

## 경험적 척도화에 의한 전기자료(Biodata)의 타당화 연구\*

강 해 린

이화여자대학교 경영학과

본 연구에서는 경험적 척도화(empirical scaling)에 의해 개발된 전기자료의 타당화 연구를 수행하였다. 65 문항으로 구성된 전기자료는 5,922명의 보험판매인력을 대상으로 실시되었으며 WAB기법에 의해 문항 채점을 위한 가중치가 산출되었다. 타당도분석 결과, 전체집단을 기초로 한 가중치는 다소 불안정한 것으로 나타났다. 따라서 전체집단을 고연령층과 저연령층의 하위집단으로 나누어 분석한 결과, 예측타당도가 증가되었으며 가중치도 안정적인 것으로 평가되었다. 기대도표와 적중률을 산출하여 전기자료가 갖는 선발도구로서의 실용성을 제시하였으며 마지막으로 경험적 척도화에 의한 전기자료 개발의 문제점과 이에 대한 연구방향에 대해 논의하였다.

인간의 행동을 설명하는 이론들은 각기 다른 관점을 취해왔지만 이들 대부분 이론들은 인간의 행동이 개인적 속성과 상황적 속성 모두와 관련성을 갖고 있다는데서 출발하고 있다(Schneider, 1987). 그런데 여기서 주목할 점은 인간의 행동을 조건지워주는 개인의 속성이 그냥 저절로 생겨나는 것은 아니라는 것이다. 그 보다는 유전적 측면과 아울러 개인의 과거경험이 일련의 행동경향을 심어주고 이는 다시 각 개인이 접하는 주변환경에 대한 지각, 선택, 반응양식 등을 조건지워주는 것이다(Buss, 1989). 심리학자들은 이러한 인간발달적 측면을 산업장면의 여러 문제들을 접근하는데 활용코자 하였으며 개인의 전기자료가 중요한 정보가 될 수 있음을 인식하였다.

개인의 전기자료를 묘사할 수 있는 형태는 다양하게 존재하지만 산업심리학자들은 인사 선발 및 배치 등과 같은 실용적 문제를 다루기 위해 점수화된 자전적 자료(scored autobiographical data), 일명 'biodata' 라는 형태에 의존하게 되었다. 전기자료의 사용은 개인의 과거 행동이나 경험이 미래에 나타날 행동이나 경험의 좋은 예측지표라는 가설에 기초한다. 그러나 이 가설은 한 개인이 과거에 행동한 그대로 미래에도 한결같이 행동한다는 것을 의미하지는 않는다. 이 가설이 뜻하는 바는 유전적 요인과 아울러 한 개인이 과거로부터 학습해 온 것으로 새로운 상황에서 특정 행동을 더 가능케 해 준다는 것이다(Mumford & Owens, 1987). 따라서 한 개인의 과거 행동이나 경험을 평가해 보면 그 사람의 미래 행동 예측이 가능하다는 것을 뜻한다.

\*. 본 논문을 주의깊게 심사해주신 익명의 심사위원들께 감사드립니다.

여기서 말하는 인간의 행동이란 산업장면에서

업무와 관련해 개인이 보이는 수행(performance)이라고 할 수 있다. 미래에 행해질 개인의 수행을 예측하기 위해 여러 가지의 선별도구가 사용되었는데 수행의 유형에 따라 예측도구의 유형이 달라진다. 예를 들어 정신능력, 기술, 숙련도와 같은 극대수행(maximum performance)이 요구되는 상황에서는 한 개인이 '무엇을 할 수 있는가'(can do)를 예측할 필요가 있으며 이를 위해 일반검사, 면접, 시뮬레이션 연습 등이 활용된다. 반면에 습관적 수행(typical performance)이란 일상적으로 접하는 상황에서 한 개인이 '무엇을 할 것인가'(will do)를 나타내며 전기자료는 극대수행보다는 바로 이러한 습관적 수행을 예측하는데 사용된다(Cronbach, 1960). 다시 말해 전기자료는 미래 상황에 나타날 한 개인의 습관적 행동을 예측하기 위해 과거 상황에서 보여준 그 개인의 습관적 행동을 측정하는 것이다. 전기자료의 예측력이 우수하다고 말하는 것은 이러한 예측변인과 준거변인의 동질성(fidelity)에 그 근거가 있다고 할 수 있다(Mitchell, 1994). 이는 바로 Wernimount와 Campbell(1968)이 언급한 바와 같이 예측에서 중요한 것은 '행동의 일관성'이며 이를 위해 준거에 대한 예측치는 행동의 '징후'가 아니라 '표본'이 되어야 한다는 것을 일깨워 주고 있다.

일반적으로 전기자료에 사용된 문항들은 과거 경험했던 사건이나 상황에 기초하고 있다. 따라서 각 개인의 답변은 회고적 자기보고의 성격을 지닌다. 과거 지나온 경험들에 대한 답변이 향후 관심의 대상인 특정 준거변인에 대한 예측력을 갖기 위해서는 답변자료를 어떻게 통합할 것인가 하는 기술적인 문제가 대두된다. 즉 전기자료의 척도화에 관한 문제이다. 전기자료를 연구하는 학자들은 다양한 척도기법들(예를 들어 경험적 척도화, 요인척도화, 하위집단

법 등)을 개발하였는데 그 중 경험적 척도화가 전통적으로 가장 오랫동안 광범위하게 사용되었으며 직무수행과 같은 준거에 대해 훌륭한 예측치를 제공하는 것으로 밝혀졌다(Ghiselli, 1973; Owens, 1976; Reilly & Chao, 1982).

경험적 척도화란 각 문항과 특정 준거와의 관련성을 토대로 문항을 선정하고 가중치를 부여하는 방법을 말한다. 이 방법에서는 문항 재검을 위한 가중치를 도출하기 위해 다양한 통계절차를 사용할 수 있는데(예를 들어 WAB, 상관분석법, 수평적비율산출법 등) WAB기법이 일반적으로 많이 사용되어 왔다(Mumford & Owens, 1987). WAB(weighted application blank)기법이란 특정준거와 관련하여 우수집단과 부진집단 사이를 경험적(통계적)으로 가장 잘 변별해 주는 문항들을 선정하고 그 변별력의 정도에 따라 가중치를 부여하는 절차이다. 각 문항에 부여된 가중치들은 합산되어 단일 점수를 구성하며 이 점수는 준거변인과 상관분석되어 그 상관정도에 따라서 전기자료의 예측력이 결정되는 것이다. 이렇게 도출된 가중치가 표집오차에 기인한 것이 아님을 확신하기 위해서는 교차타당화 절차가 요구된다.

경험적 방법에 의한 전기자료의 개발은 미국의 경우 일찌기 1910년대부터 시작되었는데 특이한 점은 주로 보험회사의 판매인력 선발을 개선시키려는 노력에서 시도되었다는 것이다. Goldsmith(1922)의 보험판매인력에 대한 실증연구는 이미 1922년에 Journal of Applied Psychology에 보고되었으며 그 이후 지금까지 다양한 영역에서 연구가 진행되었다. 예를 들면 종업원 이직(Cascio, 1976), 관리능력(Laurent, 1970), 연구능력 및 창의성(Smith, Albright, Glennon, & Owens, 1961), 종업원 정직성

(Rosenbaum, 1976), 그리고 경력개발(Klimoski, 1973) 등의 영역에서 전기자료가 예측도구로서 평가되었다. 최근에는 메타분석을 통해 타당도의 일반화 가능성이 제시되기도 하였으나(Rothstein, Schmidt, Erwin, Owens, & Sparks, 1990), Thayer(1977)와 Brown(1981) 등은 연령이나 혹은 조직의 관리행태와 같은 상황변인에 의해 전기자료의 타당도가 영향받을 수 있다고 하였다.

이와 같이 전기자료에 대한 수많은 실증연구가 보고되었음에도 불구하고 아직 우리 나라에서는 전기자료에 대한 체계적인 연구가 이루어지지 못하였다<sup>1)</sup>. 따라서 뒤늦은 감이 있으나 전통적인 WAB기법을 바탕으로 한 전기자료의 타당화연구를 시도해보는 데에 본 연구의 목적이 있다. 특히 보험산업에서 전기자료가 전통적으로 많이 활용된 점을 감안하여 본 연구에서도 우리 나라의 보험판매인력을 연구 대상으로 하였다. 아울러 Thayer(1977)가 주장한 대로 전기자료의 가중치도출에서 연령변인이 중요한 역할을 하는지도 함께 살펴보고자 한다. 이를 위해 우선 경험적 방법 즉 WAB기법을 활용하여 전기자료를 개발할 때 고려해야 할 제반적인 절차를 설명한 후 본 연구의 방법에 대해 기술하겠다.

## WAB 기법

### 준거변인 측정

경험적 방법에 의해 전기자료를 개발하는 경우 그 효과성은 얼마나 적합한 준거변인을 측정하느냐

1) 전기자료에 대한 박영석(1995)의 최근 연구는 요인구조 및 다른 검사들(인성, 적성 등)과의 관련성에 중점을 두었으며 타당도 분석은 이루어지지 않았음.

에 달렸다고 할 수 있다. 왜냐 하면 경험적 방법은 문항의 선정과 가중치 도출을 전적으로 준거와의 관련성에 의존하기 때문이다. 적합성, 신뢰성 및 실용성을 겸비한 준거변인을 개발하기 위해 필요한 첫 단계는 수행의 영역을 개념화하는 것이며 이를 위해 직무분석이 요구된다(Gatewood & Field, 1990). 그런데 전기자료 개발의 경우 정식으로 직무분석이 사용된 경우는 드물고(Fine & Cronshaw, 1994) 대신에 이직이나 생산성과 같은 객관적 자료 또는 업무 평가와 같은 주관적 자료가 주로 사용되었다.(Hunter & Hunter, 1984). 특히 경험적 방법에서는 객관적 자료(예를 들면 이직, 근속기간, 생산성 등)가 주관적 자료(예를 들면 업무평가)보다 훨씬 자주 사용되어 왔는데 기존의 연구문헌들에 따르면 객관적 준거보다 주관적 준거를 사용했을 때 전기자료의 예측력이 조금 더 높은 것으로 보고되었다(Reilly & Chao, 1982; Schmitt, Gooding, Noe, & Kirsch, 1984).

### 준거집단 구성

경험적 방법에 의한 전기자료 개발의 경우 전체 표본의 크기가 매우 중요하다. 일반적으로 표본의 크기는 문항의 수에 비례하기 때문에 문항의 수가 많으면 표본의 크기도 그만큼 커야 하는데 문항과 표본의 크기는 적어도 5:1 내지는 10:1의 비율이 바람직하다고 한다(Nunnally, 1978). 그런데 WAB 기법의 경우 문항의 각 선택지에 대해 가중치가 도출되기 때문에 전체 표본의 크기는 문항의 전체 선택지의 수에 비례해야 한다. 더군다나 가중치의 안정성 확인에 꼭 필요한 교차타당화 과정까지 고려한다면 WAB 기법의 적용은 상당히 큰 표본을 요구한다.

전체 집단의 크기와 아울러 중요한 것은 비교집

단의 구성이다. 즉 준거치의 수준에 따라 전체 조사 집단을 이분하여 비교집단(contrasting groups)을 구성하는 것이다. 만약 준거변인이 합격/불합격과 같이 이분적이라면 비교집단은 쉽게 만들어질 수 있다. 그러나 준거변인이 예를 들어 판매액과 같이 연속적이라면 전체집단을 준거치 기준으로 서열화한 후, 특정비율을 적용하여 상위집단과 하위집단을 구성해야 한다. 다양한 비율을 적용시킬 수 있겠으나 상위 27%, 하위 27%로 구성하였을 때 각 집단의 크기는 충분히 유지되면서도 두 집단의 차이를 가장 극대화시킨다고 한다(Ghiselli, Campbell, & Zedeck, 1981). 전기자료 문항의 가중치가 이들 비교집단의 문항반응 차이에서 비롯된다는 것을 감안할 때 준거집단의 구성이 매우 중요한 절차임을 강조할 필요가 있다.

### 문항개발

전기자료 문항을 개발할 때에는 준거예측에 도움이 될만한 행동적, 경험적 선행변인들을 충분히 포함시키기 위해 처음에 되도록 많은 수의 문항을 만들고 척도화 과정을 거쳐서 예측력 있는 문항들만 남게 된다. 전기자료에 대한 연구가 다양하게 진행되어 왔으나 문항의 내용에 관한 연구는 최근에 와서야 활발해 졌다. Asher(1972)와 Mael(1991)은 전기자료 문항들에 대한 분류기준을 나름대로 정리하기도 하였다.

전기자료 문항의 가장 기본적인 특성은 과거의 행동과 경험을 다룬다는 데 있으며 그 외에 의견, 태도, 가치관에 대한 문항도 일반적으로 수용되고 있다. 모든 예측도구가 그러하듯이 전기자료의 경우에도 내용적으로 직무수행의 영역과 관련된만한 속성들을 포함시키는 것이 바람직하다. 이를 위해서는

직무분석에 의한 정보가 활용된다. 그 밖에도 발달 심리학 문헌, 직무담당자와의 인터뷰, 기존에 활용되고 있는 특정문항 등과 같은 정보들이 문항개발에 유용하게 사용될 수 있다(Hogan, 1994).

Mumford와 Owens(1987)는 전기자료의 요인분석에 관한 21개의 연구들을 검토한 끝에 일반화될 만한 26개의 개념적 차원들을 발견하였다. 이들 중 여러 연구에서 빈번히 등장한 차원으로는 생활적응, 성숙정도, 학업성취도, 사회생활에서의 리더십 등을 들 수 있다. 전기자료의 경우 판매인력을 대상으로 한 연구들이 특히 많이 수행되었는데 이들 연구에서 사용된 차원은 경제적 책임감, 생활의 안정성, 가족과의 관계, 학업성적, 사회적 리더십 등이 포함된다(Baehr & Williams, 1968; Mitchell & Klimoski, 1982).

문항의 내용과 아울러 문항개발에서 중요한 것이 선택지의 형태이다. 일반적으로 중다선택형이 사용되며 선택지가 연속형인 경우 등간척도가 요구된다. 그런데 경험적 방법에서는 비연속형 선택지의 문항이 일부 포함될 때가 많다. 문항의 내용상 선택지가 정성적일 수 밖에 없는 경우를 제외한다면 선택지는 가능한 한 연속형인 것이 바람직하다고 한다(Owens & Mumford, 1987).

### 문항 가중치 산출

문항 가중치 산출은 WAB 기법의 핵심에 속한다. 상관분석법이나 회귀분석법을 사용하면 각 문항에 대해 하나의 가중치가 산출되는 데 비해 WAB 기법에서는 문항의 각 선택지마다 가중치가 산출된다. 구체적인 과정을 살펴보면, 우선 비교집단 즉

우수집단(upper criterion group)과 부진집단(lower criterion group)이 각 선택지에 대해 보인 반응빈도의 차이를 구한다.

퍼센트로 나타낸 두 집단의 빈도차이는 Strong(1926)이 만든 표(Strong's Tables of Net Weights for Differences in Percents)를 참조하면 순가중치(net weight)가 손쉽게 얻어진다(England, 1971). WAB 기법의 경우 각 선택지에 대한 가중치를 도출할 때에는 수많은 phi 계수를 일일이 계산해야 되는데 Strong(1926)은 이러한 수고를 덜어주기 위해 순가중치를 나타낸 표를 만들었다. 우수와 부진집단간의 반응빈도에서 차이를 나타내지 못하는 선택지는 제거되며 부진집단에서 더 큰 반응빈도를 보인 선택지의 경우는 부(-)의 가중치를 갖는다.

### 교차타당화

문항채점을 위한 가중치 산출이 WAB 기법처럼 전적으로 경험적(통계적) 절차에 의존하는 경우 교차타당화의 중요성이 특히 강조된다. 왜냐 하면 특정한 표본을 사용하여 개발된 가중치는 표집오차의 가능성을 배제할 수 없어서 가중치의 안정성이 우려되기 때문이다. 따라서 특정 표본을 토대로 개발된 가중치가 별도의 표본에 적용되었을 때 얼마나 안정적인지를 확인해야 한다. 교차타당화를 위해 별도의 독립적인 표본을 사용하는 것이 바람직하나 실제로 자주 사용되는 방법은 연구대상인 전체표본을 둘로 나누어 가중치를 개발하는 표본(development sample)과 교차타당화 표본(cross-validation sample)으로 구분한다. 이 때 교차타당화 표본보다 개발 표본에 1/3의 인원을 더 많이 할당하는 것이 가중치 개발을 위해 유익하다고 한다(England, 1961). 개발표본을 토대로 산출된 가중치를 교차타당화 표본에 적

용시켰을 때 발생하는 타당도계수의 감소정도(shrinkage)를 검토함으로써 가중치의 안정성을 확인할 수 있다. 교차타당화를 마친 후 실제 지원자에게 적용될 채점용 가중치는 개발표본과 교차타당화 표본을 합친 전체 표본을 사용하여 최종적으로 다시 도출한다(Pedhazur, 1982).

## 방 법

### 문항개발

전기자료에 포함될 문항은 기존에 사용되었던 문항들 중에서 선별하거나 내용을 일부 수정하기도 하였으며(Glennon et al, 1966; Klimoski & Mitchell, 1980), 또 일부는 보험판매영역에 부합되도록 새로이 만들어지기도 하였다. 이를 위해 보험판매직에 대한 기존의 직부분석자료와 함께 현재 활동하고 있는 보험판매인들과의 인터뷰 자료를 참고로 하였다. 문항을 구성함에 있어서 내용적으로 다음의 영역(차원)들을 포함시키려고 노력하였다; 학업성취, 부모의 사회적/경제적 배경, 경제적 책임, 사회적 리더십, 가족간의 인간관계, 개인생활의 안정성, 그리고 일과 관련된 경험 및 태도. 이들 영역들은 기존의 전기자료 연구에서 중요한 차원으로 언급되었고 특히 다른 판매영역(예를 들면 부동산 판매인)에서 그 차원성이 인정되었다(Klimoski & Michell, 1980). 본 연구는 이론적(rational) 접근이 아닌 경험적 방법에 의존하기 때문에 요인분석과 같은 방법을 사용하여 그 차원성을 확인하지는 않지만 내용적으로 이들 영역을 표현하려고 노력하였다. 이러한 과정을 거쳐 총 98개의 초기단계 문항을 개발하였으며 예비검사를 실

시하여 특정 선택지에 대해 지나치게 응답이 편중된 문항들을 제거한 후 65개로 축소되었다. 연구에 사용된 모든 문항은 중다선택형으로 이루어졌다.

### 신뢰도

WAB기법하에서는 검사-재검사 신뢰도 측정이 중요하게 여겨진다(Brown, 1994). 본 연구에서는 95명의 보험판매인을 대상으로 3개월의 간격을 두고 검사-재검사를 실시한 결과 .78의 신뢰도 계수를 얻었다. 지금까지 보고된 전기자료의 검사-재검사 신뢰도가 .60~.96임을 감안할 때 본 연구의 전기자료는 신뢰할 만하다고 본다(McManus & Mitchell, 1987).

### 자료수집

개발된 전기자료는 우리나라의 대형 생보사에서 활동 중인 여성판매인력을 대상으로 실시되었다. 조사지역은 전국의 영업국을 고루 포함시켰으며 1994년 6월과 7월에 걸쳐 2회로 나누어 실시되었다. 총 6057부 중 5922부가 회수되었는데 연령별로 보면 20대가 17.4%, 30대가 45%, 40대가 23.5%, 50대가 10.8%, 60대 이상이 3.3%인 것으로 나타났다. 학력은 전체 표본의 63.5%가 고졸이었으며 80.8%가 기혼자인 것으로 조사되었다.

### 준거집단 구성

보험산업의 특성상 우수 인력이란 기본적으로 보험 판매를 잘 해야 될 뿐만 아니라 기존 계약자들이 중도 파기를 하지 않도록 사후 관리가 중요시 되며 아울러 신인 판매인의 모집에도 기여해야 한다.

성공적 수행에 요구되는 이러한 요인들을 기초로 한 업적지표를 만들어 해당 기업은 이를 판매인의 보상 기준으로 삼고 있다. 따라서 보험판매직의 성공적 수행에 대한 준거 자료는 회사에서 활용하고 있는 업적지표를 사용하였다. 이 업적지표는 크게 3가지 요인 즉 신계약건수, 기계약자 유지율, 그리고 신인 설계사 증원 등으로 구성되어 매월 산출되었는데 본 연구에서는 3개월 업적지표의 평균치를 사용하였다. 가중치 도출에 필요한 준거비교집단을 구성하기 위해 전체인원(5922명)을 업적지표 기준으로 서열화시킨 다음 상위 30%(1800명)을 우수집단, 하위 30%(1800명)을 부진집단으로 정의하였다.

### 가중치 산출

전기자료 문항에 대한 가중치 산출은 앞에서 설명한 전통적인 WAB 방법을 따라서 실시하였다(England, 1971). 우선 문항의 각 선택지에 대해 우수와 부진집단이 보인 반응빈도의 차이를 퍼센트로 나타낸 다음 Strong의 표를 참조하여 순가중치(net weight)을 구했다.<sup>2)</sup> 이때 선택지의 반응빈도가 우수보다 부진집단에서 더 큰 경우 부(-)의 가중치를 갖게 된다. 가중치 산출의 한 예가 표 1에 제시되어 있다. 본 연구에서는 우수와 부진집단을 차별화시키는데 실패한 9개의 문항이 탈락되었고 그 결과 총 56문항이 타당도 분석에 사용되었다. 최종 문항들에 대한 순가중치는 모두 더해져서 각 개인마다 WAB점수가 산출되었는데 전체 집단의 WAB점수 범위는 -26~+37이었으며 평균값은 -3.12였다.

2) Strong의 표(Strong's Tables)는 Stead와 Shartle(1940)의 연구에서 처음 보고되었고 그 후 England(1961, 1971)와 Guion(1965)의 연구에서 재보고 되었다.

**표 1. 가중치 산출의 예**

당신이 사용하고 있는 카드는 몇 개입니까?	I	II	(I-II)	순가중치
0	14.2	22.8	-9	-2
1	23.8	25.0	-1	0
2	25.6	26.4	0	0
3-4	27.8	18.6	9	2
5개 이상	6.8	4.8	2	1
	100%	100%		

주. 1) I= 우수집단, II= 부진집단.  
 2) 순가중치는 Strong표를 참조함.

**분석 및 결과**

WAB점수의 준거관련타당도를 산출하기 위해 전체표본의 우수집단(1800)과 부진집단(1800) 각각에서 2/3를 개발표본, 1/3을 교차타당화 표본으로 삼았다. 먼저 개발표본(우수1200+부진 1200)을 사용하여 가중치 및 WAB점수를 산출한 후 이를 준거점수(업적지표)와 상관을 낸 결과, 타당도계수  $r=.30$ 을 얻었다. 교차타당화를 위해 개발표본에서 얻은 가중치를 교차타당화표본(우수 600+부진 600)에 적용하여 타당도를 구한 결과 교차타당도 계수  $r=.23$ 을 얻었다. 타

당도 감소의 통계적 검증 결과( $z=2.15, p<.05$ ), 개발표본의 타당도와 교차 타당도 사이에 유의미한 차이가 있는 것으로 나타났다. 즉 개발표본에서 도출된 가중치가 전체표본에 적용 되기에는 다소 불안정한 것으로 판단되었다. 따라서 WAB점수 산출에 사용된 가중치에 영향을 줄 수 있는 여러 변인들 가운데 연령을 생각하게 되었다. 이는 본 표본의 연령층이 20대부터 60대까지 매우 광범위할 뿐 아니라 이미 기존의 전기자료 연구(Thayer, 1977)에서 연령을 조절변인으로 언급하였기 때문이다. 따라서 연령을 기준으로 하위집단을 구성하여 별도의 타당도 분석을 실시한 후 이를 비교해 볼 필요성을 느꼈다.

**연령별 하위집단 구성**

20대에서 60대에까지 이르는 전체 표본을 35세를 기준으로 하여 두 집단으로 분류하였다. 35세 미만자( $n=2340$ )를 저연령집단, 35세 이상자( $n=3557$ )을 고연령집단으로 정의한 후 두 하위집단은 각자의 준거비교집단을 갖게 되었다. 즉 저연령집단의 상위 30%( $n=690$ )을 우수집단, 하위 30%( $n=690$ )을 부진집단으로 하였으며 고연령집단의 경우 역시 상위 30%( $n=1050$ )을 우수집단, 하위 30%( $n=1050$ )을 부진

**표 2. 각 하위집단의 평균값**

	저연령 집단		고연령집단	
	우수	부진	우수	부진
WAB점수	-4.9	-5.49	-.03	-5.88
업적지표	9,588	2,770	10,263	2,852
활동기간(개월)	13	15	63	45

집단으로 하였다. WAB점수를 얻기 위한 가중치 도출이 저연령집단과 고연령집단에 대해 각각 실시되었다. 저연령집단과 고연령집단의 WAB평균점수는 각각 -2.99와 -2.95였다. 각 하위집단에 대한 평균값이 표 2에 요약되어 있다.

### 타당도 분석

WAB점수와 준거점수의 상관분석을 저연령집단과 고연령집단 각각에 대해 실시하였는데 두 집단 모두  $r=.32$ 라는 타당도 계수가 산출되었다. 이는 연

때문에 하위집단에 기초한 가중치의 안정성을 보는 것은 매우 중요하다. 따라서 하위집단에서 도출된 가중치와 타당도 계수가 특정 표본에 한정된 것이 아님을 확인하기 위해 교차타당도 분석이 요구된다. 하위집단의 우수와 부진에 대해 개발 표본과 타당화 표본에 포함될 인원을 나누게 되는데 이때 개발표본에 1/3의 인원을 더 많이 할당하여 2:1의 구성 비율을 갖도록 하였다(England, 1961). 각 집단의 구체적인 인원수가 표 3에 요약되어 있다.

저연령집단과 고연령집단은 각각의 개발표본을 토대로 가중치를 산출하였고 이를 각각의 교차타당

표 3. 교차타당화를 위한 표본구성

하위집단	우수		부진	
	전체 N	표집 N	전체 N	표집 N
고연령집단	1050	D: 700 H: 350	1050	D: 700 H: 350
저연령집단	690	D: 460 H: 230	690	D: 460 H: 230

- 주. 1) D=개발표본 H=교차타당화 표본을 나타냄.  
2) 고연령집단의 개발표본은 700+700=1400, 교차타당화 표본은 350+350=700이며 저연령집단의 개발표본은 460+460=920, 교차타당화 표본은 230+230=460임.

령을 기초로 하위집단을 구성하기 전에 전체집단을 기초로 산출한 타당도 계수  $r=.27$ 에 비해 개선된 결과로 보여진다. 따라서 본 연구에서는 전체 표본을 분석대상으로 하는 것보다 저연령집단과 고연령집단의 하위집단으로 나누는 것이 더 바람직하다고 판단된다.

### 교차타당도 분석

전체집단을 기초로 가중치를 산출하였을 때 타당도 감소가 유의미하여 불안정한 것으로 밝혀졌기

때문에 하위집단에 기초한 가중치의 안정성을 보는 것은 매우 중요하다. 따라서 하위집단에서 도출된 가중치와 타당도 계수가 특정 표본에 한정된 것이 아님을 확인하기 위해 교차타당도 분석이 요구된다. 하위집단의 우수와 부진에 대해 개발 표본과 타당화 표본에 포함될 인원을 나누게 되는데 이때 개발표본에 1/3의 인원을 더 많이 할당하여 2:1의 구성 비율을 갖도록 하였다(England, 1961). 각 집단의 구체적인 인원수가 표 3에 요약되어 있다. 저연령집단과 고연령집단은 각각의 개발표본을 토대로 가중치를 산출하였고 이를 각각의 교차타당

화 표본에 적용시킨 후 타당도 계수의 축소 정도를 비교하였다. 이에 대한 결과가 표 4에 요약되었다. 고연령집단의 경우 개발 표본과 교차타당화 표본의 타당도 계수가  $r=.31$ 로 일치하여 축소부분이 발생하지 않았으므로 가중치가 매우 안정적인 것으로 판단된다. 저연령집단의 경우 개발표본의 타당도(.33)와 교차타당도(.24)가 차이를 보였지만 이는 통계적으로 유의미하지 않았다 ( $z=1.71, p<.90$ ). 따라서 저연령집단과 고연령집단 각각에서 도출된 가중치는 안정성을 갖는 것으로 판단된다.



표 4. 하위집단별 교차타당도 계수

하위집단	개발표본		교차타당화 표본		
	n	r	n	r	s <sup>a</sup>
고연령	1400	.31	700	.31	-
저연령	920	.33	460	.24	.09 <sup>b</sup>

주. a. 타당도계수의 축소(shrinkage)

b. z=1.71, p<.90

판정기준 점수 산출

WAB기법에 따라 개발된 전기자료가 선발의 목적에 사용되는 경우 거쳐야 할 마지막 단계는 예측의 정확성을 극대화시켜 주는 판정기준(cutoff) 점수를 산출하는 것이다. 일반적인 방법으로는 우수집단의 점수분포곡선을 부진집단의 분포곡선과 비교하여 두 분포의 중복이 최소가 되는 곳 즉 교차점을 판정기준 점수로 인식한다(Hogan, 1994). 그런데 본 연구에서와 같이 준거가 우수와 부진집단으로 이분화되는 경우 기대도표(expectancy chart)를 활용하면

전기자료점수

25 이상	***** 92%
15~24	***** 82%
5~14	***** 71%
-5~4	***** 52%
-15~-6	***** 37%
-25~-16	***** 22%
-25이하	** 9%

성공확률(%)

그림 1. 기대도표 (고연령 집단)

각 점수대에 대응되는 성공확률을 보여주기 때문에

준거와 전기자료점수의 연관성을 보다 쉽게 이해할 수 있다(Anastasi,1990). 그림 1과 그림 2는 각각 저연령집단과 고연령집단에 대한 기대도표를 보여주고 있다.3) 그림 1은, 예를 들어, 35세 이상의 지원자가 25점 이상을 받을 경우 성공할 확률 즉 우수집단에 포함될 확률이 92%라는 것을 의미한다. 그런데 판정기준 점수는 목표로 하는 성공확률이 얼마냐에 따라서 달라지게 되므로 지원자에 대한 성공확률을 높게 세울수록 판정기준 점수는 높아지고 성공확률을 낮출수록 판정기준 점수도 낮아질 것이다.

전기자료점수

25 이상	***** 95%
15~24	***** 74%
5~14	***** 61%
-5~4	***** 50%
-15~-6	***** 29%
-16이하	***** 19%

성공확률(%)

그림 2. 기대도표 (저연령 집단)

한편 전기자료 점수에 의한 성공 예측이 얼마나 실용성을 갖고 있는지, 알기 위해서는 하위집단에 대한 적중률을 산출하였다. 즉 전기자료 점수를 사용해

3) 기대도표를 만드는 구체적인 설명은 Anastasi(1990), Muchinsky(1993)등 여러 문헌에 소개되어 있다.

표 5. 판별결과

실제	고연령 집단		저연령 집단	
	예측		예측	
	0	1	0	1
0	676 (64.4%)	374 (35.6%)	454 (65.8%)	236 (34.2%)
1	380 (36.2%)	670 (63.8%)	235 (34.1%)	455 (65.9%)
판별의 정확성	64.10%		65.87%	

주. 0 = 부진집단, 1 = 우수집단

서 우수와 부진으로 예측한 결과가 실제로 얼마나 정확했는지를 알아보는 것이다. 이를 위해 하위집단의 평균점수(group centroid)를 기준 점수로 사용한 canonical 판별함수를 구했고 그 결과가 표 5에 요약되어 있다. 고연령집단의 경우 전체 인원(n=2100)의 64.10%가 정확하게 판별되었고 저연령집단의 경우 65.87%의 정확성을 보였다. 현재 연구 대상의 보험회사 경우에 기존 집단의 성공률(base rate)이 50% 정도인 것을 감안할 때 전기자료의 사용은 어느 정도의 유용성을 갖는 것으로 판단된다.

### 논 의

본 연구에서는 경험적 척도화 방법으로 일반적으로 많이 사용되어온 WAB기법을 바탕으로 하여 전기자료의 타당화 연구를 실시하였다. 교차타당화 과정을 거친 결과 전기자료가 보험판매인력의 선발 도구로 예측력이 있음이 제시되었다고 본다. 한편 가

중치 도출에서 연령을 기초로 하위집단을 구성하였을 때 예측력이 증가하였을 뿐 아니라 더 안정적인 것으로 나타났다. 이 연구의 결과로부터 경험적 척도화의 유용성을 다시 한번 확인할 수 있었음에도 불구하고 경험적 방법에 대한 비판 또한 꾸준히 제기 되어왔다. 따라서 경험적 방법의 문제점과 아울러 전기자료 연구와 관련된 앞으로의 방향 등에 대해 몇 가지 논의하고자 한다.

전기자료에 대한 실증연구들을 총체적으로 분석한 문헌들(Ghiselli, 1966; Reilly & Chao, 1982; Mumford & Owens, 1987)은 경험적 척도화에 의한 전기자료가 준거(예, 직무수행이나 훈련)를 예측하는데 매우 훌륭한 도구임을 인정하였다. 그러나 전기자료의 특정 문항이 왜 준거변인과 밀접한 연관성을 갖는지에 대한 이해의 틀을 제공하지 못한다는 비난을 받아왔다. 예를 들어 '유년시절 읽었던 책의 수' 혹은 '가입한 생명보험의 금액' 등의 문항이 준거, 즉 업무수행과 연관성이 있는 것으로 밝혀졌다면 왜 이러한 내용의 문항은 예측력이 있는데 다른 문항들은

준거와 연관성을 갖지 못하는지에 대한 설명을 해주지 못한다. 즉 준거에 대한 예측 능력만을 토대로 문항을 선정하는 극단적 경험주의는 문항들의 차원성이 도의시되어 이론적 해석을 취약하게 만든다고 할 수 있다. 이러한 문제점을 극복하고 전기자료를 좀 더 논리적으로 점수화하려는 노력들이 이론적 척도화(rational scaling)의 토대를 구축하였다 (Owens, 1971; Owens & Shoenfeldt, 1979). 이론적 척도화는 내적 합치도 분석이나 요인분석을 사용하여 해석가 능한 차원을 도출하기 때문에 경험적 방법에 비해 가장 큰 우위성은 심리적 이해와 이론적 가설이 가능하다는 것이다.

모든 측정의 문제에서 처럼 전기자료의 경우에도 극단적인 경험주의라든가 혹은 전적으로 이론 추구적인 접근은 모두 최적이 아닐 것이다. 인력선발과 같이 매우 실용적인 목적을 추구하는 경우에도 예측력과 아울러 그 도구개발에 최소한의 이론적 기반은 있어야 한다고 본다. 이러한 관점을 지지해 주는 최근의 연구를 보면 예측력을 극대화하면서도 전기자료의 심리적 이해를 높이기 위해 경험적 방법과 요인분석적 방법을 절충한 접근법이 매우 효용성이 높다는 것을 보여주고 있다(Dalessio & Crosby, 1991).

경험적 척도화에 대한 또 다른 비판은 경험적으로 개발된 전기자료의 일반화 가능성에 관한 것이다. 즉 경험적 채점키(scoring key)는 특정 표본을 사용해서 특정 준거에 대해 개발되었기 때문에 그 타당도는 상황에 따라 달라지고 일반화되지 못한다는 것이다(Hunter & Hunter, 1984). 따라서 여러 학자들은 전기자료의 타당도에 영향을 미치는 상황변인 내지는 조절변인의 역할에 관심을 갖게 되었다. 타당도에 영향을 미치는 요인으로는 연령, 조직관련 변인, 사

용된 준거변인 등이 거론되었다(Thayer, 1977). 특히 연령의 경우 전기자료 개발에서 중요한 의미를 갖는다고 할 수 있다. 일찌기 Kurtz(1941)는 보험판매인의 성공을 예측하는 그의 연구에서 연령의 역할이 중요함을 인식하였다. 그는 전기자료 문항과 성격검사 문항을 모두 사용하였는데 26세 이하의 집단에서는 성격검사 문항이 더 큰 가중치를 도출하였지만 26세 이상의 집단에서는 반대로 전기자료 문항들이 더 많이 기여한 것으로 보고하였다. 그런데 일부 학자들은 경험적 척도화의 일반화가능성을 전혀 불가능한 것으로 여기지 않는다. 예를들어, Rothstein과 동료들(1990)의 연구는 전기자료의 타당도 일반화가 가능하다는 것을 보여주었다. 이들은 경험적 척도화가 본질적으로 '상황적 특수성'을 갖고 있기 때문에 일반화가능성이 실패했다고 보지 않는다. 대신에 문항을 개발할 때 이론적인 직무관련성을 좀더 염두에 두며 준거와 비교집단 선정시에도 특별한 관심을 기울인다면 타당도가 일반화가능성을 갖는다고 주장하였다.

한편 전기자료와 같은 비인지적 예측도구의 개발에서 오랫동안 관심을 끌어온 것이 공인타당도와 예측타당도의 비교이다(Guion & Cranny, 1982). 즉 타당화 연구의 대상으로 현재 수행집단을 사용하느냐 아니면 지원자집단을 사용하느냐에 따라 타당도가 달라질 수 있다는 것이다. 일부 학자들은 전기자료의 공인타당도가 예측타당도를 항상 능가한다고 보고 있는데 이는 직무 경험에서의 개인차가 타당도에 영향을 미친다는 가설을 인정하기 때문이다(Hough, 1986; Hogan, 1988). 그러나 Rothstein과 동료들(1990)의 메타연구 결과는 공인타당도가 직무경험의 증가에 따라 영향을 받지 않음을 보여주고 있기 때문에 이 가설에 대한 지지는 명확치 않다고 본

다. 만일 전기자료의 공인타당도가 예측타당도보다 항상 높게 산출된다는 것이 사실이라면 본 연구를 비롯한 대부분의 타당도 계수는 지원자를 대상으로 하는 실제 선발상황에서는 낮아질 것임을 감안하여 보수적으로 평가되어야 할 것이다.

마지막으로 전기자료와 관련된 법적인 측면을 살펴보고자 한다. 강력한 고용평등법이 시행되는 나라에서는 전기자료의 사용이 법적인 문제를 야기시킬 수 있다. 경험적 척도화로 전기자료를 개발하는 경우 문항의 내용이 직무와의 관련성을 분명히 보이지 못하는 경우가 많고 이는 법적 소송의 가능성을 높인다고 할 수 있다. 물론 준거관련타당도가 직무의 관련성을 증명하는 한 방법이 될 수 있지만 법적인 방어력을 갖기 위해서는 문항의 내용 자체가 명백히 직무와 관련되었음을 보일 필요가 있다. 따라서 내용 타당화의 중요성이 부각되는데 이를 위해 직무분석 자료가 유용하게 쓰일 것이다(Pace & Schoenfeldt, 1977). 아울러 문항을 선정할 때 준거와의 예측력이나 직무관련성만을 기준으로 삼기보다는 개인의 프라이버시 침해 여부도 간과해서는 안될 것이다. 미국의 경우 예를 들어, 주거의 소유형태, 연령, 결혼 여부, 자산이나 부채의 정도 등에 대한 문항은 판매직의 경우 타당한 예측력을 지닌 것으로 밝혀졌으나 여러 주정부에 의해 그 사용이 금지되고 있다(Mitchell, 1989). 이에 비한다면 우리나라의 심리학자들은 문항개발에서 아직까지 매우 자유로운 편이지만 평등고용 관련 법령의 시행이 점점 강화되고 있는 추세임을 감안할 때 앞으로 이에 대한 배려도 있어야 할 것으로 생각된다.

이상에서 살펴보았듯이 경험적 척도화에 의한 전기자료의 개발은 많은 문제점을 안고 있지만 다각

적 측면에서 여러 연구가 활발히 진행되고 있다. 미국의 경우 약 100년 가까이 보험판매인력의 선발에 전기자료를 활용해 왔다. 현재 우리나라의 보험판매인력은 30만 정도이지만 이들의 선발은 주관적 판단에 의존해 왔으며 낮은 생산성과 높은 이직률 때문에 대량 채용, 대량 탈락의 악순환이 반복되고 있다. OECD 가입으로 보험산업이 개방되면 많은 선진 기업들과의 경쟁을 피할 수 없을 것이다. 이를 위해 보험산업의 핵심인 판매인력에 대한 체계적인 선발도구의 활용이 절실히 필요한 때이다.

## 참 고 문 헌

- 박영석(1995). 이직자의 생활사. 한국심리학회지: 산업 및 조직, 8(1), 75-102.
- Anastasi, A. (1990). *Psychological testing* (6th ed.). Macmillan Publishing Company.
- Asher, J. J. (1972). The biographical item: Can it be improved? *Personnel Psychology*, 25, 251-269.
- Baehr, M., & Williams, G. B. (1967). Underlying dimensions of personal background data and their relationship to occupational classification. *Journal of Applied Psychology*, 51, 481-490.
- Brown, S. H. (1981). Validity generalization and situational moderation in the life insurance industry. *Journal of Applied Psychology*, 66, 664-670.
- Brown, S. H. (1994). Validating biodata. In G. S. Stakes, M. D. Mumford, & W. A. Owens (Eds.), *Biodata handbooks*. Palo Alto: CPP

Books.

- Buss, A. H. (1989). Personality as traits. *Amerinan Psychologist*, 44, 1378-1388.
- Cascio, W. F. (1976). Turnover, biographical data, and fair employment practice. *Journal of Applied Psychology*, 61, 576-580.
- Cronbach, L. J.(1960). *Essentials of psychological testing*(3d ed.). New York: Harper & Row.
- Dalesio, A. T., & Crosby, M. M. (1991, April). *Validity or understanding: A comparison of three biodata scoring methods*. Paper presented at the 6th annual conference of the Society for Industrial and Organizational Psychology, st. Louis.
- England, G. W. (1961). *Development and use of weighted application blanks*. Dubuque, IA: Brown.
- England, G. W. (1971). *Development and use of weighted application blanks* (Rev. ed). Minneapolis: University of Minnesota, Industrial Relations Center.
- Fine, S. A. , & Cronshaw, S. (1994). The role of job analysis in establishing the validity of biodata. In G. S. Stokes, M. D. Mumford, & W. A. Owens (Eds.), *Biodata handbook*. Palo Alto: Cpp Books.
- Gatewood, R. D., & Feild, H. S. (1990). *Human resource selection* (2nd ed). Chicago: Dryden Press.
- Ghiselli, E. E. (1966). The validity of occupational aptitude tests. New York: Wiley.
- Ghiselli, E. E. (1973). The validity of aptitude tests in personnel selection. *Personnel Psychology*, 26, 461-477.
- Ghiselli, E. E., Campbell, J. P., & Zedeck, S. (1981). *Measurement theory for the behavioral sciences*. San Francisco: Freeman.
- Glennon, J. R., Albright, L. E., & Owens, W. A. (1966). *A Catalog of life history items*. Greensboro: Center for Creative Leadership.
- Goldsmith, D. B. (1922). The use of personal history blanks as a salesmanship test. *Journal of Applied Psychology*, 6, 149-155.
- Guion, R. M., & Cranny, C. J. (1982). A note on concurrent and predictive validity designs: A critical reanalysis. *Journal of Applied Psychology*, 67, 239-244.
- Hogan, J. B. (1988). *The influence of society desirable responding on biographical data of applicant versus incumbent samples: Implications for predictive and concurrent research designs*. Unpublished doctoral dissertation, University of Georgia, Athens.
- Hogan, J. B. (1994). Empirical keying of background data measures. In G. S. Stokes, M. D. Mumford & W. A. Owens(Eds.), *Biodata handbook*. Palo Alto: CPP Books.
- Hough, L. (1986). *Utility of temperament, biodata, and interest assessment for predicting job performance: A review and integration of the literature*. (PDRI Rep. No. 145)
- Hunter, J. E. , & Hunter, R. F. (1984). Validity and utility of alternative predictors of job performance. *Psychological Bullentin*, 96, 72-98.
- Klimoski, R., & Mitchell, T. (1980). Toward

- predicting entry and initial success in the pursuit of real estate careers: The longitudinal validation of a life history inventory. (Report Number 14). Center for Real Estate Education and Research, Ohio State University.
- Klimoski, R. J. (1973). Biographical data analysis of career patterns in engineering. *Journal of Vocational Behavior*, 3, 103-113.
- Kurtz, A. K. (1941). Recent research in the selection of life insurance salesman. *Journal of Applied Psychology*, 25, 11-17.
- Laurent, H. (1970). Cross-cultural cross-validation of empirically validated tests. *Journal of Applied Psychology*, 54, 59-62.
- Mael, F. A. (1991). A conceptual rationale for the domain and attributes of biodata items. *Personnel Psychology*, 44, 763-792.
- McManus, M. A., & Mitchell, T. W. (1987). *Test and retest reliability of the Career Profile*. Hartford: Life Insurance Marketing and Research Association. Minneapolis: Personnel Decisions Research Institute.
- Mitchell, T. (1989). *Are weighted application data valid and legal?* Paper presented at the National Assessment Conference, Minneapolis.
- Mitchell, T. W. (1994). The utility of biodata. In G. S. Stokes, M. D. Mumford, & W. A. Owens (Eds.), *Biodata handbook*. Palo Alto: Cpp Books.
- Mitchell, T. W., & Klimoski, R. J. (1982). Is it rational to be empirical? A test of methods for scoring biographical data. *Journal of Applied Psychology*, 67, 411-418.
- Muchinsky, P. M. (1993). *Psychology applied to work*. (4th ed.). Brooks/Cole.
- Mumford, M. D., & Owens, W. A. (1987). Methodology review: Principles, procedures, and findings in the application of background data measure, *Applied Psychological Measurement*, 11(1), 1-31.
- Nunnally, J. C. (1978). *Psychometric theory*. New York: McGraw-Hill.
- Owens, W. A. & Schoenfeldt, L. F. (1979). Toward a classification of persons. *Journal of Applied Psychology*, 64, 569-607.
- Owens, W. A. (1971). A quasi-actuarial basis for individual assessment. *American Psychologist*, 26, 992-999.
- Owens, W. A. (1976). Background data. In M. D. Dunnette (Ed.), *Handbook of industrial psychology*. New York: Rand-McNally.
- Pace, L. A. & Schoenfeldt, L. F. (1977). Legal concerns in the use of weighted applications. *Personnel Psychology*, 30, 159-166.
- Pedhazur, E. J. (1982). *Multiple regression in behavioral research*. New York: Holt, Rinehart & Winston.
- Reilly, R. R., & Chao, G. T. (1982). Validity and fairness of some alternative employee selection procedures. *Personnel Psychology*, 35, 1-63.
- Rosenbaum, R. W. (1976). Predictability of employee theft using weighted application

- blands. *Journal of Applied Psychology*, 61, 94-98.
- Rothstein, H. R., Schmidt, F. L., Erwin, F. W., Owens, W. A., & Sparks, C. P. (1990). Biographical data in employment selection: Can validities be made generalizable? *Journal of Applied Psychology*, 75, 175-184.
- Rothstein, H. R., Schmidt, F. L., Erwin, F. W., Owens, W. A. & Sparks, C. P. (1990). Biographical data in employment selection: Can validities be made generalizable? *Journal of Applied Psychology*, 75, 175-184.
- Schmitt, N., Gooding, R. Z., Noe, R. A., & Kirsch, M. (1984). Metaanalyses of validity studies published between 1964 and 1982 and the investigation of study characteristics. *Personnel Psychology*, 37, 407-422.
- Schneider, B. (1987). The people make the place. *Personnel psychology*, 40 (3), 437-453.
- Smith, W. J., Albright, L. E., Glennon, J. R., & Owens, W. A. (1961) The prediction of research competence and creativity from personal history. *Journal of Applied Psychology*, 45, 59-62.
- Stead, W. H., & Shartle, C. L., & Associates. (1940). *Occupational counseling techniques*. New York: American Book Co.
- Strong, E. K. (1926). An interest test for personnel managers. *Journal of Personnel Research*, 5, 194-203.
- Thayer, P. W. (1977). Somethings old, somethings new. *Personnel Psychology*, 30, 513-524.
- Wernimount, P. F., & Campbell, J. P. (1968). Signs, samples, and criteria. *Journal of Applied Psychology*, 52, 372-376.

논문 초고 접수 : 1996. 9. 12.

최종수정본 접수 : 1996. 10. 28.

## ABSTRACT

### Validation study of empirical biodata

Hye-Ryun Kang

Ewha Womans University

The purpose of the present study was to examine the validity of empirically keyed biodata. 65 biographical items were administered to 5,922 insurance saleswomen. The criterion measure of job success was obtained from company records. The empirical keying method based upon the WAB was employed to score the biodata.

The initial result showed that scoring weights developed on a total sample were somewhat unstable. Thus, the sample was divided into old versus young subgroups and the scoring keys were developed separately. The result of cross-validation revealed scoring weights of both subgroups were stable as well as having high predictability. Also, the result of discriminant analysis showed practical significance of biodata as a selection tool. Finally, problems and future directions of the empirical keying method were discussed.