면증, 사회적 역기능, 심각한 우울 등 4가지 개념에 대하여 4점 척도로 구성되어 있다. GHQ의 각 항목은 0-3점까지의 Likert 척도로 구성되어 있으나 때로는 빈도와 상관없이 증상의 유무에 대하여 0-1점으로 점수화하기도 한다(MacDowell과 Newell, 1996). 정신과적이상이 있는 경우로 판정할 수 있는 기준점은 GHQ-60이 11/12, GHQ-30과 GHQ-28은 4/5, GHQ-20은 3/4, GHQ-12는 2/3이 일반적으로 사용되고 있다(Banks, 1983).

신뢰도는 재조사법에 의한 것이 0.90이었으며, 853명을 대상으로 한 연구에서 반복법을 통한 신뢰도는 0.95였으며, GHQ-30은 0.92, GHQ-20은 0.90, GHQ-12는 0.83으로 보고되었다. 타당도 검사는 아주 다양하고 광범위하게 조사되었으며, 영국이나 호주, 스페인에서 수행된 연구결과에 의하면, 각 연구에서의 척도간의 상관관계수는 0.76-0.81로 일관성있게 나타났으며, 민감도는 81-91%, 특이도는 88-94%였다(MacDowell과 Newell, 1996).

스트레스 측정도구로서 GHQ의 평가는 타당성있게 받아들여지고 있으며(장세진 등, 1994), GHQ는 분명한 개념으로 구성되었고, 도구의 타당화에 대해서도 광범위한 연구를 통하여 입증되었으며, 이는 의료적 상담의 일부로써 유용하고 또한 일반적인 상황에서 정신 이상을 검사할 수 있는 가장 광범위하게 사용가능한 도구이다(MacDowell과 Newell, 1996).

III. 연구 방법

1. 연구 대상자

요인분석을 위한 표본 수는 일반적으로 변수의 4-5배 정도가 바람직하며 최소한 100이상이 이상적이다(이은옥, 임난영, 박현애, 1998). 본 연구에서는 서울경기 지역에 소재하는 백화점 및 호텔 등 서비스업에 종사하는 근로자를 대상으로 186명을 임의표출하였다.

2 연구 도구

PWI는 Goldberg의 GHQ-60를 기초로 하여 장세진 등(1994)이 297명을 대상으로한 탐색 연구를 통해 2단계 수정을 거쳐 45개 문항으로 재구성한 설문지이다. GHQ의 각 문항은 의문문으로 되어있으나, 재구성되면서 서술문의 형식을 취하고 있으며, 각 문항에 대하여 '전혀 그렇지 않다', '어마금 그렇다', '자주 그렇다', '항상 그렇다'의 Likert 4점 척도(0-1-2-3)로 구성되어 0점~135점의 점수를 갖는다. 스트레스 점수의 기준점(cutting point)은 63점으로 63점 이상은 고위험 스트레스 집단, 23-62점은 잠재적 스트레스 집단, 23점 미만은 건강한 집단으로 규정하였다.

PWI와의 상관계수를 통한 타당도 검정을 위해 다양한 연구에서 높은 신뢰도와 타당도를 인정받은(MacDowell과 Newell, 1996) Goldberg의 GHQ-60와 GHQ-30, GHQ-28, GHQ-20, GHQ-12를 사용하였다.

3 자료수집 방법 및 분석방법

자료수집은 1998년 10월 한달동안 이루어졌으며, 조사 대상자 210명 중 189명이 응답하였고(90%), 회수된 설문지 중 일부 응답누락 문항이 있어 분석에는 최종적으로 186명(88.6%)의 자료만을 사용하였다.

PWI의 신뢰도는 Cronbach's α , 반분법에 의한 Guttman의 신뢰도 계수를 구하였으며, PWI의 타당도는 GHQ-60, GHQ-30, GHQ-28, GHQ-20, GHQ-12와의 상관계수를 구하고 요인분석을 하였다. 요인분석은 주성분분석을 하였으며, 요인회전은 해석의 편의상 Varimax회전을 이용하였다. 요인의 수는 장세진 등(1994)의 연구에서와 같이 4개로 미리 지정하여 분석하였으며, 이 모든 통계분석은 SPSS/PC+(Window용)을 이용하였다.

IV. 연구 결과

1. PWI의 요인분석

PWI의 45개 문항의 요인분석을 실시하기 전, 모상관행렬의 단위행렬 여부를 확인하기 위해 바틀렛(Bartlett) 검정을 시행하였다(이영준, 1993). 문항 45개에 대한 단위행렬 검정 결과는 P값이 .000로서 유의수준 $\alpha \leq .01$ 에

서 단위행렬이 아니라는 충분한 증거를 보이므로, 요인 분석을 적용할 수 있었다. 검정 결과는 다음과 같다.

KMO(Kaiser-Meyer-Okin) = 0.822
Bartlett Test of Sphericity = 3473.748
Sig. = 0.000

45개 문항에 대해 주성분분석(Principle Component Analysis)을 시행하여 고유값(eigen value)이 1.0이상의 요인 13개를 추출하였다. 4개 요인으로 요인수를 지정하여 Varimax 회전하여 얻은 요인 중 고유값이 1.0보다 크고 2개 이상의 항목과의 상관계수가 0.30이상인 이상인 요인 4개를 추출하였으며, 첫 번째 요인이 총 분산의 15.9%를 설명하였고 나머지 다른 3개 요인이 총 분산의 24.6%를 설명하여 4개 요인이 총 분산의 40.5%를 설명하였다(표 1).

각 문항과 요인들과의 상관관계를 의미하는(강병서, 석기석, 오윤조, 1993) 부하량은 요인 2와 문항 28이 0.727로 가장 컸으며, 요인 3과 문항 17이 0.323으로 가장 작았다. 전반적으로는 모두 1차적인 요인부하를 보이는 요인과의 상관관계가 가장 크게 나타났다.

요인 2와 4는 장세진 등(1994)의 연구결과와 거의 일치하였으나, 요인 1과 3은 양상이 혼재되어 있었다. 즉, 요인 1에 0.3 이상의 부하량을 보이는 문항은 주로 우울증과 불안을 나타내는 문항들이 포함되었으며, Cronbach's α 는 0.893이었다. 요인 2는 사회적 역할 수행 및 자기신뢰의 문항들이 포함되었고(Cronbach's α = 0.846), 요인 3은 수면장애 및 우울이(Cronbach's α = 0.541), 요인 4에는 일반건강 및 활력 문항 등이 포함되었다(Cronbach's α = 0.781).

각 요인들 간의 유의한 차이를 보기 위하여 Kruskal-Wallis 검정을 하였으며, 요인들 간에는 유의한 차이가 있었다. 검정 결과는 다음과 같다.

Chi-Square 351.397 Significance 0.000

각 요인과 PWI와의 상관관계는 요인 1이 0.832로 가장 높았고, 그 다음은 요인 4($r=0.804$), 요인 2($r=0.751$), 요인 3($r=0.632$) 순이었으며, 요인들 간의 상관계수는 0.6이하인 0.178~0.463의 범위로 전반적으로 낮은 수준이었으며, 모든 요인들이 Cronbach's α 계수 보다 작았다(표 2).

<표 1> 요인분석 결과

문항	적재량	요인1	요인2	요인3	요인4	공유값
7	.678	-.174	.022	.021	.491	
15	.642	.032	.272	-.169	.516	
30	.638	-.227	-.001	.095	.467	
36	.629	-.040	.267	.048	.471	
4	.613	.076	.019	-.186	.416	
39	.607	-.102	.405	.047	.542	
16	.584	.020	.191	-.104	.389	
25	.576	-.225	.058	.029	.386	
44	.567	-.075	.241	.015	.385	
35	.565	-.111	-.063	-.103	.346	
6	.545	-.287	.015	.037	.381	
21	.541	-.202	-.176	-.041	.366	
38	.530	-.280	.173	-.097	.398	
3	.518	.055	-.104	-.302	.374	
40	.491	-.281	.382	.157	.491	
9	.478	-.048	.171	-.372	.398	
14	.475	.136	.352	-.182	.401	
18	.453	-.048	.119	-.011	.222	
10	.444	.033	.036	-.359	.328	
2	.432	-.112	-.080	-.42	1.383	
31	.348	-.134	.029	.340	.255	
34	.340	-.009	.114	-.156	.453	
17	.323	-.214	.236	-.059	.210	
28	.053	.727	-.054	.072	.539	
22	-.188	.725	-.029	-.001	.564	
19	-.171	.692	.128	.024	.525	
20	-.205	.641	-.033	.121	.469	
33	-.131	.582	-.135	.241	.433	
5	-.125	.576	.248	.177	.440	
24	.028	.567	-.169	-.032	.352	
27	-.096	.563	-.101	.023	.337	
37	.028	.536	-.253	.251	.416	
29	-.129	.498	-.078	.390	.423	
23	-.124	.404	-.277	-.083	.262	
45	.043	.290	-.608	.153	.479	
43	-.006	.377	-.583	.235	.537	
8	.245	.113	.567	-.028	.395	
42	.133	.317	-.517	.266	.457	
11	.402	-.050	.467	-.186	.417	
41	.253	.001	.424	.028	.244	
26	.112	.067	.367	.349	.274	
32	.230	-.287	.342	.173	.282	
13	-.193	.375	-.070	.662	.624	
12	-.171	.365	-.132	.631	.578	
1	-.190	.149	-.262	.542	.420	
고유값	7.156	5.230	3.200	2.648		
변량백분율	15.901	11.623	7.442	5.884		
누적백분율	15.901	27.523	34.636	40.520		

2. PWI의 신뢰도

PWI의 신뢰도 검정은 내적일치도를 나타내는 Cronbach's α 값과 반분법에 의한 Guttman의 신뢰도 계수를 구하였다(표 2). PWI의 요인별 Cronbach's α 는 0.541~.893의 범위에 있었으며 전체 PWI의 Cronbach's α 는 0.894이었고, Guttman 신뢰도 계수는 0.7097이었다.

3. PWI의 타당도

PWI의 타당도는 많은 연구에서 높은 신뢰도와 타당도를 인정받은 GHQ-60, GHQ-30, GHQ-28, GHQ-20, GHQ-12와의 상관관계와 요인분석 결과를 이용하여 검정하였다. PWI와 각 측정도구들과의 상관 계수는 0.744~0.905의 범위에 있었다(표 3).

V. 논 의

1. PWI의 요인분석

PWI의 요인분석은 주성분분석을 통하여 요인을 추출한 후 Varimax 회전을 하였다. 분석결과, 고유값이 1.0이상인 요인들은 13개로 나타났으며 13개 요인은 총 분산의 66.3%를 설명하였다. 이는 장세진(1993, 1994)의 연구에서 나타난 12개나, 이채용과 이종영(1996)이 고유값이 1.0이상인 요인이 13개라고 보고한 것과 유사한 결과였다. 그러나, 장세진(1993)은 4개 요인을 선택하였으며, 이채용과 이종영(1996)도 4개를 선택하였고, Vazquez-barquero, William, & Diez-Manrique (1988)도 GHQ-60에 대한 연구에서 추출된 요인 13개 중 6개 요인을 선택하여 요인회전을 하였다. 본 연구에서도 장세진 등(1994)이 제시한대로 4개 요인을 지정하여 요인회전을 하였다. 요인 2와 4는 장세진 등(1994)의 연구와 거의 일치하였으나, 요인 1과 3은 양상이 혼재되어 나타났다.

장세진 등(1994)의 연구에서 한 요인으로 추출된 우울관련 문항들이 본 연구에서는 하나의 독립된 차원으로 나타나지 않고 요인 1과 3으로 나뉘어 추출되었으며, 요인 1은 우울과 불안관련 문항들이 하나의 차원으로 나타났다. 이는 대상집단이나 표본수가 달라지는 조사마다 요인구조가 달라질 수도 있음을 의미하기도 하나(이채용과 이종영, 1996), 또한 PWI의 문항 중 우울과 불

〈표 2〉 각 요인별 신뢰도 및 PWI와의 상관관계

요 인	Guttman 신뢰도계수	PWI와의 상관관계	요인 1	요인 2	요인 3	요인 4
요인 1 - 23문항 (우울증 및 불안)	.8637	.832**	[.893]	.359**	.283**	.398**
요인 2 - 11문항 (사회적 역할수행 및 자기신뢰)	.8123	.751**		[.846]	.178*	.463**
요인 3 - 8문항 (수면장애 및 우울)	.5516	.632**			[.541]	.184*
요인 4 - 3문항 (일반건강 및 생명력)	.7649	.804*				[.781]
PWI - 45문항	.7097					

**p<0.01(2-tailed), *p<0.05(2-tailed)
[] 요인별 신뢰도

〈표 3〉 각 측정도구와 PWI와의 상관관계

	GHQ-12	GHQ-20	GHQ-28	GHQ-30	GHQ-60
PWI	.744**	.805**	.891**	.824**	.905**

**p<0.01(2-tailed)

안 관련 문항들에 대한 명확화가 필요할 것으로 보인다.

본 연구에서 고유값이 1.0 이상인 요인들이 13개였으며, 4개의 요인들이 설명할 수 있는 분산의 양이 40.5% 였는데, 이는 측정변수 45개들로부터 이론변수 4개를 추정하는 모형이 측정된 자료를 잘 설명하지 못한다고 제시한 이채용과 이종영(1996)의 연구 결과와 일치하는 것으로 보인다. 또한 GHQ는 정신질환의 최소공배수들과 관련되었으므로 논리적으로 단차원 척도인 것이 더 타당하다고 하였고(Shek, 1993), 실제로 Shek(1993)는 중국어판 GHQ-30에 대한 연구에서 GHQ의 단일 차원척도를 증명한 바 있다. 본 연구에서도 요인에 대한 Scree chart에 따르면, 요인은 4개 보다는 2개가 더 적합한 것으로 나타나 추후연구에서 요인의 수 즉, PWI의 하부개념에 대한 연구가 필요할 것으로 보인다.

따라서, PWI가 4가지 하부 개념을 측정한다고 결론을 내리기에는 아직 미약하며, PWI의 요인 분석을 통해 스트레스를 설명할 수 있는 하부 개념 설정시 주의가 필요하다고 생각된다.

2. PWI의 신뢰도

신뢰도는 측정오차의 정도를 측정하는 것으로 그 척도가 과연 여러번 동일 대상에 적용했을 때 동일한 결과

를 산출하느냐에 관한 것으로, 신뢰도를 검증하는 데는 측정도구의 안정성을 살펴보는 검사-재검사법과 도구의 내적일치도를 검사하는 반분법, Cronbach's α 계수와, 관찰자간의 신뢰도를 살펴보는 집단내 상관계수, Kendall의 일치도 계수 등이 일반적으로 사용된다(이은옥 등, 1998).

본 연구에서는 내적 일치도를 나타내는 Cronbach's α 계수와 반분법을 사용하였다. Cronbach's α 계수는 문항분석까지 동시에 실시할 수 있어 타당도를 함께 측정할 수 있는 장점이 있으나, 각 문항의 동등성을 가정하고 있기 때문에 이 가정이 깨어질 때는 실제보다 더 낮게 측정되며, 문항수가 많으면 실제보다 크게 측정되는 단점이 있다(이채용과 이종영, 1996).

반분법은 가장 고전적인 방법으로 전후반부 또는 홀짝수에 의해 두 부분으로 나누어 신뢰계수를 측정하며, 이는 척도의 두 부분이 각각 전체 척도를 대표하는 것이어야 하므로 척도 자체가 하나의 동질성을 유지하고 있어야 한다.

본 연구에서의 Cronbach's α 는 0.894로 이는 장세진(1993)의 연구에서 보고된 0.94나 이채용과 이종영(1996)의 연구에서 나타난 0.93 보다는 낮은 수준이지만, 만족할 만한 수준이었다. 반분법에 의한 신뢰도 계수는 0.7097로 이는 Goldberg가 GHQ-60에 대한 반분

법에 의한 0.95, 장세진 등(1994)의 0.88에는 미치지 못하는 수준이었으나, 역시 만족할 만한 신뢰도를 나타냈다. 이와 같은 결과는 PWI의 문항이 전반적으로 이질적 내용을 포함하지 않아서 연구 대상자의 반응을 일관성 있게 얻을 수 있는 도구임을 의미한다.

각 요인별 신뢰도는 요인 1의 α 계수가 0.893으로 가장 높았으며, 요인 3이 0.541로 가장 낮았다. 요인 3의 신뢰도 수준이 낮은 수준이기는 하나, 이상의 4개의 요인에 대한 신뢰도 검정 결과는 집단비교에서의 최소한의 신뢰도 기준인 0.50(이상일, 1994)을 충족시키고 있었다. 따라서, PWI는 집단을 대상으로 쓰일 수 있는 정도의 신뢰도를 충분히 갖춘 도구라고 생각된다.

3. PWI의 타당도

타당도란 측정도구가 측정하려는 개념의 속성을 과연 측정하고 있는지를 나타내 주는 것으로 크게 내용타당도, 구성타당도, 준거타당도 등으로 분류될 수 있다(이은옥 등, 1998). 내용타당도는 측정도구의 내용이 대표성을 띄고 있는지를 하는 문제로, 연구자의 연구목적이나 방향에 의해 영향을 받을 수 있으므로(장세진 등, 1994) 처음부터 적절한 문항을 선정하기 위하여 전문가 평가단에 의한 평가를 권하고 있다. Streiner & Norman(이치용과 이종영, 1996에 인용됨) 내용타당도를 높이기 위해서는 문항 구성후에 요인분석을 통해 잘못된 문항을 수정하거나 재구성하는 것도 제안되고 있는데, 본 연구에서의 요인분석 결과 요인부하량이 모두 0.3이상이었으나, 23, 39, 40, 42문항들은 0.35 미만의 문항들이었다. 따라서, 추후연구에서는 낮은 값을 보이는 이 문항들에 대한 문항 수정이 필요할 것으로 생각된다. 또한 장세진 등(1994)의 연구와는 달리 우울관련 문항들이 불안관련 문항들과 함께 하나의 차원으로 나타나거나, 두 요인으로도 분리되어 나타났으므로, 앞으로 불안과 우울 개념에 대한 명확한 정의가 필요하다고 생각된다.

Goldberg의 GHQ-60와 GHQ-30, GHQ-28, GHQ-20, GHQ-12와 PWI의 상관계수는 0.744~0.905의 범위에 있었다. 이는 장세진 등(1994)의 0.91~0.99 보다는 낮은 수준이나, 만족스러운 결과였다. 또한, 요인들 간의 상관계수는 전반적으로 낮은 수준이었으며, 요인간의 독립성 여부를 판단하는 절대적인 기준은 없으나 Carey와 Seibert(1993)의 연구와 김정희(1996)의 연구에서 요인들간 상관계수가 0.6이하이고 Cronbach's α 계수보다 작은 경우를 독립성의 판단 기

준으로 이용한 바 있다. 본 연구에서의 요인간 상관계수는 모두 이 기준을 충족하고 있는 것으로 보아 스트레스를 나타내는 각 요인들이 각각 별도의 독립적인 차원의 개념을 측정하고 있음을 알 수 있었다. 또한 요인 분석에서 스트레스의 모든 문항이 1차적인 요인부하를 보이는 요인과의 상관관계가 가장 큰 것으로 보아, PWI의 문항들은 수렴의 원리와 판별의 원리를 충족시키고 있는 것으로 판단되었다. 따라서 PWI는 구성타당도를 가지고 있는 것으로 생각된다.

VI. 결론 및 제언

본 연구는 정상인의 스트레스 측정을 표준화하기 위해 Goldberg의 GHQ-60을 기초로 45개 문항으로 재구성한 사회심리적 건강 측정도구(Psychosocial Well-being Index)의 신뢰도 및 타당도를 검정하기 위하여 서비스업에 종사하는 근로자 186명을 대상으로 실시하였다.

연구도구는 Goldberg의 GHQ-60을 기초로 하여 장세진 등(1994)이 297명을 대상으로한 탐색 연구를 통해 2단계 수정을 거쳐 45개 문항으로 재구성한 PWI와, PWI와의 상관계수를 통한 타당도 검정을 위해 Goldberg의 GHQ-60와 GHQ-30, GHQ-28, GHQ-20, GHQ-12를 사용하였다.

PWI의 하부개념을 확인하기 위하여 요인분석과 Kruskal-Wallis test를 이용하였으며, PWI의 신뢰도는 Cronbach's α 와 반복법에 의한 Guttman 신뢰도 계수를 구하였고, 타당도는 요인분석 결과와, 타 도구와의 상관계수를 이용하여 분석하였다.

1. 요인분석을 통해 고유값이 1.0 이상의 요인은 13개였으며, 요인의 수를 4개로 지정하여 주성분분석과 Varimax 회전법으로 분석하여 장세진 등(1994)의 결과와 비교하였다. 그 결과 요인 2, 4는 장세진 등(1994)의 결과와 거의 일치하였으나, 요인 1과 3은 양상이 혼재되어 나타났다.
2. Cronbach's α 는 0.894였으며, 반복법에 의한 Guttman 신뢰도 계수는 0.7097였다. 각 요인별 신뢰도는 전반적으로 만족할 만한 수준이었으나, 요인 3이 0.541로 낮았다. 그러나 이는 집단비교에서의 최소한의 신뢰도 기준인 0.50을 충족시키고 있었다.
3. 요인들 간의 상관계수는 전반적으로 낮은 수준이었으며, 요인들간 상관계수가 0.6이하이고 Cronbach's α 계수보다 작아 각 요인들이 각각 별도의 독립적인

차원의 개념을 측정하고 있음을 알 수 있었다. 또한 요인 분석에서 스트레스의 모든 문항이 1차적인 요인부하를 보이는 요인과의 상관관계가 가장 큰 것으로 보아, PWI의 문항들은 수렴의 원리와 판별의 원리를 충족시키고 있는 것으로 판단되었다. Goldberg의 GHQ-60와 GHQ-30, GHQ-28, GHQ-20, GHQ-12와 PWI의 상관계수는 0.744~0.905 범위에 있었다.

따라서, PWI는 집단을 대상으로 쓰일 수 있는 정도의 신뢰도를 충분히 갖춘 도구이며 구성타당도도 가지고 있는 것으로 생각된다.

본 연구의 결과를 기초로 다음과 같은 제언을 하고자 한다.

1. PWI의 우울과 불안 관련 문항이 혼재되어 나타났으므로, 이 개념들에 대한 명확화가 필요하다.
2. PWI가 4가지 하부 개념을 측정한다고 결론을 내리기에는 아직 미약하며, PWI의 요인 분석을 통해 스트레스를 설명할 수 있는 하부 개념을 설정할 때, 요인수에 대한 고려가 필요하다.

참 고 문 헌

- 강병서, 석기석, 오윤조 (1993). 통계분석을 위한 SPSS/PC+. 서울: 무역영영사.
- 고맹남 (1996). 일부 제조업 산업장 근로자의 스트레스 수준과 관련요인 연구. 가톨릭대학교 석사학위논문.
- 권남희 (1995). 산업장 근로자의 스트레스와 정신건강 관련요인. 가톨릭대학교 석사학위논문.
- 김문석 (1990). 직무스트레스에 관한 모형 설정 연구. 고려대학교 박사학위논문.
- 김성아, 감신, 김건엽, 최광서, 정경동, 김혜숙 (1997). 스트레스 측정도구로서의 Hemoglobin A1c 유용성 평가. 대한산업의학회지, 9(3), 452-458.
- 김성현, 윤진희, 박혜순 (1992). 피로와 스트레스의 관련성. 가정의학회지, 13(3), 226-232.
- 김정희 (1996). 비만 성인의 건강관련 행위 - Johnson의 행위체계 모형 적용. 서울대학교 석사학위논문.
- 백기청 (1991). 스트레스와 생활사건-방법론 및 정신질환과의 관련성을 중심으로. 정신건강연구, 10, 10-36.
- 오수성, 이광섭, 손석준, 최진수, 이정애 (1995). 일반건강측정표를 이용한 일부지역 주민의 스트레스 수준에 관한 연구. 예방의학회지, 28(1), 123-129.
- 이영희, 박정일, 구정완, 임현우 (1997). 외국인 근로자의 근로실태가 스트레스에 미치는 영향. 대한산업의학회지, 9(2), 309-319.
- 이상일 (1994). 병원의래방문 환자의 만족도 평가 및 관련 요인에 대한 연구. 예방의학회지, 27(2), 366-376.
- 이영준 (1993). SPSS/PC+를 이용한 다변량 분석. 서울: 도서출판석정.
- 이우천 (1997). 병원종사자의 직업성 스트레스에 관한 연구. 서울대학교 박사학위논문.
- 이은옥, 임난영, 박현애 (1998). 간호, 의료 연구와 통계 분석(제3판). 서울: 수문사.
- 이종목 (1989). 직무스트레스의 원인, 결과 및 대책. 서울: 성원사.
- 이종목 (1997). 직업스트레스 스트레스-과학의 이해 (대한신심스트레스학회 편), 서울: 신광출판사, 71-92.
- 이재용, 이종영 (1996). Psychosocial Well-being Index의 신뢰도 및 타당도. 예방의학회지, 29(2), 255-264.
- 장세진 (1993). 건강통계자료 수집 및 측정의 표준화 연구, 스트레스 대한예방의학회, 121-159.
- 장세진, 차봉석, 박종구, 이은경 (1994). 스트레스 측정도구의 표준화 방안. 원주의대논문집, 7(1), 21-38.
- 차봉석, 고상백, 장세진, 박창식 (1996). VDT 취급근로자의 신체적 자각증상과 정신사회적 안녕상태의 관련성. 대한산업의학회지, 8(3), 403-413.
- 정도연 (1986). 행동과학, 건강과 질병의 정신생리학적 측면. 서울: 서울대학교 출판부, 151-166.
- 차철환 (1993). 산업보건관리역사의 활동과 지침, 정신보건관리의 실제. 고려대학교 환경의학연구소, 156-171.
- Banks, M. H. (1986). Validation of the GHQ in a young community sample. Psychol. Medicine, 13, 349-353.
- Carey, R. G., & Seibert, J. H. (1993). A patient survey systems to measure quality improvement : Questionnaire reliability and validity. Medical Care, 31(9), 834-845.
- Derose, S. (1997). Demographic and psychosocial factors. In Kane, R. L.(Ed), Understanding

- health care outcomes research(pp.175-210). Gaithersburg: Aspen publishers, Inc.
- Hobbs, P., Ballinger, C. B., & Smith, A. H. W. (1983). Factor analysis and validation of the General Health Questionnaire in women : a general practice survey. British Journal of Psychiatry, 142, 257-264.
- Maciejewski, M. (1997). Generic Measures. In Kane, R. L.(Ed). Understanding health care outcomes research(pp 19-35). Gaithersburg: Aspen publishers, Inc.
- McDowell, I., & Newell, C. (1996). Measuring health : A guide to rating scales and questionnaires(2nd ed.). New York: Oxford university press, 177-237.
- Murphy, L. R. (1996). Stress Management in Work Settings : A Critical Review of the Health Effects. American Journal of Health Promotion, 11(2), 112-135.
- Sheck, D. T. L. Factor structure of the chinese version of the general health questionnaire (GHQ-30) : A confirmatory factor analysis. Journal of Clinical Psychology, 49(5), 678-684.
- Shilling, S. & Brackbill, R. M. (1987). Occupational health and safety risks and potential health consequences perceived by US workers, Public Health Rep, 102, 36-46.
- Vazquez-barquero, J.L., Williams, P., Diez-Manrique, J.F., Lequerica, J., & Arenal, A. (1988). The factor structure of the GHQ-60 in a community sample. Psychol. Medicine, 18, 211-218.

-- Abstract --

Key concept : PWI, Stress, Reliability, Validity, Factor analysis

The Reliability and Validity Test of Psychosocial Well-being Index(PWI)

Kim, Jeong Hee*

The purpose of this study is to test the reliability and validity test of PWI to utility of PWI, this newly developed by Sejin Jang which measures stress.

The subject were 186 workers in service area.

Cronbach's α and Guttman split-half coefficient is used to test the reliability of PWI. Factor analysis and the correlation of the GHQ-60, GHQ-30, GHQ-28, GHQ-20, and GHQ-12 with the PWI is used to convergent validity and discriminant validity.

The important results of this study are as follows :

Cronbach's α coefficient of data was 0.894 and Guttman split-half coefficient was 0.7097.

The PWI was classified as 13 principle component (eigenvalue>1.0). After exploring 4 factor structure according to previous study result, 4 factors was explained 40.5% out of the total variance. The factor 1 was explained 15.9% and then the rest three factor was 24.6%. Factor 2 and 4 showed good agreement but factor 1 and 3 did not. Depression-related items were classified two factors. Anxiety and depression-related items were loaded unifactor. It was not clear that the PWI was consist of 4 concepts(factors). The correlation of the GHQ-60, GHQ-30, GHQ-28, GHQ-20, and GHQ-12 with the PWI were 0.744~0.905.

According to findings of this study, the PWI showed a high degree of validity and reliability. Thus it is recommended to use the PWI in general setting for screening for stress.

In addition, it is necessary to clarify the concept of depression and anxiety. In the further study, it may be considered to the factor structure of PWI and studied to two or unidimensional factor structure.

* Part-time instructor, Doctoral Candidate, College of Nursing, Seoul National University